

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

MYRIAN RAQUEL NOGUERA SALINAS

**SHORT DATATHON PARA O DESENVOLVIMENTO
INTERDISCIPLINAR DE HABILIDADES DE ANÁLISE E
VISUALIZAÇÃO DE DADOS**

DISSERTAÇÃO DE MESTRADO

CURITIBA

2019

MYRIAN RAQUEL NOGUERA SALINAS

**SHORT DATATHON PARA O DESENVOLVIMENTO
INTERDISCIPLINAR DE HABILIDADES DE ANÁLISE E
VISUALIZAÇÃO DE DADOS**

DISSERTAÇÃO DE MESTRADO apresentada ao PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de “Mestre em Ciências” – Área de Concentração: Engenharia de Sistemas Computacionais.

Orientador: Prof. Dr. Adolfo Gustavo Serra Seca Neto

Co-orientador: Prof.^a Dra. Maria Claudia Figueiredo Pereira Emer

CURITIBA

2019

Dados Internacionais de Catalogação na Publicação

S165s Salinas, Myrian Raquel Noguera
Short Datathon para o desenvolvimento interdisciplinar de habilidades de análise e visualização de dados [recurso eletrônico] / Myrian Raquel Noguera Salinas.-- 2019.
1 arquivo texto (142 f.) : PDF ; 8,36 MB.

Modo de acesso: World Wide Web.

Texto em português com resumo em inglês.

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Computação Aplicada. Área de Concentração: Engenharia de Software, Curitiba, 2019.

Bibliografia: f. 87-91.

1. Computação - Dissertações. 2. Computação - Eventos. 3. Fraude - Detecção. 4. Sonegação fiscal. 5. Banco de dados - Gerência. 6. Visualização da informação - Análise. 7. Abordagem interdisciplinar do conhecimento. 8. Aprendizagem. 9. Grupos de trabalho. 10. Interfaces de usuário (Sistemas de computação). I. Seca Neto, Adolfo Gustavo Serra, orient. II. Emer, Maria Cláudia Figueiredo Pereira, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Computação Aplicada. IV. Título.

CDD: Ed. 23 -- 621.39

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 71

A Dissertação de Mestrado intitulada: Short Datathon para o desenvolvimento interdisciplinar de habilidades de análise e visualização de dados, defendida em sessão pública pela candidata Myrian Raquel Noguera Salinas, no dia 20 de fevereiro de 2019, foi julgada para a obtenção do título de Mestre em Computação Aplicada, área de concentração: Engenharia de Sistemas Computacionais, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Computação Aplicada.

BANCA EXAMINADORA:

Profª. Drª. Maria Cláudia Figueiredo Pereira Emer – Presidente – UTFPR

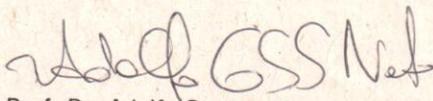
Prof. Dr. Andrey Ricardo Pimentel – UFPR

Prof. Dr. Sérgio Tadeu Gonçalves Muniz – UFPR

Prof. Dr. Gustavo Alberto Giménez Lugo – UTFPR

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 20 de fevereiro de 2019.



Prof. Dr. Adolfo Gustavo Serra S. Neto
Coordenador Adjunto do PPGCA
UTFPR - Câmpus Curitiba

Carimbo e Assinatura do(a) Coordenador(a) do Programa

Com muito carinho e afeto dedico este meu trabalho à minha família.

AGRADECIMENTOS

Inicio os meus agradecimentos com o meus orientadores, Professor Dr. Adolfo Neto e a Professora Maria Claudia Emer. A vocês meus sinceros agradecimentos por compartilhar seus conhecimentos e por me orientar e motivar nesta caminhada do conhecimento.

Agradeço a minha família, em especial meus pais, Marcos e Marcelina, e meu irmão, Marcos, pelo apoio e amor que me dedicaram.

Agradeço a Adriano Sieg, pela ajuda e compreensão, que me fortalece e impulsiona em busca dos meus objetivos. Obrigado por dividir seu tempo comigo e presentear-me com sua presença.

Meus agradecimentos aos colegas de mestrado e colegas da sala de estudos, dos quais posso citar os meus caros amigos Jean, Paulo, Marcelo, Leticia, Monica pelas parcerias durante o cumprimento das disciplinas, bem como pelas conversas nos momentos de distração. Aos professores do PPGCA da Universidade Tecnológica Federal do Paraná, que vêm abrindo portas para que seus alunos se tornem profissionais e pessoas melhores.

Agradeço à PAEC-OEA-GCUB pelo apoio financeiro, cedido por praticamente todo período como aluna do mestrado em computação aplicada.

Por fim, quero agradecer muitos outros amigos pela amizade que dura além de todos os momentos de concentração que me ajudaram a esfriar a cabeça.

RESUMO

NOGUERA SALINAS, Myrian Raquel. SHORT DATATHON PARA O DESENVOLVIMENTO INTERDISCIPLINAR DE HABILIDADES DE ANÁLISE E VISUALIZAÇÃO DE DADOS . 144 f. DISSERTAÇÃO DE MESTRADO – PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA , Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

Entender os principais problemas de fraude no mundo e interpretar os dados disponíveis para análise é um desafio atual que requer conhecimento interdisciplinar para complementar o conhecimento dos profissionais da computação. Os eventos colaborativos chamados frequentemente de *Hackathons*, *Datathons*, *Codefests*, *Hack-Days*, etc. tornaram-se relevantes em vários campos. Os exemplos incluem aprendizagem informal e colaborativa, criação de startups, artes e cultura, inovação cívica aberta, inovação corporativa, e questões sociais. Esses eventos mostram características interessantes que favorecem o ambiente de colaboração e troca de conhecimento para resolver desafios, na maioria das vezes relacionados a assuntos desconhecidos pelos estudantes e profissionais da computação. Neste trabalho, apresentamos um formato de evento chamado *Short Datathon* adaptado para atingir uma lacuna nos estudos sobre eventos colaborativos como auxílio ao desenvolvimento de habilidades de análise exploratória de dados. Avaliamos o *Short Datathon* em dois estudos de caso realizados na Universidade Tecnológica Federal do Paraná, nos quais abordamos a questão da evasão fiscal com dados reais. O primeiro estudo foi realizado durante a Semana Tecnológica com participação de 12 alunos de graduação das áreas da Computação e Mecânica. No segundo estudo participaram 8 pessoas formadas e alunos de pós-graduação nas áreas da Computação, Administração e Economia. Descrevemos como resultado deste trabalho os aspectos qualitativos dos estudos de caso e a percepção dos participantes através de questionários, destacando a contribuição do conhecimento adquirido no evento que foi favorecido pela interação interdisciplinar. Os participantes afirmaram que o evento ajudou a compreender mais sobre análise de dados e visualização e que a experiência com pessoas de outras áreas durante o evento tornou a análise de dados mais eficiente. Podemos concluir que os resultados do estudo do *Short Datathon* são ainda preliminares devido ao reduzido número de participantes, mas pelas respostas acreditamos que seja um evento promissor.

Palavras-chave: Datathon, Fraudes Internacionais, Análise de dados, Visualização de dados

ABSTRACT

NOGUERA SALINAS, Myrian Raquel. Short Datathon for the Interdisciplinary Development of Data Analysis and Visualization Skills. 144 f. DISSERTAÇÃO DE MESTRADO – PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

Understanding the main fraud problems in the world and interpreting the data available for analysis is a current challenge that requires interdisciplinary knowledge to complement the knowledge of computer professionals. Collaborative events often called *Hackathons*, *Datathons*, *Codefests*, *Hack-Days*, etc. have become relevant in various fields. Examples include informal and collaborative learning, creation of startups, arts and culture, open civic innovation, corporate innovation, and social issues. These events show interesting features that favor the collaborative environment and knowledge exchange to solve challenges, most often related to issues unknown to students and computer professionals. In this paper, we present an event format called *Short Datathon* adapted to reach a gap in collaborative event studies as an aid to the development of exploratory data analysis skills. We evaluated the Short Datathon in two case studies carried out at the Federal University of Technology - Paraná, in which we address the issue of tax evasion with real data. The first study was carried out during Technological Week with the participation of 12 undergraduate students from the Computing and Mechanics areas. In the second study, 8 college graduates and graduate students from the areas of Computing, Administration and Economics participated. We describe as a result of this work the qualitative aspects of the case studies and the perception of the participants obtained through questionnaires, highlighting the contribution of the knowledge acquired in the event that was favored by the interdisciplinary interaction. Participants stated that the event helped them to understand more about data analysis and visualization and that the experience with people from other areas during the event made data analysis more efficient. We can conclude that the results of the Short Datathon study are still preliminary because of the small number of participants, but by the answers we believe it to be a promising event.

Keywords: Datathon, International Fraud, Data Analysis, Data Visualization

LISTA DE FIGURAS

FIGURA 1	– Tipos de Fraude de dados. Fonte: (LENZ, 2015)	21
FIGURA 2	– Tipos de Crimes Financeiros. Fonte: (GOTTSCHALK, 2010)	23
FIGURA 3	– Amostra da Rede Financeira Internacional. Fonte: (VITALI et al., 2011) .	29
FIGURA 4	– Rede de Empresas Offshore. Fonte: (ZHUHADAR; CIAMPA, 2017)	31
FIGURA 5	– Blocos de Construção para o Paraíso Fiscal. Fonte: (JONES et al., 2018) .	33
FIGURA 6	– Quantidade de <i>Hackathons</i> realizados no mundo no ano 2016. Fonte: <i>Site Hackathon</i>	34
FIGURA 7	– Estrutura típica de um <i>Hackathon</i> . Fonte: (KOMSSI et al., 2015)	35
FIGURA 8	– Fases do Projeto de Pesquisa. Fonte: Autoria Própria	44
FIGURA 9	– Fases da realização do Questionário. Fonte: Autoria Própria	44
FIGURA 10	– Estrutura típica de um <i>Hackathon</i> . Fonte: (KOMSSI et al., 2015)	51
FIGURA 11	– Estrutura Geral de um <i>Short Datathon</i> . Fonte: Autoria Própria	51
FIGURA 12	– Processo de Planejamento do <i>Short Datathon</i> . Fonte: Autoria Própria	53
FIGURA 13	– Imagem dos participantes. Fonte: Autoria Própria	63
FIGURA 14	– Gráfico gerado pelo Equipe Sistemas de Computação. Fonte: Autoria Própria	65
FIGURA 15	– Gráfico gerado pelo Equipe CA. Fonte: Autoria Própria	66
FIGURA 16	– Gráfico gerado pelo Equipe CA. Fonte: Autoria Própria	66
FIGURA 17	– Atividade principal dos participantes. Fonte: Autoria Própria	67
FIGURA 18	– Participação em eventos <i>Hackathon e Datathon</i> . Fonte: Autoria Própria .	68
FIGURA 19	– Resultados das afirmações do questionário do Estudo de Caso Piloto. Fonte: Autoria Própria	70
FIGURA 20	– Gráfico gerado pelo Equipe SEGA Data. Fonte: Autoria Própria	74
FIGURA 21	– Gráfico gerado pelo Equipe SEGA Data. Fonte: Autoria Própria	75
FIGURA 22	– Gráfico gerado pelo Equipe Olhos de Lince. Fonte: Autoria Própria	76
FIGURA 23	– Atividade principal dos participantes. Fonte: Autoria Própria	77
FIGURA 24	– Participação em eventos <i>Hackathon e Datathon</i> . Fonte: Autoria Própria .	77
FIGURA 25	– Resultados das afirmações do questionário. Fonte: Autoria Própria	78
FIGURA 26	– Primeira página do questionário aplicado no " <i>From Ladies to Ladies</i> " ...	93
FIGURA 27	– Segunda página do questionário aplicado no " <i>From Ladies to Ladies</i> " ...	94
FIGURA 28	– Tipos de eventos de que você já participou. Fonte: Autoria Própria	95
FIGURA 29	– Fatores que poderiam limitar sua participação num evento qualquer. Fonte: Autoria Própria	95
FIGURA 30	– Questionário aplicado no GGJ 2017	96
FIGURA 31	– Questionário aplicado no GGJ 2017	97
FIGURA 32	– Questionário aplicado no GGJ 2017	98
FIGURA 33	– Questionário aplicado no GGJ 2017	99
FIGURA 34	– Questionário aplicado no GGJ 2017	100
FIGURA 35	– Questionário aplicado no GGJ 2017	101
FIGURA 36	– Participou em um (ou mais) dos eventos listados abaixo. Fonte: Autoria Própria	102
FIGURA 37	– Primeira página do questionário aplicado na Comunidade <i>Hackathon</i> Brasil	104

FIGURA 38	– Segunda página do questionário aplicado na Comunidade <i>Hackathon</i> Brasil	105
FIGURA 39	– Tipos de Hackathon que você já participou. Fonte: Autoria Própria	106
FIGURA 40	– Participaria de um Hackathon focado em temas como: Dívida pública, evasão fiscal, etc.?. Fonte: Autoria Própria	106
FIGURA 41	– Primeira página do questionário aplicado no II Fórum Gestão e Economia	108
FIGURA 42	– Segunda página do questionário aplicado no II Fórum Gestão e Economia	109
FIGURA 43	– Área de trabalho ou estudo. Fonte: Autoria Própria	110
FIGURA 44	– Conhece ou ouviu falar do evento Hackathon. Fonte: Autoria Própria	110
FIGURA 45	– Primeira página do questionário aplicado no Datathon UFRGS	112
FIGURA 46	– Segunda página do questionário aplicado no Datathon UFRGS	113
FIGURA 47	– Terceira página do questionário aplicado no Datathon UFRGS	114
FIGURA 48	– Frequência de participação em eventos <i>Hackathon</i> . Fonte: Autoria Própria	115
FIGURA 49	– Frequência de participação em eventos <i>Datathon</i> . Fonte: Autoria Própria	115
FIGURA 50	– Participantes do Hackathon Compras Públicas. Fonte: Site da Secretaria da Fazenda	119
FIGURA 51	– Techo do texto publicado sobre o evento. Fonte: Site da Secretaria da Fazenda	119
FIGURA 52	– Vencedores do <i>Datathon</i> UFRGS. Fonte: Site do evento	121
FIGURA 53	– Primeira página do questionário aplicado	124
FIGURA 54	– Segunda página do questionário aplicado	125
FIGURA 55	– Terceira página do questionário aplicado	126
FIGURA 56	– Área principal de trabalho ou estudo dos participantes. Fonte: Autoria Própria	127
FIGURA 57	– Principal atividade dos participantes. Fonte: Autoria Própria	127
FIGURA 58	– Faixa Etária dos participantes. Fonte: Autoria Própria	128
FIGURA 59	– Tipos de eventos de que você já participou. Fonte: Autoria Própria	128
FIGURA 60	– Genero dos participantes. Fonte: Autoria Própria	129
FIGURA 61	– Nível de participação em eventos Hackathon. Fonte: Autoria Própria	129
FIGURA 62	– Nível de participação em eventos Datathon. Fonte: Autoria Própria	130
FIGURA 63	– Área principal de trabalho ou estudo. Fonte: Autoria Própria	131
FIGURA 64	– Principal atividade dos participantes. Fonte: Autoria Própria	131
FIGURA 65	– Faixa Etária dos participantes. Fonte: Autoria Própria	132
FIGURA 66	– Tipos de eventos de que você já participou. Fonte: Autoria Própria	132
FIGURA 67	– Genero dos participantes. Fonte: Autoria Própria	132
FIGURA 68	– Nível de participação em eventos Hackathon. Fonte: Autoria Própria	133
FIGURA 69	– Nível de participação em eventos Datathon. Fonte: Autoria Própria	133
FIGURA 70	– Convite realizado no Sympla para o Estudo de Caso Piloto	136
FIGURA 71	– Convite realizado no Sympla para o segundo Estudo de Caso	137
FIGURA 72	– Convite realizado no Facebook para o Estudo de Caso Piloto	138
FIGURA 73	– Divulgação realizada no grupo Hackathons Brasil na rede social Facebook	139
FIGURA 74	– Criação da página Datathon de Análise de Dados de Evasao Fiscal na rede social Facebook	139
FIGURA 75	– Painel de gestão de alcance de publicações realizadas na página criada no Facebook	140
FIGURA 76	– Divulgação realizada no site Eventos Tech	141
FIGURA 77	– Divulgação realizada no site da Semana Technologica	142
FIGURA 78	– Imagem do cartaz impresso para divulgação nos painéis da UTFPR Curitiba	143
FIGURA 79	– Convite enviados a grupos de <i>Whatsapp</i>	144

LISTA DE TABELAS

TABELA 1	– Programação típica de um <i>Hackathon</i> . Fonte: (PERNG et al., 2017)	35
TABELA 2	– Artigos sobre eventos <i>Datathon</i> e os temas abordados. Fonte: Autoria Própria	39
TABELA 3	– Resumo de aspectos positivos e negativos dos Hackathons. Fonte: Autoria própria	41
TABELA 4	– Resumo dos questionários aplicados. Fonte: Autoria Própria	46
TABELA 5	– Comparativo entre Maratona de Programação de Compras Públicas e o <i>Datathon</i> da UFRGS. Fonte: Autoria Própria	48
TABELA 6	– Características diferentes no <i>Short Datathon</i> . Fonte: Autoria Própria	50
TABELA 7	– Aba com resultados processados. Fonte: (ALSTADSAETER et al., 2017)	54
TABELA 8	– Aba com dados não processados. Fonte: (ALSTADSAETER et al., 2017)	54
TABELA 9	– Critérios de avaliação do Software. Fonte: Autoria Própria	55
TABELA 10	– Cronograma de atividades do <i>Short Datathon</i> Piloto. Fonte: Autoria Própria	57
TABELA 11	– Seções do Questionário. Fonte: Autoria Própria	60
TABELA 12	– Resumo de Propostas Iniciais. Fonte: Autoria Própria	63
TABELA 13	– Tabela de Retrospectiva. Fonte: Autoria Própria	67
TABELA 14	– Cronograma de atividades do Segundo <i>Short Datathon</i> . Fonte: Autoria Própria	71
TABELA 15	– Resumo de Propostas Iniciais. Fonte: Autoria Própria	72
TABELA 16	– Resumo das características dos Estudos de Casos. Fonte: Autoria Própria	80
TABELA 17	– Avaliações das afirmações na escala de Likert dos Estudos de Caso. Fonte: Autoria Própria	80
TABELA 18	– Avaliações das afirmações na escala de Likert. Fonte: Autoria Própria	81
TABELA 19	– Resultados das Afirmações. Fonte: Autoria Própria	116
TABELA 20	– Tabela de valores Likert do Estudo de Caso Piloto. Fonte: Autoria Própria	130
TABELA 21	– Tabela de valores Likert do segundo Estudo de Caso. Fonte: Autoria Própria	134

LISTA DE SIGLAS

PPGCA	Programa de Pós-Graduação em Computação Aplicada
PAEC	Programa de Alianças para a Educação e a Capacitação
OEA	Organização dos Estados Americanos
GCUB	Grupo Coimbra de Universidades Brasileiras
RAE	Real Academia Espanhola
SIFI	Systemically Important Financial Institutions
EUA	Estados Unidos da América
EY	Ernst and Young
PwC	PricewaterhouseCoopers
UFRGS	Universidade Federal do Rio Grande do Sul
FIPE	Fundação Instituto de Pesquisa Econômica
AFRESP	Associação dos Agentes Fiscais de Rendas do Estado de São Paulo
BID	Banco Interamericano de Desenvolvimento
Sebrae	Serviço Brasileiro de Apoio às Micro e Pequenas Empresas
ONG	Organização Não-Governamental
PIB	Produto Interno Bruto
IPC	Índice de Percepção de Corrupção
Pnad	Pesquisa Nacional por Amostra de Domicílios
IBGE	Instituto Brasileiro de Geografia e Estatística
FGV	Fundação Getúlio Vargas
IBPT	Instituto Brasileiro de Planejamento e Tributação
BEC	Bolsa Eletrônica de Compras

SUMÁRIO

1 INTRODUÇÃO	15
1.1 MOTIVAÇÃO E JUSTIFICATIVA	16
1.2 OBJETIVOS	17
1.2.1 Objetivo Geral	17
1.2.2 Objetivos específicos	17
1.3 ESTRUTURA DO TRABALHO	18
2 FUNDAMENTAÇÃO TEÓRICA	19
2.1 FRAUDE	19
2.1.1 Tipos de Fraude	20
2.1.2 Tipos de Crimes Financeiros	22
2.1.3 Principais Fatos que afetam os Governos	27
2.2 ESTRUTURA INTERNACIONAL DA EVASÃO FISCAL	28
2.2.1 Gigantes Financeiros	28
2.2.2 Paraísos fiscais ou Jurisdições <i>Offshore</i>	30
2.2.3 Os <i>Big 4</i>	31
2.3 <i>HACKATHON</i> E EVENTOS DE ATIVISMO COM FOCO EM FRAUDES	32
2.3.1 <i>Hackathon</i> e <i>Datathon</i>	33
2.3.2 Participação Pública e <i>Citizensourcing</i>	35
2.4 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO	37
3 TRABALHOS RELACIONADOS	38
3.1 UM PANORAMA SOBRE <i>HACKATHONS</i> E <i>DATATHONS</i>	38
3.2 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO	41
4 MÉTODO	43
4.1 FASES DO PROJETO DE PESQUISA	43
4.1.1 Levantamento Inicial com Questionários	43
4.1.1.1 Identificar participantes	45
4.1.1.2 Elaboração do questionário	45
4.1.1.3 Gerenciar aplicação do questionário	45
4.1.1.4 Analisar resultados	46
4.1.2 Exploração de Eventos	47
4.1.3 Desenvolvimento do formato <i>Short Datathon</i>	49
4.1.4 Planejamento do Evento <i>Short Datathon</i>	52
4.1.4.1 <i>Dataset</i>	52
4.1.4.2 <i>Software</i> de Visualização de Dados	54
4.2 ESTUDO DE CASO	56
4.2.1 Estudo de Caso Piloto	58
4.2.2 Coleta de Evidências	58
4.2.3 Análise de Evidências	59
4.3 QUESTIONÁRIO <i>SHORT DATATHON</i>	59
4.4 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO	60
5 RESULTADOS E DISCUSSÃO	62

5.1	RELATO DO ESTUDO DE CASO PILOTO	62
5.1.1	Apresentação dos trabalhos do Estudo de Caso Piloto	63
5.1.2	Retrospectiva	66
5.1.3	Resultado de Questionário Aplicado	67
5.1.4	Considerações finais do Piloto do Datathon	68
5.1.5	Considerações Iniciais para o Segundo Estudo de Caso	69
5.2	RELATO DO SEGUNDO ESTUDO DE CASO	71
5.2.1	Apresentação dos trabalhos	73
5.2.2	Retrospectiva	75
5.2.3	Resultado de Questionário Aplicado	76
5.2.4	Considerações finais do <i>Short Datathon</i>	77
5.3	RESULTADOS GERAIS	79
5.4	DISCUSSÃO	81
5.5	CONSIDERAÇÕES SOBRE ESTE CAPÍTULO	83
6	CONCLUSÃO	84
6.1	CONTRIBUIÇÕES	86
6.2	TRABALHOS FUTUROS	86
	REFERÊNCIAS	87
	Apêndice A – QUESTIONÁRIOS DE LEVANTAMENTO INICIAL	92
A.1	QUESTIONÁRIO APLICADO NO “ <i>FROM LADIES TO LADIES</i> ”	92
A.1.1	Resultado do Questionário aplicado no “ <i>From Ladies to Ladies</i> ”	95
A.2	QUESTIONÁRIO APLICADO NO <i>GLOBAL GAME JAM 2017</i>	96
A.3	QUESTIONÁRIO APLICADO NA COMUNIDADE <i>HACKATHON BRASIL</i>	103
A.3.1	Resultado do Questionário aplicado na Comunidade <i>Hackathon Brasil</i>	106
A.4	QUESTIONÁRIO APLICADO NO II FÓRUM GESTÃO E ECONOMIA DA UTFPR 107	
A.4.1	Resultado do Questionário aplicado no II Fórum Gestão e Economia da UTFPR ...	110
A.5	QUESTIONÁRIO APLICADO NO DATATHON UFRGS	111
A.5.1	Resultado do Questionário aplicado no Datathon UFRGS	115
A.5.2	Considerações sobre o Levantamento com Questionário	116
	Apêndice B – EXPLORAÇÃO DE EVENTOS <i>HACKATHON</i> E <i>DATATHON</i>	118
B.1	RELATÓRIO DO <i>HACKATHON</i> COMPRAS PÚBLICAS SP	118
B.2	RELATÓRIO DO 1º <i>DATATHON</i> UFRGS	120
B.2.1	Artigo publicado pela imprensa da UFRGS sobre o 1º <i>Datathon</i> UFRGS	122
	Apêndice C – QUESTIONÁRIOS APLICADOS NOS ESTUDOS DE CASO DO <i>SHORT</i> <i>DATATHON</i>	123
C.1	QUESTIONÁRIO APLICADO NO ESTUDO DE CASO PILOTO E NO SEGUNDO ESTUDO DE CASO	123
C.2	RESULTADOS DO QUESTIONÁRIO APLICADO NO <i>SHORT DATATHON</i> PILOTO 127	
C.3	RESULTADOS DO QUESTIONÁRIO APLICADO NO SEGUNDO ESTUDO DE CASO DO <i>SHORT DATATHON</i>	131
	Apêndice D – DIVULGAÇÃO DOS <i>SHORT DATATHON</i>	135
D.1	CONVITE NA PLATAFORMA DE EVENTOS SYMPLA	136
D.2	DIVULGAÇÃO EM REDES SOCIAIS	138
D.3	OUTROS TIPOS DE DIVULGAÇÃO	141

D.3.1 <i>Site</i> do Eventos Tech	141
D.3.2 <i>Site</i> da Semana Technologica	142
D.3.3 Cartazes em Painéis da UTFPR Curitiba	143
D.3.4 Convite em Grupos de <i>Whatsapp</i>	144

1 INTRODUÇÃO

A fraude é tão antiga quanto a própria humanidade e pode ter uma variedade ilimitada de formas diferentes e pode ser identificada em muitas áreas no mundo, incluindo a fraude financeira, fraude artística, fraude arqueológica e a fraude científica (LENZ, 2015). Os avanços na detecção e prevenção não ficam atrás. Numerosas ferramentas e técnicas foram desenvolvidas e aplicadas nos diferentes cenários como ajuda na redução e controle dos fraudes.

Numerosos fatos que podem ser caracterizados como fraude, desde fraudes com cartões de crédito até fatos mais complexos, como evasão fiscal. A evasão fiscal gera um problema significativo para os governos e priva de recursos públicos para investir em políticas sociais. As empresas multinacionais e gigantes financeiros na busca para maximizar seus lucros e poder expandir-se internacionalmente para ampliar seus mercados, geralmente implementam estratégias com o objetivo de reduzir sua carga fiscal, desenvolver atividades que favorecem seu crescimento e gerar o máximo de lucro (DOWBOR, 2017). Somente no ano de 2016 o governo brasileiro estimou suas perdas fiscais em 571,5 bilhões de reais segundo estudo publicado¹.

A evasão fiscal é fundamentalmente difícil de estudar porque não existe uma única fonte de dados que capture tudo e é necessário de conhecimento sobre conceitos de manipulação de dados e também conceitos relacionados as áreas de economia, direito internacional, estatística e outras especialidades. Esse processo inicia no momento de realizar pesquisas, procurar os dados de fontes oficiais, a interpretação deles, a análise, o pré-processamento até a visualização representativa da informação. Uma série de atividades complexas que requer conhecimento interdisciplinar.

Na atualidade, grandes quantidades de dados estão se tornando cada vez mais disponíveis por meio de repositórios de dados abertos, bem como empresas e governos que coletam dados para melhorar a tomada de decisões e a eficiência. Consequentemente, há uma necessidade de aumentar a capacidade de análise de dados. A ciência de dados é uma área relativamente nova está evoluindo rapidamente, junto com as ferramentas necessárias para executar

¹<http://archive.is/L5Yjb>. Último acesso: 10/03/2018

a análise. Para atender às necessidades dos estudantes e profissionais que aprendem habilidades de análise de dados, propomos um formato curto de *Hackathon* que chamamos de *Short Datathon*.

Os *Hackathons* são eventos geralmente relacionados à computação do tipo focados no problema que levam a novos protótipos de ideias em um curto período de tempo (BRISCOE; MULLIGAN, 2014), mas o importante deste tipo de evento é que não é necessário que todos os integrantes da equipe sejam programadores e certamente não é necessário ser da área de computação, o que favorece a interação interdisciplinar. *Hackathons* variam em objetivos e temas, mas são frequentemente semelhantes estruturalmente e em suas características (KOMSSI et al., 2015). Apesar de terem se originado em comunidades de tecnologia, os *Hackathons* já foram amplamente adotados por diversas organizações.

1.1 MOTIVAÇÃO E JUSTIFICATIVA

Na busca por estudos de detecção de fraudes que enfocam fatos que afetam níveis mais altos de prejuízo em governos como a evasão fiscal e o sistema de dívida pública, foram encontradas pesquisas interessantes realizadas a partir dos dados do vazamento dos *Panamá Papers* no início de 2016 (ALSTADSAETER et al., 2017) (ZHUHADAR; CIAMPA, 2017). Esses arquivos, totalizando 2,6 *terabytes*, continham mais de 11,5 milhões de registros financeiros e legais listados pelo prestador de serviços corporativo panamenho Mossack Fonseca. Informações sobre aproximadamente 320.000 empresas *offshore* foram expostos junto com os nomes de 220.000 empresas. Estas empresas *offshore* foram supostamente *shell corporations* ou empresas fantasmas que foram projetadas para obscurecer a evasão fiscal ilegal e sigilosa de indivíduos e corporações.

Os documentos foram distribuídos e analisados por cerca de quatrocentos jornalistas em 107 órgãos de comunicação social em mais de oitenta países, além uma equipe de profissionais da computação. As primeiras notícias sobre o caso, assim como 149 documentos, foram publicadas em 3 de abril de 2016 (ZHUHADAR; CIAMPA, 2017).

Para estimar o tamanho e a distribuição da evasão fiscal em três países nórdicos, os autores de outro estudo combinaram dados de auditorias aleatórias, anistias fiscais, o dados do vazamento dos *Panamá Papers* e do Banco HSBC da Suíça com os registros de riqueza (ALSTADSAETER et al., 2017). Normalmente este tipo de análise de dados precisa de conhecimentos específicos e o engajamento de pessoas para criar um espaço de intercâmbio dinâmico e colaborativo com o objetivo de analisar um problema. Essas características estão presentes

num evento *Hackathon*.

Recentemente, um tipo de evento tomou protagonismo pelas suas características. O *Hackathon* palavra composta por "*hack*" (programar de forma excepcional) e "*marathon*" (maratona) são eventos geralmente relacionados à computação do tipo focados no problema que levam a novos protótipos de ideias em um curto período de tempo (BRISCOE; MULLIGAN, 2014). Uma versão interessante deste evento é o *Datathon* refere-se à aplicação do modelo *Hackathon* à análise de dados (ANSLOW et al., 2016)

Esses eventos evoluíram de maneira cada vez mais exclusiva na colaboração de curto prazo e despertaram interesse dos pesquisadores em explorar outros aspectos, como intercâmbios, colaborações, vínculos comunitários, aprendizado e aprimoramento de habilidades. Identificamos uma lacuna nos estudos desses eventos colaborativos como auxílio no desenvolvimento de habilidades de análise e visualização de dados nos estudantes e profissionais da computação.

Existem também algumas desvantagens em relação aos eventos *Hackathon* identificados na literatura como: a frustração dos não-desenvolvedores, a questão do não-aprendizado devido à competitividade e o vies gênero. Além disso, outras causas mencionadas em *blogs* e portais de notícias como: o alto custo de investimento, a necessidade de muitas pessoas para organizar o evento e o fator de fadiga e stress por causa da duração do evento.

Neste trabalho será realizado uma análise qualitativa por meio de um estudo de caso e uma análise quantitativa através de questionários aplicados aos participantes.

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

O objetivo deste estudo é criar e avaliar o formato de evento *Short Datathon* para exposição a conceitos de manipulação de dados e desenvolvimento de habilidades de análise e visualização de dados.

1.2.2 OBJETIVOS ESPECÍFICOS

- Avaliar a percepção de compreensão de conceitos de manipulação de dados, além da análise e visualização de dados de alunos e de profissionais da computação sobre fraudes internacionais por meio do *Short Datathon*;
- Avaliar, conforme a opinião dos participantes, se o *Short Datathon* facilitou a compre-

ensão sobre análise e visualização de dados;

- Avaliar, conforme a opinião dos participantes, se o *Short Datathon* incentiva o *Citizen-sourcing*.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está dividido em seis capítulos, organizados da seguinte forma: Introdução, neste Capítulo 1, com uma breve contextualização do tema, a motivação, justificativa e objetivos.

A fundamentação teórica (Capítulo 2) apresentamos as principais definições de fraude e os tipos de fraudes. Também descrevemos a estrutura internacional de uma das fraudes mais importantes: a Evasão Fiscal. Mencionamos os atores e o papel dos Paraísos Fiscais.

Além disso, descrevemos os eventos de ativismo e *Hackathons* como atividades de participação pública e *citizensourcing*. Os *Hackathons* oferecem oportunidades de desenvolvimento de habilidades técnicas e também outras como o trabalho em equipe, comunicação, colaboração interdisciplinar e outros. Pontos favoráveis para o nosso estudo.

No Capítulo 3 apresentamos o Estado da Arte, evidenciando trabalhos relacionados aos eventos *Hackathons* e *Datathons*.

No Capítulo 4, está descrito o planejamento e metodologia da proposta detalhadamente neste trabalho menciona as fases do projeto de pesquisa, a descrição da exploração de eventos *Hackathons* e *Datathons* e o planejamento do estudo de caso.

Os relatos dos Estudos de Caso e os resultados descrevemos no Capítulo 5.

Por fim, o Capítulo 6 conclui o documento apresentando as conclusões gerais deste trabalho e propõe possibilidades de diferentes abordagens para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Apresentamos os principais conceitos necessários, além de uma contextualização para melhor compreensão deste trabalho. Neste capítulo abordamos os diferentes conceitos de fraude e a sua classificação segundo vários autores. Além disso, descrevemos as estruturas que afetam os governos como os Bancos, Empresas multinacionais e Organizações Internacionais e os tipos de influência que exercem.

Por fim, descrevemos os tipos de atividades ou eventos que são realizados com o objetivos de promover o ativismo com foco em fraudes como *Hackathon e Datathon*.

2.1 FRAUDE

Algumas definições da palavra fraude encontradas nos principais dicionários de diversas línguas mencionam aspectos básicos; outras abrangem diversas atividades consideradas fraude. O Dicionário Oxford define a fraude como "Engano culpado ou criminoso destinado a resultar em ganho financeiro ou pessoal."¹ No dicionário online da RAE² a fraude é considerada uma "ação contrária à verdade e à justiça, que prejudica a pessoa contra quem é cometida. Ato tendente a contornar uma disposição legal em detrimento do Estado ou de terceiros. Crime cometido pela pessoa encarregada de supervisionar a execução de contratos públicos, ou de alguns privados, confabulando com a representação de interesses opostos"³.

Outro dicionário consultado foi a Larousse da língua francesa com a seguinte definição: "Fraude é o ato desonesto feito intencionalmente ao violar a lei ou regulamentos"⁴. No di-

¹Fraud: Wrongful or criminal deception intended to result in financial or personal gain. Oxford Dictionary <https://en.oxforddictionaries.com/definition/fraud>. Último acesso: 10/03/2018. Tradução para o português realizada pela autora.

²Real Academia Espanhola. Último acesso: 29/04/2018

³Fraude: 1. m. Acción contraria a la verdad y a la rectitud, que perjudica a la persona contra quien se comete. 2. m. Acto tendente a eludir una disposición legal en perjuicio del Estado o de terceros. 3. m. Der. Delito que comete el encargado de vigilar la ejecución de contratos públicos, o de algunos privados, confabulándose con la representación de los intereses opuestos. <http://www.rae.es/>. Último acesso: 29/04/2018. Tradução para o português realizada pela autora.

