

MindTrack: Modelando Dados de Testes Cognitivos

Luís F. Laguardia

^aFundação Getúlio Vargas, Rio de Janeiro, Brasil

1. Introdução

Recentemente na disciplina de Visualização de Dados participei do desenvolvimento do MindTrack, um site de testes cognitivos baseado em jogos digitais. A ideia por si só não é nova: diversos sites famosos, como Human Benchmark e CogniFit já oferecem esse serviço. O nosso diferencial é principalmente a abordagem na visualização dos resultados: entregamos ao usuário uma análise gratuita ao mesmo tempo clara e detalhada, ao passo que os outros serviços geralmente pecam em um (ou mais de um) desses aspectos.

Como todo trabalho de graduação, porém, o tempo disponível para realizá-lo foi ordens de grandeza menor do que o tempo necessário para realmente aproveitar todo o seu conteúdo e, dessa forma, não pudemos ir muito além da exibição dos dados coletados. Isso permitiu que começássemos a suspeitar de alguns padrões interessantes nos resultados dos testes, mas sem nenhum embasamento matemático/estatístico que desse algum tipo de certeza. Dessa forma, meu objetivo com este trabalho é aplicar os conceitos de modelagem estatística vistos em aula para explorar algumas das questões que surgiram diante dos dados obtidos.

1.1. Experimento

De início, é preciso explicar sobre quais são os dados disponíveis, como os testes cognitivos foram realizados e como foi dada a coleta desses dados.

Para todo usuário novo, era necessário passar por um formulário de criação de conta. Neste formulário, lhe eram apresentadas questões fixas: idade, sexo, região onde nasceu, experiência com jogos eletrônicos, além de outras cujas respostas não vieram a ser utilizadas.

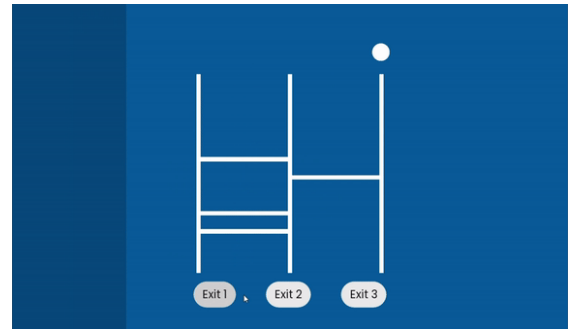


Figura 1: Exemplo de teste de raciocínio. É apresentado um labirinto ao usuário, e ele deve marcar qual saída a bolinha tomará, sabendo que ela virá em toda oportunidade. Nesse caso, a resposta é a “Saída 2”.

Na próxima etapa, antes de cada teste, era necessário responder um questionário curto sobre fatores que suspeitávamos que influenciariam nos resultados: quantas horas de sono o usuário teve, qual era o seu estado de humor e qual era a sua frequência de exercícios físicos na semana.

Por fim, eram realizados cinco testes digitais para diferentes áreas da cognição na seguinte ordem: atenção, coordenação, memória, percepção e raciocínio. A ordem é arbitrária, e a explicação do embasamento científico por trás dos jogos - isto é, a correlação entre o desempenho no jogo e a capacidade cognitiva na respectiva área - vai além do escopo deste trabalho, mas pode ser encontrada na seção de referências - Luís Laguardia, Tiago Barraudas, Maisa Fraiz (2023).

Os testes foram publicados na internet e, portanto, o ambiente de testagem foi diversificado e pouco controlado. Grande parte dos indivíduos tomou o teste de forma assistida durante a apresentação do trabalho, enquanto outra parte considerável tomou de forma autônoma através do site publicado. Ao todo, lidamos com os resultados de 64

testes completos nas modelagens aqui desenvolvidas.

1.2. Perguntas de Interesse

Algumas perguntas surgiram de preceitos já estabelecidos, como se existia correlação entre desempenho cognitivo e frequência de exercícios físicos. Outras surgiram a partir da observação dos dados e percepção de padrões. Neste trabalho, iremos abordar as seguintes perguntas:

- a É viável modelar os resultados de cada teste cognitivo como uma distribuição gaussiana?
- b Existe correlação de desempenho entre os testes, isto é, ir bem no teste de uma área indica tendência a ir bem no teste de outra área?
- c Existe correlação entre a quantidade de sono e o desempenho nos testes?
- d Existe correlação entre a experiência prévia com jogos eletrônicos e o desempenho nos testes?
- e Existe correlação entre a região onde o usuário nasceu e o desempenho nos testes?
- f A partir dos itens anteriores, é possível prever com acurácia aceitável o resultado de um teste?

