

Análise dos Resultados do TreasureHunt de 2024

Ricardo de la Rocha Ladeira

26 de novembro, 2024

1 Introdução

Este documento registra os resultados e a análise estatística sobre os dados de desempenho nas competições, a confiabilidade dos questionários aplicados e os resultados dos questionários relativos às competições geradas pelo TreasureHunt em 2024. O objetivo é manter a análise dos dados e sua interpretação no mesmo documento, para minimizar a desconexão entre esses dois aspectos. A interpretação dos dados se deu com base no que consta na referência abaixo:

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977.

Neste ano o jogo foi aplicado em uma turma no IFC campus Blumenau:

- turma de Segurança Computacional (Optativa), ministrada em 2024/2 para alunos da 8ª fase do curso de Bacharelado em Ciência da Computação, chamada aqui de **SUP**.

Nota: este relatório foi baseado no documento *Análise da Confiabilidade*, de Rafael R. Obelheiro, data: 11 de dezembro de 2017.

2 Desempenho da Turma

Média de acertos: 5,90

Desempenho da turma: 73,75%

Desempenho da turma com o easteregg: 65,56%

Quantidade total de submissões: 62

Quantidade total de submissões com o easteregg: 76

Quantidade total de submissões corretas: 59

Taxa de submissões corretas: 95,16%

Taxa de submissões corretas com o easteregg: 77,63%

2.1 Discussão dos Resultados

Considerando que, na opinião do professor organizador da atividade, o conjunto de exercícios selecionado possui dificuldade superior se comparado aos conjuntos de anos anteriores, o desempenho da turma foi *muito bom*.

O easteregg mais uma vez não foi resolvido, mas alguns alunos chegaram perto de resolvê-lo. Por outro lado, o problema envolvendo a técnica de esteganografia com senha foi resolvido pela primeira vez, o que representa grande avanço.

A quantidade de submissões incorretas total foi considerada alta. São 17 envios de *flags* incorretas, o que poderia significar que é necessário reforçar as regras do jogo. No entanto, ao analisar os dados na tabela de submissão, percebe-se que dos 17 envios, 14 referem-se a tentativas de acerto do easteregg. Assim, desconsiderando o easteregg, a taxa de submissões corretas foi *muito alta*.

3 Análise de Comportamentos Suspeitos e Ataques Potenciais

3.1 Respostas Repetidas:

Resposta 'TreasureHunt{Security}':

Usuário 3 enviou a resposta TreasureHunt{Security} para o problema 9 às 2024-11-22 20:33:09

Usuário 2 enviou a resposta TreasureHunt{Security} para o problema 9 às 2024-11-22 20:37:05

Usuário 4 enviou a resposta TreasureHunt{Security} para o problema 9 às 2024-11-22 21:24:53

Resposta 'TreasureHunt{dhi2uh39}':

Usuário 6 enviou a resposta TreasureHunt{dhi2uh39} para o problema 9 às 2024-11-22 20:23:16

Usuário 3 enviou a resposta TreasureHunt{dhi2uh39} para o problema 9 às 2024-11-22 21:00:21

Usuário 4 enviou a resposta TreasureHunt{dhi2uh39} para o problema 9 às 2024-11-22 21:21:20

3.2 Logins Suspeitos:

Máquina 191.52.7.89 logou com os usuários:

3 às 2024-11-22 18:46:02 2024-11-22 20:13:54 2024-11-22 21:00:04

Máquina 191.52.7.91 logou com os usuários:

4 às 2024-11-22 18:45:59 2024-11-22 21:19:20

Máquina 191.52.7.103 logou com os usuários:

5 às 2024-11-22 18:46:06 2024-11-22 20:29:45

Máquina 191.52.7.98 logou com os usuários:

6 às 2024-11-22 18:46:05 2024-11-22 20:09:53 2024-11-22 20:46:53

Máquina 191.52.7.94 logou com os usuários:

7 às 2024-11-22 18:46:46 2024-11-22 19:45:44

Máquina 191.52.7.104 logou com os usuários:

9 às 2024-11-22 18:46:41 2024-11-22 21:28:19

Máquina 191.52.7.95 logou com os usuários:

11 às 2024-11-22 18:46:46 2024-11-22 20:04:59

Nenhum usuário enviou muitas respostas (quantidade de questões * 2).