⁴Fraude: Acte malhonnête fait dans L'intention en contrevenant à la loi aux règlements.

cionário italiano Treccani a definição de fraude é mais abrangente ”1. Ato ou comportamento dirigido para danificar por engano um direito de outros: roubar, obter, subtrair com fraude; cometer fraude. em detrimento de alguém (ou mesmo uma instituição ou Estado); descobrir, relatar fraude. Em direito privado, queremos dizer fraude por qualquer comportamento injusto em prejuízo dos outros, mais sério do que má-fé. No direito penal, o termo fraude indica uma série de comportamentos caracterizados por modalidades evasivas previstas como crime pelo código ou por leis especiais; fraude no comércio, crime contra indústria e comércio cometido por aqueles que, no exercício de uma atividade comercial, ou uma loja pública, entregam ao comprador uma coisa por outra, ou que tem origem, fonte, qualidade ou quantidade diferente daquela declarada ou acordada; fraude imposto, comportamento realizado por meios artificiais (por exemplo, alegações de documentos falsificados, emissão ou utilização de faturas relacionadas com transações inexistentes, não faturação, etc.) e destinado a evasão fiscal; em particular, fraude à lei, violação de uma regra imperativa da lei, não explicitamente em sua formulação literal, mas implícita e indiretamente, combatendo assim o conteúdo real da lei”⁵.

No dicionário Priberam da língua portuguesa define que a fraude é ”1. Más artes que causam dolo. 2. Má-fé. 3. Candonga. 4. Contrabando”⁶. Em Direito Penal Brasileiro, fraude é obter, para si ou para outrem, vantagem ilícita, em prejuízo alheio, induzindo ou mantendo alguém em erro, mediante artifício, ardil, ou qualquer outro meio fraudulento⁷.

2.1.1 TIPOS DE FRAUDE

Pode-se revelar melhor a complexidade da fraude de dados ao classificá-la em quatro tipos. Todos eles são impulsionados por três características invariantes do tempo das sociedades,

<http://www.larousse.fr/dictionnaires/francais/fraude/35120>. Último acesso: 29/04/2018. Tradução para o português realizada pela autora.

⁵Frode: fròde (ant. fròda) s. f. [lat. *fraus fraudis*]. 1. Atto o comportamento diretto a ledere con l’inganno un diritto altrui: carpire, ottenere, sottrarre con la f.; commettere f. in danno di qualcuno (o anche di un ente, di un’istituzione, dello stato); scoprire, denunciare una frode. In diritto privato, s’intende per frode ogni comportamento sleale in pregiudizio di altri, più grave della mala fede. In diritto penale, il termine frode indica una serie di condotte caratterizzate da modalità elusive previste come reato dal codice o dalle leggi speciali; frode in commercio, delitto contro l’industria e il commercio commesso da chi, nell’esercizio di un’attività commerciale, o di un pubblico spaccio, consegna all’acquirente una cosa per un’altra, o che abbia comunque origine, provenienza, qualità o quantità diversa da quella dichiarata o pattuita (se la cosa è un alimento, si ha la f. alimentare); f. fiscale, comportamento realizzato con modalità artificiose (per es., allegazione di documenti contraffatti, emissione o utilizzazione di fatture relative a operazioni inesistenti, omesse fatturazioni, ecc.) e finalizzato all’evasione delle imposte; in partic., frode alla legge, violazione di una norma imperativa della legge, non esplicitamente nella sua formulazione letterale, ma implicitamente e indirettamente, contrastando cioè l’effettivo contenuto della norma. <http://www.treccani.it/vocabolario/frode/>. Último acesso: 29/04/2018. Tradução para o português realizada pela autora.

⁶<https://www.priberam.pt/dlpo/>. Último acesso: 29/04/2018.

⁷<https://www.jusbrasil.com.br/topicos/10617301/artigo-171-do-decreto-lei-n-2848-de-07-de-dezembro-de-1940>. Último acesso: 10/03/2018.

dos negócios e dos seres humanos: poder, glória e dinheiro (LENZ, 2015). Uma visão geral sobre os quatro tipos de fraude de dados é apresentada na Figura 1.

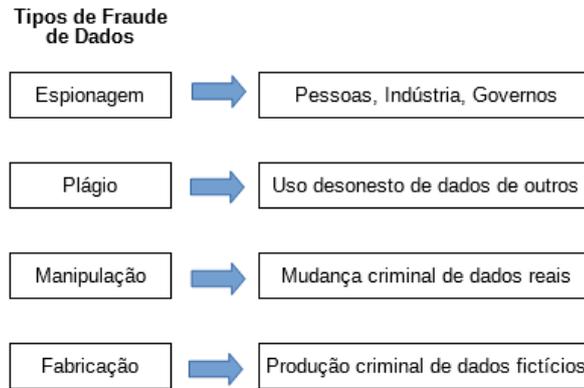


Figura 1: Tipos de Fraude de dados. Fonte: (LENZ, 2015)

- **Espionagem:** De uma perspectiva histórica, o espião está relacionado a populações antigas e modernas e suas ações militares contra outros países. A espionagem é muitas vezes parte de um esforço institucional por parte de um governo ou uma preocupação comercial. No entanto, o termo tende a associar-se com inimigos potenciais ou reais para fins militares. Na vida diária dos consumidores também temos espiões que consideram o comércio como seu domínio para a fraude de dados por exemplo a fraude do cartão que consiste no roubo de cartão ou o mau uso da conta bancária ou do número de identificação pessoal (PIN). Portanto, temos a seguinte subclassificação dentro da espionagem (LENZ, 2015): Roubo, Duplicação de geração, *Skimming* e *Phishing*.
- **Plágio:** O plágio é considerado o uso desonesto de dados de uma segunda parte. Tem muitas facetas e existe em muitos campos da vida humana, como artes, negócios e ciência. Alguns dos casos mais conhecidos são Pirataria de documentos (“Falsificação de Galileo”) e a dissertação de doutorado do ex-ministro alemão da defesa, K.-T. von Guttenberg. Ambos os casos ocorreram na Alemanha, e a falsificação foi comprovada em 2012 e 2011, respectivamente (LENZ, 2015).
- **Manipulação:** Manipulação de dados é a mudança desonesta do conteúdo de dados próprios ou de terceiros existentes, independentemente de o conteúdo ser encapsulado em documentos de texto ou não, ou seja, tabelas, diagramas ou fotos (LENZ, 2015).
- **Fabricação:** Inventar dados e resultados de pesquisa e depois gravá-los ou publicá-los (LENZ, 2015).

2.1.2 TIPOS DE CRIMES FINANCEIROS

Outra classificação (GOTTSCHALK, 2010) e (GOTTSCHALK; SOLLI-SÆTHER, 2011) foi proposta que abrange fatos relacionados a crimes financeiros. O crime financeiro geralmente envolve fraude e é realizado por meio de cheque, fraude de cartão de crédito, fraude hipotecária, fraude médica, fraude corporativa, fraude em conta bancária, fraude de ponto de venda, fraude monetária e fraude de assistência médica, e envolve atos como informações privilegiadas, violações tributárias, propinas, peculato, roubo de identidade, ataques cibernéticos, lavagem de dinheiro e engenharia social. Tipos de crimes financeiros podem variar tanto quanto as organizações criminosas e os criminosos envolvidos. A oportunidade surge quando uma fraqueza num procedimento foi descoberta.

Várias atividades ilegais podem ocorrer nos setores comercial e público. Enquanto houver pontos fracos que possam ser explorados para ganho, empresas e outras organizações, bem como indivíduos privados, serão aproveitados (PICKETT, 2002). No estudo (GOTTSCHALK, 2010) sugere uma estrutura dos principais tipos e categorias, como amostra na Figura 2.

- **Fraude do consumidor:** Estas são tentativas de coagir por meio do engano os consumidores a pagar por mercadorias não recebidas ou mercadorias que são abaixo do padrão, ou a preços ou taxas inflacionados. Também é um termo usado no sentido oposto, em que o consumidor é fraudulento. Um exemplo frequente deste tipo de fraude é aquela definida como um engano deliberado contra uma companhia de seguros com o objetivo de ganho financeiro (GOTTSCHALK; SOLLI-SÆTHER, 2011), (ABDALLAH et al., 2016) e (AHMED et al., 2016)
- **Fraude de cartão de crédito:** O cartão de crédito é muito popular e desempenha um papel importante no comércio eletrônico e na área de transações monetárias *online*. Como resultado do crescente uso do cartão de crédito, os fraudadores tentam encontrar mais oportunidades para cometer fraudes que podem causar grandes perdas para os titulares de cartões e bancos. Alguns exemplos são mencionados em estudos recentes (ABDALLAH et al., 2016), (KOU et al., 2004) e (AHMED et al., 2016).

A fraude de cartão de crédito é caracterizada por dois tipos: fraude *offline* e fraude *online*. A fraude *offline* é cometida usando um cartão físico roubado. Na maioria dos casos, a instituição que emite o cartão pode bloqueá-lo antes de ser usado de forma fraudulenta. A fraude *online* é cometida via *web*, compras por telefone. Apenas os detalhes do cartão são necessários no momento da compra (KOU et al., 2004).

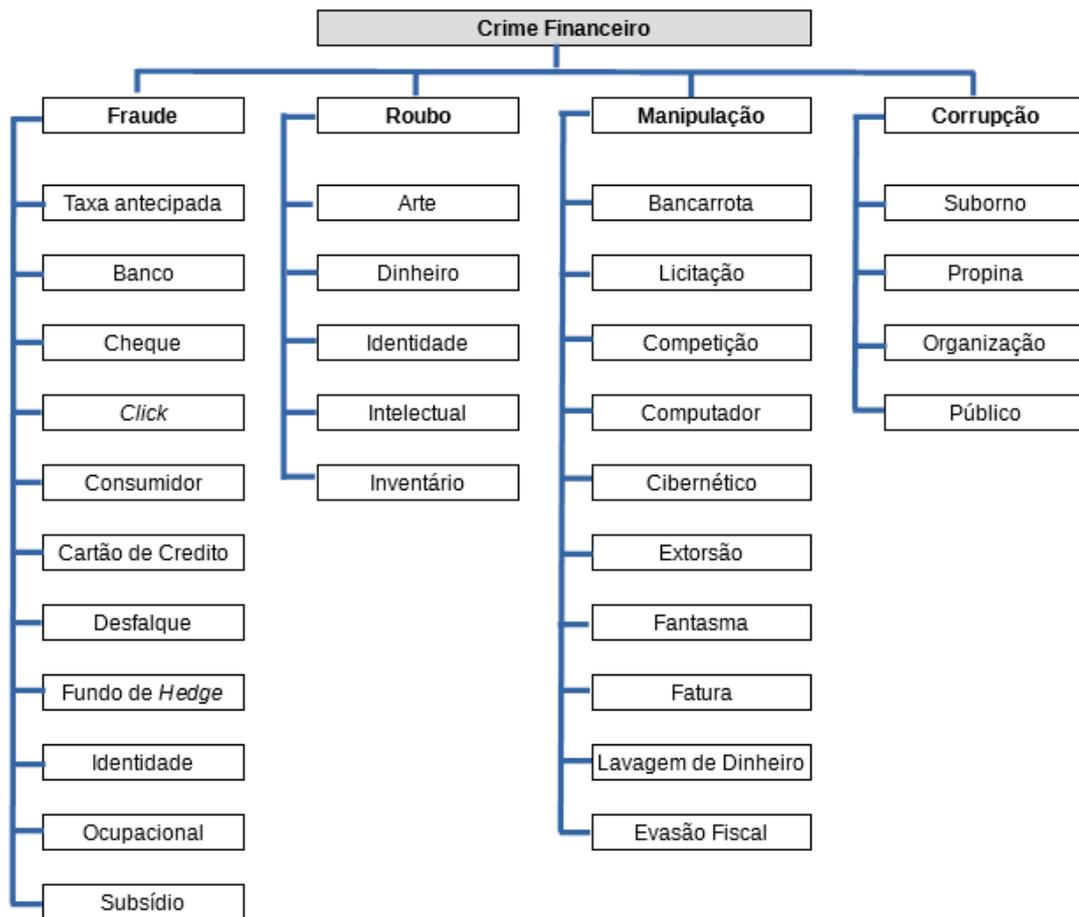


Figura 2: Tipos de Crimes Financeiros. Fonte: (GOTTSCHALK, 2010)

- **Fraudes externas:** Por exemplo, fraudadores de taxa antecipada tentam garantir uma comissão pré-paga para um acordo que nunca é cumprido ou um trabalho que nunca é feito (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Click:** Ocorre quando um indivíduo ou um *software* clica de forma fraudulenta e reiterada em um anúncio *online* sem a intenção real de conhecer mais sobre o produto ou serviço publicado ou fazer uma compra (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Fraude de Cheques:** Quando um cheque da empresa é roubado, alterado ou falsificado, ele pode ser desviado para uma pessoa não autorizada que acessa os fundos e depois fecha a conta ou simplesmente desaparece (PICKETT, 2002).
- **Fraude de Identidade:** Fraudes feitos por meio de roubos de identidade, usando dados pessoais, como número do seguro social, endereço e data de nascimento (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Desfalque:** Apropriação fraudulenta para benefício pessoal de propriedade ou dinheiro

confiado por outro. O ator primeiro toma posse da propriedade com a permissão do proprietário (GOTTSCHALK; SOLLI-SÆTHER, 2011).

- **Fraude Hipotecária:** Forma específica de fraude financeira que se refere à manipulação de documentos de propriedade ou hipoteca. Este fato visa deturpar o valor de uma propriedade com a finalidade de influenciar um credor para financiar um empréstimo (WEST; BHATTACHARYA, 2016).
- **Fraude Bancária:** A fraude bancária é definida como quem executa um esquema ou artifício para defraudar uma instituição financeira, ou para obter qualquer um dos fundos, créditos, ativos, valores mobiliários ou outros bens de propriedade ou sob custódia ou controle de uma instituição financeira, por meio de falsos ou fraudulentos pretextos, representações ou promessas (NGAI et al., 2011).
- **Fraude Ocupacional:** Uso de sua ocupação para enriquecimento pessoal por meio do uso indevido deliberado ou da aplicação incorreta dos recursos ou ativos da organização empregadora (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Taxa Antecipada:** As vítimas são abordadas por carta, fax ou *e-mail* sem contato prévio. Os endereços das vítimas são obtidos de diretórios de telefone e *e-mail*, revistas de negócios, revistas e jornais. Descreve a necessidade de retirar fundos de algum país, geralmente a recuperação de fundos contratuais, embarques de petróleo bruto ou herança de reis ou governadores (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Fraude de Subsídio:** Refere-se a crimes cometidos quando os subsídios do governo são concedidos. Uma pessoa ou uma empresa pode fornecer informações incorretas ao solicitar subsídios do governo ou usar os subsídios contrários às intenções e acordos (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Fraude de Fundo *Hedge*:** O fundo de *Hedge* é definido como um investimento comum que é organizado e administrado por uma empresa de gestão profissional e não amplamente disponível ao público (GOTTSCHALK; SOLLI-SÆTHER, 2011). Os gestores de fundos muitas vezes investem uma quantia considerável de sua própria riqueza nos fundos que administram. Eles tendem a se recusar a discutir suas estratégias de negociação porque não quer que os concorrentes imitem seus movimentos.
- **Roubo de Dinheiro:** Quando alguém a quem foi confiado dinheiro o desvia para uso pessoal. Outro mecanismo é o *skimming*, quando o dinheiro é retirado antes de entrar nos livros. A lapidação é uma técnica pela qual o roubo de dinheiro ou cheques é encoberto

pelo uso de recibos posteriores, de modo que a lacuna nos fundos não seja percebida (PICKETT, 2002).

- **Crime de Propriedade Intelectual:** O crime de propriedade intelectual é uma séria preocupação financeira para fabricantes de automóveis, fabricantes de bens de luxo, empresas de mídia e empresas farmacêuticas. O crescente valor da propriedade intelectual na produção de riqueza tem sido espelhado por sua crescente vulnerabilidade ao crime (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Roubo de Identidade:** Este é um crime em que um impostor obtém informações importantes, tais como números de identificação da previdência social, números de carteira de habilitação ou cartão de crédito para se passar por outra pessoa (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Roubo de Arte:** Trata-se de crime de arte envolvendo roubo por assalto, fraude (engano, imitação, falsificação e atribuição falsa) e pode envolver lavagem de dinheiro (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Roubo de inventário:** Apropriação fraudulenta para benefício pessoal dos bens de propriedade de uma empresa (PICKETT, 2002).
- **Oferta de Licitação:** Quando um fornecedor recebe uma vantagem injusta para derrotar uma competição aberta para um determinado contrato, um fornecedor pode receber informações extras para fazer lances baixos, mas depois arrecadar mais receita por meio de muitas variações do contrato (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Faturas Infladas:** Uma empresa inflaciona suas contas sem o consentimento do pagador, que pode ser um cliente (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Funcionários Fantasmas:** Gerar nomes extras na folha de pagamento de uma empresa e desviando os fundos para uma conta bancária especificamente configurada para esse golpe (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Crimes relacionados com computadores:** Alguns exemplos deste tipo de crimes incluem sabotagem, pirataria de software e roubo de dados pessoais (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Moeda falsa e Lavagem de Dinheiro:** A falsificação de moeda e a lavagem de dinheiro têm o potencial de desestabilizar as economias nacionais e ameaçar a segurança global, pois essas atividades são às vezes usadas por terroristas e outros criminosos perigosos

para financiar suas atividades ou ocultar seus lucros (GOTTSCHALK; SOLLI-SÆTHER, 2011).

Segundo (WEST; BHATTACHARYA, 2016) a lavagem de dinheiro é um método utilizado para inserir receitas obtidas de empreendimentos ilícitos em negócios válidos.

- **Evasão Fiscal:** Definida como a tentativa intencional de contornar a lei fiscal, a fim de reduzir ilegalmente a responsabilidade tributária de uma pessoa (GOTTSCHALK; SOLLI-SÆTHER, 2011). A forma legal é chamada de anulação fiscal, enquanto ilegal é conhecida como evasão de impostos (ABRANTES; FERRAZ, 2016). A probabilidade de praticar a evasão fiscal está diretamente ligada à probabilidade de sua detecção, razão pela qual o governo está constantemente procurando maneiras mais eficientes e efetivas de enfrentar esse problema.
- **Crime Cibernético:** Os ataques à infraestrutura de segurança cibernética de organizações comerciais podem ter vários objetivos. Um objetivo perseguido pelos criminosos é obter acesso não autorizado às informações confidenciais (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Bancarrota:** Trata-se de atos criminosos cometidos em conexão com processos de bancarrota ou liquidação. Uma pessoa que pede bancarrota ou uma empresa que entrou em liquidação pode ocultar ativos após o início do processo, evitando assim que os credores recebam suas reivindicações (GOTTSCHALK; SOLLI-SÆTHER, 2011).
- **Crime de Competição:** Colabora e influencia os preços, os lucros e os descontos, bem como a colaboração de concurso e compartilhamento de mercado. Os regulamentos de proibição nas leis de competição priorizam a colaboração de cartéis, na qual os participantes do mercado de uma determinada indústria colaboram para limitar a concorrência (GOTTSCHALK; SOLLI-SÆTHER, 2011).

O cartel é um acordo de cooperação entre empresas que buscam manter (entre elas) a cota de produção do mercado, determinando os preços e limitando a concorrência. No cenário de órgãos governamentais, podem ser definidos como a participação em processos de aquisição de todas as empresas de um grupo específico (RALHA; SILVA, 2012). A maioria dessas redes complexas apresenta comunidades em sua estrutura. Revelar essas comunidades é altamente relevante em muitos campos de estudo.

- **Corrupção:** Definida como a oferta, solicitação, recebimento ou aceitação de uma vantagem relacionada a uma posição, cargo ou tarefa e abrange uma ampla gama de atividades ilegais, como propinas, peculato e extorsão. A corrupção pública é o abuso

do poder confiado pelos líderes políticos para ganhos privados. O efeito negativo da corrupção enfraquece todos os esforços para melhorar a governança e promover o desenvolvimento. Além disso, a corrupção pode desempenhar um papel importante na construção de organizações criminosas. Os empreendedores criminosos podem expandir suas atividades ilegais subornando as autoridades locais (GOTTSCHALK; SOLLI-SÆTHER, 2011).

2.1.3 PRINCIPAIS FATOS QUE AFETAM OS GOVERNOS

Considerando o que foi mencionado na seção anterior, descrevemos dois fatos específicos que constituem um dano imenso a qualquer governo e nações. A manipulação da dívida pública e a evasão fiscal.

As perdas dos governos com a evasão fiscal são muito significativas. A estimativa de perdas fiscais por causa de fraudes no setor público do Reino Unido entre 2012 e 2013 foi de 15,4 bilhões de libras esterlinas, enquanto que o governo brasileiro no ano 2016 estimou suas perdas perdas fiscais em 571,5 bilhões de reais segundo estudo publicado.⁸

Nos últimos trinta anos, os lucros registrados pelas maiores empresas do mundo multiplicam por mais de três em termos reais, passando de 2 bilhões de dólares em 1980 a 7,2 bilhões de dólares em 2013. A existência de paraísos fiscais facilitaram este incremento nos lucros empresariais [Oxfam, 2016].

A evasão fiscal é fundamentalmente difícil de estudar porque não existe uma única fonte de informação que capture tudo. Alguns estudos como (ALSTADSAETER et al., 2017) combinam auditorias aleatórias, anistias fiscais e vazamentos de instituições financeiras *offshore* com os registros de riqueza para estimar o tamanho e a distribuição da evasão fiscal em lugares específicos como Noruega, Suécia e Dinamarca.

Outro estudo (RAD; SHAHBAHRAMI, 2016) analisa, projeta e implementa um sistema para identificar contribuintes de alto risco e fornecer um modelo para prever o valor da taxaço dos contribuintes para os próximos anos, para que ele desempenhe o papel de sistema de assistência. Primeramente, os peritos fiscais devem emitir as notificações de avaliação com montantes realistas durante a avaliação e a auditoria fiscal. O comportamento financeiro suspeito é detectado usando a variância e o desvio padrão médio, em seguida, os dados que existem na base tributária são usados para classificar os contribuintes usando o campo coeficiente de trabalho. Foi possível descobrir os níveis de risco e a previsão da taxaço estudando as

⁸<http://archive.is/L5Yjb>. Último acesso: 10/03/2018

volatilidades e a detecção de desvios nos valores pagos pelos contribuintes durante os últimos anos.

A dívida pública é um dos instrumentos para trazer recursos ao Estado, possibilitando investimentos. No entanto, algumas pesquisas realizadas por cidadãos em vários países mostram que grande parte das chamadas dívidas públicas - internas e externas - não respondem exatamente a esse critério, uma vez que foram geradas sem contrapartida em bens, serviços ou benefícios para a comunidade. Além disso, crescem devido ao efeito de custos abusivos, condições precárias, refinanciamentos sucessivos, entre outras estratégias, que geram a autogeração contínua de novas dívidas, por meio de um mecanismo que exige recursos constantes para pagamento de juros altos, comissões e despesas, enquanto que o saldo da dívida continua aumentando (CARNEIRO, 2003).

Os Governos, fruto do endividamento público com gigantes privados, viraram reféns e tornaram-se incapazes de regular este sistema financeiro em favor dos interesses da sociedade (MORIN, 2015). No final, o custo da dívida pública é transferido diretamente para a sociedade, em particular para os mais pobres, pagando impostos altos sobre tudo o que consomem, e pela ausência ou insuficiência de serviços públicos para aqueles que têm direito-saúde, educação, assistência, pensões- e também, entregando bens públicos através de privatizações e a exploração ilimitada de seus recursos naturais, com danos ambientais, ecológicos e sociais irreparáveis (CARNEIRO, 2003).

2.2 ESTRUTURA INTERNACIONAL DA EVASÃO FISCAL

Como visto nos pontos 2.1.1 e 2.1.2, existe uma variedade de formas de fraude aplicadas a diferentes cenários e que envolvem tipos de atores de acordo com as características de cada fato. Há fatos complexos que envolvem inúmeros atores, às vezes em total anonimato e outros atores conhecidos que atuam apoiados pelo sistema e suas articulações. Nesta seção descrevemos alguns desses casos.

2.2.1 GIGANTES FINANCEIROS

São considerados gigantes financeiros aos conglomerados de empresas. Estes gigantes controlam milhares de empresas, em dezenas de países, abarcando vários setores de atividade econômica.

A estrutura da rede de controle destas corporações afeta a concorrência no mercado global e a estabilidade financeira. A investigação da arquitetura da rede de propriedade inter-

nacional, juntamente com o cálculo do controle mantido por cada participante global, apresenta que as corporações transnacionais formam uma gigantesca estrutura de gravata-borboleta e que uma grande parte do controle flui para um pequeno núcleo unido de instituições financeiras (VITALI et al., 2011). Esse núcleo pode ser visto como uma "superentidade" econômica que levanta novas questões importantes tanto para pesquisadores quanto para formuladores de políticas.

A pesquisa realizada pelo Instituto Federal Suíço de Pesquisa Tecnológica identificou os 147 grupos que controlam 40% do sistema corporativo mundial, sendo 75% deles bancos (VITALI et al., 2011). As 29 corporações financeiras classificadas no SIFI (*Systemically Important Financial Institutions*) trabalham cada uma com um capital consolidado médio de US\$ 1.82 trilhões para os bancos e US\$ 0,61 trilhões para as seguradoras analisadas (VITALI et al., 2011).

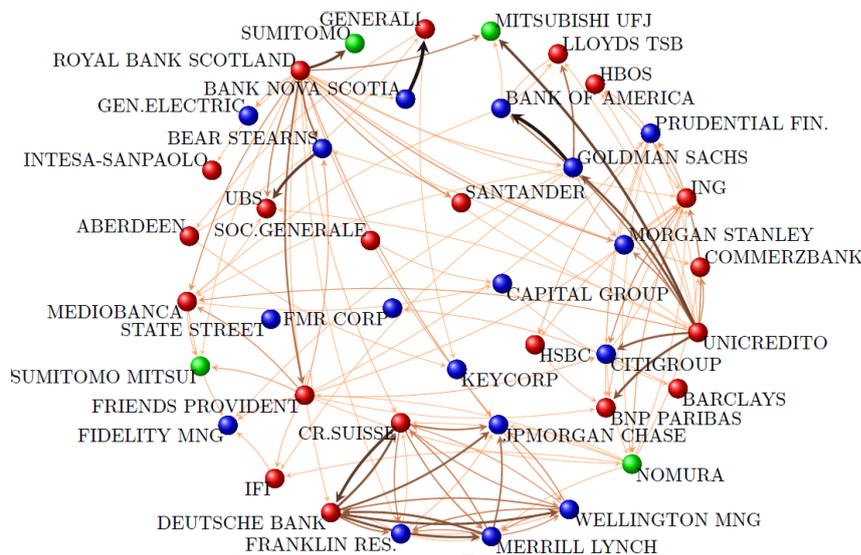


Figura 3: Amostra da Rede Financeira Internacional. Fonte: (VITALI et al., 2011)

Os gigantes financeiros concentraram grande poder de controle através de fusões e incorporações entre empresas. Essa situação gera numerosos problemas de governança, fraudes, corrupção e desorganização que ultrapassam fronteiras e todo tipo de legislação (DOWBOR, 2017).

O gigantismo corporativo gerou uma inoperância interna preocupante, que explica em boa parte o fato de estarem rigorosamente todas pagando gigantescas multas com condenações que variam desde o desrespeito aos direitos humanos a fraudes sistêmicas na área financeira bem como enganação pura e simples de clientes, como se constata com o Deutsche Bank e outros condenados em 2016 (DOWBOR, 2017).

2.2.2 PARAÍÇOS FISCAIS OU JURISDIÇÕES *OFFSHORE*

Os paraísos fiscais são jurisdições com baixa tributação que oferecem aos investidores oportunidades para evitar impostos. Exemplos desses paraísos fiscais incluem a Irlanda e Luxemburgo na Europa, Hong Kong e Cingapura na Ásia, e várias nações insulares do Caribe nas Américas. Jurisdições com baixa tributação também são comuns dentro dos países, assumindo a forma de zonas econômicas especiais na China, estados com baixos impostos e zonas empresariais nos Estados Unidos, e regiões subnacionais favorecidas pelos impostos incluindo Alemanha Oriental, sul da Itália, leste do Canadá e outras (DESAI et al., 2006).

Uma das principais características das jurisdições *offshore* é que elas se destinam a transações financeiras para não residentes, entidades ou indivíduos, nos quais os regulamentos que normalmente são impostos ao banco territorial não são aplicados a bancos *offshore*. E para atividades não bancárias ou mercado de ações, geralmente são aplicadas condições mais favoráveis do que o setor bancário. Favorecem o segredo em operações comerciais e bancárias, além de que as obrigações fiscais são inexistentes ou inferiores ao país de residência das entidades ou indivíduos (DOWBOR, 2017).

Os paraísos fiscais favorecem os gigantes financeiros e privam os países em que operam de bilhões de dólares por ano em receitas fiscais. A evasão fiscal dessas corporações gera grandes consequências e repercussões econômicas. De acordo com a Oxfam⁹, a evasão das multinacionais significa que os países pobres deixam de receber pelo menos US\$ 100 bilhões por ano em impostos [Oxfam, 2016].

De acordo com o índice de sigilo financeiro da *Tax Justice Network*¹⁰, a Suíça continua a ser o maior paraíso fiscal, seguido pelos EUA, as Ilhas Cayman, Hong Kong, Cingapura, Luxemburgo, Alemanha, Taiwan, Emirados Árabes Unidos (Dubai) e Guernsey.

No início de 2016, um enorme volume de arquivos foi vazado por partes desconhecidas. Esses arquivos chamados de *Panama Papers* totalizaram 2,6 *terabytes* que continham mais de 11,5 milhões de registros financeiros e legais listados pelo prestador de serviços corporativo panamenho Mossack Fonseca. Informações sobre aproximadamente 320.000 empresas *offshore* foram expostas junto com os nomes de 220.000 empresas. Estas empresas *offshore* foram supostamente *shell corporations* que foram projetadas para obscurecer a evasão fiscal ilegal e sigilosa de indivíduos e corporações (ZHUHADAR; CIAMPA, 2017).

Neste estudo alguns padrões interessantes foram descobertos. Estes resultados podem

⁹<https://www.oxfam.org/>. Último acesso: 10/03/2018

¹⁰<https://www.taxjustice.net/>. Último acesso: 10/03/2018

sejam transferidos para fora das jurisdições onde a atividade econômica real ocorre, resultando em uma perda de receita global de cerca de US\$ 130 bilhões por ano (COBHAM; JANSKÝ, 2017).

As firmas de contabilidade consideradas *Big 4* são Deloitte¹², EY¹³, KPMG¹⁴ e PwC¹⁵. Elas desempenham um papel importante não apenas nos serviços contábeis que fornecem para as empresas multinacionais globais, mas na prestação mais ampla de serviços financeiros, desde assessoria fiscal à formação de empresas.

O estudo realizado por (JONES et al., 2018) demonstra uma clara evidência de uma correlação entre o uso de paraísos fiscais e o uso do *Big 4* é. Além disso, também fornecem evidências de causalidade - descobriram que assumir uma firma de contabilidade *Big 4* está associado a pelo menos 2,9% maior taxa de crescimento no número de filiais de paraísos fiscais durante o período estudado em comparação com aquelas empresas que não usam uma firma de contabilidade *Big 4*.

Além disso, mostra a estrutura conceitual subjacente representada por uma pirâmide e se baseia no que chamamos de "blocos de construção para a intensidade do paraíso fiscal". Argumentam que tanto a vantagem específica do país (CSA) quanto a vantagem específica da empresa (FSA) são críticas para as empresas investirem no exterior e utilizarem subsidiárias de paraísos fiscais. Portanto, na base da pirâmide, mostram que FSAs e CSAs são igualmente importantes uma para as outras. Como todas as empresas têm algum nível de FSA que são de natureza não financeira e não vinculam a localização, elas podem ser transferidas para o exterior e re combinadas com FSAs vinculados à localização para criar vantagem competitiva. A propriedade desses FSAs tem que ser transferida para locais de paraísos fiscais para que as empresas evitem as imperfeições do mercado causadas pelo complexo código tributário internacional (JONES et al., 2018). Neste ponto, as empresas multinacionais começarão a estabelecer uma estratégia que alinha suas subsidiárias com subsidiárias em locais não-fiscais (Ver Figura 5).

2.3 HACKATHON E EVENTOS DE ATIVISMO COM FOCO EM FRAUDES

Os diversos casos e situações apresentados neste estudo mostram que os efeitos desses eventos são ruins, pois uma menor arrecadação na receita corresponde à menor quantidade de

¹²<https://www2.deloitte.com/br/pt.html>. Último acesso: 10/03/2018

¹³<http://www.ey.com/>. Último acesso: 10/03/2018

¹⁴<https://home.kpmg.com/br/pt/home.html>. Último acesso: 10/03/2018

¹⁵<https://www.pwc.com.br/>. Último acesso: 10/03/2018

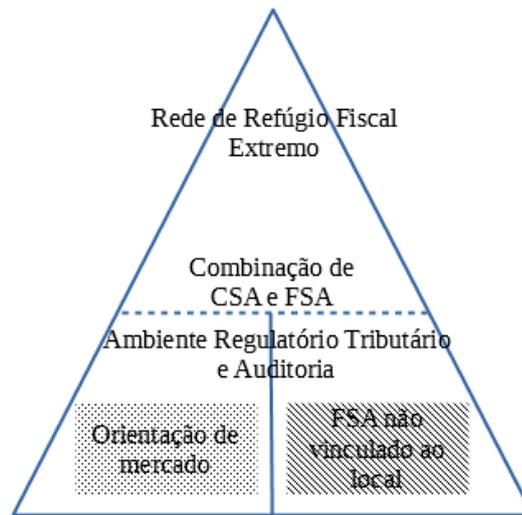


Figura 5: Blocos de Construção para o Paraíso Fiscal. Fonte: (JONES et al., 2018)

recursos públicos disponíveis para serviços de saúde, educação, acesso à moradia, meio ambiente e direitos humanos em geral. O quadro é ainda pior para as populações pobres dos países pobres, que são as que menos recursos possuem para enfrentar as situações de adversidade. Na maioria dos casos, um maior ativismo público é necessário.

Um tipo de atividade que tomou protagonismo nos últimos anos são os eventos que reúnem pesquisadores, estudantes e profissionais de diversas áreas com interesse pela área tecnológica e por desenvolvimento de soluções de combate à corrupção como *Hackathon*, *Fiscalize.me*, *Code for Brazil* e outros (BRITO et al., 2014) e (GAMA, 2017). Estes eventos visam promover o ativismo digital que envolve o cidadão na busca por soluções tecnológicas que auxiliem no processo de fiscalização da instituição, ou que estimulem a participação da sociedade no controle dos gastos públicos.

2.3.1 HACKATHON E DATATHON

A palavra *Hackathon* "hack" (programar de forma excepcional) e "marathon" (maratona). Embora o termo tem sido usado pela primeira vez em 1999, as origens dos encontros focados em criar algo novo no lugar que está hospedando remonta à década de 1970 de grupos de programação (BRISCOE; MULLIGAN, 2014). Só no ano 2016 foram realizadas 3450 *Hackathon* em todo o mundo segundo o site *Hackathon.com*¹⁶ (Figura 6).