2. Métodos

Começando pelo item (a): é viável modelar os resultados dos testes como gaussianas? Para essa pergunta, utilizamos o teste de Kolmogorov-Smirnov. Essa técnica não foi explicada em aula, mas sua interpretação é relativamente simples: seja um conjunto de dados unidimensional. É possível criar a CDF desse conjunto. Para cada item desse conjunto, calcula-se a distância da CDF nesse ponto até à CDF da distribuição almejada. Ao fim, retorna-se o valor da maior distância calculada, a estatística KS. Na maioria dos pacotes de estatística, ela é retornada juntamente a seu p-valor, e é comum definirmos que p-valores inferiores a 0.05 descartam a hipótese nula, isto é, indicam que o

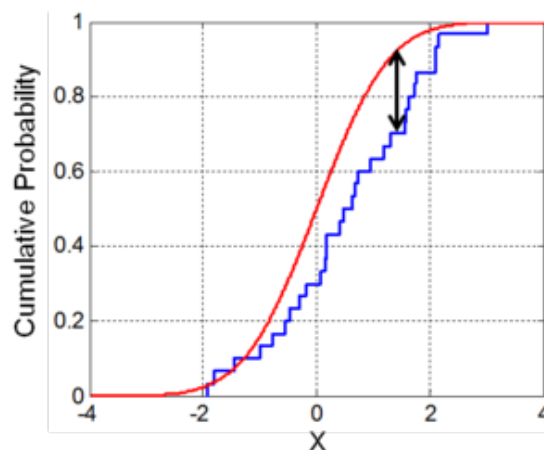


Figura 2: A faixa preta indica a estatística KS para a distribuição da imagem.

conjunto de dados não pertence à distribuição em questão - Wikipedia contributors (2023b).

Para o item (b), queremos responder se existe correlação entre os resultados dos testes. Para isso, utilizamos duas técnicas: A primeira e mais simples trata-se de calcular a matriz de correlação dos resultados. Ela é interessante pois dá uma estimativa rápida do resultado que queremos encontrar, mas é muito simples para abstrairmos a noção de incerteza sobre os dados. Para isso, utilizamos a segunda técnica: Para cada jogo i , modelamos seus resultados como uma variável dependente, e calculamos sua regressão linear a partir dos resultados de cada outro jogo j . Essa técnica nos permite calcular o coeficiente angular que relaciona as duas variáveis, retornando - tal qual a matriz de correlação - uma medida para quão atreladas duas variáveis estão; mas também nos permite calcular o intervalo de confiança para esse coeficiente angular, sanando o problema da incerteza.

Para calcular os intervalos de confiança, tivemos que optar por utilizar a técnica de estimação por máxima verossimilhança, isto é, EMV. Optamos por utilizar EMV e não Bayes pois, embora as medidas de incerteza da estimação bayesiana sejam mais interpretáveis, elas são inerentemente dependentes da priori escolhida para modelar os dados. Como os jogos foram criados pela nossa equipe e tiveram seus sistemas de pontuação assinalados de maneira relativamente arbitrária, era complexo encontrar alguma distribuição para os resultados que

fizesse sentido antes da observação dos dados. Na verdade, como veremos com os resultados da questão (a) na seção seguinte, é complexo assinalar uma distribuição para os resultados mesmo depois da observação dos dados. Desenvolvemos uma breve discussão a respeito disso na seção final.