3.3 Discussão dos Resultados

Todas as senhas identificadas como repetidas foram tentativas de resposta do exercício 9 (easteregg), que não foi resolvido. As tentativas dizem respeito a expressões que aparecem na interface web, por isso não podem ser consideradas como compartilhamento de resposta.

Alguns usuários fizeram login mais de uma vez, mas a atividade não foi identificada como suspeita, sendo inclusive realizada do mesmo computador.

4 Análise dos Questionários

Nesta seção analisamos os questionários aplicados, iniciando pela subseção de dados de entrada, onde listamos os dados brutos obtidos através das respostas dos jogadores nos questionários pré- e pós-competição. Os dados foram mapeados diretamente em números de 1 a 5, conforme Escala de Likert, onde 1 corresponde à menor resposta possível de ser sinalizada (muito baixa, muito desmotivador, discordo totalmente) e 5 corresponde à maior (muito alta, muito motivador, concordo totalmente). A segunda subseção traz a análise da confiabilidade do questionário. A terceira e última subseção descreve os resultados obtidos.

4.1 Dados de Entrada

Todos os dados de entrada referentes aos questionários (pré- e pós-competição) respondidos por estudantes que participaram do TreasureHunt em 2024 estão no *data frame* `quest`:

Todos os registros são referentes à turma *SUP*.

4.2 Análise da Confiabilidade dos Questionários

4.2.1 Satisfação

As questões relativas à satisfação são 1.1, 1.2, 1.3, 1.5 e 1.7. Serão consideradas apenas as respostas do questionário pós-competição.

O coeficiente α de Cronbach é $\alpha = 0,761719$, o que indica confiabilidade substancial.

4.2.2 Conscientização

As questões relativas à conscientização são 1.6 e 2.1. Serão consideradas apenas as respostas do questionário pós-competição.

O coeficiente α de Cronbach é $\alpha = 0,707865$, o que indica confiabilidade substancial.

4.2.3 Dificuldade

As questões relativas ao nível de dificuldade do desafio são 3.1 e 3.2, presentes apenas no questionário pós-competição.

O coeficiente α de Cronbach é $\alpha = 0,902386$, o que indica confiabilidade quase perfeita.

4.2.4 Motivação

As questões relativas à competitividade e composição de problemas como fatores de motivação são 4.1 e 4.2, presentes apenas no questionário pós-competição.

O coeficiente α de Cronbach é $\alpha = 0,038462$, o que indica confiabilidade pequena.

4.3 Resultados dos Questionários

Questão	Pré	Pós	Evolução	Significativa?
1.1	4,0	4,4	0,4	Não ($p = 0,422572$)
1.2	4,3	4,3	0,0	Não ($p = 1,000000$)
1.3	4,4	4,0	-0,4	Não ($p = 0,343436$)
1.4	2,8	2,8	0,0	Não ($p = 1,000000$)
1.5	4,8	4,7	-0,1	Não ($p = 0,591051$)
1.6	2,5	2,3	-0,2	Não ($p = 0,692633$)
1.7	4,7	4,4	-0,3	Não ($p = 0,434380$)
2.1	3,2	2,9	-0,3	Não ($p = 0,576312$)

Questão	Pré	Pós	Evolução	Significativa?
3.1	-	3,4	-	-
3.2	-	3,9	-	-
4.1	-	3,7	-	-
4.2	-	3,7	-	-

4.4 Discussão dos Resultados

Não houve diferença estatisticamente significativa em todos os aspectos observados por meio das respostas dos questionários. Os resultados em termos absolutos apresentaram pequena piora. Essa piora contrasta com o desempenho da turma, que foi muito bom.