Hackathons são "eventos de programação de computador focados no problema" que levam a novos protótipos de ideias em um curto período de tempo (BRISCOE; MULLIGAN,

¹⁶<https://www.hackathon.com/>. Último acesso: 11/12/2018

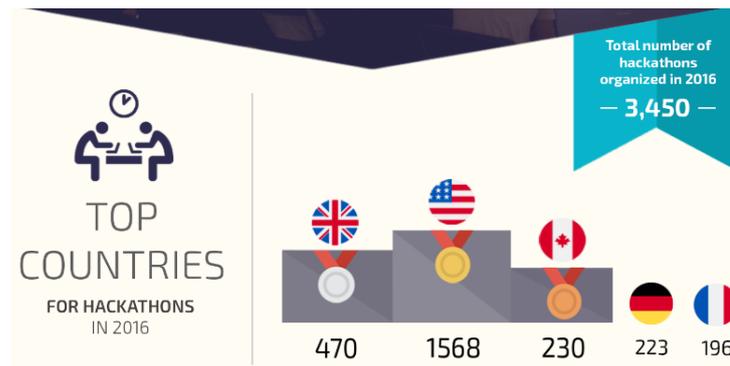


Figura 6: Quantidade de *Hackathons* realizados no mundo no ano 2016. Fonte: Site *Hackathon*

2014). Esse tipo de evento envolve muitas vezes equipes multidisciplinares de tecnólogos, *designers* e especialistas no assunto que trabalham juntos para um objetivo (IRANI, 2015). Os participantes podem desenvolver uma ideia, codificar um protótipo e lançar sua ideia para potencialmente ganhar apoio empresarial e ajuda financeira para o desenvolvimento de produtos. *Hackathons* variam em objetivos e temas, mas são frequentemente semelhantes estruturalmente e em suas características (KOMSSI et al., 2015).

Um modelo geral de organização para *Hackathons* corporativos divide este evento em três fases (Ver Figura 7). Uma fase *pré-hackathon* é dedicada à criação de ideias e à construção de equipes. Uma segunda fase consiste em um desenvolvimento direto de 1 a 3 dias. Durante esse tempo, a equipe se concentra no desenvolvimento de uma versão protótipo de sua ideia, que é demonstrada no final do evento para um público que irá julgá-lo com base nos critérios de competição. Em uma fase *pós-hackathon*, há uma decisão de continuar ou não com o desenvolvimento do que foi alcançado (GAMA, 2017).

Detalhamos o cronograma ou programação típica de atividades num *Hackathon* na Tabela 1 conforme mencionado em (PERNG et al., 2017) na qual as atividades são distribuídas durante os 3 dias sem interrupção. Existem também outros tipos de *Hackathons* que podem durar semanas ou até meses chamados *Semesterathons* e *Summerathons* dependendo do objetivo do evento (AVALOS et al., 2017).

Recentemente, um novo fenômeno pode ser observado, os chamados *Datathons* transferem o conceito de *Hackathon* para o desafio em relação à análise de dados. O intervalo de tarefas pode ser diverso, desde o pré-processamento de dados e *cleasing* até a mineração de dados e visualização de dados (ANSLOW et al., 2016).

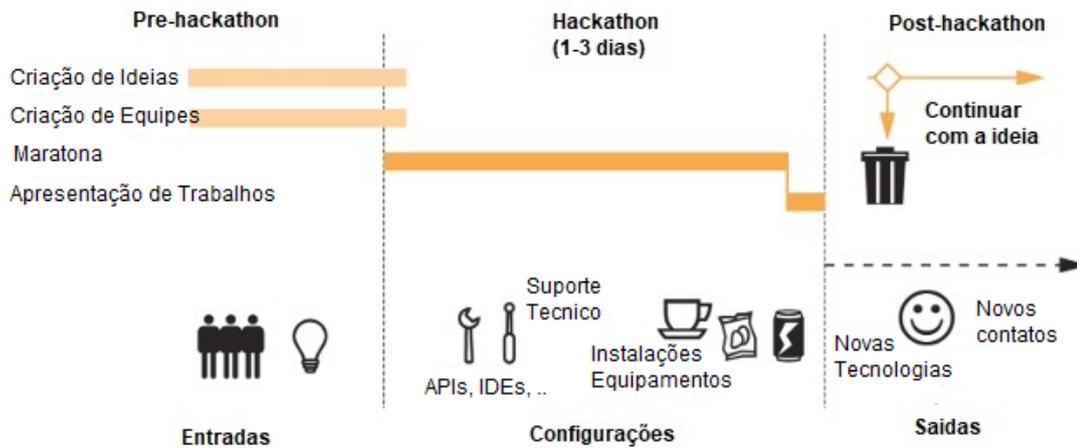


Figura 7: Estrutura típica de um *Hackathon*. Fonte: (KOMSSI et al., 2015)

Tabela 1: Programação típica de um *Hackathon*. Fonte: (PERNG et al., 2017)

Momento /Dia	Sexta	Sabado	Domingo
Manhã	Participantes registrados preparando ideias para seus projetos (não obrigatórios)	Comece a trabalhar no projeto: esclarecimento de ideias e especificação de requisitos técnicos.	Pode ter a rodada final de mentores ou workshops (por exemplo, para apresentação).
Tarde	Participantes começando a chegar e se registrando para o evento por volta das 16h15m	Organizadores apresentando oficinas técnicas e mentores disponíveis para conselhos.	O desenvolvimento do projeto termina por volta das 15-17h. As equipes se preparam para a apresentação do projeto.
Noite	Apresentação dos patrocinadores do evento. Participantes lançando idéias. Formando equipes e conhecendo os membros da equipe. Alocação de tarefas.	Continuar com o desenvolvimento do projeto por 24 horas.	Apresentação do projeto e julgamento dos resultados

2.3.2 PARTICIPAÇÃO PÚBLICA E *CITIZENSOURCING*

Participação pública é qualquer atividade que envolva aqueles que são afetados por uma decisão governamental em um processo de tomada de decisão e busque respondê-la¹⁷. O conceito de participação pública não é novo. Foi usado pela primeira vez na Grécia antiga, mas só se tornou um termo formalizado no final dos anos 1960 (CARPINI et al., 2004).

A participação pública passou por mudanças significativas ao longo dos anos. Vários processos foram experimentados e testados (CORNWALL, 2008). Os usos iniciais da participação do público eram para obter adesão em vez de *insights* para o problema mas desenvolveu uma escada de participação cidadã que esclareceu a tensão no poder entre os cidadãos e a elite governante.

¹⁷<https://www.iap2.org/>. Último Acesso: 08/01/2019

A participação de hoje varia entre participação para "incentivos materiais" e "participação interativa", e está se movendo lentamente em direção à auto-mobilização. Isso significa que os objetivos vão desde a obtenção de recursos e conhecimentos externos até os participantes que participam do processo de tomada de decisão (CORNWALL, 2008). Nos anos 90, a pressão pública por mais responsabilidade social nas corporações foi o resultado de importantes questões éticas em uma força de trabalho globalizada. As corporações começaram a se engajar com seus *stakeholders* e criaram um departamento de responsabilidade social corporativa (CSR). Esse tipo de participação foi passivo. Era uma maneira de mostrar aos interessados que eles se importavam, mas geralmente não incorporavam seus pontos de vista sobre o problema ao que estavam fazendo (SILLANPÄÄ; KOIVULA, 2010).

Outro conceito que define indivíduos reunidos provendo ideias para a resolução de um problema em comum é o *crowdsourcing*. Os projetos de *crowdsourcing* existem há bastante tempo, porém o que hoje é conhecido como tal só pôde ser praticado a partir do uso da Internet no final dos anos 1990 e expansão da conexão de alta velocidade nos anos 2000 (BRABHAM, 2012). Esse termo é definido como sendo um modelo de produção online de distribuição de problemas e soluções que impulsiona uma inteligência coletiva para atender objetivos específicos de uma organização seja ela qual for. Tal termo foi primeiramente identificado por (HOWE, 2008) como sendo um neologismo gerado a partir das palavras *crowd* (multidão) e *sourcing* (terceirização).

A partir desta contextualização é importante notar que dentro do contexto do setor público o *crowdsourcing* se torna *citizensourcing*, substituindo *crowd* por *citizen* (cidadão) (LINDERS, 2012). Ou seja, o setor público busca os cidadãos para prover opiniões e ideias para a resolução de problemas que possui e assim conseguir uma maior quantidade de alternativas.

Hackathons são exemplos de métodos contemporâneos de engajamento que começaram com as motivações do setor privado e evoluíram para os ideais do engajamento do setor público. Quando o método de engajamento público se encaixa no contexto do problema, ele pode educar os cidadãos sobre questões cívicas, aumentar a tolerância e compreensão para as perspectivas dos outros, melhorar a alfabetização de sistemas e demonstrar a validade do processo democrático enquanto informa o processo de tomada de decisões (CARPINI et al., 2004).

Numerosas empresas, universidades e instituições públicas organizam *Hackathons* voltados para setores específicos, como um determinado grupo demográfico ou para apoiar causas educacionais, de saúde ou de desenvolvimento econômico (AVALOS et al., 2017). Além disso, os *Hackathons* cívicos estão se tornando mais populares entre os governos em todo o mundo,

onde esses eventos são projetados para incentivar os cidadãos a desenvolver uma solução inovadora usando dados abertos do governo (JOHNSON; ROBINSON, 2014).

2.4 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO

Ao longo desta fundamentação, é possível observar que existem numerosos tipos de fraude que vão desde o plágio, fraudes ocorridas durante alguma operação com cartão de crédito ou nos leilões online, até aqueles fatos que impactam a administração de um governo como a corrupção, a evasão fiscal, a lavagem de dinheiro, o sistema da dívida pública, etc.

Foram descritas algumas estruturas que afetam os governos como os Bancos, Empresas multinacionais e Organizações Internacionais. O mecanismo de pressão ou formas de influência são os investimentos, financiamento de campanhas políticas e suborno. O resultado dessa pressão são acordos que beneficiam os interesses do setor influente e também contratos superfaturados, privatização de serviços públicos, facilidade de evasão de impostos e aumento na dívida pública.

Como consequência gera numerosos problemas sociais que afetam diretamente o cidadão. Serviços básicos de saúde, educação, e outros são negligenciados. A redução na arrecadação de impostos é equilibrada com mais impostos ou menos gastos em áreas essenciais. Esses acontecimentos refletem-se no desconforto da população gerando instabilidade social e movimentos de ativismo.

O ativismo tem sido uma manifestação de desconforto da população. O acesso às informações governamentais através da política de dados abertos e o uso da tecnologia promoveram um maior papel da informação no processo de decisão. Isso se relaciona diretamente às práticas de *Hackathon* uma vez que estas se utilizam para seu auxílio na resolução de problemas.

O capítulo que se segue mostrará trabalhos relacionados sobre modelos e técnicas de detecção de fraudes e estudos realizados sobre *Hackathons e Datathons*.

3 TRABALHOS RELACIONADOS

Esta seção de trabalhos relacionados está dividida em duas partes. Uma primeira parte aborda o mapeamento da literatura sobre modelos e técnicas de mineração de dados para detecção de fraudes, com ênfase especial em estudos voltados para problemas de fraude no setor público. A segunda parte desta seção apresenta estudos sobre eventos *Hackathon e Datathon*.

3.1 UM PANORAMA SOBRE *HACKATHONS* E *DATATHONS*

Os *Hackathons* são eventos que envolvem pessoas em pequenos grupos para desenvolver uma solução em um período de tempo limitado, normalmente com duração de 1 a 3 dias (KOMSSI et al., 2015). Eles ganharam popularidade nas empresas de tecnologia nos anos 2000 como uma forma de promover a codificação exploratória, a geração de novas ideias e a prototipagem com baixo risco (CARR; LASSITER, 2016). Os *Hackathons* são bons para colaboração, experimentação e aprendizado. De acordo com o objetivo, podemos agrupá-los em algumas categorias como:

- **Inovação:** Todos os *Hackathons* são uma maneira de desafiar os participantes a apresentarem um produto inovador (TRAINER et al., 2016), (FREY; LUKS, 2016), (KARLSEN; LØVLIE, 2017), (ALMISHARI et al., 2017), (ROSELL et al., 2014) e (DINTER et al., 2017);
- **Pedagogia:** Alguns *Hackathons* estão focados na aprendizagem e são uma forma de as pessoas aprenderem mais (NANDI; MANDERNACH, 2016), (KARLSEN; LØVLIE, 2017), (ANSLOW et al., 2016);
- **Envolvimento da comunidade:** *Hackathons* são maneiras de envolver uma comunidade por alguns dias em circunstâncias intensivas e focadas (TRAINER et al., 2016) e (MUNRO, 2015);
- **Recrutamento:** *Hackathons* fornecem ótimas oportunidades para procurar emprego e

procurar talentos do lado da empresa¹ (OLSON et al., 2017) e (FLORES et al., 2018).

Um *Datathon* é tipo de evento *Hackathon* em que as pessoas se reúnem durante um determinado período de tempo, para trabalhar em problemas com um conjunto de dados específico (ANSLOW et al., 2016). O aumento da quantidade de dados abertos com informações mais transparentes de governos, organizações e agências, conseqüentemente, despertou o interesse dos profissionais em explorar esses dados em eventos focados a temas específicos. Alguns exemplos de eventos *Datathon* (NETO et al., 2018), (ABOAB et al., 2016), (KOHLI et al., 2017) são mencionados na Tabela 2:

Tabela 2: Artigos sobre eventos *Datathon* e os temas abordados. Fonte: Autoria Própria

Evento	Tema	Fonte	Referencia
Canadian Open Data Experience (CODE)	Emprego juvenil, oportunidades de negócios ou vida saudável	Portal de transparência de Canada	[Anslow et al. 2016]
Primeiro Datathon Brasileiro em Terapia Intensiva	Análise de dados sanitários	Centro de Informação Médica para Cuidados Intensivos III (MIMIC-III)	[Neto et al. 2018] [Aboab et al. 2016]
SIIM Hackathon	Saúde	Dados clínicos e de imagens de câncer (TCIA)	[Kohli et al. 2017]

Os *Hackathons* foram se desenvolvendo em um formato cada vez mais exclusivo de colaboração de curto prazo e despertaram o interesse de pesquisadores em explorar outros aspectos, por exemplo como esses eventos atendem às necessidades de seus respectivos participantes e organizadores (PE-THAN et al., 2019). Fizeram estudos empíricos de 10 *Hackathons* realizadas por comunidades científicas, uma corporação e universidades e mencionaram que o evento pode atingir vários objetivos como enriquecer as redes sociais, facilitar a aprendizagem colaborativa e o desenvolvimento da força de trabalhos, se forem projetados com cuidado.

Também foram identificados vínculos sociais ou comunitários fortalecidos como principais resultados de eventos *Hackathon*. No estudo de (NANDI; MANDERNACH, 2016) os autores realizaram observações e estudaram o *feedback* da pesquisa e as evidências digitais de um *Hackathon* acadêmico. Eles descrevem diferentes modelos de colaboração entre equipes, como o evento serviu como uma plataforma informal de aprendizado para habilidades não ensinadas frequentemente em salas de aula universitárias e como proporcionou oportunidades para desenvolver habilidades de comunicação e fazer contatos com patrocinadores da indústria.

No levantamento de varias observações de *Hackathons* científicos realizado pelo (TRAINER et al., 2016). Os autores consideram os intercâmbios entre os resultados sociais e técnicos

¹ <https://archive.is/8Bkrd>. Último acesso: 08/01/2019

em eventos de *Hackathon*. Suas observações revelam que através de comunidades e tipos de eventos distintos, as estratégias para alcançar fins técnicos variam, mas o uso de *collocation* ou reunir pessoas *in situ* para fortalecer os laços sociais ou comunitários são semelhantes.

Outro estudo identifica a construção da comunidade como um ponto central dos *Hackathons* (MÖLLER et al., 2014). Eles consideram os *Hackathons* na biologia computacional como um meio de abordar as limitações que as conferências acadêmicas tradicionais impõem às oportunidades de colaboração. Eles enfatizam que as colaborações e laços comunitários mais fortes que emergem de *hackathons* e eventos similares favorecem os resultados.

Outro aspecto abordado num estudo foi a aprendizagem e aprimoramento de habilidades essenciais em ciência e análise de dados durante o evento *Datathon* (ANSLOW et al., 2016). Eles relatam sua experiência de quatro *Datathons* que envolviam alunos e membros da comunidade reunidos para resolver problemas desafiadores com dados abertos e dados de organizações sem fins lucrativos.

Segundo (CAPAY et al., 2017) a aprendizagem efetiva é vista como quando uma pessoa progride através de um ciclo de quatro etapas: tendo uma experiência concreta seguida pela observação e reflexão sobre essa experiência, que leva à formação de conceitos abstratos (análise) e generalizações (conclusões) que são usados para testar hipóteses em situações futuras, resultando em novas experiências. Por tanto, este tipo de eventos poderiam auxiliar no processo de aprendizagem por meio da experiência.

Outros autores mencionaram as desvantagens do ponto de vista dos participantes desses eventos. A seguir, algumas das desvantagens identificadas:

- **Frustração de não-desenvolvedores:** Segundo (KARLSEN; LØVLIE, 2017) se é verdade que todos podem participar de *Hackathons*, também é verdade que são eventos onde o produto é feito (aplicativo). Os participantes não desenvolvedores são geralmente abordados nas fases inicial e final, enquanto na fase central de codificação eles sentem que não estão fazendo nada;
- **Não aprende:** Os seguintes autores (RICHTERICH, 2017) e (BOWEN, 2017) mencionam que os *Hackathons* também podem ser vistos como eventos altamente competitivos, focados no tempo. Características que, segundo eles, não são uma maneira de os participantes aprenderem;
- **Questões de gênero:** O campo de TI tem um viés de gênero para o qual as mulheres não pertencem à tecnologia. Isso representa um impedimento para as mulheres participarem

de eventos de tecnologia e é novamente uma limitação para as possibilidades de aprendizagem. Isso também vale para o *Hackathons*, sendo tradicionalmente cenários dominados por homens (RICHTERICH, 2017), (KOMSSI et al., 2015) e (RICHARD et al., 2015).

Tabela 3: Resumo de aspectos positivos e negativos dos Hackathons. Fonte: Autoria própria

Aspectos Positivos	Aspectos Negativos
Enriquece as redes sociais, facilita a aprendizagem colaborativa e o desenvolvimento da força de trabalho [Pe-Than et al., 2019].	Frustração de não-desenvolvedores [Karlsen and Løvlie, 2017].
Plataforma informal de aprendizado para desenvolvimento de habilidades como a comunicação [Nandi and Mandernach, 2016].	Não aprende, altamente competitivos [Richterich, 2017] e [Bowen, 2017].
Fortalece os laços sociais ou comunitários [Trainer et al., 2016].	Questões de gênero [Richterich, 2017], [Komssi et al., 2015] e [Richard et al., 2015].
Permite a construção da comunidade e oportunidades de colaboração [Möller et al., 2014].	
Aprendizagem e aprimoramento de habilidades essenciais em ciência e análise de dados [Anslow et al., 2016].	
Aprendizagem efetiva [Capay et al., 2017]	

3.2 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO

Segundo abordado anteriormente os modelos e técnicas mineração de dados auxiliam na identificação e análise de possíveis fraudes de quase qualquer tipo desde as fraudes aplicadas no cartão de crédito até as fraudes financeiras mais complexas como a lavagem de dinheiro.

Hoje, grandes quantidades de dados disponíveis nos repositórios de dados abertos das empresas e instituições governamentais constituem um grande avanço na área de transparência e fornecem oportunidades emergentes de análise de dados. Mas o problema central que se impõe é como lidar com a complexidade, diversidade e o gigantesco volume das informações ali contidas. Somente alguns poucos especialistas, profundos conhecedores das regras de negócio de um segmento e simultaneamente profundos conhecedores da estrutura da base de dados, são capazes de propor e realizar consultas analíticas que evidenciem uma hipótese específica sendo analisada. Por conseguinte, é necessário aumentar a capacidade de análise de dados nas pessoas que com interesse no nesses dados.

Os estudos mencionados sobre *Hackathons* e *Datathons* relatam sua experiência auxiliando nos processos de aprendizado técnico, desenvolvimento de *soft skills*, engajamento, construção de rede e aprimoramento de habilidades de análise de dados. Portanto elaboramos uma proposta de formato *Short Datathon* para avaliar por meio do estudo de caso o desenvolvimento de habilidades de análise e visualização de dados sobre fraudes. Acreditamos que reduzir o tempo de atividade, torná-lo não competitivo e desenvolver tarefas mais simples utilizando um aplicativo de visualização de dados mais amigável, pode atrair um público mais diversificado e resolver as desvantagens mencionadas pelos autores. Além disso, o uso de bancos de dados focados em problemas atuais e de alto impacto, como a evasão fiscal, poderia envolver os participantes e promover a conscientização social.

O capítulo que se segue mostrará a metodologia desenvolvida nesta dissertação.

4 MÉTODO

Este capítulo tem como objetivo elucidar a maneira como a pesquisa foi realizada em todas as suas etapas. Desta forma, dividimos em partes para que em cada uma seja apresentada os métodos e técnicas utilizadas. A seção inicial trataremos sobre as fases do projeto geral da pesquisa. Na seção seguinte, mostraremos o planejamento do estudo de caso, o procedimento da coleta de evidências e o análise das evidências. Após isso, mostraremos o planejamento do questionário aplicado no final de cada Estudo de Caso.

4.1 FASES DO PROJETO DE PESQUISA

A Figura 8 representa as atividades executadas durante o desenvolvimento do projeto de pesquisa. Uma vez definido o tema de pesquisa, realizamos o processo de mapeamento da literatura disponível para desenvolver o estado de arte. Concluída esta fase, realizamos o levantamento de dados iniciais por meio de questionários aplicados em quatro eventos e uma comunidade específica em redes sociais.

Além disso, realizamos uma exploração de eventos *in situ* de um *Hackathon* e de um *Datathon*. A partir dos dados relevados nas atividades mencionadas anteriormente foi possível iniciar o planejamento e execução do piloto. Na seguinte fase de levantamento analisamos os resultados gerados do piloto e, por fim o desenvolvimento do estudo de caso formalmente dito com o análise dos resultados.

4.1.1 LEVANTAMENTO INICIAL COM QUESTIONÁRIOS

Uma das razões pelas quais um questionário pode ser interessante para a pesquisa, é que este instrumento permite, de maneira eficiente, coletar dados de um número significativo de pessoas com um esforço mínimo. Porém foram criados vários questionários para o levantamento do interesse dos participantes de diferentes eventos.

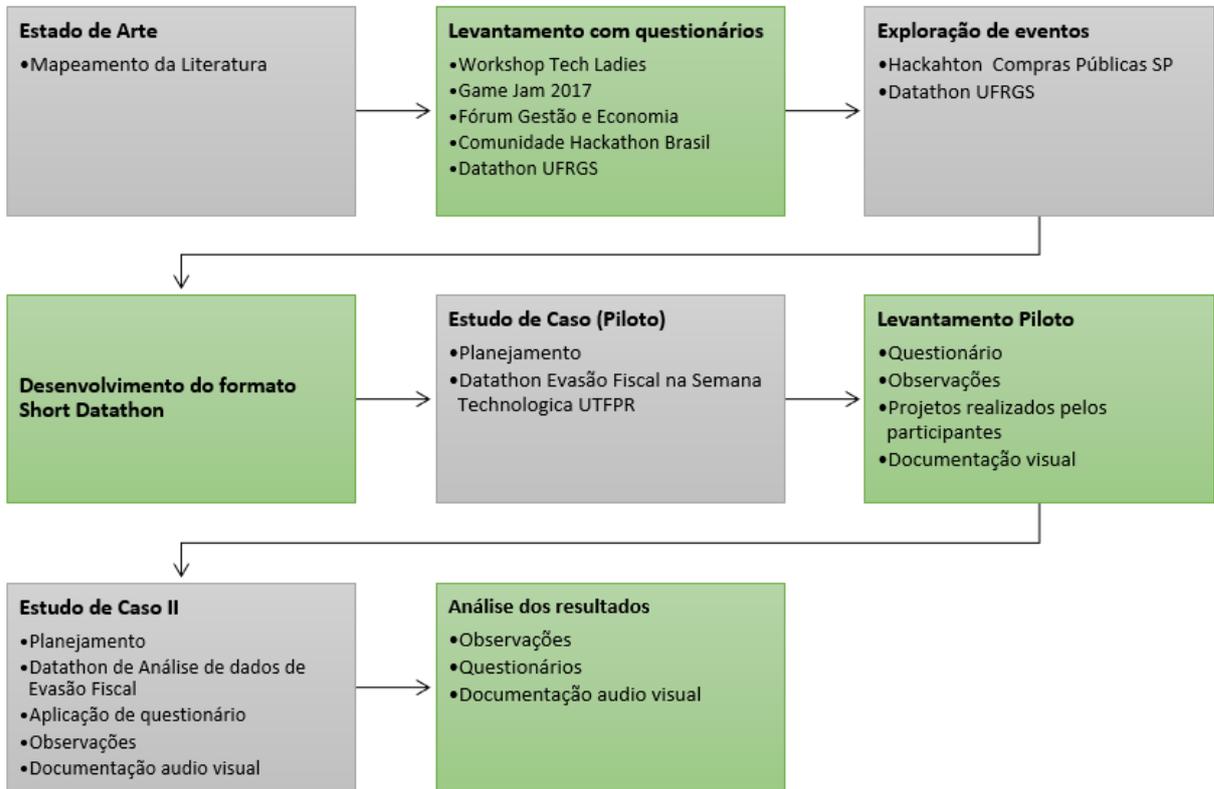


Figura 8: Fases do Projeto de Pesquisa. Fonte: Autoria Própria

Neste trabalho de levantamento definimos várias etapas durante a aplicação dos questionários que apresentamos na Figura 9.

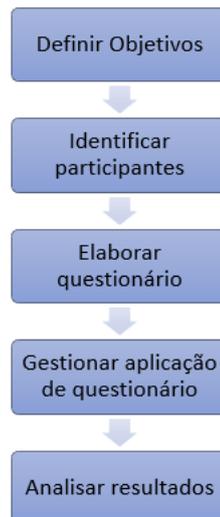


Figura 9: Fases da realização do Questionário. Fonte: Autoria Própria

4.1.1.1 IDENTIFICAR PARTICIPANTES

Um problema comum ao realizar uma pesquisa é encontrar os respondentes corretos e recolher respostas suficientes para que você tenha dados relevantes. Portanto, nossa principal preocupação foi encontrar os respondentes certos, cuja resposta fosse valiosa o suficiente para atingir o objetivo (LEEUEW, 2012).

Este trabalho teve o foco de revelar o interesse de um público diverso como mulheres, estudantes e profissionais de outras áreas, além de pessoas já acostumadas a eventos *Hackathon* e *Datathon*.

4.1.1.2 ELABORAÇÃO DO QUESTIONÁRIO

Em alguns casos adotamos a criação de um questionário impresso em folhas pela facilidade de obter resposta no momento, especialmente quando o contato é *in situ*. Também foram preparados questionários *online* para abranger a outras pessoas geograficamente distantes. Na elaboração foi prevista uma seção inicial que apresentou os objetivos da pesquisa e os termos de consentimento de utilização dos dados coletados.

Os questionários foram criados conforme critérios mencionados por (LEEUEW, 2012) e são compostos por uma seção de questões para coleta de dados demográficos. Outra seção é constituída por afirmações que devem ser classificadas de acordo com os itens apresentados, que correspondem a uma escala de cinco pontos de tipo Likert: 1 = Discordo fortemente, 2 = Discordo, 3 = Indiferente, 4 = Concordo e 5 = Concordo fortemente. Uma escala tipo Likert é constituída por questões que o respondente além de concordar ou não, apresenta o grau de intensidade das respostas (LEEUEW, 2012).

Para validação dos questionários, foram apresentados a dois professores e orientadores do Programa de Pós graduação em Computação Aplicada para discussão da temática e redação dos itens elaborados. Os itens foram corrigidos e modificados de acordo com as sugestões dos professores e novamente apresentados para apreciação final. Após obter a versão final dos questionários, estes foram aplicados aos participantes dos diferentes eventos.

4.1.1.3 GERENCIAR APLICAÇÃO DO QUESTIONÁRIO

Os questionários foram gerenciados de forma separada e considerados segundo o momento em que foram aplicados. O levantamento com questionários teve três etapas (Tabela 4):

- Primeira Etapa: Etapa em que o escopo do projeto ainda não foi definido. O objetivo do questionário foi identificar os tipos de eventos que as pessoas frequentam e o seu interesse em eventos do tipo *Hackathon*;
- Segunda Etapa: O escopo do projeto foi parcialmente definido e precisávamos identificar o interesse em um *Hackathon* em tópicos específicos com uma audiência interdisciplinar. Os questionários foram aplicados no Fórum de Gestão e Economia da UTFPR com participantes das áreas de Administração, Economia, Finanças e outros. Também foi encaminhado o questionário para as pessoas que já participaram de *Hackathons*;
- Terceira Etapa: Numa terceira etapa de levantamento o questionário foi criado e aplicado aos participantes do Datathon UFRGS para identificar a percepção sobre o evento. O escopo do projeto foi delimitado.

Tabela 4: Resumo dos questionários aplicados. Fonte: Autoria Própria

Eventos/Comunidades	Etapa de Levantamento	Forma de aplicação	Objetivos
Workshop "From ladies to ladies" das Tech Ladies	Primeira	Formulário físico	Identificar: Tipos de eventos que frequentam. Motivos que limitam sua participação. Interesse em eventos Hackathon.
Global Game Jam 2017	Primeira	Formulário físico e <i>online</i>	Identificar: Tipos de eventos que frequentam. Interesse em eventos Hackathon. Motivos de participação num GGJ.
Fórum de Gestão e Economia da UTFPR	Segunda	Formulário físico	Identificar: Tipos de eventos que frequentam. Interesse em eventos Hackathon sobre evasão fiscal ou dívida pública.
Comunidade Hackathons Brasil	Segunda	Formulário <i>online</i>	Identificar: Tipos de Hackathons que frequentam. Interesse em eventos Hackathon sobre evasão fiscal ou dívida pública.
Datathon UFRGS	Terceira	Formulário físico e <i>online</i>	Identificar: Tipos de eventos que frequentam. Percepção sobre a sua participação no Datathon UFRGS.

4.1.1.4 ANALISAR RESULTADOS

A análise dos resultados foi baseada nas respostas que recebemos através do questionário *online* e nas folhas. As respostas foram cuidadosamente analisadas a fim de obter os resultados precisos com base na pesquisa. A principal preocupação foi em interpre-

tar a informação de forma errada, o que definitivamente não serviria ao nosso propósito de investigação (LEEuw, 2012). No Apêndice A apresentamos os diferentes questionários e os resultados.

A aplicação dos questionários permitiram identificar aspectos a considerar para continuar o projeto de formato de evento:

1. As mulheres que preencheram o questionário tem preferência em participar de eventos do tipo Palestras e Conferências pela flexibilidade no horário;
2. As mulheres gostam de temas relacionados a aspectos sociais e ativismo;
3. Os eventos do tipo maratona que envolvem *games* tem um público bem variado quanto à idade, na maioria homens. Além disso, a preferência de participação é somente nesse tipo de eventos. Outros aspectos importantes foram levantamos por meio das afirmações na escala de Likert. A importancia da colaboração entre pessoas, a diversão e a experiência através do trabalho em equipe foram ressaltados.
4. As pessoas que frequentam os eventos do tipo *Hackathon* e preencheram o questionário são na maioria do gênero masculino. Elas participaram de *Hackathons* tanto privados como civicos/sociais e manifestaram interesse em temas como evasão fiscal e dívida pública.
5. As pessoas de outras áreas de conhecimento como Economia e Adminitração não costumam frequentar eventos do tipo *Hackathon* mas manifestaram interesse em temas como evasão fiscal e dívida pública.
6. As pessoas da área da Estatística e Matemática Aplicada que participaram de um *Datathon* gostaram do evento e acharam fácil de compreender. Além disso, manifestaram que a experiência com pessoal de outras áreas torna a análise mais eficiente.

4.1.2 EXPLORAÇÃO DE EVENTOS

O trabalho de pesquisa teve uma fase de exploração de eventos na qual realizamos acompanhamentos *in situ* de um *Hackathon* e de um *Datathon* com o propósito de levantar questões específicas de estrutura e organização.

Participamos na 1ª Maratona de Programação de Compras Públicas¹ realizado nos dias 5 e 6 de maio de 2018. Iniciativa da Secretaria da Fazenda, com o patrocínio da Fundação Insti-

¹<http://hackathoncompraspublicas.fazenda.sp.gov.br/>. Último acesso: 10/06/2018

tuto de Pesquisa Econômicas (FIPE), da Associação dos Agentes Fiscais de Rendas do Estado de São Paulo (AFRESP) e da Accenture, e o apoio do Banco Interamericano de Desenvolvimento (BID) e do Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (Sebrae).

O evento foi competitivo e tiveram 30 horas para desenvolver, com a base de dados da Bolsa Eletrônica de Compras (BEC), soluções que possam aperfeiçoar o ciclo de suprimentos e informar a sociedade sobre o que e como o Estado compra, além do nível de eficiência das transações em relação aos preços praticados no mercado.

Participamos também no Primeiro *Datathon* da UFRGS² nos dias 15, 16 e 17 de outubro de 2018. Foi uma maratona de análise de dados em que as equipes procuraram alternativas para que dados eleitorais sejam disponibilizados com transparência e acessibilidade a qualquer cidadão.

Algumas considerações dos eventos explorados mencionamos na Tabela 5. Mais detalhes sobre os eventos descritos no Apêndice B.

Tabela 5: Comparativo entre Maratona de Programação de Compras Públicas e o Datathon da UFRGS. Fonte: Autoria Própria

Características	Maratona de Programação de Compras Públicas	Datathon da UFRGS
Duração	30 horas	7 horas
Formato de duração	30 horas corridas	1º dia: 4 horas 2º dia: 3 horas
Dataset	Bolsa Eletrônica de Compras (BEC/SP)	Dados eleitorais 2018 do 1º turno
Quantidade de inicial de equipes	16	4
Quantidade de equipes que ficaram até o final	10	2
% Permanência das equipes	62,5%	50%
Tipo de evento	Competitivo	Competitivo
Teve mentores?	Sim	Sim
Forneceram lanches, almoço, jantar, café, outros?	Sim	Apenas lanches e café
Ferramentas Software	As equipes escolheram	Utilizaram o Rstudio, sugerido pela organização.
Infraestrutura oferecida	5 Salas de aula com cadeiras e tomadas. Conexão de Internet.	1 Sala de aula com computadores, cadeiras e tomadas. Conexão de Internet.

²<https://www.ufrgs.br/datathon/index.html>. Último acesso: 19/11/2018

4.1.3 DESENVOLVIMENTO DO FORMATO *SHORT DATATHON*

Realizamos uma adaptação do formato de *Hackathon* focado em dados, ou seja *Datathon*, para atingir nosso objetivo através da exploração, pré-processamento e visualização de dados.

A estrutura típica de um *Hackathon* começa com a ideação e a formação de equipes (Figura Comparativa 10). Essas atividades podem ser organizadas *on line* ou pessoalmente, antes do hackathon ou quando ele começar. Segundo (KOMSSI et al., 2015) coletar e desenvolver idéias de antemão deixa mais tempo para a programação. Os participantes se organizam em pequenas equipes com base nas ideias ou tecnologias que lhes interessam.

Os *Hackathons* têm um tempo de início e de término definido, entre os quais as equipes se concentram apenas na criação de uma versão protótipo de sua ideia. Os organizadores fornecem o que as equipes precisam para que possam programar sem interrupção. Por exemplo, uma conexão de Internet estável e APIs apropriadas, também fornecem comida, café, bebidas energéticas, doces e às vezes até um lugar para dormir. O suporte técnico ativo no local também ajuda as equipes a resolver problemas técnicos mais rapidamente (KOMSSI et al., 2015).

Depois de muitas horas de programação, as equipes têm alguns minutos para demonstrar seu protótipo na frente de um público. A maioria das demonstrações mostra apenas um pequeno número de recursos de trabalho; o objetivo principal é demonstrar o conceito e seu valor. Após as celebrações da adrenalina e do pós-hackathon, uma decisão deve ser tomada: a ideia vale a pena continuar, ou deve ser abandonada. Em alguns casos, ideias promissoras podem receber financiamento para desenvolvimento adicional (KOMSSI et al., 2015).