Finalmente, para avaliar a bondade do ajuste dos modelos criados no item (b), utilizamos duas técnicas: R^2 e MAE. O R^2 foi escolhido por ser amplamente utilizado em aula e muito familiar no contexto estatístico. Sua interpretabilidade é simples e direta, e ele ajuda a passar uma ideia rápida de quão bem aquele modelo se ajusta aos dados. Para este trabalho, porém, ele tem uma desvantagem clara: é muito sensível a outliers, dado que depende dos erros quadráticos. Isso é um grande problema, pois nosso dado é bastante difuso e grande parte dos pontos não pode ser facilmente estimada a partir dos dados da variável independente. Para contornar esse problema, exibimos também o valor do erro absoluto médio, ou MAE (mean absolute error, em inglês). Este tem a desvantagem de não ter um valor padronizado - isto é, um MAE de 500 pode ser bom para um dado modelo, mas ruim para outro -, mas como nossos modelos serão similares, ele será útil para comparações entre os modelos.

Para responder às questões (c), (d) e (e), vamos utilizar técnicas similares às descritas anteriormente. É importante notar que os resultados aqui obtidos serão terão estatísticas de ajuste consideravelmente piores, porém, dado que as variáveis preditivas foram agrupadas de modo a se tornarem booleanas. Por exemplo, no item (c), onde queremos correlacionar sono e resultado nos testes, transformamos a coluna de horas de sono de forma que a escala original, que ia assumir os valores “ ≤ 6 ”, “7”, “8”, “9” e “ ≥ 10 ” na coluna “sono saudável”, que é verdadeira se o indivíduo dormiu de 7-9 horas de sono e falsa caso contrário. Fizemos essa transformação pois notamos que certos grupos de dados se comportam de maneira muito parecida, como ficará claro na análise exploratória da seção a seguir.

Por fim, no item (f) usaremos todas as conclusões obtidas nos itens anteriores para modelar o

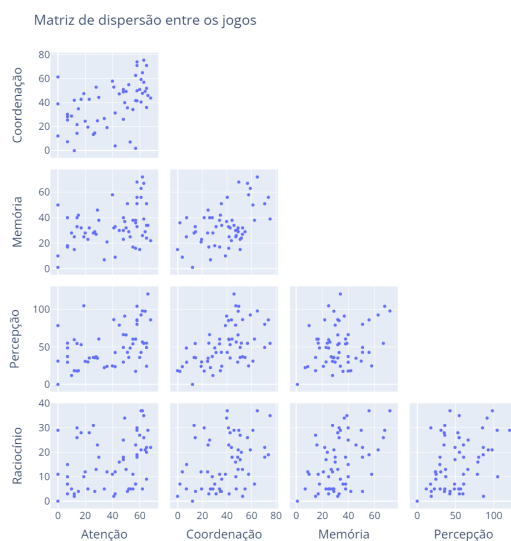


Figura 3: Matriz de dispersão dos resultados dos jogos.

resultado de cada jogo baseado nos resultados dos outros jogos e nas variáveis que se mostraram relevantes. Os resultados dos jogos são todos não-negativos, então os modelaremos usando um GLM com família Gamma e função de link logarítmica, pois ela respeita o espaço de valores da Gamma. Construiremos modelos considerando diferentes combinações de variáveis e, por isso, avaliaremos esses modelos usando a métrica AIC vista em aula, na intenção de balancear a eficácia e a complexidade das nossas previsões.

3. Resultados

Vamos começar fazendo uma rápida análise exploratória sobre os nossos dados. De início, o plot mais natural é a matriz de dispersão, onde cada ponto é a pontuação de um usuário em dois testes específicos. Ela nos dá os primeiros indícios de que existe alguma correlação entre as pontuações nos jogos, como nos jogos de coordenação com memória e percepção (vide Figura 3).

Além disso, também gostaríamos de visualizar os dados de cada teste individualmente, para detectar rapidamente se eles formam alguma distribuição conhecida (vide Figura 4).

Ainda, é necessário visualizar os dados adicionais de horas de sono e experiência com jogos. Para isso, fizemos os boxplots nas Figuras 5 e 6.

3.1. Item a

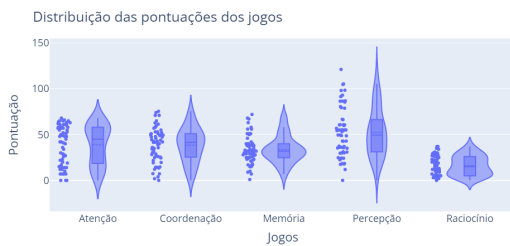


Figura 4: Dispersão individual dos resultados.