Entre as questões presentes somente no questionário pós-competição, as respostas sugerem que o conjunto de exercícios era razoavelmente difícil e exigia tempo levemente alto para resolvê-los, devido aos valores entre neutro e concordo parcialmente nesses quesitos.

As questões sobre a motivação com a competitividade e com a existência de exercícios compostos também obtiveram resposta entre neutro e motivador, indicando que esses fatores não foram vistos como negativos pela turma.

4.4.1 Histórico

A tabela abaixo apresenta os valores obtidos calculando os índices gerais de α de Cronbach nas competições de cada ano.

Fator	2017	2018	2019	2020	2021	2022	2023	2024
Satisfação	0,78	0,92	0,76	0,65	0,77	0,00	0,64	0,76
Conscientização	0,60	0,44	0,68	0,39	0,67	0,00	0,14	0,71
Dificuldade	0,72	0,60	0,75	0,93	0,73	0,00	0,69	0,90
Motivação	0,55	0,86	0,61	0,00	0,86	0,00	0,22	0,04

Obs. 1: os valores de 2017 dizem respeito a três turmas (BCC, TADS e FIC).

Obs. 2: os valores de 2018 dizem respeito à turma de BCC da UDESC.

Obs. 3: os valores de 2019 dizem respeito à turma TADS, do IFC, e à turma de BCC, da UDESC.

Obs. 4: os valores de 2020 dizem respeito a somente uma turma do IFC, em uma edição remota da competição.

Obs. 5: os valores de 2021 dizem respeito a duas turmas, uma de ensino superior e uma de ensino médio.

Obs. 6: os valores de 2022 ainda não foram calculados e por esse motivo estão zerados. Em 2022 a competição aconteceu em duas turmas, uma de ensino superior e uma de ensino médio.

Obs. 7: os valores de 2023 dizem respeito a somente uma turma do IFC, em uma edição presencial da competição, com uma turma de ensino superior.

O conjunto de exercícios utilizados e os contextos de aplicação são diferentes, portanto, a tabela serve apenas para registrar os valores.

5 Conclusão

A tabela abaixo mostra o resumo das conclusões que podem ser extraídas da análise dos dados:

Resultado estatístico	Conclusão
α de Cronbach igual a 0,76 em satisfação	As respostas obtidas mostraram que as questões relativas a este atributo foram consistentes substancialmente.
α de Cronbach = 0,71 em conscientização	As respostas obtidas mostraram que as questões relativas a este atributo obtiveram consistência substancial.
α de Cronbach igual a 0,90 em dificuldade	As respostas obtidas mostraram que as questões relativas a este atributo obtiveram consistência quase perfeita.
α de Cronbach = 0,04 em motivação	As respostas obtidas mostraram que as questões relativas a este atributo obtiveram consistência pequena.
Diferenças não significativas nas questões ($p > 0.05$)	Em números absolutos, os estudantes pioraram a percepção sobre a preparação para participação de competições de Segurança Computacional, com exceção das questões 1.1, 1.2 e 1.4. No entanto, as diferenças não foram estatisticamente significativas e não é possível afirmar estatisticamente que há diferença entre as respostas.
73,75% de desempenho da turma	A quantidade de acertos obtida pela turma, com um conjunto de exercícios considerado difícil pelo organizador da atividade, representa um resultado extremamente positivo.
95,16% de submissões corretas	A precisão nas respostas, se excetuadas as tentativas de resolução do easteregg (exercício extra) indica que as regras do jogo foram compreendidas pelos estudantes.

Com relação aos resultados dos questionários, não houve diferença estatisticamente significativa entre as respostas dos questionários pré- e pós-competição. Embora a amostra seja pequena e possa limitar os resultados, os indicadores com pequenos decréscimos entre as percepções pré- e pós-competição serão reavaliados para aplicação da competição em 2025 e discutidos com a turma.

Espera-se que em 2025 seja possível aplicar o jogo novamente em uma turma de curso superior.