Seguindo a estrutura típica do *Hackathon*, o formato *Short Datathon* tem como atividade prévia a divulgação. O *Short Datathon* também tem um tempo de início e fim definidos, entre os quais as equipes se concentram apenas na análise e visualização de dados, mas considerando o nível de conhecimento sobre um tópico específico, prevemos uma contextualização. Os organizadores fornecem o que as equipes precisam para que possam realizar as atividades de análise e visualização de dados como: uma conexão de Internet estável, infraestrutura e banco de dados de teste. Também fornecem lanches e café.

Depois de várias horas de trabalho, as equipes têm alguns minutos para demonstrar o gráfico gerado do análise na frente de um público. Com relação à duração, supomos que um formato curto de 6 horas semelhante ao observado no *Datathon* da UFRGS pode ser bem-sucedido se combinarmos as tarefas e ferramentas corretas. Além disso, queríamos alcançar um público mais variado, especialmente pessoas que não costumam participar de longas horas de

Tabela 6: Características diferentes no *Short Datathon*. Fonte: Autoria Própria

Características	Hackathon	Short Datathon	Motivo
Duração	24-48 horas	6 horas	O tempo foi avaliado para desenvolver as atividades e atingir o objetivo de exposição a conceitos de manipulação e análise de dados
Atividades prévias	Ideação e a formação de equipes	Divulgação	A ideação e a formação de equipes são realizados no dia do evento porque o tempo da maratona é suficiente.
Escolha de Software	Livre	Definida/Livre	Ambas opções foram avaliadas durante os estudos de caso mas os participantes decidiram utilizar o software definido.
Break e lanches	Fornecida pela Organização	Apenas café e salgadinhos	Os eventos tiveram um intervalo de almoço e a sala de trabalho foi fechada.
Trabalho interdisciplinar	As pessoas que não são da área da computação participam ativamente na ideação e na apresentação final.	As pessoas que não são da área da computação participam integralmente.	As atividades foram planejadas para o trabalho conjunto das pessoas da computação com as pessoas das outras áreas. As pesquisas sobre conceitos e dados e a utilização da Software para testar dados foi constante e integral.
Equipe Organizadora	Geralmente composto pelos mentores, os facilitadores, os voluntários, os juizes e outros.	Um especialista sobre o tópico para realizar a contextualização e outras pessoas que possam ajudar com a divulgação e a logística.	A configuração do evento é simples mas se a participação é maior serão necessárias mais pessoas no dia do evento para ajudar as equipes.

desenvolvimento em equipe, característica dos *Hackathons*.

Fizemos uma comparação gráfica da estrutura típica do *Hackathon* com a estrutura do *Short Datathon* na Figura 10 para visualizar melhor as semelhanças e as diferenças.

As tarefas do *Short Datathon* foram divididas em 6 etapas (Figura 11):

- **Apresentação Inicial:** Foi prevista uma contextualização de aproximadamente 45 minutos para aprofundar temas como *Datathon*, paraísos fiscais, criação de uma empresa *offshore*, o que são as empresas intermediárias, características das empresas e contas criadas em paraísos fiscais, vantagens e desvantagens de uma empresa *offshore*, uma descrição sobre o vazamento dos *Panama Papers* e casos no Brasil de uso de paraísos fiscais. Realizamos uma atividade de quebra-cabeça para explicar o vazamento dos *Panamá Papers* e incentivar a interação entre os participantes;
- **Trabalho Inicial:** Foram preparados dados de teste para capacitar os participantes no uso da ferramenta utilizando os dados de Zucman (ALSTADSAETER et al., 2017). Também fornecemos uma pasta compartilhada com materiais de ajuda, como: um vídeo explicativo

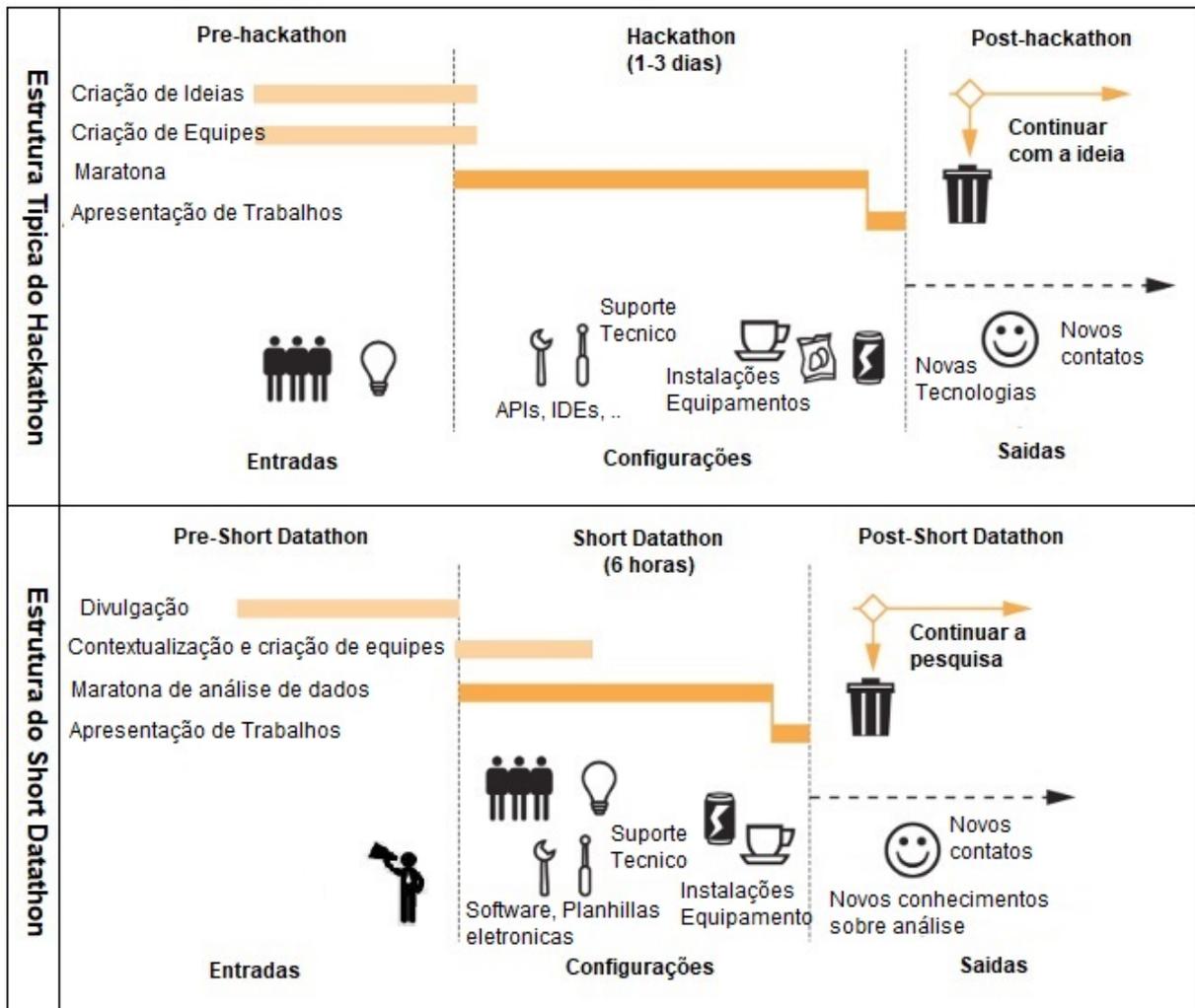


Figura 10: Estrutura típica de um Hackathon. Fonte: (KOMSSI et al., 2015)

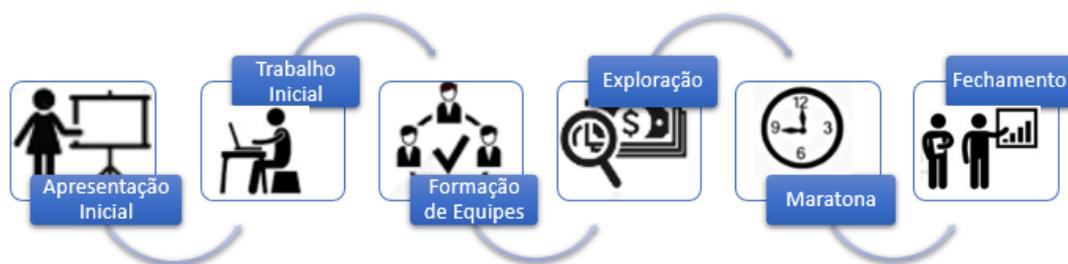


Figura 11: Estrutura Geral de um Short Datathon. Fonte: Autoria Própria

sobre o pré-processamento de dados, manual do usuário do Gapminder, dados prontos para gerar gráficos, links para fontes de dados sugeridas e outros³;

- Formação de Equipes: Os participantes se dividiram em equipes de até 5 membros. A escolha dos membros de cada equipe foi livre;

³<http://bit.ly/datathonutfpr>. Último Acesso: 08/01/2019

- Exploração: As equipes exploram dados de várias fontes de dados;
- Maratona: As equipes trabalharam com foco no pré-processamento de dados e geração de gráficos no software Gapminder;
- Fechamento: No final do dia, as equipes apresentam o trabalho e fechamos o evento.

4.1.4 PLANEJAMENTO DO EVENTO *SHORT DATATHON*

O planejamento do evento *Short Datathon* foi uma das atividades mais importantes para o desenvolvimento dos Estudos de Caso. Começamos com as anotações realizadas nos eventos *Hackathon e Datathon* exploradas pois não tínhamos experiência nesse tipo de eventos, depois definimos as etapas do planejamento e o cronograma do evento.

Aproveitamos a Semana Tecnológica⁴ para realizar o estudo de caso piloto do *Short Datathon* e solicitamos o usufruto do laboratório B201 com capacidade para 25 pessoas, além de computadores, cadeiras e projetor. O evento foi incorporado como minicurso na grade da Semana Tecnológica e divulgado no seu site.

A Figura 12 mostra as etapas do planejamento. Definimos o tema e as fontes a serem utilizadas para a análise e a visualização dos dados. Como foi mencionado anteriormente, utilizamos as planilhas eletrônicas com os dados de Zucman (ALSTADSAETER et al., 2017).

4.1.4.1 *DATASET*

Considerando as diferentes fraudes internacionais estudadas no Capítulo 2, a evasão fiscal foi escolhida para validar o formato por meio de estudos de caso. A justificativa para escolha foi o uso dos dados sobre evasão fiscal de três países escandinavos disponibilizados pelos autores do estudo (ALSTADSAETER et al., 2017). Nesse estudo eles combinaram microdados vazados de instituições financeiras em paraísos fiscais com dados de auditorias aleatórias, anistias e dados de registros de toda a população para estudar o tamanho e a distribuição da evasão fiscal nesses países. Os dados utilizados e gerados do estudo foram disponibilizados livremente no *site* do professor Gabriel Zucman⁵ no formato de planilhas eletrônicas (.xlsx) como segue:

- Apêndice A: Renda, riqueza e suas distribuições na Escandinávia;

⁴<http://archive.is/MjktD>. Último Acesso: 08/01/2019

⁵<http://gabriel-zucman.eu/leaks/>. Último Acesso: 08/01/2019

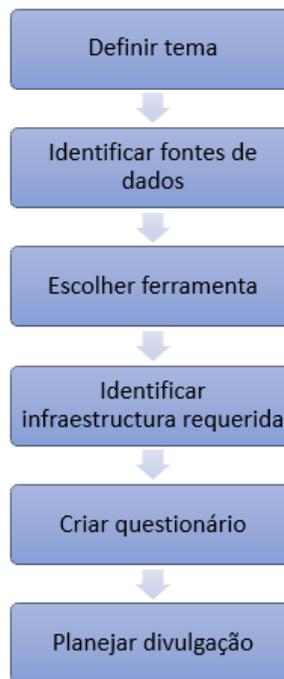


Figura 12: Processo de Planejamento do *Short Datathon*. Fonte: Autoria Própria

- Apêndice B: Desigualdade de riqueza na Noruega;
- Apêndice C: Desigualdade de riqueza na Suécia;
- Apêndice D: Desigualdade de riqueza na Dinamarca;
- Apêndice E: Vazamento do HSBC na Suíça;
- Apêndice F: Vazamento de documentos do Panamá;
- Apêndice G: Participantes da anistia fiscal;
- Apêndice H: Evasão fiscal em auditorias aleatórias;
- Apêndice I: Estoque macro de riqueza em paraísos fiscais;
- Apêndice J: Lacunas fiscais na distribuição.

Cada planilha contém várias abas de dados processados, dados não processados e gráficos estatísticos gerados pelos autores. Por exemplo, a Tabela 7 mostra o Apêndice E na aba "T.E2" com os resultados de dados processados com fórmulas. A Tabela 8 mostra o mesmo arquivo na aba "Data" com dados não processados.

Uma das planilhas eletrônicas que utilizamos para o trabalho de teste foi o Apêndice E do HSBC *Private Bank Switzerland*, a subsidiária suíça do gigante bancário HSBC. Em 2007,

Tabela 7: Aba com resultados processados. Fonte: (ALSTADSAETER et al., 2017)

	B	C	D	E	F	G	H	I	J	K	L
Table E.2: Probability to be in the HSBC leak (all matched accounts)											
	Number of matched households				% of population				Memo: Excluding HSBC wealth for ranking		
	Scandinavia	Sweden	Norway	Denmark	Scandinavia	Sweden	Norway	Denmark	Scandinavia		
P0-50	57	29	7	20	0,0%	0,0%	0,0%	0,0%	90	0,0%	
P50-90	123	71	17	43	0,0%	0,0%	0,0%	0,0%	159	0,0%	
P90-95	55	32	2	13	0,0%	0,0%	0,0%	0,0%	40	0,0%	
P95-99	110	71	5	36	0,0%	0,0%	0,0%	0,0%	101	0,0%	
P99-99.5	38	24	3	12	0,1%	0,1%	0,0%	0,1%	36	0,1%	
P99.5-99.9	79	44	5	26	0,2%	0,2%	0,0%	0,2%	51	0,1%	
P99.9-P99.95	20	13	3	12	0,4%	0,5%	0,2%	0,8%	14	0,3%	
P99.95-P99.99	28	8	2	12	0,7%	0,4%	0,2%	1,0%	21	0,5%	
P99.99-P100	10	7	1	2	0,9%	1,4%	0,4%	0,7%	8	0,8%	
Total	520	299	45	176	0,005%	0,006%	0,002%	0,006%	520	0,005%	
Notes: The counts for scandinavia do not add up to Sweden + Norway + Denmark because wealth bins are defined relatively to each economy (P99.99-100 is the top 0.01% for Scandinavia as a whole, and people in the top 0.01% in Norway are not necessarily in the top 0.01% of Scandinavia)											

Tabela 8: Aba com dados não processados. Fonte: (ALSTADSAETER et al., 2017)

Country name	Amounts	# clients associated	% have nationality	# client accounts	#client accounts active 2007	Avg per client (2007)	% total wealth	% # connected clients	% # active 2007 clients	% Avg	iso3
Switzerland	31.200	11.235	0,14	13.557	6.078	5.133.268	12,5%	10,0%	12,5%	1,00025244	CHE
United Kingdom	21.700	8.844	0,48	8.593	4.227	5.133.665	8,7%	7,9%	8,7%	1,0003298	GBR
Venezuela	14.800	1.138	0,43	696	372	39.784.946	5,9%	1,0%	0,8%	7,7523701	VEN
United States	13.400	4.183	0,23	5.026	2.280	5.877.193	5,4%	3,7%	4,7%	1,14521143	USA
France	12.500	9.187	0,51	8.545	3.499	3.572.449	5,0%	8,2%	7,2%	0,69611628	FRA
Israel	10.000	6.554	0,49	6.222	2.803	3.567.606	4,0%	5,8%	5,8%	0,69517257	ISR
Italy	7.500	7.499	0,2	5.687	2.388	3.140.704	3,0%	6,7%	4,9%	0,61198766	ITA
Bahamas	7.000	202	0,27	418	244	28.688.525	2,8%	0,2%	0,5%	5,59015611	BHS
Brazil	7.000	8.667	0,55	5.549	1.896	3.691.983	2,8%	7,7%	3,9%	0,71940828	BRA
Belgium	6.300	3.002	0,53	3.425	1.230	5.121.951	2,5%	2,7%	2,5%	0,99804738	BEL
Saudi Arabia	5.800	1.504	0,44	1.330	793	7.313.997	2,3%	1,3%	1,6%	1,42518266	SAU
Lebanon	4.800	2.998	0,46	2.672	1.346	3.566.122	1,9%	2,7%	2,8%	0,69488334	LBN
Netherlands	4.600	654	0,46	711	340	13.529.412	1,8%	0,6%	0,7%	2,63629883	NLD
Germany	4.400	2.106	0,58	1.992	845	5.207.101	1,8%	1,9%	1,7%	1,01463932	DEU
Cayman Islands	4.300	212	0	206	179	24.022.346	1,7%	0,2%	0,4%	4,68091923	CYM
India	4.100	1.668	0,51	1.403	748	5.481.283	1,6%	1,5%	1,5%	1,06806573	IND
BVI	4.000	677	0	819	690	5.797.101	1,6%	0,6%	1,4%	1,12960504	VGB
Canada	3.900	1.859	0,48	1.825	772	5.051.813	1,6%	1,7%	1,6%	0,98438056	CAN
Argentina	3.500	3.625	0,58	2.191	860	4.069.767	1,4%	3,2%	1,8%	0,79302215	ARG
Egypt	3.500	700	0,22	806	410	8.536.585	1,4%	0,6%	0,8%	1,66341231	EGY
Ireland	3.500	350	0,51	353	207	16.908.213	1,4%	0,3%	0,4%	3,29468138	IRL
Turkey	3.500	3.105	0,41	2.711	1.720	2.047.944	1,4%	2,9%	2,5%	0,55403103	TUR

um funcionário do HSBC extraiu os registros internos completos dos 30.412 clientes desse banco, uma grande parte dos quais estava evadindo impostos. A segunda planilha (Apêndice F) utilizada contém dados do vazamento dos “Panama Papers”.

4.1.4.2 SOFTWARE DE VISUALIZAÇÃO DE DADOS

Um desafio especial foi definir a ferramenta. Normalmente na maioria dos evento *Hackathon* a escolha da ferramenta fica a critério dos participantes mas neste caso foi necessario definir uma ferramenta pelos seguintes motivos:

- Preparar a infraestrutura de trabalho com as ferramentas instaladas nos computadores.

Considerando a variedade de aplicativos disponíveis para o trabalho com dados, seria complicado disponibilizar a maioria destes nos computadores;

- Atrair pessoas com pouco ou nenhum conhecimento de programação de computadores como os estudantes e profissionais de outras áreas como Economia, Administração, Estatística, etc;

Alguns dos critérios de busca de uma ferramenta adequada para utilizar no evento foram: que permita importar dados em formatos xls, xlsx e csv, que permita criar gráficos com várias opções de visualização e interface intuitiva, que permita uma curva de aprendizagem reduzida e de distribuição livre. Avaliamos as cinco ferramentas apresentadas na Tabela 9:

Tabela 9: Critérios de avaliação do Software. Fonte: Autoria Própria

Software	Intuitivo (Nível)	Tempo de Aprendizagem	Licenciamento	Importa dados xls, xlsx, cvs	Precisa de conhecimento de programação	Observação
LibreOffice	Alto	Curto	Livre	Sim	Não	Opções de gráficos limitados
Orange	Médio	Longo	Livre	Sim	Não	Ferramenta para preprocessar e minerar dados
R	Baixo	Longo	Livre	Sim	Sim	Necessário conhecimento de programação
Gapminder	Alto	Curto	Livre	Sim	Não	Contém opções de gráficos dinâmicos para series temporais
Tableau	Médio	Médio	Proprietário	Sim	Não	Tem uma versão de avaliação por 30 dias

Os valores de cada critério foram atribuídos de acordo com a experiência do pesquisador. A ferramenta que atendeu aos critérios mencionados foi o Gapminder⁶, um *software* para visualização de estatísticas referentes ao desenvolvimento humano. Criado por Hans Rosling, Gapminder e mantido por uma fundação sem fins lucrativos (ONG) que visa o desenvolvimento global, fornecendo uma ferramenta para visualizar os dados econômicos, de saúde ou educacionais de mais de 150 países. A ferramenta permite explorar dados de vários provedores de dados, como o Banco Mundial, o Eurostat e outros através de listas de bolhas, mapas, barras ou gráficos de linhas.

Definida a ferramenta e as fontes de dados, realizamos um levantamento dos recursos tecnológicos como segue:

- Instalação do aplicativo Gapminder nos computadores;

⁶<https://www.gapminder.org/>. Último Acesso: 08/01/2019

- Instalação de um processador de planilhas eletrônicas;
- Criação de um repositório ou pasta compartilhada para disponibilizar documentos de referência e as planilhas no *Google Drive*⁷;
- Conexão com Internet para realizar pesquisas e acessar a pasta compartilhada no *Google Drive*;
- Projetor para realizar as apresentações;
- Preparação de um vídeo do processo de pré-processamento dos dados para consulta e auxílio dos participantes.

A divulgação do primeiro *Short Datathon* (Estudo de Caso Piloto) foi realizada principalmente pela organização da Semana Tecnológica, também em redes sociais e via convites em grupos de *Whatsapp*. No segundo evento *Short Datathon* (Segundo Estudo de Caso) acrescentamos os canais de divulgação com a criação de uma página do evento na rede social *Facebook*. Nesta página realizamos a publicação regular de conteúdo relacionado ao tema e também publicamos convites em grupos específicos como *Hackathons Brasil* com 1740 membros. Outro tipo de divulgação foi a impressão de cartazes colocados em 5 painéis da UTFPR Curitiba e o envio de e-mail para todos os alunos. Detalhamos mais informação no Apêndice D.

Com relação a como desenvolver a atividade *Short Datathon* de princípio ao final definimos um cronograma de atividades (Tabela 10). Apresentamos as tarefas de cada uma das 6 etapas do Short Datathon com o tempo aproximado e os materiais necessários para desenvolver cada tarefa.

4.2 ESTUDO DE CASO

O formato *Short Datathon* proposto nesta dissertação foi utilizado no contexto de dois estudos de caso em um ambiente real. O estudo de caso é uma metodologia para estudar fenômenos contemporâneos em seu ambiente natural (RUNESON; HÖST, 2008). É uma estratégia empírica que se aplica quando desejamos obter conhecimento sobre algum assunto com a intenção de melhorá-lo (PER HOST, 2012).

Embora haja uma tendência na engenharia de software para a realização de estudos empíricos, as abordagens que usam estudos de caso estão em constante crescimento (RUNESON; HÖST, 2008).

⁷<http://bit.ly/datathonutfpr>. Último Acesso: 08/01/2019

Tabela 10: Cronograma de atividades do *Short Datathon* Piloto. Fonte: Autoria Própria

	Etapas	Tarefas	Detalhe	Tempo Aproximado	Materiais
Manhã 09h - 12h	Apresentação Inicial	Contextualização	Falar sobre casos específicos de sonegação no Brasil.	10 minutos	Apresentação de trechos de jornais num slide
		Explicação de conceitos	Explicar conceitos sobre evasão fiscal e uso de paraísos fiscais.	20 minutos	Slides explicativos, Projetor
		Dinâmica com Quebra cabeça	Preparar atividade com quebra cabeça sobre o vazamento dos Panamá <i>Papers</i> . Formar equipes de 3 pessoas.	15 minutos	Quebra cabeça impresso em folhas simples ou <i>online</i>
	Trabalho Inicial	Exploração com dados de teste	Explorar os dados disponibilizados pelo Prof. Zucman em planilhas eletrônicas e Gapminder.	15 minutos	Aplicativo de planilhas eletrônicas, Planilhas eletrônicas, Gapminder
		Geração de gráficos	Gerar gráficos de exemplo.	10 minutos	Aplicativo LibreOffice, Gapminder
		Tirar dúvidas	Tirar dúvidas	5 minutos	
	Equipes	Divisão em equipes	Separação em equipes de 2 até 5 pessoas. A escolha dos participantes do equipe será livre.	5 minutos	
	Exploração	Preparação de proposta	Discutir sobre os dados que serão explorados com o auxílio da planilha.	30 minutos	Modelo de proposta na pasta compartilhada
		Exploração de dados em equipes	Explorar os dados disponibilizados pelo Prof. Zucman.	1 hora	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder
		Aplicação de questionário	Entregar questionário antes do almoço	10 minutos	Questionário online ou físico (Folhas impressas)
Tarde 14h - 16h	Maratona	Maratona	Explorar dados, preprocessar e gerar gráficos	3 horas, 30 minutos	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder
	Fechamento	Apresentação dos trabalhos	Apresentação dos gráficos gerados por cada equipe.	20 minutos	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder
		Retrospectiva	Avaliar o evento por meio de uma técnica adaptada dos métodos ágeis	10 minutos	Post-it, quadro

Para guiar este estudo de caso, as seguintes perguntas foram respondidas:

- **Objetivo:** O que se deve alcançar? R: Obter opiniões e informações sobre a influência da participação em *Short Datathons* no desenvolvimento de habilidades de análise e visualização de dados.
- **O caso:** O que está sendo estudado? R: Um formato *Short Datathon* destinado a desenvolver a capacidade de análise de dados nos participantes.
- **Teoria:** Qual a referência que queremos provar ou discutir? R: A hipótese levantada inicialmente: “O *Short Datathon* é um formato de evento adequado para o intercâmbio de conhecimentos sobre fraudes internacionais e desenvolvimento de habilidades de análise e visualização de dados, além de incentivar o engajamento”
- **Questões de pesquisa:** O que se deseja saber? R: Se o formato *Short Datathon* proposto é interessante, de fácil compreensão e auxilia no desenvolvimento de habilidades de análise e visualização de dados.

- Método: Como os dados serão coletados? R: Por um questionário aplicado a participantes do *Short Datathon*. Durante o evento os participantes serão observados. A observação pode gerar dados que são aceitos no modelo de estudo de caso. Os trabalhos gerados pelas equipes durante o *Short Datathon*.
- Estratégia de seleção: Como procurar informações? R: Em estudos de caso, informações podem ser selecionadas de forma única e separada. Diferente de questionários e experimentos tradicionais, em que realizamos uma média de todas as respostas. Para o trabalho, consideraremos tanto a média de dados quanto informações pontuais que considerarmos importantes.

Em seguida detalhamos o protocolo adotado, o estudo de caso piloto, os mecanismos para coletar evidências e a análise das evidências.

4.2.1 ESTUDO DE CASO PILOTO

A preparação final para fazer a coleta de dados é a realização de um estudo piloto. O estudo de caso auxilia os pesquisadores na hora de aprimorar os planos para a coleta de dados tanto em relação ao conteúdo dos dados quanto aos procedimentos que devem ser seguidos (YIN, 2001).

Com relação à necessidade ou não de se realizar um estudo piloto, Bailer et al. (2011) afirmam que, apesar de serem tomados todos os possíveis cuidados na fase de planejamento da pesquisa, é no momento do delineamento e implementação do piloto que falhas antes imperceptíveis podem vir à tona. Tendo isto em vista realizamos um estudo de caso piloto em nosso projeto de *Short Datathon*.

4.2.2 COLETA DE EVIDÊNCIAS

Existem três formas para coleta de informação de um trabalho científico que podem estar presentes em um estudo de caso (RUNESON; HÖST, 2008):

- *Survey*: Coleta de dados por meio de questionários;
- Experimentação: Investigação de um assunto que não possui as fronteiras bem definidas. O assunto é investigado em seu próprio ambiente e a coleta de dados é realizada da forma mais variada possível;

- Pesquisa de ação: O avaliador, além de atuar como pesquisador sobre o assunto que está avaliando, também pode influenciar o meio e os indivíduos da pesquisa. Geralmente é o modelo de estudo de caso adotado na Engenharia de Software, pois os autores desenvolvem a pesquisa desejando realizar uma melhoria sobre o assunto estudado.

4.2.3 ANÁLISE DE EVIDÊNCIAS

A análise das evidências foca-se na escolha do modelo de análise, no estabelecimento de um procedimento de análise dos dados, na definição do tipo de triangulação e na captura de *feedback* dos participantes. O modelo de análise refere-se ao tipo de método que será utilizado para analisar os dados do caso (GIL, 2009). Já o procedimento de análise estabelece as etapas utilizadas para analisar os dados do caso.

Diversos autores como (YIN, 2001) e (PARÉ, 2004) recomendam a triangulação para aumentar a validade interna do Estudo de Caso. Conforme afirma (PATTON, 1999), os tipos de triangulação são: de dados, quando envolve mais de uma fonte de dados; de teoria, quando se conjuga mais de uma explicação; de pesquisadores, quando envolve mais de um pesquisador; e de métodos, quando trama mais de um método.

Neste trabalho foi adotado a triangulação de dados que envolve as seguintes fontes: o documento de proposta inicial, os gráficos gerados pelas equipes ao final do *Short Datathon* e o resultado dos questionários aplicados e o respaldo audiovisual por meio dos vídeos das apresentações das equipes.

4.3 QUESTIONÁRIO *SHORT DATATHON*

Neste estudo empregamos também questionários com uma estrutura de coleta de informações demográficas e do tipo escala *Likert* de cinco pontos para reunir dados (LEEuw, 2012). A população alvo foram os participantes da *Short Datathon*.

O questionário da pesquisa foi dividido em quatro seções (Tabela 11):

1. Informações demográficas sobre os respondentes;
2. Nível de participação em eventos;
3. Percepção sobre o *Short Datathon*. Para medir a percepção, a escala *Likert* de cinco pontos foi utilizada para refletir o nível de percepção dos respondentes em relação às afirmações.

4. A última seção foi para comentários adicionais, onde os respondentes foram convidados a inserir qualquer *feedback* ou considerações sobre o evento de forma livre, que pode ser usada para acompanhamento de esclarecimentos, se necessário.

Tabela 11: Seções do Questionário. Fonte: Autoria Própria

Seções	Questões / Afirmações
Informações demográficas	Qual é sua área de trabalho ou estudo?
	Qual é sua atividade principal?
	Qual é sua faixa etária?
	Genero
Nível de participação em eventos	Quais são os tipos de eventos de que você já participou?
	Qual é a frequência com que você participa em eventos <i>Hackathon</i> ?
	Qual é a frequência com que você participa em eventos <i>Datathon</i> ?
Percepção sobre o <i>Short Datathon</i>	Gostei de participar do <i>Datathon</i>
	O <i>Datathon</i> me ajudou a compreender mais sobre análise de dados e visualização
	Tive facilidade em entender a dinâmica do <i>Datathon</i>
	Compreendi o funcionamento das ferramentas utilizadas no <i>Datathon</i>
	Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente
	O <i>Datathon</i> promoveu momentos de cooperação entre as pessoas que participaram
	A interação com pessoal de outras áreas do meu <i>time</i> foi boa e contribuiu positivamente no resultado final
Acredito que a experiência com pessoal de outras áreas num evento <i>Datathon</i> , torna o análise de dados mais eficiente	
Comentários Adicionais	Outras considerações sobre o <i>Datathon</i> (Críticas ou sugestões)

Para validação dos questionários, estas foram apresentadas a dois professores e orientadores do Programa de Pós-graduação em Computação Aplicada para discussão da temática e redação dos itens elaborados. Os itens foram corrigidos e modificados de acordo com as sugestões dos professores e novamente apresentados para apreciação final. Após obter a versão final dos questionários, estes foram aplicados aos participantes dos estudos de caso.

Os questionários aplicados e os resultados podem ser encontrados no Apêndice C.

4.4 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO

Neste capítulo apresentamos o método proposto para atingir nosso objetivo. O trabalho desenvolveu uma série de tarefas iniciais para delimitar o escopo como a aplicação de questionários de levantamento e a exploração de eventos. Além disso, detalhamos o planejamento do Estudo de Caso, a coletas de evidências e a análise das mesmas para validar o estudo.

O capítulo que se segue mostrará o desenvolvimento do formato *Short Datathon* proposto nesta dissertação.

5 RESULTADOS E DISCUSSÃO

5.1 RELATO DO ESTUDO DE CASO PILOTO

Realizamos o convite ao evento chamado inicialmente "Datathon de Análise de dados sobre evasão fiscal" à comunidade acadêmica da UTFPR, e demais interessados em análise de dados dentro das atividades da Semana Tecnológica¹.

A data do primeiro evento foi o dia 23/10/2018. Iniciamos as atividades às 09:20 horas com a presença de 12 participantes. Considerando o escasso conhecimento sobre o tema, a pesquisadora realizou uma introdução sobre a atividade *Datathon* e os temas seguintes utilizando *slides*:

- Paraísos fiscais ou jurisdições *offshore*,
- Criação de uma empresa *offshore*,
- Empresas Intermediárias,
- Características das empresas e contas em paraísos fiscais,
- Vantagens e desvantagens de uma empresa *offshore*,
- O vazamento dos Panamá *Papers*: Para abordar esse tema a pesquisadora realizou uma atividade de quebra-cabeça em equipes para motivar a interação entre os participantes, depois relatou como aconteceu o vazamento de documentos.
- O uso da Ferramenta Gapminder: Uma parte essencial do evento foi o treino sobre a utilização da ferramenta. Foi realizado um passo a passo do funcionamento da ferramenta e o pré-processamento dos dados para sua visualização no aplicativo.

Os participantes se dividiram em equipes depois do treino inicial para a exploração respectiva e a preparação da proposta. Na seguinte Tabela 12 apresentamos as equipes e suas propostas:

¹<http://semanatecnologica.dainf.ct.utfpr.edu.br/index.php> Último acesso: 19/11/2018

Tabela 12: Resumo de Propostas Iniciais. Fonte: Autoria Própria

Nome da Equipe	Quantidade de Participantes	Resumo da Proposta
COFFEE-DATA	3	Análise entre indicadores de troca de divisas entre países e <i>offshores</i>
GINICATORS	5	Análise do índice de GINI e de impostos com dados de evasão fiscal
Sistemas de Computação	4	Análise sobre relação de liberdade econômica e social e contas criadas em empresas <i>offshore</i>

Depois do intervalo de almoço todos os participantes da equipe Sistemas de Computação estiveram presentes para continuar com seu projeto. Somente um participante da equipe “COFFEE-DATA” e uma participante da “GINICATORS” voltaram para continuar mas decidiram trabalhar juntos num novo projeto e mudar de nome pelo “CA”. Isso pode ter acontecido porque na programação da Semana Tecnológica, o *Short Datathon* foi publicado como duas atividades separadas (uma pela manhã e outra à tarde). Então houve pessoas que se inscreveram para outros cursos ou conferências no período da tarde.

**Figura 13: Imagem dos participantes. Fonte: Autoria Própria**

5.1.1 APRESENTAÇÃO DOS TRABALHOS DO ESTUDO DE CASO PILOTO

Apresentamos os trabalhos das equipes “Sistemas de Computação” e “CA”. A primeira equipe em apresentar o trabalho foi “Sistemas de Computação” sobre a visualização das quantidades de clientes do Banco HSBC por países (uma das planilhas dos dados de Zucman que contém dados sobre o vazamento do Banco HSBC da Suíça), combinados com o Índice de Percepção de Corrupção (IPC) e os valores da receita fiscal (% PIB).

O objetivo da equipe foi visualizar dados do ano 2007 e combinar com a quantidade de clientes identificados como brasileiros com contas na Suíça e gerar um gráfico no Gapminder de maneira a procurar algum tipo de padrão.

A equipe utilizou três fontes de dados:

- Índice de Percepção de Corrupção (IPC) do *Transparency International*;
- Receita Fiscal (% PIB) do *World Bank*;
- Quantidade de clientes do Banco HSBC obtidas dos dados de Zucman.

No eixo "X" os participantes colocaram os dados de Receita Fiscal (% PIB) e no eixo "Y" os dados de quantidade total de clientes com contas no Banco HSBC por país. Identificaram o nível de Percepção de Corrupção por intervalos de cores, mais perto do vermelho são países com níveis altos de percepção de corrupção e mais perto do verde o contrario, menos percepção de corrupção (Figura 14).