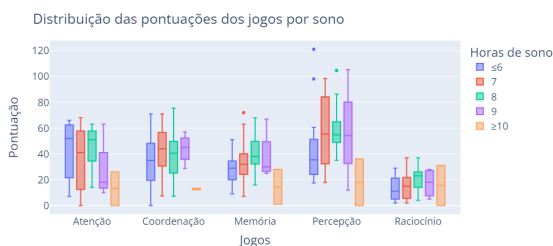


Figura 5: Boxplot de pontuação nos testes por horas de sono.

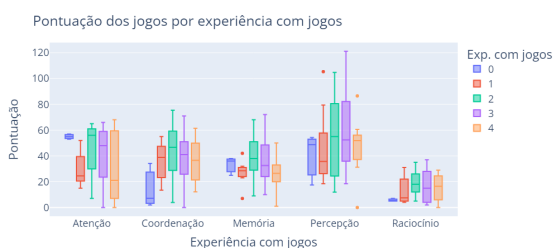


Figura 6: Boxplot de pontuação nos testes por nível de experiência com jogos eletrônicos.

Nosso primeiro objetivo é verificar se as pontuações nos testes podem ser aproximadas por alguma distribuição conhecida. Essa é uma propriedade interessante pois diversos testes cognitivos famosos, como o teste de QI, são meticulosamente desenhados para que seus resultados sejam aproximados por uma gaussiana - Wikipedia contributors (2023a). Isso nos daria a possibilidade de uma definição simples para testes futuros e nos permitiria usar uma ferramental muito maior de métodos estatísticos. Utilizando o teste de Kolmogorov-Smirnov para uma distribuição gaussiana, conseguimos os resultados da Tabela 1.

A partir disso, e observando quais P-valores se mantêm acima da faixa de 0.05, podemos manter a hipótese de que apenas os testes de Coordenação e Percepção advêm de distribuições normais. Com uma análise um pouco mais atenta, ainda é possível notar que, embora o teste de percepção tenha superado o valor 0.05, ele superou por uma margem ínfima centesimal, o que indica que esse valor tem grandes chances de mudar diante de uma nova amostragem. Sendo assim, concluímos que é difícil dizer que qualquer teste além do de Coordenação pode ter sua distribuição aproximada por uma normal e, descartamos essa possibilidade para o desenvolvimento deste trabalho.

Dois fatores importantes podem ser notados a partir disso, porém. O primeiro é que esses resultados poderiam ser preditos com certa facilidade a partir de uma leitura atenciosa da Figura 4. Todos as plotagens de violinos assumem formatos bem disformes, com exceção de exatamente a do teste de Coordenação. Dessa forma, fica evidente a importância de uma análise exploratória bem executada para o desenvolvimento de uma pesquisa profunda e reveladora sobre os dados. O segundo fator é que os únicos testes que obtiveram P-valor adequado foram os testes cuja pontuação permitia valores decimais, e não somente inteiros. Isso indica que a transformação do sistema de pontuações dos jogos em pesquisas futuras pode ser um grande passo em direção à aproximação por uma gaussiana.

Teste	Estatística KS	P-valor
Atenção	0.1548	0.0018
Coordenação	0.0737	0.6082
Memória	0.1202	0.0398
Percepção	0.1148	0.0585
Raciocínio	0.1185	0.0446

Tabela 1: Resultados da aplicação do teste de Kolmogorov-Smirnov aos dados de cada teste.

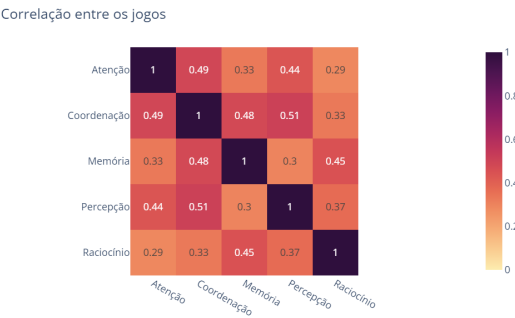


Figura 7: Matriz de correlação dos resultados dos jogos.