A interpretação dos gráficos geraram as seguintes observações:

1. A Suíça não pode ser considerada dentro da análise porque o Banco HSBC foi desse país, então é natural ter muitos clientes suíços.
2. Brasil aparece como um dos países com mais clientes com contas na Suíça e sua percepção de corrupção é bastante alta.
3. O gráfico foi separado em dois grupos de países: aqueles que tem a partir de 5000 clientes e os que tem menor do que essa quantidade de clientes. A maioria dos países se concentra abaixo a faixa dos 5000 e os valores da receita e o nível de percepção de corrupção é heterogêneo.

A segunda equipe a apresentar foi a equipe "CA" com uma proposta de visualização do produto interno bruto (PIB) e o Índice de Percepção de Corrupção (IPC) com os investimentos estrangeiros das 10 maiores economias do mundo num intervalo anual de tempo desde 2012 até 2017.

A equipe utilizou três fontes de dados:

- Produto interno bruto total (Total PIB) do *World Bank*;
- Investimento estrangeiro direto do *World Bank*;



Figura 14: Gráfico gerado pelo Equipe Sistemas de Computação. Fonte: Autoria Própria

- Índice de Percepção de Corrupção (IPC) do *Transparency International*.

No eixo "X" colocaram os dados de Produto interno bruto total (Total PIB) e no eixo "Y" os dados de Investimento estrangeiro direto. Os participantes identificaram o nível de Percepção de Corrupção por intervalos de cores: mais perto do vermelho são países com níveis altos de percepção de corrupção e mais perto do verde o contrário, menor percepção de corrupção.

O indicador de Investimento estrangeiro direto são as entradas líquidas de investimento para adquirir uma participação de gestão duradoura (10% ou mais do capital votante) em uma empresa que opera em uma economia diferente da do investidor.

A interpretação dos gráficos geraram as seguintes observações:

1. O gráfico dinâmico permite ver as mudanças acontecidas nos países durante o intervalo de tempo com relação ao nível de investimento e o total do PIB anual (Figura 15).
2. Foram marcados quatro países para visualizar melhor sua evolução durante esse período de tempo. Segundo o gráfico, a Alemanha teve uma diminuição no investimento estrangeiro mas o nível de PIB aumentou. A Índia teve um comportamento semelhante no período de 2012 a 2014, também se observa um aumento no PIB. Os Estados Unidos teve um aumento no PIB e o investimento foi mantido durante esses anos. Embora Brasil não tenha apresentado uma mudança significativa com o valor do PIB, os investimentos se mantiveram (Ver Figura 16).



Figura 15: Gráfico gerado pelo Equipe CA. Fonte: Autoria Própria

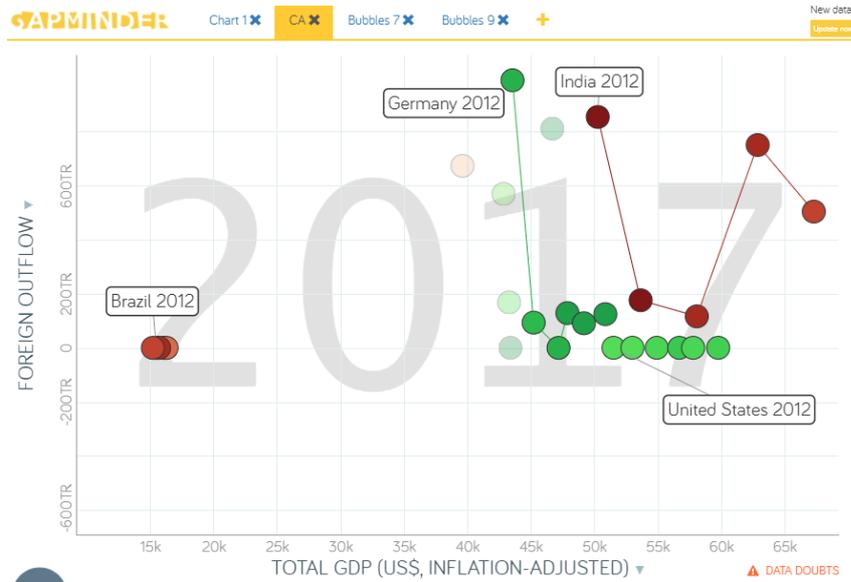


Figura 16: Gráfico gerado pelo Equipe CA. Fonte: Autoria Própria

5.1.2 RETROSPECTIVA

Realizamos a avaliação do evento com os 6 participantes que ficaram até o final. Utilizamos uma adaptação da técnica da retrospectiva do Scrum. Foi entregue a cada participante dois *post-it*, um para escrever aspectos de que “gostaram” no evento e no outro aspectos do que “não gostaram”. O quadro da sala foi dividido em duas partes “Gostei” e “Não gostei” para que cada participante colasse o *post-it* com os seus comentários.

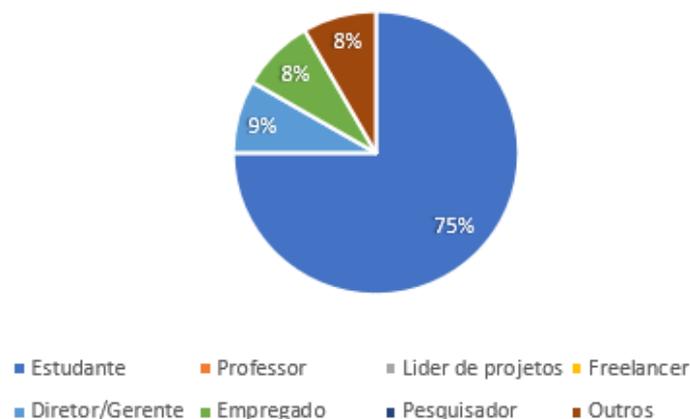
Os comentários foram transcritos na seguinte Tabela 13:

Tabela 13: Tabela de Retrospectiva. Fonte: Autoria Própria

Gostei	Não gostei
1-Desenvolve a habilidade de trabalho em equipe. 2-Foi possível aprender muitas coisas durante o processo de produção do trabalho. 3-Também serve para mostrar a importância desse tipo de evento. 4-Tema 5-Ferramenta 6-Dá uma noção de como funciona e como usar os dados. 7-Ferramenta muito boa. 8-Tema foi bem interessante. 9-Local adequado. 10-Apresentação e discussão finais foram interessantes, teve colaboração mesmo com pouca gente. 11-Programação boa. 12-Ferramenta amigável. 13-Dados relevantes e disponíveis. 14-Assunto interessante, atual. 15-Trabalhar em equipe. 16-Trabalhar com dados reais. 17-Tema complexo porém interessante conhecer sobre economia.	1-Poucos estudantes participaram. 2-Tempo curto para realizar o trabalho. 3-O evento deveria ter uma divulgação mais pesada mesmo estando dentro da semana tecnológica. 4-Pouco tempo. 5-Pouco tempo. 6-Dados ruins. 7-Ferramenta limitada. 8-Os mouses são ruins. 9-O tema limitou as análises apesar de ser interessante mas não é um grande problema. 10-Estrutura dos dados nos arquivos. Perdemos tempo juntando vários arquivos diferentes.

5.1.3 RESULTADO DE QUESTIONÁRIO APLICADO

O questionário foi aplicado ao final da primeira etapa do evento (antes do intervalo de almoço). No total responderam 12 participantes. Conforme o resultado dos questionários a maioria dos participantes é da área de Computação. As pessoas que responderam com a opção “Outros (mencionar):” especificaram que são da área de Estatística (1 pessoa) e Engenharia Mecânica (1 pessoa). Como atividade principal prevalece a atividade de “Estudante”.

**Figura 17: Atividade principal dos participantes. Fonte: Autoria Própria**

A faixa etária dos participantes se concentra entre 20-29 anos e os homens tiveram

maior participação neste evento com uma porcentagem de 92%.

Com relação à frequência de participação em eventos *Hackathon e Datathon* a maioria mencionou que foi a sua primeira participação.

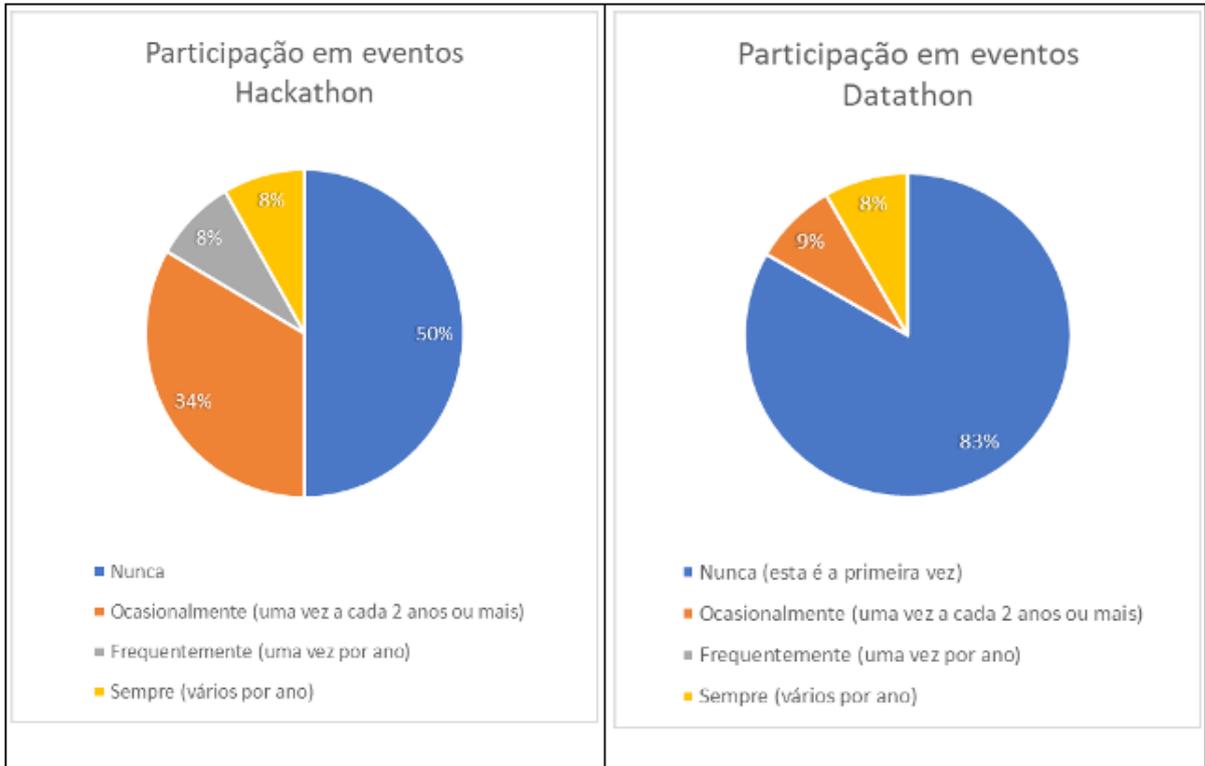


Figura 18: Participação em eventos *Hackathon e Datathon* . Fonte: Autoria Própria

Esta seção do questionário foi desenvolvido com afirmações e opções de escolha do tipo escala de Likert. Segundo os resultados pode-se visualizar a prevalência dos valores “Concordo” e “Concordo fortemente” em cada afirmação.

5.1.4 CONSIDERAÇÕES FINAIS DO PILOTO DO DATATHON

Mencionamos algumas considerações importantes sobre o piloto na seguinte lista:

- O Datathon foi realizado numa terça-feira;
- O evento Datathon foi realizado dentro da Semana Tecnológica, o que ajudou na divulgação e no gerenciamento da inscrição. As inscrições foram realizadas no *site* da Semana Tecnológica;
- A divulgação do evento foi reforçada através da plataforma de eventos Sympla e na rede social Facebook;

- O evento teve o acompanhamento e o auxílio do Professor Adolfo Neto;
- Durante todo o evento a pesquisadora auxiliou as equipes;
- Quantidade inicial de inscritos: 17 pessoas;
- Quantidade de participantes durante o evento: 16 pessoas;
- Quantidade de participantes que ficaram até o final: 6 pessoas;
- O evento teve um intervalo para almoço de 12 a 14 horas. Durante este intervalo a sala onde se realizou foi fechada e os participantes, por conta própria, foram almoçar;
- Os participantes do evento receberam um certificado com a quantidade de horas de atividade, que pode ser usado pelos alunos da UTFPR para solicitarem créditos de Atividades Complementares;
- A solicitação de credenciais de visitantes foi necessário para o uso dos computadores dos participantes externos à universidade;
- Todos os arquivos gerados pelos participantes foram salvos na pasta compartilhada.

5.1.5 CONSIDERAÇÕES INICIAIS PARA O SEGUNDO ESTUDO DE CASO

Nesta segunda edição incorporamos algumas mudanças motivadas por observações feitas no piloto do evento por exemplo:

- Premiação para os melhores trabalhos;
- O tempo de duração passou de 6 a 8 horas;
- O tema do desafio principal foi mais específico;
- Solicitamos doações de produtos não perecíveis como ingresso;
- Criação de uma página no Facebook para divulgação;
- Foi aberta a possibilidade de que cada participante possa trazer seu próprio computador com as ferramentas que achar melhor.

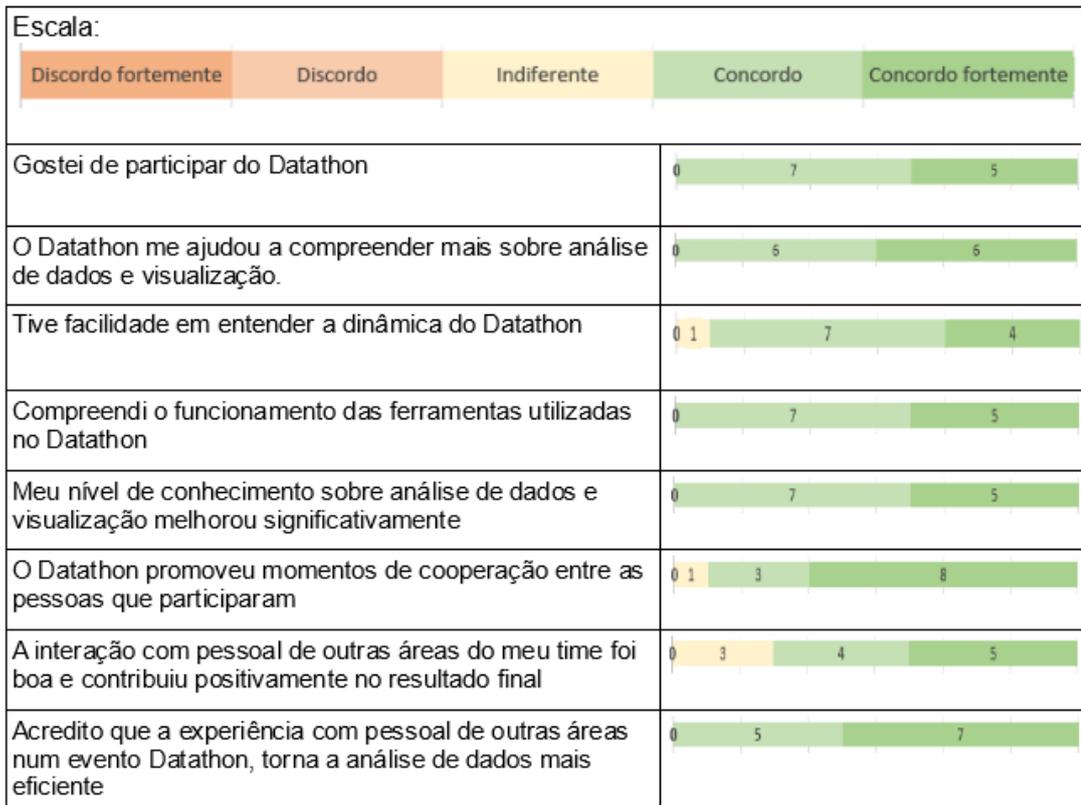


Figura 19: Resultados das afirmações do questionário do Estudo de Caso Piloto. Fonte: Autoria Própria

Solicitamos formalmente o uso da mesma sala para o segundo evento (Sala B201) o qual foi concedida pela universidade. Esta possibilidade de utilizar a mesma infraestrutura constitui uma vantagem importante pelos seguintes motivos: aplicativo já instalado em todos os computadores, os computadores e o projetor já foram testados, a conexão a internet é boa, tem ar condicionado no caso de precisar.

O convite foi realizado a comunidade acadêmica da UTFPR, e demais interessados em análise de dados. Foi criada uma página na rede social de Facebook com comentários e fotos do evento anterior e publicação de informações sobre o novo evento. Também foram colocados cartazes nos painéis da universidade.

Algumas situações identificadas que poderiam ser consideradas como riscos para o evento:

- Pouco tempo para divulgação, no total 3 semanas;
- A divulgação não teve o respaldo do primeiro evento que estava incluído dentro da grade de um evento maior conhecido pela comunidade acadêmica da UTFPR;
- O dia escolhido para o segundo evento, sábado, com o intuito de atrair outro perfil de

participantes além de estudantes;

- Não teve cobertura de suporte *helpdesk*.

O tema do desafio foi mais específico porque na primeira edição ficou muito genérico o qual dificultou a compreensão inicial dos participantes. O desafio de análise e visualização de dados desta segunda edição teve o foco no que representa para o Brasil a sonegação .

Tabela 14: Cronograma de atividades do Segundo Short Datathon. Fonte: Autoria Própria

	Etapas	Tarefas	Detalhe	Tempo Aproximado	Materiais
Manhã 09h - 12h	Apresentação Inicial	Contextualização	Falar sobre casos específicos de sonegação no Brasil.	10 minutos	Apresentação de trechos de jornais num <i>slide</i>
		Explicação de conceitos	Explicar conceitos sobre evasão fiscal e uso de paraísos fiscais.	20 minutos	<i>Slides</i> explicativos, Projetor
		Dinâmica com Quebra cabeça	Foi preparado um quebra cabeça <i>online</i> , o <i>link</i> está disponível na pasta compartilhada.	15 minutos	<i>Link</i> na pasta compartilhada https://www.puzzlesjunior.com/desafio-puzzle-de-panama-papers_5beec3b4b1f32.html
	Trabalho Inicial	Exploração com dados de teste	Explorar os dados disponibilizados pelo Prof. Zucman em planilhas eletrônicas e Gapminder	15 minutos	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder, Projetor
		Geração de gráficos	Gerar gráficos de exemplo.	10 minutos	Aplicativo LibreOffice, Gapminder, Projetor
		Tirar dúvidas	Tirar dúvidas	5 minutos	
	Equipes	Divisão em equipes	Separação em equipes de 2 até 4 pessoas. A escolha dos participantes do equipe será livre.	5 minutos	
	Exploração	Preparação de proposta	Discutir sobre os dados que serão explorados com o auxílio da planilha.	30 minutos	Modelo de proposta na pasta compartilhada
		Exploração de dados em equipes	Explorar os dados disponibilizados pelo Prof. Zucman e outras fontes de dados.	1 hora	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder
		Aplicação de questionário	Entregar questionário antes do almoço	10 minutos	Questionário online ou físico (Folhas impressas)
			INTERVALO 12H - 14H		
Tarde 14h - 19h	Maratona	Maratona	Explorar dados, preprocessar e gerar gráficos	4 horas	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder
	Fechamento	Apresentação dos trabalhos	Apresentação dos gráficos gerados por cada equipe.	1 hora	Aplicativo LibreOffice, Planilhas eletrônicas, Gapminder, Projetor
		Retrospectiva	Avaliar o evento por meio de uma técnica adaptada dos métodos ágeis	10 minutos	Post-it, quadro

5.2 RELATO DO SEGUNDO ESTUDO DE CASO

O evento foi o dia 24/11/2018 e iniciou as 09:15 horas com a presença de 8 participantes. Também teve o acompanhamento de uma participante do evento anterior que ofereceu sua ajuda como mentora da ferramenta para as equipes.

Considerando o escasso conhecimento sobre o tema, a pesquisadora realizou uma introdução sobre a atividade *Datathon* e os temas seguintes utilizando *slides*:

- Paraísos fiscais ou jurisdições *offshore*,
- Criação de uma empresa *offshore*,
- Empresas Intermediárias,
- Características das empresas e contas em paraísos fiscais,
- Vantagens e desvantagens de uma empresa *offshore*,
- O vazamento dos Panamá *Papers*: Para abordar esse tema a pesquisadora realizou uma atividade de quebra-cabeça *online* em equipes para motivar a interação entre os participantes, depois relatou como aconteceu o vazamento de documentos.
- O uso da Ferramenta Gapminder: Uma parte essencial do evento foi o treino sobre a utilização da ferramenta. Foi realizado um passo a passo do funcionamento da ferramenta e o processamento dos dados para sua visualização no aplicativo.

Os participantes se dividiram em equipes depois do treino inicial para a exploração respectiva e a preparação da proposta. Na seguinte Tabela 15 apresentamos as equipes e suas propostas:

Tabela 15: Resumo de Propostas Iniciais. Fonte: Autoria Própria

Nome da Equipe	Quantidade de Participantes	Resumo da Proposta
SEGA Data	4	Análise da informalidade no Brasil comparada a inflação.
Marfim	2	A influência da política fiscal governamental nos Paraísos Fiscais.
Olho de Lince	2	Evasão fiscal diante das principais funções do governo federal.

Depois do intervalo de almoço, todos os participantes da equipe “Olho de Lince” estiveram presentes para continuar com seu projeto. Uns dos participantes da equipe “SEGA Data” não conseguiu voltar mas uma pessoa nova que chegou a tarde para o evento foi acrescentada à equipe. Nenhum dos participantes da equipe “Marfim” voltou para continuar com seu projeto.

5.2.1 APRESENTAÇÃO DOS TRABALHOS

Foram apresentados os trabalhos das equipes "SEGA Data" e "Olho de Lince" para as pessoas presentes nesse momento. O professor Adolfo Neto foi o avaliador dos trabalhos.

A equipe "SEGA Data" foi a primeira a apresentar o trabalho com uma análise da informalidade no Brasil comparada à outras variáveis como taxa de desemprego, sonegação empresarial e o índice de preços ao consumidor amplo num intervalo anual de 2001 até 2014. Na descrição do problema eles mencionaram que a informalidade no Brasil é uma das causas da sonegação de impostos. É comum no país a prestação de serviços por profissionais liberais sem formalização, esse recurso podem ou não passar pelo sistema financeiro, é uma tendência desses preferirem receber em espécie. As variáveis e as fontes de dados utilizadas foram:

- Grau de Informalidade I (GI): Índice representado pelo quociente entre a soma dos empregados sem carteira (ESC) mais os trabalhadores por conta própria (TCP) e a soma dos trabalhadores protegidos (TP) mais empregados sem carteira (ESC) mais trabalhadores por conta própria (TCP). Como vemos na equação 1:

$$GI = \frac{(ESC + TCP)}{(TP + ESC + TCP)} \quad (1)$$

- Grau de Informalidade II (GII): Índice representado pelo quociente entre a soma dos empregados sem carteira (ESC) mais trabalhadores por conta própria (TCP) mais não-remunerados (NR) e a soma dos trabalhadores protegidos (TP) mais empregados sem carteira (ESC) mais trabalhadores por conta própria (TCP) mais não-remunerados (NR) mais empregadores (EM). Como vemos na equação 2:

$$GII = \frac{(ESC + TCP + NR)}{(TP + ESC + TCP + NR + EM)} \quad (2)$$

- Grau de Informalidade III (GIII): Índice representado pelo quociente entre a soma dos empregados sem carteira (ESC) mais trabalhadores por conta própria (TCP) e a soma dos trabalhadores protegidos (TP) mais empregados sem carteira (ESC) mais trabalhadores por conta própria (TCP) mais empregadores (EM). Como vemos na equação 3:

$$GIII = \frac{(ESC + TCP)}{(TP + ESC + TCP + EM)} \quad (3)$$

- Taxa de Desemprego (TD): Percentual das pessoas que procuraram, mas não encontraram ocupação profissional remunerada entre todas aquelas consideradas ativas no mercado de trabalho, grupo que inclui todas as pessoas com 10 anos ou mais de idade que estavam

procurando ocupação ou trabalhando na semana de referência da Pesquisa Nacional por Amostra de Domicílios (Pnad/IBGE)

- Sonegação Empresarial (SE): Representada pela divisão da sonegação pela soma entre a sonegação e a arrecadação empresarial.
- Índice de Preços ao Consumidor Amplo (IPCA): Este índice tem como objetivo medir a inflação de um conjunto de produtos e serviços comercializados no varejo, referentes ao consumo pessoal das famílias. Esta faixa de renda foi criada com o objetivo de garantir uma cobertura de 90% das famílias pertencentes às áreas urbanas de cobertura do Sistema Nacional de Índices de Preços ao Consumidor - SNIPC

As fontes de dados utilizadas foram: da Fundação Getulio Vargas (FGV), do Instituto de Pesquisa Econômica Aplicada (IPEADATA)², do Instituto Brasileiro de Geografia e Estatística (IBGE) e do Instituto Brasileiro de Planejamento e Tributação (IBPT).

No eixo "Y" foram apresentados os índices mencionados e sua evolução num intervalo anual de 2001 até 2014.



Figura 20: Gráfico gerado pelo Equipe SEGA Data. Fonte: Autoria Própria

A visualização dos dados mostrou que quanto menor o índice de informalidade da economia, menor a proporção de sonegação.

A segunda equipe a apresentar foi a equipe "Olho de Lince" com a visualização de dados na Figura 22 sobre a despesa da União entre os anos 2007 e 2018 em saúde, educação, segurança pública e previdência comparados com o valor estimado pela Receita Federal de sonegação anual. Embora os gastos com os serviços públicos sejam elevados, os membros da

²<http://www.ipeadata.gov.br/Default.aspx>. Último acesso: 25/11/2018

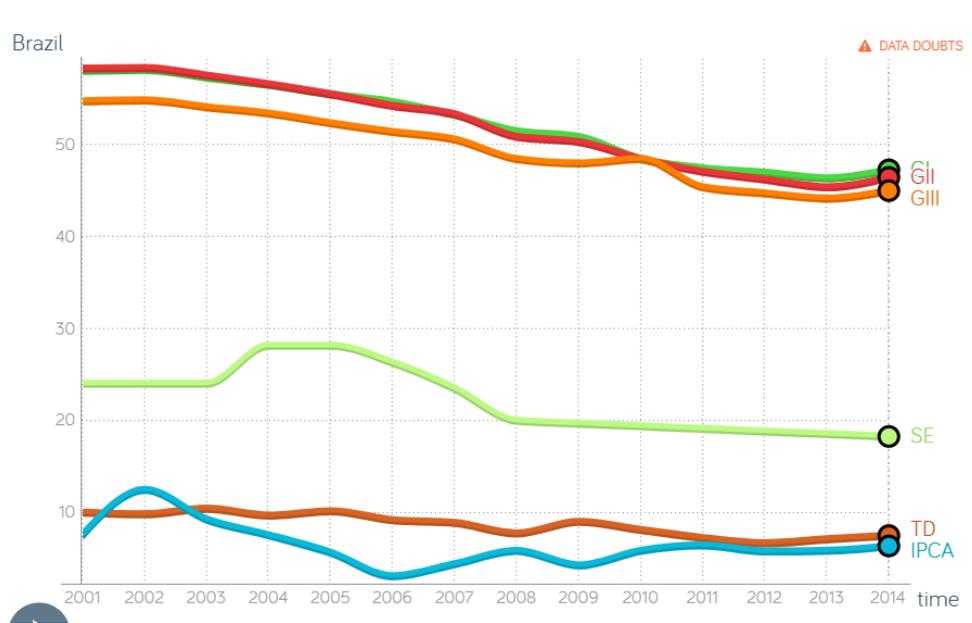


Figura 21: Gráfico gerado pelo Equipe SEGA Data. Fonte: Autoria Própria

equipe concluíram que a evasão fiscal poderia suprir os gastos da União. A fonte de dados utilizada foi do Tesouro Nacional do Ministerio da Fazenda³. Para obter um valor estimativo do montante sonegado em moeda local (real) a equipe fez uma pesquisa nas estimativas realizadas pela Receita Federal e fez uma projeção por ano.

No eixo Y foram apresentados os valores em moeda local (real) das despesas em saúde, educação, segurança pública, previdencia e o valor estimativo de sonegação num intervalo de tempo de 2007 até 2018.

5.2.2 RETROSPECTIVA

Realizamos a avaliação do evento por meio de um formulário *online* e enviado por email aos participantes. Recebemos as 5 respostas abaixo:

- *Acredito que a experiência foi válida! Faço uma sugestão para a temática de meio ambiente para um próximo evento. Buscar parcerias com ONGs, órgãos públicos para poder ampliar esse evento abrindo espaço para uma maior participação da sociedade como um todo.*
- *Excelente evento, de extrema relevância para a formação acadêmica e prática dos participantes.*
- *O tema é bem interessante*

³<http://www.tesouro.fazenda.gov.br/>. Último acesso: 25/11/2018



Figura 22: Gráfico gerado pelo Equipe Olhos de Lince. Fonte: Autoria Própria

- *O tema é muito interessante e deveria continuar sendo abordado. Porque em um país como o Brasil a evasão fiscal é enorme e eu acredito que através da informática podemos identificar todos os desvios, sejam fiscais, lavagem de dinheiro, etc,*
- *Tema ótimo.*

5.2.3 RESULTADO DE QUESTIONÁRIO APLICADO

O questionário foi aplicado ao final da primeira etapa do evento (antes do intervalo de almoço). No total responderam 7 participantes. Conforme o resultado dos questionários 4 participantes são da área de Computação e os demais são das áreas de Economia e Contabilidade.

Como atividade principal prevalece a atividade de “Estudante” mas houve uma diversidade maior que o estudo de caso piloto com a participação de consultores, empregados e *freelancers* (Figura 23).

A faixa etária dos participantes se concentra entre 20 e 29 anos mas houve um grupo de pessoas entre os 40 e 49 anos. Nesta edição do evento não houve participação feminina.

A maioria dos participantes mencionou maior participação em Palestras e Conferências. Além disso mencionaram que foi sua primeira participação num evento Hackathon e Datathon (Ver Figura 24).



Figura 23: Atividade principal dos participantes. Fonte: Autoria Própria

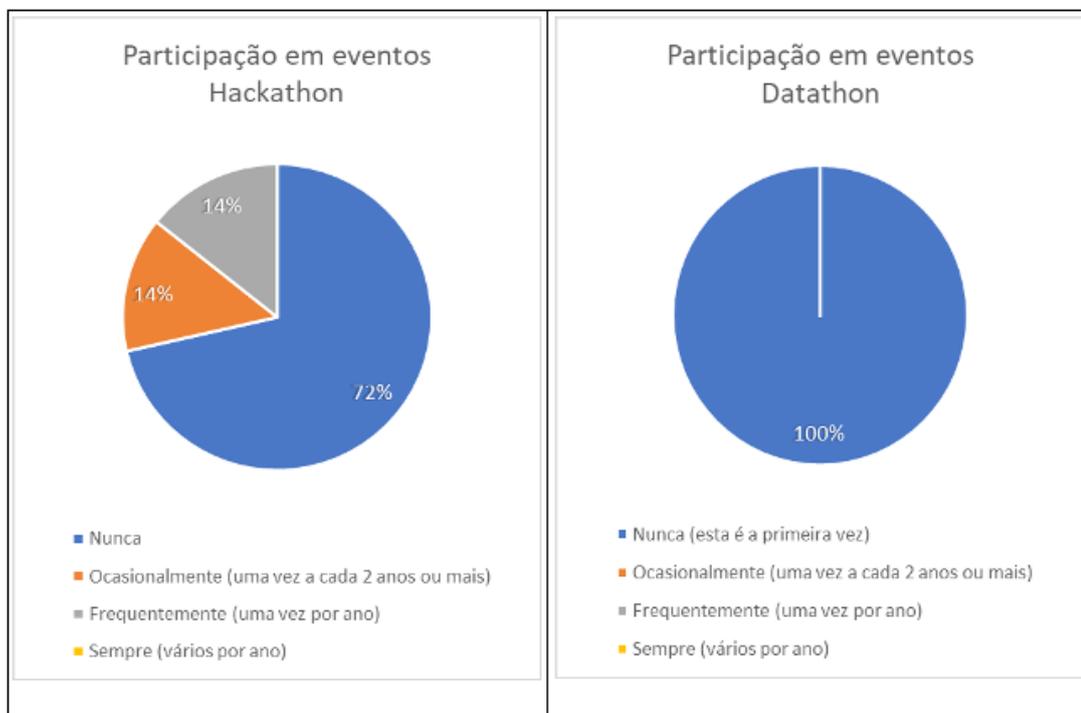


Figura 24: Participação em eventos *Hackathon* e *Datathon* . Fonte: Autoria Própria

Esta seção do questionário foi desenvolvido com afirmações e opções de escolha do tipo escala de Likert. Segundo os resultados pode-se visualizar a prevalência dos valores “Concordo” e “Concordo fortemente” em cada afirmação (Figura 25).

5.2.4 CONSIDERAÇÕES FINAIS DO *SHORT DATATHON*

Mencionamos algumas considerações importantes sobre o Datathon na seguinte lista:

- O *Short Datathon* foi realizado num sábado;

de Atividades Complementares;

- A solicitação de credenciais de visitantes foi necessário para o uso dos computadores dos participantes externos à universidade;
- Todos os arquivos gerados pelos participantes foram salvos na pasta compartilhada.

5.3 RESULTADOS GERAIS

Apresentamos uma tabela geral do estudo de caso piloto e do segundo estudo de caso para comparar as características de cada um (Tabela 16). A primeira característica apresentada é o total de participantes. O estudo de caso piloto teve mais participantes possivelmente por ter sido realizado durante a Semana Tecnológica da UTFPR. Também o dia do evento pode ter influenciado a participação.

O porcentual de participantes que ficaram até o final no segundo estudo de caso foi 62,5% em comparação com 50% do estudo de caso piloto. Ambos os estudos tiveram a participação de pessoas de outras áreas mas não conseguimos atingir as mulheres. Acreditamos que a baixa participação das mulheres foi por não realizarmos a divulgação adequada. Uma das atividades que poderia ter sido realizada era o envolvimento de grupos e redes de empoderamento feminino na organização ou divulgação.

A duração do evento foi definida inicialmente em 6 horas mas, a pedido dos participantes do estudo de caso piloto, foi estendido a 8 horas para o segundo evento. O segundo estudo de caso finalizou antes do horário planejado, ou seja, os participantes terminaram o desafio em 6 horas.

A faixa etária dos participantes do estudo de caso piloto foi homogênea, ou seja pessoas de 20 a 29 anos. No segundo estudo de caso participaram pessoas da faixa etária de 20 a 29 anos e de 30 a 39 anos.

O tempo de divulgação do estudo de caso piloto foi maior em contraste com o segundo estudo de caso que teve apenas 3 semanas de divulgação.

Combinamos os resultados dos questionários aplicados no evento num único gráfico para apreciar a distribuição dos níveis de percepção dos participantes (Figura 17). Além disso, mostramos o porcentual de satisfação por nível:

- Concordo fortemente: 44,73%
- Concordo: 51,97%

Tabela 16: Resumo das características dos Estudos de Casos. Fonte: Autoria Própria

Características	Estudo de Caso Piloto	Segundo Estudo de Caso
Total participantes	12	8
Total de participantes que ficaram até o final	6	5
Total de estudantes ou profissionais da computação	9	5
Total de estudantes ou profissionais de outras áreas	3	3
Total de participantes que responderam o questionário	12	7
Total de mulheres	1	0
Dia do evento	Terça-feira	Sábado
Duração do evento	6 horas	8 / 6 horas
Quantidade de equipes formadas	3	3
Quantidade de equipes que finalizaram o desafio	2	2
Faixa etária	Homogêneo (20-29)	Heterogêneo
Tempo de divulgação	6 semanas	3 semanas

- Indiferente: 3,28%
- Discordo: 0,0%
- Discordo fortemente: 0,0%

Tabela 17: Avaliações das afirmações na escala de Likert dos Estudos de Caso. Fonte: Autoria Própria

Escala:	
Gostei de participar do Datathon	
O Datathon me ajudou a compreender mais sobre análise de dados e visualização.	
Tive facilidade em entender a dinâmica do Datathon	
Compreendi o funcionamento das ferramentas utilizadas no Datathon	
Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente	
O Datathon promoveu momentos de cooperação entre as pessoas que participaram	
A interação com pessoal de outras áreas do meu time foi boa e contribuiu positivamente no resultado final	
Acredito que a experiência com pessoal de outras áreas num evento Datathon, torna a análise de dados mais eficiente	

Conforme descrito na seção 4.1.1, na escala de Likert os respondentes não apenas respondem se concordam ou não com as afirmações, mas também informam qual seu grau de concordância ou discordância. Atribuímos um valor de 1 a 5 a cada resposta a partir da qual é calculada a média ponderada para cada item, baseando-se na frequência das respostas. Desta forma foi obtida a Avaliação Média (Tabela 18). Para obter o Total de Avaliação Média somamos o valor da Avaliação Média de cada item e dividimos pela quantidade de afirmações, ou seja oito.

Então, considerando a escala de Likert de cinco pontos, em que 5 representa total satisfação, o questionário aplicado apresentou resultados na Avaliação Média Total de 4,41.

Tabela 18: Avaliações das afirmações na escala de Likert. Fonte: Autoria Própria

Afirmações	Avaliação Média
Gostei de participar do Datathon	4,421052632
O Datathon me ajudou a compreender mais sobre análise de dados e visualização	4,473684211
Tive facilidade em entender a dinâmica do Datathon	4,368421053
Compreendi o funcionamento das ferramentas utilizadas no Datathon	4,421052632
Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente	4,315789474
O Datathon promoveu momentos de cooperação entre as pessoas que participaram	4,526315789
A interação com pessoal de outras áreas do meu time foi boa e contribuiu positivamente no resultado final	4,210526316
Acredito que a experiencia com pessoal de outras áreas num evento Datathon, torna o análise de dados mais eficiente	4,578947368
Avaliação Média Total	4,414473684

5.4 DISCUSSÃO

Realizamos uma comparação com a literatura consultada e identificamos questões importantes de mencionar. Nossos estudos de caso, como o estudo de (NANDI; MANDERNACH, 2016), mostraram uma colaboração constante entre os participantes de cada equipe favorecendo o resultado final. O compartilhamento de conhecimento sobre um tema atual com conceitos pouco conhecidos foram fatores que também engajaram e motivaram ao participantes

Como menciona no estudo (MÖLLER et al., 2014) sobre *Hackathons* e *Codefests*. Os eventos *Short Datathon* também podem suprir aquelas limitações que as conferências e palestras

tradicionais impõem às oportunidades de colaboração e ação. O evento poderia auxiliar como uma plataforma informal para desenvolvimento de habilidades de análise e visualização de dados.

Além do mencionado anteriormente, o formato *Short Datathon* poderia ser utilizado como apoio em estratégias de aprendizagem por meio da experiência considerando as características mencionadas pelo (CAPAY et al., 2017) que menciona que acontece a aprendizagem efetiva quando uma pessoa progride através de um ciclo de quatro etapas: tendo uma experiência concreta seguida pela observação e reflexão sobre essa experiência, que leva à formação de conceitos abstratos (análise) e generalizações (conclusões) que são usados para testar hipóteses em situações futuras, resultando em novas experiências.

Considerando as desvantagens identificadas da estrutura típica do *Hackathon* (KOMSSI et al., 2015) mencionadas na literatura apresentamos algumas soluções com o desenvolvimento do formato *Short Datathon*:

- Para tratar a questão da frustração de não-desenvolvedores mencionado em (KARLSEN; LØVLIE, 2017) desenvolvemos uma série de tarefas combinadas com um software simples e intuitivo que permitiram a participação das pessoas de outras áreas de forma integral desde o princípio até o final;
- A questão de não-aprendizagem (RICHTERICH, 2017) e (BOWEN, 2017) foi resolvida fazendo o evento ser não competitivo. Os participantes conseguem aproveitar melhor o evento e criar um ambiente de compartilhamento de conhecimentos;
- As questões de gênero (RICHTERICH, 2017), (KOMSSI et al., 2015) e (RICHARD et al., 2015). Acreditamos que a pouca participação das mulheres foi pela falta de envolvimento e divulgação adequados. Uma das atividades que poderia ter sido realizada é o envolvimento de grupos e redes de empoderamento feminino na organização ou divulgação.

Além disso, listamos outras desvantagens não mencionadas nos estudos:

- Custo de investimento alto: Na estrutura típica do *Hackathon* (24 a 48 horas) é necessário uma organização e logística maior, um local onde as pessoas possam virar a noite, lanches e refeições e outros. A vantagem principal do formato *Short Datathon* é a duração. Os participantes não precisam virar a noite e os lanches podem ser fornecidos pela organização ou ficar por conta dos participantes;

- Equipe Organizadora: Um evento de *Hackathon* típico precisa de mais pessoas pela complexidade mencionada no item anterior. Além de prever pessoas que serão juízes (muitas vezes pessoas de outras cidades) para avaliar os projetos apresentados. O formato *Short Datathon* precisa de uma pessoa para realizar a contextualização e outra para ajudar na divulgação e na logística. Acreditamos que segundo a quantidade de participantes seria necessário mais pessoas para ajudar as equipes;
- Duração do evento: Uma questão não mencionada na literatura sobre o *Hackathon* é o fator fadiga nestes eventos. Existem vários *blogs*⁴ que descrevem comentários como por exemplo: "*Não é algo que repetiríamos com frequência por motivos de: é extremamente cansativo*". Outro portal de notícias⁵ menciona o seguinte comentário: "*Foi muito cansativo, foram horas sem dormir e tivemos alguns problemas no decorrer do processo que geraram stress, mas foram contornados*". O formato *Short Datathon* tem a duração adequada para o objetivo proposto sem gerar fadiga nem stress.

5.5 CONSIDERAÇÕES SOBRE ESTE CAPÍTULO

Neste capítulo, vimos os resultados obtidos nos estudos de caso em duas oportunidades diferentes. Realizamos a interpretação de todas as respostas para os dois questionários. As equipes que ficaram até o final dos eventos conseguiram apresentar seus trabalhos. Durante os eventos, observamos uma interação constante entre os membros das equipes.

Dentre as principais respostas, os participantes mencionaram que o evento ajudou a entender mais sobre análise e visualização de dados, o formato foi compreensível e promoveu momentos de cooperação entre as pessoas. Eles também afirmaram que a interação com pessoas de outras áreas de suas equipes foi boa e contribuiu positivamente para o resultado final.

⁴<https://descola.org/drops/hackathon-24-horas-que-valem-por-dias/>. Último Acesso: 08/01/2019

⁵<https://www.castrolanda.coop.br/noticia/sistema-de-monitoramento-integrado-ganha-hackathon-agroleite-23415>. Último Acesso: 08/01/2019

6 CONCLUSÃO

Numerosos fatos podem ser caracterizados como fraude, desde fraudes com cartões de crédito até fatos mais complexos, como evasão fiscal. A evasão fiscal gera um problema significativo para os governos e os priva de recursos públicos para investir em políticas sociais. Além disso, a evasão fiscal é difícil de estudar porque não existe uma única fonte de informação que capture tudo e é necessário conhecimento sobre conceitos de manipulação de dados e também conceitos relacionados as áreas de economia, direito internacional, estatística e outras especialidades.

Os eventos do tipo *Hackathon* tomaram protagonismo pelas suas características mas apresentam várias desvantagens como: a frustração que experimentam as pessoas não-desenvolvedoras, a questão de não-aprendizado devido à competitividade e o fato que são eventos com prevalência de público masculino (viés de gênero). Além disso, outras causas mencionadas em *blogs* e portais de notícias como: o alto custo de investimento, a necessidade de muitas pessoas para organizar o evento e o fator de fadiga e stress por causa da duração do evento.

No presente trabalho propusemos um formato de evento que chamamos de *Short Datathon* para exposição a conceitos de manipulação de dados e desenvolvimento de habilidades de análise e visualização de dados. Preparamos um cronograma de atividades para o trabalho com dados reais sobre evasão fiscal.

Realizamos dois estudos de caso para avaliar o formato *Short Datathon* nos quais os participantes foram expostos a conceitos de manipulação de dados e conceitos sobre evasão fiscal através de uma contextualização inicial para motivar a pesquisa e a análise de outras fontes de dados relacionados ao tema. Os participantes realizaram pesquisas na Internet e compartilharam conhecimentos durante todo o evento. Aplicamos questionários para levantar a percepção dos participantes e validar nossa hipótese.

O primeiro estudo de caso foi realizado com a participação de 12 estudantes de graduação da UTFPR. Como retorno, foram gerados 2 gráficos dinâmicos sobre os dados que eles escolheram e pré-processaram. Além disso, mencionaram no questionário que o *Short Datathon* ajudou

a compreender mais sobre análise de dados e visualização, o formato foi compreensível e promoveu momentos de cooperação entre as pessoas que participaram. Os participantes afirmaram também que a interação com pessoal da suas equipes foi boa e contribuiu positivamente no resultado final.

No segundo estudo de caso incorporamos algumas mudanças levantadas no primeiro estudo por exemplo: premiação para os melhores trabalhos, o tempo de duração passou de 6 a 8 horas, o tema do desafio principal foi mais específico, doações de produtos não perecíveis foram solicitadas como ingresso. Outro aspecto adicionado foi ampliar a divulgação. Também foi aberta a possibilidade de que cada participante pudesse trazer seu próprio computador com as ferramentas que considerasse melhor.

No segundo estudo de caso participaram 8 pessoas de graduação e pós-graduação das áreas da Computação, Economia e Administração. Como resultado do segundo estudo de caso, foram gerados 2 gráficos dinâmicos sobre os dados que eles escolheram e pré-processaram. Além disso, mencionaram no questionário que o *Short Datathon* ajudou-os a compreender mais sobre análise de dados e visualização, o formato foi compreensível e promoveu momentos de cooperação entre as pessoas que participaram. Os participantes afirmaram também que a interação com pessoas de outras áreas da suas equipes foi boa e contribuiu positivamente no resultado final.

Portanto, acreditamos que o formato *Short Datathon* poderia ser uma primeira etapa de engajamento e introdução de um plano mais abrangente de detecção de fraudes internacionais colaborativo. Uma segunda etapa do plano poderia envolver técnicas de mineração de dados ou *machine learning* com o uso de outras ferramentas de *software*.

Entre as limitações deste estudo, podemos citar que: 1) apenas dois estudos de caso foram aplicados; 2) apenas um tópico específico foi utilizado; 3) cada estudo de caso teve uma amostra de pessoas reduzida; e 4) os dois estudos de caso foram realizados dentro da UTFPR sem patrocínio de organizações ou empresas; 5) não alcançamos a participação das mulheres como esperávamos.

6.1 CONTRIBUIÇÕES

Durante a realização deste trabalho, desenvolvemos um formato de evento *Short Datathon* como resultado da pesquisa.

Também foi aceito um artigo sobre *Short Datathon*¹ no *12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE 2019)*².

- SALINAS, M.N.; EMER, M.C.F.P.; NETO, A.G.S.S. Short datathon for the interdisciplinary development of data analysis and visualization skills. In: *12th International Workshop on Cooperative and Human Aspects of Software Engineering - CHASE 2019*

6.2 TRABALHOS FUTUROS

Em trabalhos futuros propõe-se testar o formato numa amostra maior. Isso iria exigir mais pessoas na organização para acompanhar as equipes. Também poderia-se abordar outros temas e o uso de outras ferramentas para a análise dos dados. Poderia-se incorporar mineração de dados e aprendizado de máquina.

A outra abordagem bem interessante seria uma *Hackathon Gamer* que tenha como desafio prototipagem de um jogo sério com foco em fraude internacional.

Esta atividade poderia ser abordada também com foco na participação das minorias como mulheres, pessoas com deficiência, grupos LGBTI, etc.

¹<https://arxiv.org/abs/1903.07539>. Último Acesso: 11/03/2019

²<http://www.chaseresearch.org/workshops/chase2019>. Último Acesso: 06/02/2019

REFERÊNCIAS

ABDALLAH, A.; MAAROF, M. A.; ZAINAL, A. Fraud detection system: A survey. **Journal of Network and Computer Applications**, Elsevier BV, v. 68, p. 90–113, jun 2016.

ABOAB, J. et al. A “datathon” model to support cross-disciplinary collaboration. **Science Translational Medicine**, American Association for the Advancement of Science (AAAS), v. 8, n. 333, p. 333ps8–333ps8, apr 2016.

ABRANTES, P. C.; FERRAZ, F. Big data applied to tax evasion detection: A systematic review. In: **2016 International Conference on Computational Science and Computational Intelligence (CSCI)**. [S.l.]: IEEE, 2016.

AHMED, M.; MAHMOOD, A. N.; ISLAM, M. R. A survey of anomaly detection techniques in financial domain. **Future Generation Computer Systems**, Elsevier BV, v. 55, p. 278–288, feb 2016.

ALMISHARI, S. et al. The rise of hackathon-led innovation in the MENA region: Visualizing spatial and temporal dynamics of time-bounded events. In: **Social Computing and Social Media. Applications and Analytics - 9th International Conference, SCSM 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II**. [s.n.], 2017. p. 367–377. Disponível em: <https://doi.org/10.1007/978-3-319-58562-8_28>.

ALSTADSAETER, A.; JOHANNESSEN, N.; ZUCMAN, G. Tax evasion and inequality. **SSRN Electronic Journal**, Elsevier BV, 2017.

ANSLOW, C. et al. Datathons. In: **Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE 16**. [S.l.]: ACM Press, 2016.

AVALOS, M. et al. Hackathons, semesterathons, and summerathons as vehicles to develop smart city local talent that via their innovations promote synergy between industry, academia, government and citizens. In: **2017 International Smart Cities Conference (ISC2)**. [S.l.]: IEEE, 2017.

BOWEN, L. M. The limits of hacking composition pedagogy. **Computers and Composition**, Elsevier BV, v. 43, p. 1–14, mar 2017.

BRABHAM, D. C. Motivations for participation in a crowdsourcing application to improve public engagement in transit planning. **Journal of Applied Communication Research**, Informa UK Limited, v. 40, n. 3, p. 307–328, aug 2012.

BRISCOE, G.; MULLIGAN, C. Digital innovation: The hackathon phenomenon. In: **Creativeworks London Working Paper No. 6, May 2014, Retrieved April 12, 2016**. [S.l.]: <http://www.creativeworkslondon.org.uk/wp-content/uploads/2013/11/Digital-Innovation-The-Hackathon-Phenomenon1.pdf>, 2014.

BRITO, K. dos S. et al. Using parliamentary brazilian open data to improve transparency and public participation in brazil. In: **Proceedings of the 15th Annual International Conference on Digital Government Research - dg.o 14**. [S.l.]: ACM Press, 2014.

CAPAY, M.; SKALKA, J.; DRLIK, M. Computer science learning activities based on experience. In: **2017 IEEE Global Engineering Education Conference (EDUCON)**. [S.l.]: IEEE, 2017.

CARNEIRO, M. L. F. **Auditoria da Dívida Externa: Questão de Soberania**. Contraponto, 2003. ISBN 858591050X. Disponível em: <<https://www.amazon.com/Auditoria-D>

CARPINI, M. X. D.; COOK, F. L.; JACOBS, L. R. PUBLIC DELIBERATION, DISCURSIVE PARTICIPATION, AND CITIZEN ENGAGEMENT: A review of the empirical literature. **Annual Review of Political Science**, Annual Reviews, v. 7, n. 1, p. 315–344, may 2004.

CARR, S. J.; LASSITER, A. Big data, small apps: Premises and products of the civic hackathon. In: **Springer Geography**. [S.l.]: Springer International Publishing, 2016. p. 543–559.

COBHAM, A.; JANSKÝ, P. Measuring misalignment: The location of US multinationals' economic activity versus the location of their profits. **Development Policy Review**, Wiley, jul 2017.

CORNWALL, A. Unpacking participation: models, meanings and practices. **Community Development Journal**, Oxford University Press (OUP), v. 43, n. 3, p. 269–283, jun 2008.

DESAI, M. A.; FOLEY, C. F.; HINES, J. R. The demand for tax haven operations. **Journal of Public Economics**, Elsevier BV, v. 90, n. 3, p. 513–531, feb 2006.

DINTER, B.; KOLLWITZ, C.; FRITZSCHE, A. Teaching data driven innovation – facing a challenge for higher education. In: **Proceedings of the 23rd Americas Conference on Information Systems (AMCIS)**. [S.l.: s.n.], 2017.

DOWBOR, L. A **Era do Capital Improdutivo (Em Portugues do Brasil)**. [s.n.], 2017. ISBN 978-85-69536-11-6. Disponível em: <<https://www.amazon.com/Era-Capital-Improdutivo-Portugues-Brasil/dp/8569536119?SubscriptionId=AKIAIOBINVZYXZQZ2U3Atag=chimbori05-20linkCode=xm2camp=2025creative=165953creativeASIN=8569536119>>.

FLORES, M. et al. How can hackathons accelerate corporate innovation? In: **IFIP Advances in Information and Communication Technology**. [S.l.]: Springer International Publishing, 2018. p. 167–175.

FREY, F. J.; LUKS, M. The innovation-driven hackathon. In: **Proceedings of the 21st European Conference on Pattern Languages of Programs - EuroPlop 16**. [S.l.]: ACM Press, 2016.

GAMA, K. Preliminary findings on software engineering practices in civic hackathons. In: **2017 IEEE/ACM 4th International Workshop on CrowdSourcing in Software Engineering (CSI-SE)**. [S.l.]: IEEE, 2017.

GIL, A. **Estudo de Caso - Fundamentação Científica; Subsídios para Coleta e Análise de Dados; Como Redigir o Relatório**. [S.l.]: Atlas, 2009.

GOTTSCHALK, P. **Investigation and Prevention of Financial Crime: Knowledge Management, Intelligence Strategy and Executive Leadership**. Routledge, 2010. ISBN 978-1-4094-0331-9. Disponível em: <<https://www.amazon.com/Investigation-Prevention-Financial-Crime-Intelligence/dp/1409403319?SubscriptionId=0JYN1NVW651KCA56C102tag=techkie-20linkCode=xm2camp=2025creative=165953creativeASIN=1409403319>>.

GOTTSCHALK, P.; SOLLI-SÆTHER, H. Financial crime in business organizations: an empirical study. **Journal of Financial Crime**, Emerald, v. 18, n. 1, p. 76–92, jan 2011.

HOWE, J. **Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business**. 1. ed. New York, NY, USA: Crown Publishing Group, 2008. ISBN 0307396207, 9780307396204.

IRANI, L. Hackathons and the making of entrepreneurial citizenship. **Science, Technology, & Human Values**, SAGE Publications, v. 40, n. 5, p. 799–824, apr 2015.

JOHNSON, P.; ROBINSON, P. Civic hackathons: Innovation, procurement, or civic engagement? **Review of Policy Research**, Wiley, v. 31, n. 4, p. 349–357, jul 2014.

JONES, C.; TEMOURI, Y.; COBHAM, A. Tax haven networks and the role of the big 4 accountancy firms. **Journal of World Business**, Elsevier BV, v. 53, n. 2, p. 177–193, feb 2018.

KARLSEN, J.; LØVLIE, A. S. ‘you can dance your prototype if you like’: independent filmmakers adapting the hackathon. **Digital Creativity**, Informa UK Limited, v. 28, n. 3, p. 224–239, jul 2017.

KOHLI, M. et al. Creation and curation of the society of imaging informatics in medicine hackathon dataset. **Journal of Digital Imaging**, Springer Nature, v. 31, n. 1, p. 9–12, jul 2017.

KOMSSI, M. et al. What are hackathons for? **IEEE Software**, Institute of Electrical and Electronics Engineers (IEEE), v. 32, n. 5, p. 60–67, sep 2015.

KOU, Y. et al. Survey of fraud detection techniques. In: **IEEE International Conference on Networking, Sensing and Control, 2004**. [S.l.]: IEEE, 2004.

LEEuw, E. D. de. **International Handbook of Survey Methodology**. [S.l.]: Routledge, 2012.

LENZ, H. J. Data fraud detection: A first general perspective. In: **Enterprise Information Systems**. [S.l.]: Springer International Publishing, 2015. p. 14–35.

LINDERS, D. From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. **Government Information Quarterly**, Elsevier BV, v. 29, n. 4, p. 446–454, oct 2012.

MöLLER, S. et al. Community-driven development for computational biology at sprints, hackathons and codefests. **BMC Bioinformatics**, Springer Nature, v. 15, n. Suppl 14, p. S7, 2014.

MORIN, F. **L’hydre mondiale: L’oligopole bancaire (French Edition)**. Lux Éditeur, 2015. ISBN 9782895966784. Disponível em: <<https://www.amazon.com/Lhydre-mondiale-Loligopole-bancaire-French-ebook/dp/B00W2DWSU8?SubscriptionId=0JYN1NVW651KCA56C102tag=techkie-20linkCode=xm2camp=2025creative=165953creativeASIN=B00W2DWSU8>>.

MUNRO, D. Hosting hackathons a tool in retaining students with beneficial side effects. **J. Comput. Sci. Coll.**, Consortium for Computing Sciences in Colleges, USA, v. 30, n. 5, p. 46–51, maio 2015. ISSN 1937-4771. Disponível em: <<http://dl.acm.org/citation.cfm?id=2752981.2752994>>.

NANDI, A.; MANDERNACH, M. Hackathons as an informal learning platform. In: **Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE 16**. [S.l.]: ACM Press, 2016.

NETO, A. S. et al. First brazilian datathon in critical care. **Revista Brasileira de Terapia Intensiva**, GN1 Genesis Network, v. 30, n. 1, 2018.

NGAI, E. et al. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. **Decision Support Systems**, Elsevier BV, v. 50, n. 3, p. 559–569, feb 2011.

OLSON, K. R. et al. Health hackathons: theatre or substance? a survey assessment of outcomes from healthcare-focused hackathons in three countries. **BMJ Innovations**, BMJ Specialist Journals, v. 3, n. 1, p. 37–44, 2017. ISSN 2055-8074. Disponível em: <<https://innovations.bmj.com/content/3/1/37>>.

PARÉ, G. Investigating information systems with positivist case research. **Communications of the Association for Information Systems**, Association for Information Systems, v. 13, 2004.

PATTON, M. Q. Enhancing the quality and credibility of qualitative analysis. **Health services research**, v. 34, p. 1189–1208, dez. 1999. ISSN 0017-9124.

PE-THAN, E. P. P. et al. Designing corporate hackathons with a purpose: The future of software development. **IEEE Software**, Institute of Electrical and Electronics Engineers (IEEE), v. 36, n. 1, p. 15–22, jan 2019.

PER HOST, B. R. M. H. A. R. R. **Software Engineering**. John Wiley & Sons, 2012. ISBN 1118104358. Disponível em: <https://www.ebook.de/de/product/16222197/per_runeson_host_b_jorn_rgnell_martin_host_austen_rainer_software>

PERNG, S.-Y.; KITCHIN, R.; DONNCHA, D. M. Hackathons, entrepreneurship and the passionate making of smart cities. Center for Open Science, apr 2017.

PICKETT. **Financial Crime Investigation**. John Wiley & Sons, 2002. ISBN 0471203351. Disponível em: <https://www.ebook.de/de/product/3669336/pickett_financial_crime_investigation.html>.

RAD, M. S.; SHAHBAHRAMI, A. Detecting high risk taxpayers using data mining techniques. In: **2016 2nd International Conference of Signal Processing and Intelligent Systems (ICSPIS)**. [S.l.]: IEEE, 2016.

RALHA, C. G.; SILVA, C. V. S. A multi-agent data mining system for cartel detection in brazilian government procurement. **Expert Systems with Applications**, Elsevier BV, v. 39, n. 14, p. 11642–11656, oct 2012.

RICHARD, G. T. et al. StitchFest. In: **Proceedings of the 46th ACM Technical Symposium on Computer Science Education - SIGCSE 15**. [S.l.]: ACM Press, 2015.

RICHTERICH, A. Hacking events. **Convergence: The International Journal of Research into New Media Technologies**, SAGE Publications, p. 135485651770940, may 2017.

ROSELL, B.; KUMAR, S.; SHEPHERD, J. Unleashing innovation through internal hackathons. **2014 IEEE Innovations in Technology Conference**, p. 1–8, 2014.

RUNESON, P.; HÖST, M. Guidelines for conducting and reporting case study research in software engineering. **Empirical Software Engineering**, Springer Nature, v. 14, n. 2, p. 131–164, dec 2008.

SALINAS, M. N.; EMER, M. C. F. P.; NETO, A. G. S. S. Short datathon for the interdisciplinary development of data analysis and visualization skills.

SILLANPÄÄ, A.; KOIVULA, T. Mapping conflict research: A bibliometric study of contemporary scientific discourses¹. **International Studies Perspectives**, Oxford University Press (OUP), v. 11, n. 2, p. 148–171, may 2010.

TRAINER, E. H. et al. How to hackathon: Socio-technical tradeoffs in brief, intensive collocation. In: **Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW 16**. [S.l.]: ACM Press, 2016.

VITALI, S.; GLATTFELDER, J. B.; BATTISTON, S. The network of global corporate control. **PLoS ONE**, Public Library of Science (PLoS), v. 6, n. 10, p. e25995, oct 2011.

WEST, J.; BHATTACHARYA, M. Intelligent financial fraud detection: A comprehensive review. **Computers & Security**, Elsevier BV, v. 57, p. 47–66, mar 2016.

YIN, R. K. **Estudo De Caso (Em Portuguese do Brasil)**. Bookman, 2001. ISBN 85-7307-852-9. Disponível em: <<https://www.amazon.com/Estudo-Caso-Em-Portuguese-Brasil/dp/8573078529?SubscriptionId=AKIAIOBINVZYXZQZ2U3Atag=chimbóri05-20linkCode=xm2camp=2025creative=165953creativeASIN=8573078529>>.

ZHUHADAR, L.; CIAMPA, M. Leveraging learning innovations in cognitive computing with massive data sets: Using the offshore panama papers leak to discover patterns. **Computers in Human Behavior**, Elsevier BV, dec 2017.

APÊNDICE A – QUESTIONÁRIOS DE LEVANTAMENTO INICIAL

Na seção 4.1.1 deste trabalho menciona a fase de levantamento inicial por meio de questionários. Os mesmos tiveram o foco de levantar o interesse de um público diverso como mulheres, estudantes e profissionais de outras áreas, além de pessoas já acostumadas a eventos *Hackathon e Datathon*. A seguir, apresentamos um resumo dos diferentes questionários aplicados.

A.1 QUESTIONÁRIO APLICADO NO "FROM LADIES TO LADIES"

A seguir está o questionário completo que aplicamos na segunda edição do *From Ladies to Ladies*¹, no dia 09 de dezembro de 2017. Este evento esteve composto por três frentes: palestras, *workshops* e mesas redondas; que discutam temas alinhados aos propósitos da *Tech Ladies Brasil*², principalmente o empoderamento feminino na tecnologia (Figuras 26 e 27).

¹<http://archive.is/vSTq1>. Último acesso: 19/11/2018

²<http://www.techladies.com.br/>. Último acesso: 19/11/2018



Qual é o tipo de empresa em que você trabalha? (Escolher só uma opção)

Tecnologia	<input type="checkbox"/>
Financeira	<input type="checkbox"/>
Comercio	<input type="checkbox"/>
Saúde	<input type="checkbox"/>
Industrial	<input type="checkbox"/>
Telecomunicações	<input type="checkbox"/>
Telefonia	<input type="checkbox"/>
Setor público	<input type="checkbox"/>
Educação	<input type="checkbox"/>
Outros (mencionar).....	<input type="checkbox"/>

Qual é sua atividade principal (Escolher só uma opção, a principal)

Desenvolvedor/Programador	<input type="checkbox"/>
Professor de universidade ou ensino médio	<input type="checkbox"/>
Líder de projetos	<input type="checkbox"/>
Freelancer	<input type="checkbox"/>
Diretor/Gerente	<input type="checkbox"/>
Estudante	<input type="checkbox"/>
Pesquisador	<input type="checkbox"/>
Estagiário	<input type="checkbox"/>
Outros (mencionar).....	<input type="checkbox"/>

Qual é sua faixa etária? (Escolher só uma opção)

15-19 anos	<input type="checkbox"/>
20-29 anos	<input type="checkbox"/>
30-39 anos	<input type="checkbox"/>
40-49 anos	<input type="checkbox"/>
50 e mais	<input type="checkbox"/>

Quais são os tipos de eventos de que você já participou? (Pode escolher mais de uma opção)

Palestras/Conferencias	<input type="checkbox"/>
Maratona de Programação	<input type="checkbox"/>
Hackathon	<input type="checkbox"/>
Hackathon cívicos/sociais (smart city, anti corrupção, etc.)	<input type="checkbox"/>
Outros eventos de ativismo como Fiscalize.me	<input type="checkbox"/>
Meetups	<input type="checkbox"/>
Outros (mencionar).....	<input type="checkbox"/>

Sexo

Especificar:.....

Qual é a frequência que você participa em eventos (congressos, conferências, Meetups, etc.)? (Escolher só uma opção)

Nunca (esta é primeira vez)	<input type="checkbox"/>
Ocasionalmente(uma vez a cada 2 anos ou mais)	<input type="checkbox"/>
Frequentemente (uma vez por ano)	<input type="checkbox"/>
Sempre (vários por ano)	<input type="checkbox"/>

Figura 26: Primeira página do questionário aplicado no "From Ladies to Ladies"



Quais são as motivações para participar do evento: From Ladies to Ladies? (Pode marcar várias opções)

- Temas abordados no evento são de meu interesse
- Lugar do evento é atraente
- Possibilidade de fazer novos contatos
- Conheço a organização e me identifico com ela
- Possibilidade de ganhar brindes
- Outros (mencionar).....

Quais são os fatores que poderiam limitar sua participação num evento qualquer? (Pode marcar várias opções)

- Custo do evento
- Lugar do evento fica longe
- Os organizadores do evento não são conhecidos
- Horários não acessíveis
- Duração do evento é extensa
- Tenho filhos pequenos
- Outros (mencionar).....

Se você ainda não participou em algum evento de ativismo como Hackathons cívicos/sociais ou outros similares. Você gostaria de participar?

- Sim
- Não

Se já participou em algum evento de ativismo como Hackathons cívicos/sociais ou outros similares. Você gostou de participar?

- Sim
- Não

Quais são as motivações para participar de eventos de Ativismo cívico/social? (Pode marcar várias opções)

- Temas abordados no evento são de meu interesse
- Possibilidade de usar meu conhecimento para abordar temas de interesse social
- Possibilidade de fazer novos contatos
- Conheço a organização e me identifico com ela
- Ser reconhecido pelo ativismo social
- Não tenho interesse em participar deste tipo de eventos
- Outros (mencionar).....

Figura 27: Segunda página do questionário aplicado no "From Ladies to Ladies"

A.1.1 RESULTADO DO QUESTIONÁRIO APLICADO NO "FROM LADIES TO LADIES"

O principal objetivo da aplicação de questionário neste evento foi descobrir o nível de interesse em diferentes tipos de eventos, além de aproveitar a presença majoritária de mulheres e identificar possíveis fatores que poderiam limitar sua participação num evento qualquer.

Foram consultadas 34 participantes, com prevalência de mulheres pela natureza do evento. A maioria respondeu que participam de palestras e conferências (Figura 28). Os itens com menor quantidade de escolha foram *Hackathon*, *Hackathon* cívicos/sociais, maratona de programação e outros eventos de ativismo. Além disso, mencionaram que os fatores que poderiam limitar sua participação num evento qualquer são o horário e o custo (Figura 29).

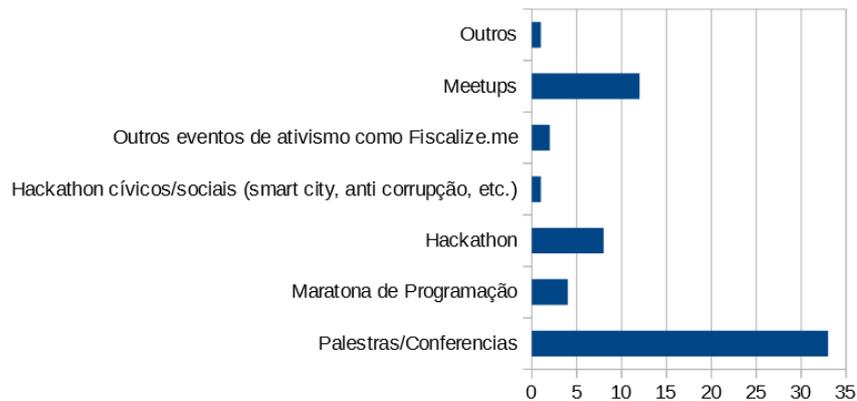


Figura 28: Tipos de eventos de que você já participou. Fonte: Autoria Própria

Outras das questões abordadas no questionário foi o interesse em participar num *Hackathon* cívico/social e a maioria respondeu que gostaria de participar. Com relação ao motivo especificaram que o tema é do seu interesse, gostam da possibilidade de compartilhar o seu conhecimento e fazer novos contatos.



Figura 29: Fatores que poderiam limitar sua participação num evento qualquer. Fonte: Autoria Própria

A.2 QUESTIONÁRIO APLICADO NO *GLOBAL GAME JAM* 2017

A seguir está o questionário completo aplicado no *Global Game Jam*³ (GGJ) (Figuras 30, 31, 32, 33, 34 e 35).

Global Game Jam - Questionário

Prezado amigo, o objetivo deste questionário é um estudo sobre métodos colaborativos de desenvolvimento de softwares. Com estes estudo será possível obter um perfil dos participantes de eventos com o Game Jam.

Este questionário está dividido em duas partes, a primeira é para podermos ter um perfil dos participantes do evento, e a segunda parte para obter a sua percepção sobre o evento.

Ao terminar, basta clicar em enviar.

Muito obrigado por sua colaboração.

Questionário 1 - Perfil do participante

1. Qual sua idade?

2. Sexo

3. Qual sua atual ocupação? (Trabalha, estudo, ambos ou a procura de novas oportunidades)

4. Já tem experiência com desenvolvimento de software? Marque uma ou mais alternativas.

Check all that apply.

- Sim, como Hobby.
 Sim, em Empresa.
 Sim, na academia (Faculdade, ensino técnico, cursos).
 Não.

5. Assinale o item que mais se aplicam ao seu grau de experiência em programação:

Mark only one oval.

- Iniciante.
 Básico.
 Intermediário.
 Avançado.
 Ainda não tenho experiência.
-

Figura 30: Questionário aplicado no GGJ 2017

³<http://www.ggjcw.com/2016/10/ggjcw2017/>. Último acesso: 19/11/2018

6. Já utilizou, ou utiliza, alguma ferramenta para desenvolvimento? Qual? (Marque uma ou mais).

Check all that apply.

- Python.
- Java.
- C, C++, C#.
- PHP.
- JavaScript.
- Perl.
- Ruby.
- Assembly.
- Unreal
- Unity
- Nenhum.
- Other: _____

7. Já teve alguma experiência com: (Marque uma ou mais)

Check all that apply.

- Programação em pares.
- Desenvolvimento orientado a testes.
- Métodos ágeis de programação.
- Refatoração de código.
- Nenhum.

8. Você já participou em um (ou mais) dos eventos listados abaixo?

Check all that apply.

- Coderetreats.
- Coding Dojo.
- Hackathons.
- Maratona de Programação.
- Olimpíada de Informática.
- Game Jam
- Meetups sobre alguma linguagem de programação ou ferramentas de programação.
- Nenhum.

Figura 31: Questionário aplicado no GGJ 2017

9. Qual foi (ou quais foram) suas funções na sua equipe neste Game Jam?

10. Já participou de outros Games Jam? Caso tenha participado, quantos?

Questionário 2 - Percepção sobre o evento

11. Gostei de participar do Game Jam.

Mark only one oval.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

12. Game Jam me ajudou a aprender mais sobre programação em geral.

Mark only one oval.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

13. Tive facilidade em entender a dinâmica do Game Jam.

Mark only one oval.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo

Figura 32: Questionário aplicado no GGJ 2017

14. **Compreendi o funcionamento das linguagens de programação utilizadas no Game Jam.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

15. **Apreendi sobre programação por meio do Game Jam.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

16. **Meu nível de conhecimento sobre programação melhorou significativamente.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

17. **O Game Jam promoveu momentos de cooperação entre as pessoas que participaram.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

18. **Eu me diverti neste Game Jam.**

Mark only one oval.

Figura 33: Questionário aplicado no GGJ 2017

19. **Eu recomendaria o Game Jam para meus conhecidos.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

20. **Gostaria de participar de um Game Jam novamente.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

21. **A experiência no Game Jam vai contribuir para meu desempenho em desenvolvimento de software.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

22. **Compreendi como funciona a programação em pares e/ou em grupo.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

Figura 34: Questionário aplicado no GGJ 2017

23. **Acredito que a programação em pares, e/ou em grupo, torna o desenvolvimento mais eficiente.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

24. **Game Jam está adequado à minha forma de aprender.**

Mark only one oval.

- Discordo fortemente
- Discordo
- Indiferente
- Concordo
- Concordo fortemente

Figura 35: Questionário aplicado no GGJ 2017

O GGJ é um evento mundial no qual desenvolvedores de jogos e interessados reúnem-se em locais ao redor do mundo, formam grupos e cada grupo colabora na criação de um jogo em 48 horas a partir de um tema comum. Uma das sedes do evento foi na Pontifícia Universidade Católica do Paraná.

O objetivo da aplicação de questionário neste evento também foi levantar o nível de interesse em diferentes tipos de eventos focado em outro público alvo.

No total 27 participantes responderam ao questionário com prevalência de 81,5% de homens. Com relação ao tipo de evento de que participam com frequência mencionaram que são *Game Jam* e *Hackathon* (Figura 36).

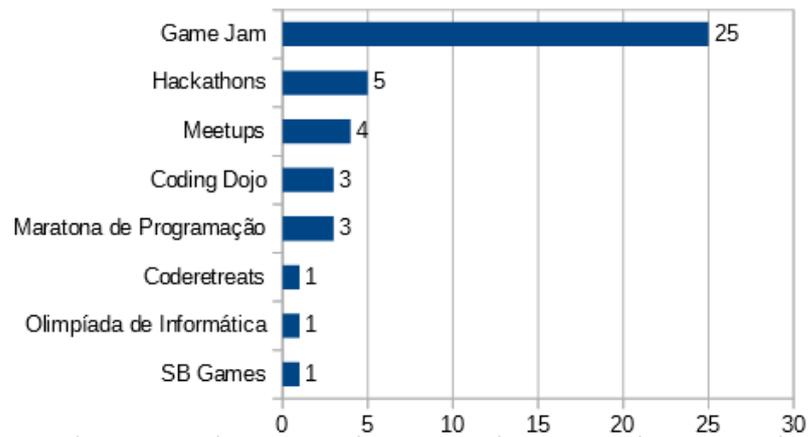


Figura 36: Participou em um (ou mais) dos eventos listados abaixo. Fonte: Autoria Própria

Incorporamos uma série de questões do tipo escala de *Likert* de 5 pontos para obter informação adicional. Nesta escala os respondentes de um questionário não apenas respondem se concordam ou não com as afirmações, mas também informam qual é o grau de concordância ou discordância.

Como resultado conseguimos respostas bem interessantes, como por exemplo a maioria concordam fortemente com as seguintes afirmações:

- O *Game Jam* promoveu momentos de cooperação entre as pessoas que participaram.
- Eu me diverti neste *Game Jam*.
- Eu recomendaria o *Game Jam* para meus conhecidos.
- A experiência no *Game Jam* vai contribuir para meu desempenho em desenvolvimento de software.

- Acredito que a programação em pares, e/ou em grupo, torna o desenvolvimento mais eficiente.

A.3 QUESTIONÁRIO APLICADO NA COMUNIDADE *HACKATHON* BRASIL

Com o intuito de estender a consulta para outras pessoas, preparamos um questionário *online* e pesquisamos algumas comunidades que tinham relação com eventos do tipo *Hackathon*. Encontramos uma comunidade com 1740 membros na rede social *Facebook* chamada *Hackathons* Brasil que tem como objetivo centralizar informações de Hackathons do Brasil inteiro e solicitamos a publicação do questionário para os membros (Figuras 37 e 38).

Hackathon sobre Dívida Pública, Evasão Fiscal, etc.

O PPGCA-UTFPR, através do grupo de pesquisa de Engenharia de Software, agradece sua atenção e seu auxílio ao avanço nesta pesquisa.

O objetivo deste estudo é investigar questões relacionadas a eventos de ativismo cívico social. Esses eventos visam estimular o engajamento cívico, reunir designers de software, desenvolvedores e organizadores comunitários para resolver as necessidades de suas comunidades e mostrar o que é possível obter em ganhos para a sociedade ao usar a tecnologia .

OK

1. Qual é sua faixa etária? (Escolher só uma opção)

- 15-19 anos 40-49 anos
 20-29 anos 50 e mais
 30-39 anos

2. Sexo

- Masculino
 Feminino
 Outros

3. Qual é sua atividade principal (Escolher só uma opção, a principal)

- Desenvolvedor/Programador Diretor/Gerente
 Professor de universidade ou ensino médio Estudante
 Líder de projetos Pesquisador
 Freelancer Estagiário
 Outro (especifique)

Figura 37: Primeira página do questionário aplicado na Comunidade *Hackathon* Brasil

4. Quais são os tipos de eventos de que você já participou? (Pode escolher mais de uma opção)

- Palestras/Conferencias Coderetreats
 Maratona de Programação Coding Dojo
 Hackathon Nenhum
 Meetups
 Outro (especifique)

5. Quais são os tipos de **Hackathon** que você já participou? (Pode escolher mais de uma opção)

- Game Jam Hackathon Privados (Organizados por empresas privadas)
 Hackathon cívicos/sociais (Smart City, Anti Corrupção, Saúde, etc) Nenhum
 Outro (especifique)

6. Participaria de uma Hackathon focado em temas como: Dívida pública, evasão fiscal, etc.?

- Sim
 Não
 Não sei

7. Se respondeu NÃO. Por que? (Pode escolher mais de uma opção)

- Temas abordados NÃO são de meu interesse Acho muito difícil de compreender
 Temas abordados NÃO são de meu conhecimento (Economia, Finanças, Políticas Públicas etc)
 Outro (especifique)

Figura 38: Segunda página do questionário aplicado na Comunidade *Hackathon* Brasil

A.3.1 RESULTADO DO QUESTIONÁRIO APLICADO NA COMUNIDADE *HACKATHON* BRASIL

Recebemos um total de 52 respostas no período de tempo de 07/05/2018 a 30/07/2018. Uma porcentagem de 73,08% de pessoas do gênero masculino e 26,92% feminino com prevalência de desenvolvedores.

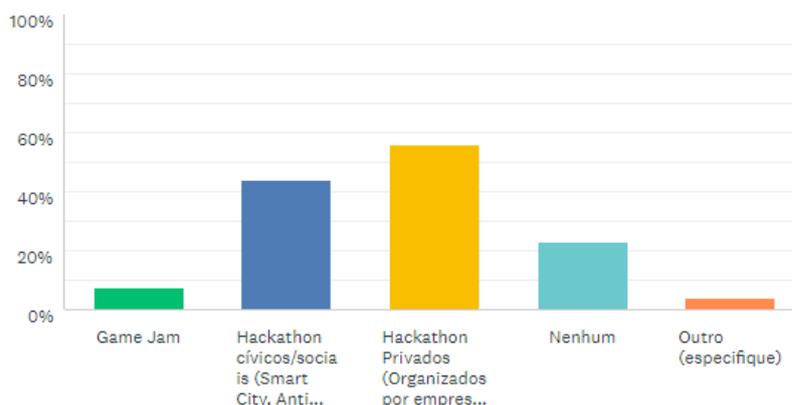


Figura 39: Tipos de Hackathon que você já participou. Fonte: Autoria Própria

Como esperávamos a maioria já tinha participado de eventos *Hackathon* mas especificamente aqueles organizados por empresas privadas. Também se visualiza 44,23% de pessoas que participou de *Hackathons* cívicos/sociais (*Smart City*, Anti Corrupção, Saúde, etc) (Figura 39). Na última pergunta consultamos o interesse num evento *Hackathon* focado a temas como Dívida Pública, Evasão Fiscal, etc.: 84,62% dos respondentes do questionário têm interesse em participar (Figura 40).

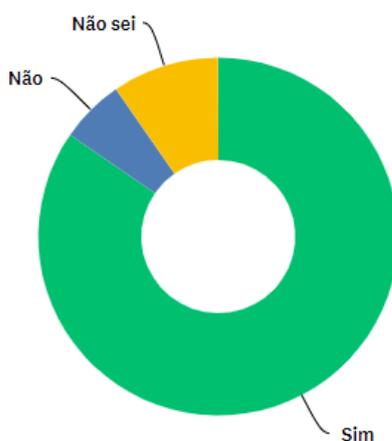


Figura 40: Participaria de um Hackathon focado em temas como: Dívida pública, evasão fiscal, etc.?. Fonte: Autoria Própria

A.4 QUESTIONÁRIO APLICADO NO II FÓRUM GESTÃO E ECONOMIA DA UTFPR

O II Fórum Gestão e Economia da UTFPR⁴ foi outro evento aproveitado para aplicar o questionário. Nos dias 22, 23 e 24 de agosto de 2018 foram realizadas uma série de atividades com a finalidade de promover maior integração e interação do conhecimento na área de Gestão e Economia (Figuras 41 e 42).

⁴<http://archive.is/hzFfX>. Último acesso: 19/11/2018

**Prezados,**

O PPGCA-UTFPR, através do grupo de pesquisa de Engenharia de Software, solicita sua atenção e seu auxílio nesta pesquisa.

O objetivo deste estudo é investigar questões relacionadas a eventos de ativismo cívico social. Esses eventos visam estimular o engajamento cívico, reunir desenvolvedores e profissionais de diversas áreas para resolver as necessidades de suas comunidades e mostrar o que é possível obter em ganhos para a sociedade ao usar a tecnologia.

O uso dos registros efetuados durante a pesquisa é estritamente limitado à mesma e o anonimato dos participantes será preservado.

Qual é sua área de trabalho ou estudo?
(Pode escolher mais de uma opção)

- Economia
- Administração
- Comércio
- Contabilidade
- Direito
- Comunicação
- Educação
- Computação
- Outros (mencionar).....

Qual é sua atividade principal (Escolher só uma opção, a principal)

- Estudante
- Professor
- Líder de projetos
- Freelancer
- Diretor/Gerente
- Empregado
- Pesquisador
- Outros (mencionar).....

Qual é sua faixa etária? (Escolher só uma opção)

- 15-19 anos
- 20-29 anos
- 30-39 anos
- 40-49 anos
- 50 e mais

Quais são os tipos de eventos de que você já participou? (Pode escolher mais de uma opção)

- Palestras/Conferências
- Meetups
- Hackathon públicos/privados
- Outros (mencionar).....

Gênero

- Feminino
- Masculino
- Prefiro não informar
- Prefiro me descrever.....

Qual é a frequência com que você participa em eventos Palestras, Conferências, Meetups, etc.)? (Escolher só uma opção)

- Nunca (esta é primeira vez)
- Ocasionalmente(uma vez a cada 2 anos ou mais)
- Frequentemente (uma vez por ano)
- Sempre (vários por ano)

Figura 41: Primeira página do questionário aplicado no II Fórum Gestão e Economia



Você conhece ou ouviu falar do evento Hackathon?

- Sim
 Não

Se respondeu SIM à pergunta anterior: Você gostaria de participar de algum evento que trate do tema Evasão Fiscal e suas consequências?

- Sim
 Não

Se você tem interesse em colaborar com esta pesquisa deixe seu e-mail :

Assinatura do participante

Figura 42: Segunda página do questionário aplicado no II Fórum Gestão e Economia

A.4.1 RESULTADO DO QUESTIONÁRIO APLICADO NO II FÓRUM GESTÃO E ECONOMIA DA UTFPR

O público alvo do evento foram estudantes e profissionais das áreas de Administração, Contabilidade, Economia, Gestão Pública e outros convidados. O interesse em aplicar o questionário com pessoas de outras áreas foi para explorar uma possível integração de áreas de conhecimento num evento *Hackathon*, considerando temas (Dívida Pública, Evasão Fiscal, etc.) estreitamente relacionados.

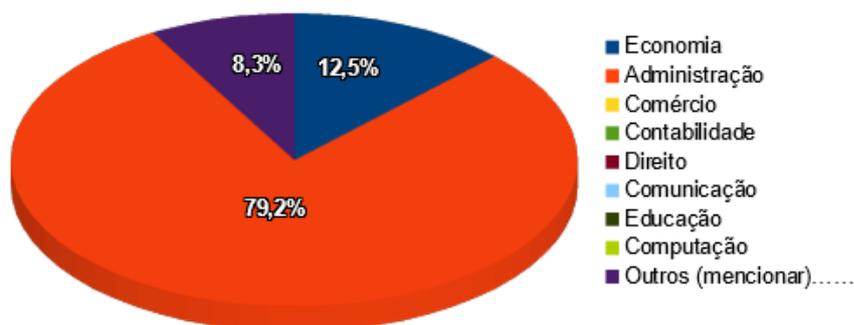


Figura 43: Área de trabalho ou estudo. Fonte: Autoria Própria

De um total de 24 respondentes, a maioria mencionou ser da área de Administração e Economia, com prevalência de estudantes (Figura 43). O tipo de evento que normalmente frequentam são as Palestras e Conferências mas 42% mencionou conhecer ou ouvir falar sobre o evento *Hackathon* (Figura 44). Dessas pessoas 64% indicaram que gostariam participar de uma *Hackathon* sobre temas Dívida Pública, Evasão Fiscal, etc.

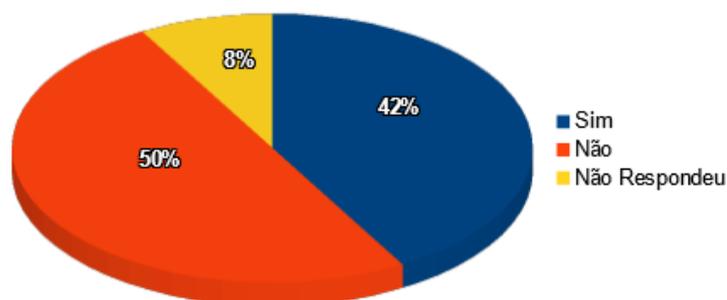


Figura 44: Conhece ou ouviu falar do evento Hackathon. Fonte: Autoria Própria

A.5 QUESTIONÁRIO APLICADO NO DATATHON UFRGS

O questionário também foi aplicado no 1º *Datathon* da UFRGS⁵ realizado nos dias 15, 16 e 17 de outubro de 2018 (Figuras 45, 46 e 47).

⁵<https://www.ufrgs.br/datathon/index.html>. Último acesso: 19/11/2018

**Prezados,**

O PPGCA-UTFPR, através do grupo de pesquisa de Engenharia de Software, solicita sua atenção e seu auxílio nesta pesquisa.

O objetivo deste estudo é investigar questões relacionadas a eventos de ativismo cívico social. Esses eventos visam estimular o engajamento cívico, reunir desenvolvedores e profissionais de diversas áreas para resolver as necessidades de suas comunidades e mostrar o que é possível obter em ganhos para a sociedade ao usar a tecnologia.

O uso dos registros efetuados durante a pesquisa é estritamente limitado à mesma e o anonimato dos participantes será preservado.

Qual é sua área de trabalho ou estudo?
(Pode escolher mais de uma opção)

- Economia
- Administração
- Comércio
- Contabilidade
- Direito
- Comunicação
- Educação
- Computação
- Outros (mencionar).....

Qual é sua atividade principal (Escolher só uma opção, a principal)

- Estudante
- Professor
- Líder de projetos
- Freelancer
- Diretor/Gerente
- Empregado
- Pesquisador
- Outros (mencionar).....

Qual é sua faixa etária? (Escolher só uma opção)

- 15-19 anos
- 20-29 anos
- 30-39 anos
- 40-49 anos
- 50 e mais

Quais são os tipos de eventos de que você já participou? (Pode escolher mais de uma opção)

- Palestras/Conferências
- Meetups
- Hackathon públicos/privados
- Datathon públicos/privados
- Outros (mencionar).....

Gênero

- Feminino
- Masculino
- Prefiro não informar
- Prefiro me descrever.....

Qual é a frequência com que você participa em eventos Hackathon? (Escolher só uma opção)

- Nunca
- Ocasionalmente(uma vez a cada 2 anos ou mais)
- Frequentemente (uma vez por ano)
- Sempre (vários por ano)

Figura 45: Primeira página do questionário aplicado no Datathon UFRGS

Qual é a frequência com que você participa em eventos Datathon? (Escolher só uma opção)

- Nunca (esta é primeira vez)
 Ocasionalmente (uma vez a cada 2 anos ou mais)
 Frequentemente (uma vez por ano)
 Sempre (vários por ano)

Percepção sobre um Datathon (Escolher só uma opção de cada afirmação)

<p>1- Gostei de participar do Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>2- Participar deste Datathon me ajudou a compreender mais sobre análise de dados e visualização.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>
<p>3- Tive facilidade em entender a dinâmica do Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>4- Compreendi o funcionamento das ferramentas utilizadas no Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>
<p>5- Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>6- Participar deste Datathon promoveu momentos de cooperação entre as pessoas que participaram.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>

Figura 46: Segunda página do questionário aplicado no Datathon UFRGS



7- A interação com pessoas de outras áreas do meu time foi boa e contribuiu positivamente no resultado final.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

8- Acredito que a experiência com pessoas de outras áreas num evento Datathon torna a análise de dados mais eficiente.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

Outras considerações sobre o Datathon:

Se você tem interesse em colaborar com esta pesquisa a través de uma entrevista deixe seu e-mail :

--

Assinatura do participante

Figura 47: Terceira página do questionário aplicado no Datathon UFRGS

A.5.1 RESULTADO DO QUESTIONÁRIO APLICADO NO DATATHON UFRGS

O questionário foi aplicado no dia da apresentação dos projetos, porém foram respondidos pelos participantes que estiveram presentes (no total 6). Foi enviado por e-mail o questionário online para as pessoas restantes que não estiveram no dia e 3 delas responderam.

O evento contou com a participação de estudantes de graduação e pós graduação da matemática, estatística e ciência da computação, divididos em quatro equipes. Com relação a frequência de participação em eventos Hackathon e Datathon a maioria mencionou que foi sua primeira participação (Figuras 48 e 49).



Figura 48: Frequência de participação em eventos *Hackathon*. Fonte: Autoria Própria

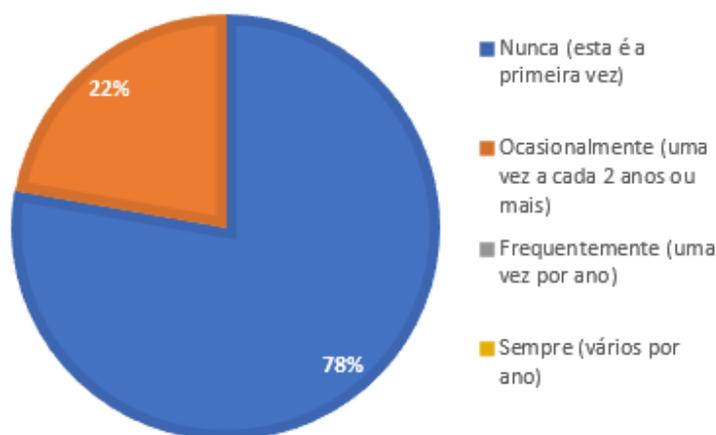


Figura 49: Frequência de participação em eventos *Datathon*. Fonte: Autoria Própria

Uma seção do questionário foi desenvolvida com afirmações e opções de escolha do tipo escala de *Likert*. Segundo os resultados pode-se visualizar a prevalência dos valores “Concordo” e “Concordo fortemente” em cada afirmação (Tabela 19).

Tabela 19: Resultados das Afirmações. Fonte: Autoria Própria

Escala:					
Discordo fortemente	Discordo	Indiferente	Concordo	Concordo fortemente	
Gostei de participar do Datathon	0		9		
O Datathon me ajudou a compreender mais sobre análise de dados e visualização.	0	2		7	
Tive facilidade em entender a dinâmica do Datathon	0		4	5	
Compreendi o funcionamento das ferramentas utilizadas no Datathon	0		4	5	
Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente	0	1		6	2
O Datathon promoveu momentos de cooperação entre as pessoas que participaram	0		4	5	
A interação com pessoal de outras áreas do meu time foi boa e contribuiu positivamente no resultado final	0		5	4	
Acredito que a experiência com pessoal de outras áreas num evento Datathon, torna a análise de dados mais eficiente	0		4	5	

A.5.2 CONSIDERAÇÕES SOBRE O LEVANTAMENTO COM QUESTIONÁRIO

A aplicação dos questionários permitiram identificar aspectos a considerar para continuar o projeto de formato de evento:

- 1.As mulheres que preencheram o questionário tem preferência em participar de eventos do tipo Palestras e Conferências pela flexibilidade no horário;
- 2.As mulheres gostam de temas relacionados a aspectos sociais e ativismo;
- 3.Os eventos do tipo maratona que envolvem *games* tem um público bem variado quanto à idade, na maioria homens. Além disso, a preferência de participação é somente nesse tipo de eventos. Outros aspectos importantes foram levantamos por meio das afirmações na escala de Likert A importancia da colaboração entre pessoas, a diversão e a experiência através do trabalho em equipe foram ressaltados.
- 4.As pessoas que frequentam os eventos do tipo *Hackathon* e preencheram o questionário são na maioria do gênero masculino. Elas participaram de *Hackathons* tanto privados

como civicos/sociais e manifestaram interesse em temas como evasão fiscal e dívida pública.

- 5.As pessoas de outras áreas de conhecimento como Economia e Administração não costumam frequentar eventos do tipo *Hackathon* mas tem interesse em temas como evasão fiscal e dívida pública.
- 6.As pessoas da área da Estatística e Matemática Aplicada que participaram de um *Datathon* gostaram do evento e acharam fácil de compreender. Além disso, manifestaram que a experiência com pessoal de outras áreas torna a análise mais eficiente.

APÊNDICE B – EXPLORAÇÃO DE EVENTOS *HACKATHON* E *DATATHON*

Como foi mencionado na seção 4.1 o trabalho de pesquisa teve uma fase de exploração de eventos na qual realizamos acompanhamentos *in situ* de um *Hackathon* e de um *Datathon*. A seguir, apresentamos um relatório de cada evento.

B.1 RELATÓRIO DO *HACKATHON* COMPRAS PÚBLICAS SP

Nos dias sábado e domingo, 5 e 6 de maio de 2018 foi realizada uma maratona de tecnologia na sede da Escola Fazendária do Estado de São Paulo (Fazesp). Durante aquele final de semana, startups e jovens da área de tecnologia e inovação tiveram 30 horas para desenvolver, com a base de dados da Bolsa Eletrônica de Compras (BEC), soluções que possam aperfeiçoar o ciclo de suprimentos e informar a sociedade sobre o que e como o Estado compra, além do nível de eficiência das transações em relação aos preços praticados no mercado.

A 1ª Maratona de Programação de Compras Públicas¹ foi uma iniciativa da Secretaria da Fazenda, com o patrocínio da Fundação Instituto de Pesquisa Econômicas (Fipe), da Associação dos Agentes Fiscais de Rendas do Estado de São Paulo (Afresp) e da Accenture, e o apoio do Banco Interamericano de Desenvolvimento (BID) e do Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (Sebrae). Dos inscritos foram selecionados 81 participantes, divididos em grupos, composto por três e cinco integrantes maiores de 18 anos (Figura 50). Foto 2

Realizamos observações *in situ* neste evento sobre a organização de *Hackathon*. Foi publicado uma notícia sobre o evento no Portal da Fazenda na qual menciona a participação da pesquisadora no evento mas é necessário considerar um erro no texto (Figura 51). A pesquisadora não é bolsista do Tribunal de Contas do Paraná.

Os desafios propostos neste evento foram os seguintes:

¹<http://hackathoncompraspublicas.fazenda.sp.gov.br/>. Último acesso: 10/06/2018

²<http://archive.is/MbtSi>. Último acesso: 19/11/2018



Figura 50: Participantes do Hackathon Compras Públicas. Fonte: Site da Secretária da Fazenda



A estudante de mestrado paraguaia e bolsista do Tribunal de Contas do Paraná (TJ-PR), Mirian Nogueira, aproveitou a maratona hacker para conhecer um pouco mais da metodologia adotada pela Bolsa Eletrônica de Compras. Segundo ela, a proposta do TJ-PR é desenvolver algo parecido. "Acho muito legal que as pessoas se envolvam e comecem a participar, entender aqueles dados", elogia. "Não adianta ter os dados abertos sem saber o que fazer com eles. Esse tipo de atividade dá para o cidadão se envolver ainda mais, especialmente os profissionais da área de economia e tecnologia".

Figura 51: Techo do texto publicado sobre o evento. Fonte: Site da Secretária da Fazenda

- Fornecedores: Criar solução para acesso a ofertas de compras a empresas interessadas em participar dos processos de compras do Estado, notadamente voltado a micro e pequenas empresas, incluindo informações sobre histórico de preços praticados em compras semelhantes do Estado, georeferenciamento das ofertas e notificação push sobre ofertas de compras abertas.
- Sociedade: Criar solução para propiciar acesso a informações sobre compras praticadas pelo Estado, em diferentes regiões, comparando diferentes compras com preços de mercado, etc.
- Prefeituras: Criar solução que permita às Prefeituras terem acesso a informações sobre compras e preços praticados pelo Estado na localidade e municípios próximos, incluindo acesso a base de preços referenciais do Estado, com possibilidade de comparação com preços praticados pela própria Prefeitura e pelo mercado da região.
- Estado: Criar solução que permita analisar a base de dados das compras públicas do Estado e identificar processos de compras pouco competitivos, processos de compras rea-

lizados não observando a sazonalidade da oferta, possibilidades de sinergias em compras semelhantes praticadas por diferentes órgãos, disparidade de preços em relação ao mercado, etc.

Foram 10 equipes e o desafio delas era desenvolver, com a base de dados da Bolsa Eletrônica de Compras (BEC/SP), soluções que possam aperfeiçoar o ciclo de suprimentos e informar a sociedade sobre o que e como o Estado compra, e o nível de eficiência em relação aos preços praticados no mercado. Os 10 projetos foram avaliados por uma banca formada por profissionais de tecnologia e representantes da Fazenda que escolheram as quatro melhores propostas a partir de critérios como impacto, coerência, viabilidade, inovação e originalidade.

A equipe vencedora foi InfoHack, com o projeto InfoPreços – BEC, solução direcionada as prefeituras que auxilia na compra de produtos pelo menor preço, por meio de pesquisas de dados quantitativos entre fornecedores e os valores praticados em outros estados.

B.2 RELATÓRIO DO 1º *DATATHON* UFRGS

O Departamento de Estatística da UFRGS, em comemoração dos 40 anos do curso de Bacharelado em Estatística realizou o Primeiro *Datathon* da UFRGS³ nos dias 15, 16 e 17 de outubro de 2018. Uma maratona de análise de dados em que as equipes propunham alternativas para que dados eleitorais sejam disponibilizados com transparência e acessibilidade a qualquer cidadão.

Aproveitamos este evento para realizar observação *in situ* das interações dos participantes nas equipes e do processo de compartilhamento de conhecimento. O evento contou com a participação de estudantes de graduação e pós graduação da matemática, estatística e ciência da computação, divididos em quatro equipes.

Foram apresentados os trabalhos das equipes "Mitos e Verdades" e "What If" para a comunidade acadêmica e a imprensa da UFRGS. As outras equipes não estiveram presentes no final. A equipe "Mitos e Verdades" foi a primeira a apresentar o trabalho sobre a visualização dos votos do primeiro turno por nível de escolaridade. Os gráficos em mapas divididos por estados representaram especificamente os votos para os candidatos a presidência concorrentes para o segundo turno e permitiram ver que a maioria das pessoas com maior nível de escolaridade votaram no candidato Jair Bolsonaro, enquanto o nível de escolaridade dos votantes do candidato Haddad foi heterogêneo ressaltando uma maioria de pessoas com escolaridade baixa.

³<https://www.ufrgs.br/datathon/index.html>. Último acesso: 19/11/2018

A segunda equipe a apresentar foi a equipe "What If" com uma proposta mais elaborada e publicada em site próprio⁴. O aplicativo tem como objetivo facilitar a visualização dos dados de Receita e Despesas dos candidatos e partidos concorrentes nas Eleições 2014. Também facilita a pesquisa por um candidato ou partido específico. Uns dos gráficos principais representa a relação receita e vitória. Mostra que candidatos que gastaram mais em campanhas políticas tiveram mais chances de vitória.

A equipe escolhida como vencedora pela banca avaliadora é o "What If" por atingir melhor o objetivo do *Datathon* (Figura 52).



Figura 52: Vencedores do *Datathon* UFRGS. Fonte: Site do evento

⁴<https://sulzbach.shinyapps.io/datathon/>. Último acesso: 19/11/2018

B.2.1 ARTIGO PUBLICADO PELA IMPRENSA DA UFRGS SOBRE O 1º DATATHON UFRGS

10 | JORNAL DA UNIVERSIDADE | JANEIRO/FEVEREIRO DE 2019

CÂMPUS



Maratona de dados

Ciência *Evento desafia estudantes e coloca em evidência a importância da aprendizagem ativa*

Alunos oriundos de diferentes cursos – Matemática, Estatística, Informática – e de diferentes níveis – graduação, mestrado – estavam em silêncio enquanto o professor Markus fala: “No logo do nosso Instituto [de Matemática e Estatística] dá para ver um ‘Cezinho’ ali de Ciência da Computação. Todas as áreas participaram”, brinca. Apesar de exaustos após 48 horas de trabalho quase ininterrupto, os participantes aguardavam atentamente a definição do I Datathon UFRGS. O evento ocorreu de 15 a 17 de outubro de 2018 dentro da agenda de atividades da IX Semana Acadêmica da Estatística (Semana Acadêmica da Estatística) e foi ocasião para comemorar os 40 anos do bacharelado em Estatística. Os alunos não eram os únicos que estavam ansiosos – nos dias anteriores, uma única dúvida ecoava nos corredores do departamento: será que a nossa empreitada vai ser bem-sucedida?

O termo datathon é uma junção de duas palavras da língua inglesa: data (dados) e marathon (maratona). É literalmente uma maratona de análise de dados em que os participantes se reúnem em grupos e têm horas ou dias para apresentar uma iniciativa inovadora, normalmente dentro de um assunto predeterminado. Neste caso, para aproveitar o tema em evidência na época, o tópico definido foi eleições. “Não houve tanto financiamento de campanha esse ano; a conversa sobre

eleições começou um pouco tarde demais, e é um momento um tanto complicado. Então pensamos: vamos tomar essa direção”, explica Markus Stein, professor do departamento de Estatística. As propostas desenvolvidas pelas quatro equipes participantes, formadas por afinidade na largada da maratona e que tinham de 2 a 4 integrantes, abarcaram desde a procura pelo perfil do eleitor de determinado candidato até a busca por inconsistências entre o desejo de renovação do executivo e as eleições legislativas.

A equipe vencedora analisou as receitas e os gastos de candidatos e partidos. Eles perceberam que a grande maioria dos candidatos recebe um repasse baixo dos partidos e uma minoria recebe um valor muito acima da média, o que sugere que os partidos preferem apostar alto em poucos candidatos a dividir o fundo eleitoral de forma equânime. “A gente consegue ver que a tendência de candidatos que gastam bastante serem eleitos permanece. Com isso, chegamos a uma conclusão puramente visual de que é muito difícil alguém que tem um gasto baixo conseguir se eleger”, explica o integrante Cristiano Sulzbach, formado em estatística e mestrando em Ciência da Computação.

O grupo criou um aplicativo que facilita a visualização dos dados por meio de gráficos e tabelas. Os membros da equipe explicaram o objetivo da ideia para os presentes, que ocupavam

uma das salas de aula do Instituto: “A quantidade de dados disponíveis está cada vez maior. Como confiar e ter acesso a esses dados? Será que vão conseguir entender e interpretar aquela informação? Queremos permitir que uma pessoa leiga consiga saber o que o partido está pensando, com quem, etc., no sentido de simplificar ao máximo, para que qualquer pessoa consiga ter acesso a essa informação. O Tribunal Superior Eleitoral (TSE) dá acesso a uma tabela gigante; muitas pessoas não sabem nem baixar”. O aplicativo está disponível no link <https://sulzbach.shinyapps.io/datathon/> (para acessar, deve-se usar o usuário datathon e a senha datathon).

Myrian Nogueira, mestranda em Computação Aplicada pela Universidade Tecnológica Federal do Paraná (UTFPR), considera que a sanção, em 2011, da Lei de Acesso à Informação – a qual regulamenta o direito constitucional de se conhecerem as informações públicas –, apesar de representar um avanço essencial, ainda é insuficiente, caso não haja quem analise os dados. “É uma obrigação que as instituições e órgãos governamentais publicarem suas informações, mas é um monte de dados que você não tem como entender. O pessoal do governo fala: ‘eu já publiquei, agora você que se vire para entender, eu já fiz minha parte’. Esses dados precisam ser processados, trabalhados”, observa.

Myrian, que é paraguaia, veio a Porto Alegre especialmente para acompanhar o I Datathon UFRGS. Seu trabalho de pós-graduação investiga eventos que envolvam métodos de aprendizagem ativa, integração entre os diferentes campos do saber e engajamento cívico, como os hackathons – maratonas de programação – e os datathons. “O pessoal compartilha conhecimento, e o aprendizado fica com cada um após o evento. É muito interessante a interação com o pessoal de outras áreas, como aconteceu aqui no Datathon. É necessário não só um cientista de dados, mas também alguém que entenda desses dados, que dê um sentido a eles”, argumenta. Uma maior participação de alunos de outros departamentos ficou definida como uma meta a ser perseguida pelas próximas edições. Markus sugeriu: “Por que não adicionar ao nosso logotipo um ‘E’ de engenharia, de economia? Acho que podemos agrupar todos que gostam de análise de dados e têm vontade de desenvolver soluções simples para a sociedade”.

Também professor de Estatística, Rodrigo dos Reis tem participado de iniciativas do tipo hackathon há mais de cinco anos e considera que esses eventos estão cada vez mais alinhados com as exigências do mercado de trabalho. Ele cita empresas cujos processos seletivos são pequenos datathons ou hackathons.

“O mercado quer esse tipo de coisa: uma construção de ideias de uma maneira muito rápida e que dê resultado.” Ele não economizou elogios aos alunos: “É raro chegar em um ambiente e ver essa qualidade. Eu tenho visto startups se formando, crescendo e ganhando corpo com ideias muito mais simples do que as que foram apresentadas. É um movimento que já surge grande, já começa bem em termos de profundidade, e eu gostaria muito que tivesse um segundo, um terceiro, um quarto”, empolga-se.

Gabriel Holmer, aluno do quarto semestre de Estatística, reflete que o datathon representa uma oportunidade para aprender mais conceitos de forma superintensiva. “Eu, por exemplo, nem sabia fazer um mapa ontem, e a gente apresentou uns ali”, reflete. Questionado sobre se a maratona tinha sido exaustiva, ele desabafa: “Bah, tô morrendo de sono”.

Entretanto, o trabalho duro compensou. “Por que não utilizar a maratona em sala de aula como avaliação? A gente pode conviver muito e entender a evolução de cada um. De repente, em vez de três provas, podemos ter três datathons ao longo do semestre”, sugeriu, em tom de brincadeira, o professor Markus Stein, sendo aclamado pelos alunos.

Henrique Moretto
estudante do 8.º semestre
de Jornalismo da UFRGS

APÊNDICE C – QUESTIONÁRIOS APLICADOS NOS ESTUDOS DE CASO DO *SHORT DATATHON*

Conforme mencionado na seção 4.3 o trabalho de pesquisa empregou também questionários. A seguir está o questionário que aplicamos em participantes dos *Short Datathon* e os resultados obtidos.

C.1 QUESTIONÁRIO APLICADO NO ESTUDO DE CASO PILOTO E NO SEGUNDO ESTUDO DE CASO

Nesta seção apresentamos o questionário aplicado no Estudo de Caso Piloto e no segundo Estudo de Caso (Figuras 53, 54 e 55).

**Prezados,**

O PPGCA-UTFPR, através do grupo de pesquisa de Engenharia de Software, solicita sua atenção e seu auxílio nesta pesquisa.

O objetivo deste estudo é investigar questões relacionadas a eventos de ativismo cívico social. Esses eventos visam estimular o engajamento cívico, reunir desenvolvedores e profissionais de diversas áreas para resolver as necessidades de suas comunidades e mostrar o que é possível obter em ganhos para a sociedade ao usar a tecnologia.

O uso dos registros efetuados durante a pesquisa é estritamente limitado à mesma e o anonimato dos participantes será preservado.

Qual é sua área de trabalho ou estudo?
(Pode escolher mais de uma opção)

- Economia
- Administração
- Comércio
- Contabilidade
- Direito
- Comunicação
- Educação
- Computação
- Outros (mencionar).....

Qual é sua atividade principal (Escolher só uma opção, a principal)

- Estudante
- Professor
- Líder de projetos
- Freelancer
- Diretor/Gerente
- Empregado
- Pesquisador
- Outros (mencionar).....

Qual é sua faixa etária? (Escolher só uma opção)

- 15-19 anos
- 20-29 anos
- 30-39 anos
- 40-49 anos
- 50 e mais

Quais são os tipos de eventos de que você já participou? (Pode escolher mais de uma opção)

- Palestras/Conferências
- Meetups
- Hackathon públicos/privados
- Datathon públicos/privados
- Outros (mencionar).....

Gênero

- Feminino
- Masculino
- Prefiro não informar
- Prefiro me descrever.....

Qual é a frequência com que você participa em eventos Hackathon? (Escolher só uma opção)

- Nunca
- Ocasionalmente(uma vez a cada 2 anos ou mais)
- Frequentemente (uma vez por ano)
- Sempre (vários por ano)

Figura 53: Primeira página do questionário aplicado

Qual é a frequência com que você participa em eventos Datathon? (Escolher só uma opção)

- Nunca (esta é primeira vez)
 Ocasionalmente(uma vez a cada 2 anos ou mais)
 Frequentemente (uma vez por ano)
 Sempre (vários por ano)

Percepção sobre um Datathon (Escolher só uma opção de cada afirmação)

<p>1- Gostei de participar do Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>2- Participar deste Datathon me ajudou a compreender mais sobre análise de dados e visualização.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>
<p>3- Tive facilidade em entender a dinâmica do Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>4- Compreendi o funcionamento das ferramentas utilizadas no Datathon.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>
<p>5- Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>	<p>6- Participar deste Datathon promoveu momentos de cooperação entre as pessoas que participaram.</p> <p> <input type="checkbox"/> Discordo fortemente <input type="checkbox"/> Discordo <input type="checkbox"/> Indiferente <input type="checkbox"/> Concordo <input type="checkbox"/> Concordo fortemente </p>

Figura 54: Segunda página do questionário aplicado



7- A interação com pessoas de outras áreas do meu time foi boa e contribuiu positivamente no resultado final.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

8- Acredito que a experiência com pessoas de outras áreas num evento Datathon torna a análise de dados mais eficiente.

- Discordo fortemente
 Discordo
 Indiferente
 Concordo
 Concordo fortemente

Outras considerações sobre o Datathon:

Se você tem interesse em colaborar com esta pesquisa a través de uma entrevista deixe seu e-mail :

--

Assinatura do participante

Figura 55: Terceira página do questionário aplicado

C.2 RESULTADOS DO QUESTIONÁRIO APLICADO NO *SHORT DATATHON* PILOTO



Figura 56: Área principal de trabalho ou estudo dos participantes. Fonte: Autoria Própria

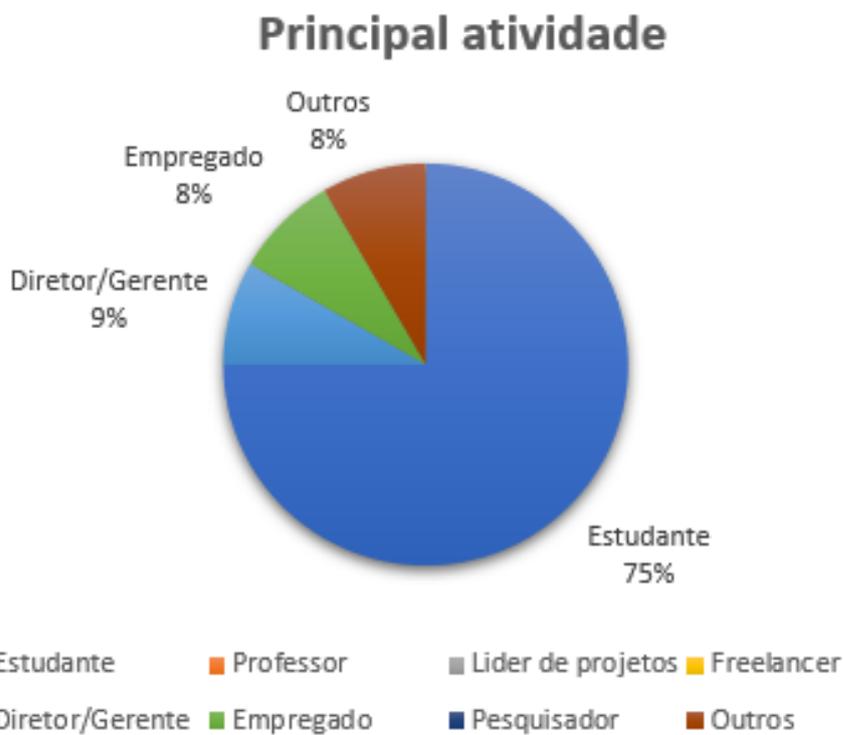


Figura 57: Principal atividade dos participantes. Fonte: Autoria Própria



Figura 58: Faixa Etária dos participantes. Fonte: Autoria Própria



Figura 59: Tipos de eventos de que você já participou. Fonte: Autoria Própria

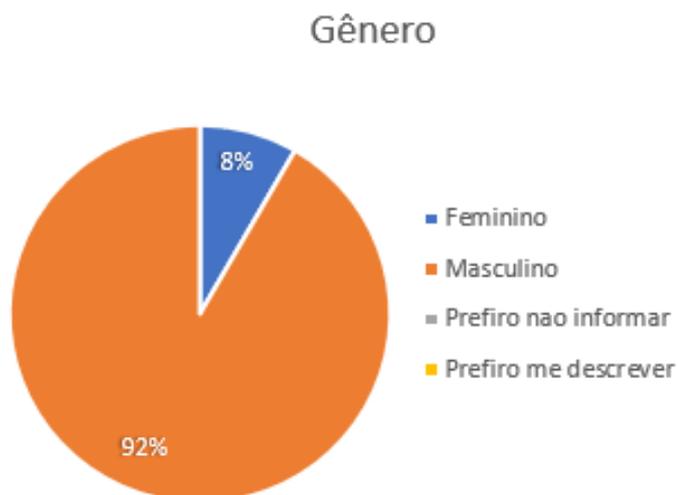


Figura 60: Gênero dos participantes. Fonte: Autoria Própria

Participação em eventos Hackathon



Figura 61: Nível de participação em eventos Hackathon. Fonte: Autoria Própria

Participação em eventos Datathon

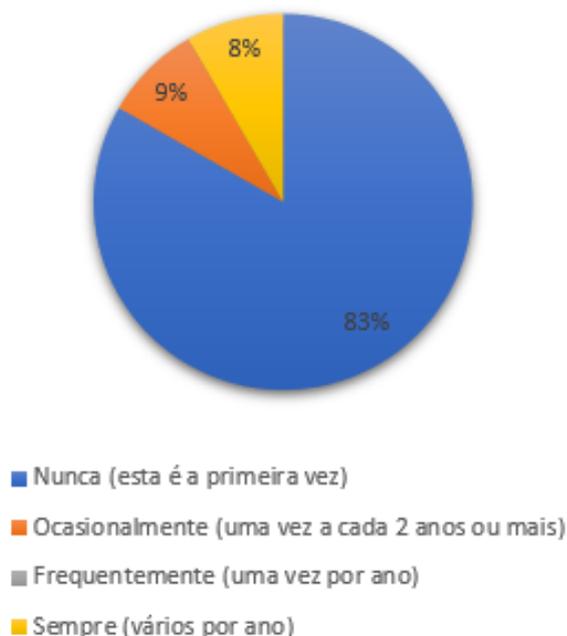


Figura 62: Nível de participação em eventos Datathon. Fonte: Autoria Própria

Tabela 20: Tabela de valores Likert do Estudo de Caso Piloto. Fonte: Autoria Própria

Escala:	
Discordo fortemente	Concordo fortemente
Gostei de participar do Datathon	0 7 5
O Datathon me ajudou a compreender mais sobre análise de dados e visualização.	0 6 6
Tive facilidade em entender a dinâmica do Datathon	0 1 7 4
Compreendi o funcionamento das ferramentas utilizadas no Datathon	0 7 5
Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente	0 7 5
O Datathon promoveu momentos de cooperação entre as pessoas que participaram	0 1 3 8
A interação com pessoal de outras áreas do meu time foi boa e contribuiu positivamente no resultado final	0 3 4 5
Acredito que a experiência com pessoal de outras áreas num evento Datathon, torna o análise de dados mais eficiente	0 5 7

C.3 RESULTADOS DO QUESTIONÁRIO APLICADO NO SEGUNDO ESTUDO DE CASO DO *SHORT DATATHON*



Figura 63: Área principal de trabalho ou estudo. Fonte: Autoria Própria

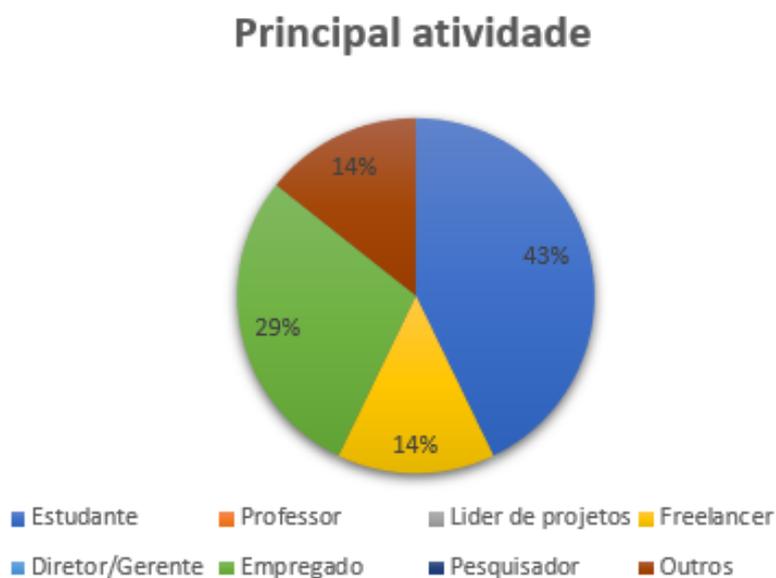


Figura 64: Principal atividade dos participantes. Fonte: Autoria Própria

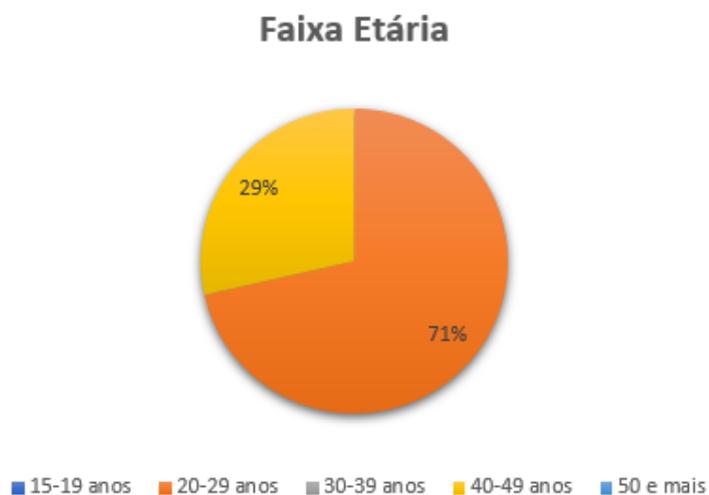


Figura 65: Faixa Etária dos participantes. Fonte: Autoria Própria



Figura 66: Tipos de eventos de que você já participou. Fonte: Autoria Própria



Figura 67: Gênero dos participantes. Fonte: Autoria Própria

Participação em eventos Hackathon

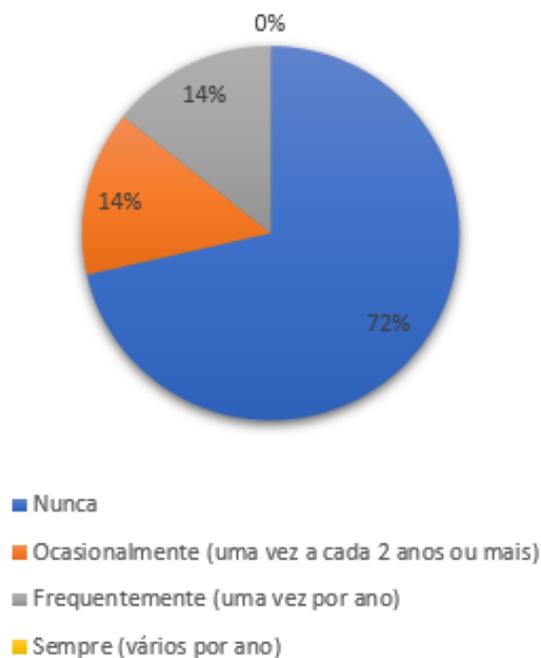


Figura 68: Nível de participação em eventos Hackathon. Fonte: Autoria Própria

Participação em eventos Datathon

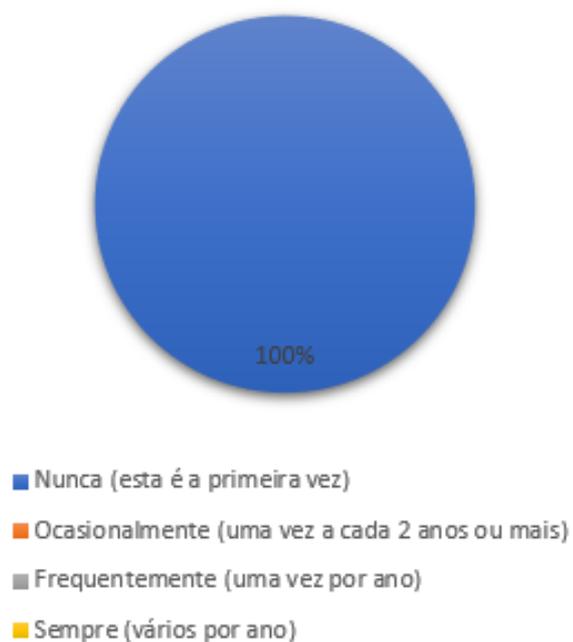


Figura 69: Nível de participação em eventos Datathon. Fonte: Autoria Própria

Tabela 21: Tabela de valores Likert do segundo Estudo de Caso. Fonte: Autoria Própria

Escala:				
Discordo fortemente	Discordo	Indiferente	Concordo	Concordo fortemente
Gostei de participar do Datathon	0	3	3	0
O Datathon me ajudou a compreender mais sobre análise de dados e visualização.	0	4	3	0
Tive facilidade em entender a dinâmica do Datathon	0	3	4	0
Compreendi o funcionamento das ferramentas utilizadas no Datathon	0	4	3	0
Meu nível de conhecimento sobre análise de dados e visualização melhorou significativamente	0	6	1	0
O Datathon promoveu momentos de cooperação entre as pessoas que participaram	0	4	3	0
A interação com pessoal de outras áreas do meu time foi boa e contribuiu positivamente no resultado final	0	5	2	0
Acredito que a experiência com pessoal de outras áreas num evento Datathon, torna o análise de dados mais eficiente	0	3	4	0

APÊNDICE D – DIVULGAÇÃO DOS *SHORT DATATHON*

Neste apêndice apresentamos os tipos de divulgação e as ferramentas utilizadas na promoção dos eventos *Short Datathon*.

D.1 CONVITE NA PLATAFORMA DE EVENTOS SYMPLA

Sympla é uma ferramenta *online* para gestão de eventos e venda de ingressos. Através da tecnologia, possibilita que organizadores possam criar, organizar, gerir e divulgar eventos, proporcionando ao participante uma variedade de opções para escolher. Esta ferramenta foi utilizada para os dois eventos.

🔒 https://www.sympla.com.br/datathon-de-analise-de-dados-de-evasao-fiscal__367756#info

Datathon de Análise de Dados de Evasão Fiscal

📍 UTFPR - Sede Centro - Curitiba, PR
🕒 23 de outubro de 2018, 09h-17h

DESCRIÇÃO DO EVENTO

ENSALAMENTO CONFIRMADO: SALA B201

Uma **Datathon** é uma Hackathon com foco em dados. A partir de um tema ou desafio é preciso formar um time, pode ser de colegas ou de pessoas que você vai conhecer no próprio evento. Times bons tem pessoas de perfil diverso: estudantes ou profissionais nas áreas de TI, design, comunicação/marketing, gestão/administração, exatas, todos que estejam dispostos a trabalhar em equipe são bem-vindos.

O evento acontecerá na **Semana Tecnológica da UTFPR** e convidamos a todos os alunos e profissionais a participar e compartilhar conhecimentos. Durante essa jornada analisaremos um tema pouco comum: Evasão Fiscal.

Programação estimada

- 9h Introdução e Apresentação
- 9h30 Início das Atividades
- 12h Almoço
- 14h Retorno às Atividades
- 16h30 Apresentação dos trabalhos
- 17h Encerramento

Instruções específicas para quem deseja participar

- Não precisa trazer seu computador, usaremos os computadores que teremos na sala.
- Não precisa ter conhecimento de programação.
- Devemos ter algum lanche, mas nada muito significativo. Se você quer comer algo significativo antes do almoço, traga seu lanche.
- O evento prevê um intervalo para almoço de 12 a 14 horas. O almoço fica a cargo dos participantes.

Vagas limitadas!

Inscrições: <http://semanatecnologica.dainf.ct.utfpr.edu.br/minicursos.php>

Links sobre o tema:

<https://www.youtube.com/watch?v=kOZ1hra5bXE>
<http://gabriel-zucman.eu/leaks/>

Pasta Compartilhada para os participantes:
<https://drive.google.com/open?id=1k3iwGIHtS5CtPIKXnK9DTLrRnFE6EPLxO>

Figura 70: Convite realizado no Sympla para o Estudo de Caso Piloto

🔒 https://www.sympla.com.br/datathon-de-analise-de-dados-de-evasao-fiscal-ii_400096#info

Datathon de Análise de Dados de Evasão Fiscal II

📍 UTFPR - Sede Centro - Curitiba, PR

🕒 24 de novembro de 2018, 09h-19h

DESCRIÇÃO DO EVENTO

Um **Datathon é uma Hackathon** com foco em dados. Equipes serão formadas antes ou durante o evento para enfrentar um desafio.. Times bons tem pessoas de perfis diferentes: estudantes ou profissionais nas áreas de computação, design, comunicação/marketing, gestão/administração, exatas, entre outras. Todos que estejam dispostos a trabalhar em equipe são **bem-vindos**.

Nesta segunda edição do evento convidamos a todos os alunos e profissionais a participar e compartilhar conhecimentos. Durante essa jornada analisaremos o tema: Evasão Fiscal.

----- SALA B-201, BLOCO B, DAINF -----

Programação prevista

- 9h00 Introdução e Apresentação
- 9h30 Início das Atividades
- 12h00 Almoço
- 13h30 Retorno às Atividades
- 18h00 Apresentação dos Trabalhos e Premiação
- 19h Encerramento

Instruções específicas para quem deseja participar

- A sala disporá de computadores para serem utilizados durante o evento, mas é permitido trazer seu próprio computador com as ferramentas que achar melhor.
- Não é necessário ter conhecimentos em programação de computadores.
- Devemos ter algum lanche, mas nada muito significativo. Se você quer comer algo significativo antes do almoço, traga seu lanche.
- O evento prevê um intervalo para almoço de 12h00 a 13h30 horas. O almoço fica a cargo dos participantes.
- Todos que participarem do evento receberão um certificado com a quantidade de horas, que pode ser usado pelos alunos da UTFPR para solicitarem créditos de Atividades Complementares.
- O melhor trabalho receberá premiação.

Vagas limitadas!

A inscrição no evento é gratuita! Só pedimos a cada participante para doarem 1 kg de alimento não perecível - não vale sal ;)

Links sobre o tema:

<https://www.youtube.com/watch?v=kOZ1hra5bXE>

<http://gabriel-zucman.eu/leaks/>

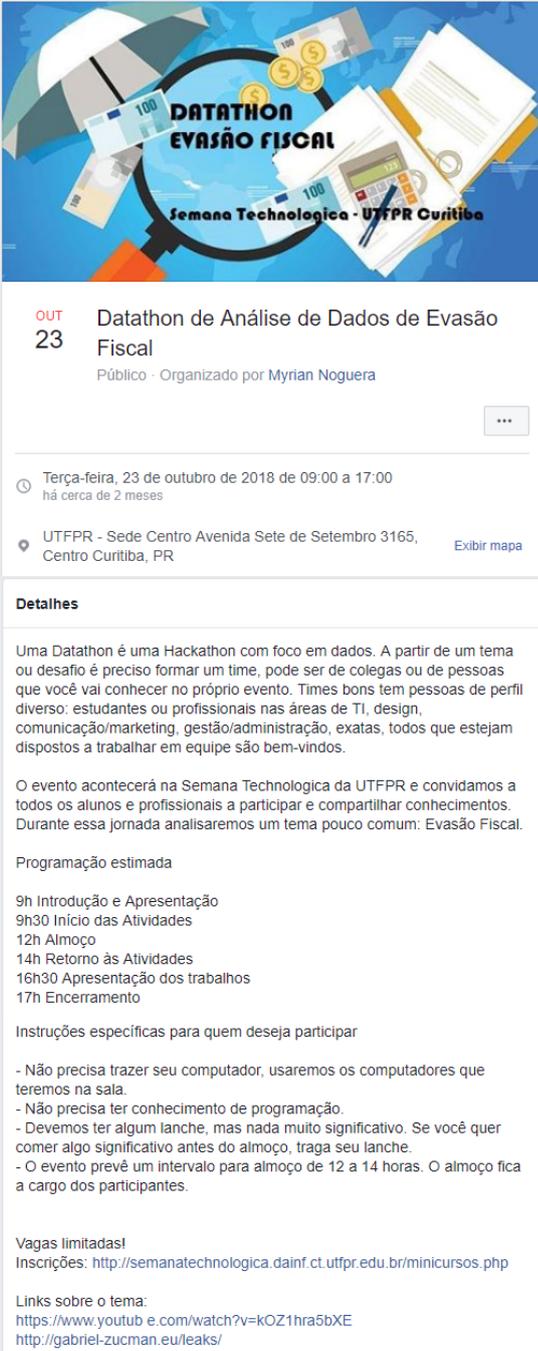
<https://www.youtube.com/watch?v=Miqb4qr-9LA>

Link de pasta compartilhada: <https://drive.google.com/drive/folders/1k3iwGIHt55CtPIKXnK9DTLrNfE6EPLxO>

Figura 71: Convite realizado no Sympla para o segundo Estudo de Caso

D.2 DIVULGAÇÃO EM REDES SOCIAIS

A rede social *Facebook* também oferece numerosas opções de divulgação. Para o primeiro *Short Datathon* criamos um evento na rede mencionada e compartilhamos com os contatos. No segundo evento exploramos outras funções como criar uma página para compartilhar conteúdo sobre a primeira edição e fazer o convite para a segunda edição do *Short Datathon*.



OUT 23 Datathon de Análise de Dados de Evasão Fiscal
Público · Organizado por Myrian Nogueira

Terça-feira, 23 de outubro de 2018 de 09:00 a 17:00
há cerca de 2 meses

UTFPR - Sede Centro Avenida Sete de Setembro 3165, Centro Curitiba, PR [Exibir mapa](#)

Detalhes

Uma Datathon é uma Hackathon com foco em dados. A partir de um tema ou desafio é preciso formar um time, pode ser de colegas ou de pessoas que você vai conhecer no próprio evento. Times bons tem pessoas de perfil diverso: estudantes ou profissionais nas áreas de TI, design, comunicação/marketing, gestão/administração, exatas, todos que estejam dispostos a trabalhar em equipe são bem-vindos.

O evento acontecerá na Semana Tecnológica da UTFPR e convidamos a todos os alunos e profissionais a participar e compartilhar conhecimentos. Durante essa jornada analisaremos um tema pouco comum: Evasão Fiscal.

Programação estimada

9h Introdução e Apresentação
9h30 Início das Atividades
12h Almoço
14h Retorno às Atividades
16h30 Apresentação dos trabalhos
17h Encerramento

Instruções específicas para quem deseja participar

- Não precisa trazer seu computador, usaremos os computadores que teremos na sala.
- Não precisa ter conhecimento de programação.
- Devemos ter algum lanche, mas nada muito significativo. Se você quer comer algo significativo antes do almoço, traga seu lanche.
- O evento prevê um intervalo para almoço de 12 a 14 horas. O almoço fica a cargo dos participantes.

Vagas limitadas!
Inscrições: <http://semanatecnologica.dainf.ct.utfpr.edu.br/minicursos.php>

Links sobre o tema:
<https://www.youtube.com/watch?v=kOZ1hra5bXE>
<http://gabriel-zucman.eu/leaks/>

Figura 72: Convite realizado no Facebook para o Estudo de Caso Piloto

Hackathons Brasil
Grupo público

Sobre

Discussão

Membros

Eventos

Vídeos

Fotos

Arquivos

Pesquisar neste grupo

Atalhos

- Datathon de Análisis... 20+
- OMG
- Hackathons Brasil
- ENG COMP - UTFPR... 1
- Hendyla Capiatá 20+
- Bazar On Line Curi... 20+

Ver mais

Myrian Noguera compartilhou um link.
15 de novembro de 2018

Ocorrerá na UTFPR no dia 24/11/18 o nosso segundo "Datathon: Análise de Dados de Evasão Fiscal", um evento voltado aos(as) alunos(as) e profissionais estatísticos, matemáticos, cientistas da computação, cientistas políticos, jornalistas, publicitários, designers, programadores, hackers e todas as pessoas criativas engajadas. 🤖💻
Será permitida a inscrição individual ou em equipes de no máximo quatro integrantes. O melhor trabalho receberá premiação!!!
Mais informações no link!! 👍



SYMPLA.COM.BR
Datathon de Análise de Dados de Evasão Fiscal II
Sábado, 24 de novembro de 2018, 09h-19h

Figura 73: Divulgação realizada no grupo Hackathons Brasil na rede social Facebook

<https://www.facebook.com/datathonutfpr/>

Pesquisar

Myrian Página inicial Criar

Página Caixa de entr... 1 Notificações 22 Informações Ferramentas ... Central de An... Configurações Ajuda




Datathon de Análise de Dados de Evasão Fiscal
@datathonutfpr

Página inicial

Curtir Seguir Compartilhar ...

Enviar mensagem

Figura 74: Criação da página Datathon de Análise de Dados de Evasão Fiscal na rede social Facebook

■ Alcance: orgânico/pago ▼ ■ Cliques em publicações ■ Reações, comentários e compartilhamentos ■ i ▼						
Publicado	Publicação	Tipo	Direcionamento	Alcance i	Envolvimento	
5/12/2018 08:34	 Datathon de Análise de Dados de Evasão Fiscal compartilhou uma publicação.			13 	0	0
29/11/2018 12:22	 Artigo bem interessante sobre o Datathon realizado pelos professores do Departamento de			17 	0	1
25/11/2018 16:28	 Galera! Um breve relato do que foi a primeira versão do Datathon na Semana Tecnologica da			21 	4	2
24/11/2018 20:48	 Fotos da publicação de Datathon de Análise de Dados de Evasão Fiscal			17 	18	0
24/11/2018 20:47	 Fotos da publicação de Datathon de Análise de Dados de Evasão Fiscal			273 ■	73	33
24/11/2018 18:32	 Rolou um Datathon muito bacana!!!			14 	7	1
23/11/2018 08:01	 Finalmente chegou o grande dia do nosso evento, o 2º Datathon de Análise de Dados de Evasão			13 	0	0
22/11/2018 14:30	 Datathon de Análise de Dados de Evasão Fiscal compartilhou uma publicação.			13 	0	0
21/11/2018 15:53	 Venha participar!! Vamos discutir e trabalhar sobre dados relacionados a este tema bem importante.			16 	1	1
20/11/2018 22:38	 Pessoal!! O evento é no sábado 24 📅 Faça sua inscrição 🙌🙌			26 	0	2
19/11/2018 09:31	 Vídeo realizado pelo Professor Ladislau Dowbor que fala sobre os Paraísos Fiscais e esclarece			115 ■	0	2
16/11/2018 08:40	 Datathon - Semana Tecnologica			49 	21	0
15/11/2018 09:09	 8% da riqueza financeira do mundo é mantida em Offshore, custando pelo menos US\$ 200 bilhões			46 	2	1
14/11/2018 19:01	 Tem curiosidade ou quer conhecer mais sobre esses temas? Venha participar do Datathon na			38 	0	0
14/11/2018 18:38	 Agora está mais fácil enviar uma mensagem para a Página Datathon de Análise de Dados de			32 	0	0
14/11/2018 18:22	 Ocorrerá no dia 24/11/18 o segundo "Datathon: Análise de Dados de Evasão Fiscal", um evento			795 ■	54	53

Figura 75: Painel de gestão de alcance de publicações realizadas na página criada no Facebook

D.3 OUTROS TIPOS DE DIVULGAÇÃO

Outras alternativas de divulgação utilizadas foram a publicação em outros *sites* como: Eventos Tech¹ principal canal no Brasil, centralizador de agendas de eventos do setor tecnológico. Também fizemos a divulgação no *site* da Semana Tecnológica da UTFPR.

Uma opção de divulgação bastante utilizada são os grupos de *Whatsapp*. Enviamos convites a grupos relacionados a temas como *Machine Learning*, *Data Mining*, redes de ativismo de mulheres na tecnologia e ativismo social como o Núcleo da Auditoria Cidadã da Dívida. Além disso, afixamos cartazes nos painéis da UTFPR.

D.3.1 SITE DO EVENTOS TECH



Figura 76: Divulgação realizada no site Eventos Tech

¹<https://eventostech.com.br/index.php/event/datathon-de-analise-de-dados-de-evasao-fiscal/>. Último acesso: 11/11/2018

D.3.2 SITE DA SEMANA TECHNOLOGICA

ⓘ Não seguro | semanatechnologica.dainf.ct.utfpr.edu.br/minicursos.php

23/10

Data	Local	Nome
09:00	Sala B106	Curso de Latex (Parte 1)- César Manuel Vargas Benitez - Professor DAELN
09:00	Sala B108	Programando em BASH - Automatizando Tarefas no Linux - Matheus Alexandre Fonseca e Lucas Ricardo Marques- PETCOCE
09:00	Sala B201	Datathon de Análise de Dados de Evasão Fiscal (Parte 1) - Adolfo Neto e Myrian Raquel Nogueira Salinas - Professor DAINF
https://www.sympla.com.br/datathon-de-analise-de-dados-de-evasao-fiscal__367756		
14:00	Sala B201	Datathon de Análise de Dados de Evasão Fiscal (Parte 2) - Adolfo Neto e Myrian Raquel Nogueira Salinas - Professor DAINF
https://www.sympla.com.br/datathon-de-analise-de-dados-de-evasao-fiscal__367756		
18:30	Auditório	Cabeamento Estruturado Metálico e Óptico - Valcimar Baglioli - Instrutor Oficial Furukawa
18:30	Sala B202	Impressora 3D - João Alberto Fabro - Professor DAINF

Figura 77: Divulgação realizada no site da Semana Technologica

D.3.3 CARTAZES EM PAINÉIS DA UTFPR CURITIBA



**DATATHON
EVA\$ÃO FISCAL**

UTFPR - Curitiba

24/11
Sala B-201

Hackathon com foco em dados

Evento voltado a alunos(as) e profissionais de Computação, Estatística, Matemática, Economia, Design, entre outras áreas.

✓ Não é necessário ter conhecimentos em programação de computadores.

O melhor trabalho receberá premiação!!!



Facebook: <https://www.facebook.com/datathonutfpr/>

Figura 78: Imagem do cartaz impresso para divulgação nos painéis da UTFPR Curitiba

D.3.4 CONVITE EM GRUPOS DE WHATSAPP

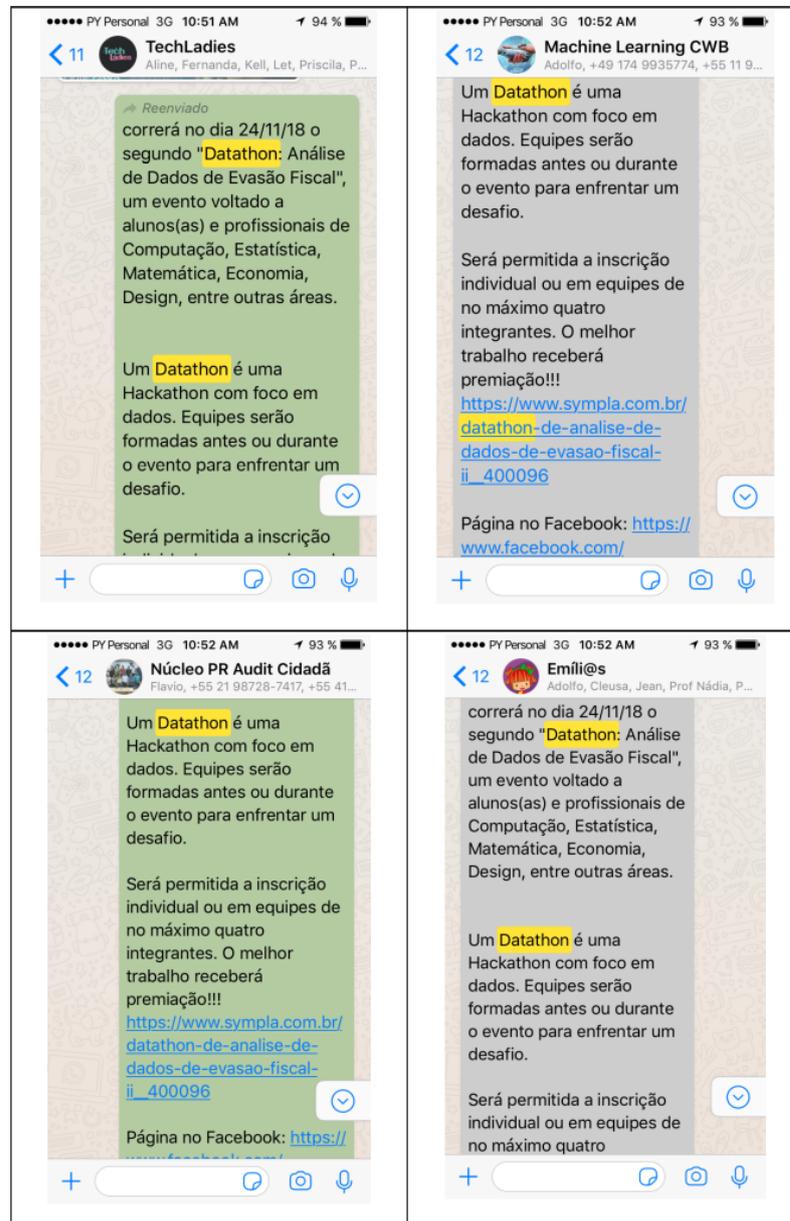


Figura 79: Convite enviados a grupos de Whatsapp