3.2. Item b

Também nos propomos a calcular a correlação entre as pontuações nos testes. A Figura 3 nos dá indícios de que todos os testes estão positivamente correlacionados, ou seja, que ir bem em um teste significa uma tendência para ir bem nos demais. A matriz de correlação na Figura 7 evidencia esse fenômeno ainda mais destacando as correlações mais fortes, como entre os dois testes rápidos coordenação e percepção.

Apesar disso, nenhuma delas nos dá uma noção de incerteza e, para isso, construímos um modelo OLS padrão entre cada par de pontuações dos testes. Os resultados obtidos seguem na Tabela 2.

É interessante perceber que os intervalos de confiança nos revelam dados que realmente não conseguíamos observar na matriz de correlação. Por exemplo, os dois pares cujas previsões têm intervalos de confiança mais curtos - isto é, que o modelo tem mais certeza das estimativas para os coeficientes angulares - são os pares (Memória, Percepção) e (Coordenação, Percepção), que são, respectivamente, um dos menores valores e o maior dos valores na matriz de correlação. Além disso, a maior parte dos pares tem uma estimativa para o coeficiente angular que supera a do (Coordenação,

Teste 1	Teste 2	Coef. Angular (95% de conf.)
Atenção	Coordenação	0.5626 ± 0.26
Atenção	Memória	0.4747 ± 0.35
Atenção	Percepção	0.3486 ± 0.19
Atenção	Raciocínio	0.5923 ± 0.51
Coordenação	Memória	0.5899 ± 0.28
Coordenação	Percepção	0.3553 ± 0.16
Coordenação	Raciocínio	0.5935 ± 0.44
Memória	Percepção	0.1678 ± 0.14
Memória	Raciocínio	0.6458 ± 0.33
Percepção	Raciocínio	0.9486 ± 0.62

Tabela 2: Valores dos coeficientes angulares das regressões lineares da pontuação de cada teste para cada outro teste.

Teste 1	Teste 2	R2	MAE
Atenção	Coordenação	0.2409	15.1599
Atenção	Memória	0.1122	17.5977
Atenção	Percepção	0.1902	16.2106
Atenção	Raciocínio	0.0847	18.0289
Coordenação	Memória	0.2276	13.8147
Coordenação	Percepção	0.2596	13.0211
Coordenação	Raciocínio	0.1118	14.1125
Memória	Percepção	0.0885	11.5138
Memória	Raciocínio	0.2024	10.8935
Percepção	Raciocínio	0.1389	20.801

Tabela 3: Medidas de bondade de ajuste de cada um dos modelos da Tabela 2.

Percepção), o maior valor na matriz de correlação. Isso é um bom indício que os valores da matriz estão sendo fortemente afetados pela quantidade de incerteza nos dados, e que uma nova amostra de dados possivelmente geraria uma matriz completamente diferente.

3.3. Item c

Neste item, estamos interessados em analisar a correlação entre a quantidade de sono do usuário e pontuação nos testes. Uma diversidade de estudos anteriores já indicou que níveis de sono muito anormais têm impactos negativos no desempenho cognitivo, tanto para sono em deficiência quanto para sono em excesso - SciELO (2023). Nesse sentido, essa análise serve principalmente para medirmos o quanto os dados obtidos nesse trabalho es-

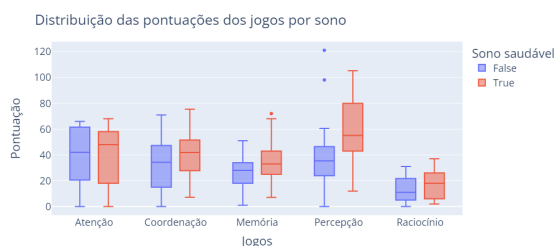


Figura 8: Box plots das pontuações com relação à coluna de sono transformada.

tão alinhados aos dados da população geral. É interessante notar análise exploratória nos dá sinais promissores: uma rápida leitura da Figura 5 indica que esse padrão se repete em nossos resultados.

Porém, um problema crucial nos impede de apenas fazermos uma regressão com os dados brutos: algumas das classes de respostas têm muito mais dados que as outras. Mais exatamente, enquanto a classe de "7 horas de sono" tem 23 pontos de dados, a classe de "10 ou mais horas de sono" tem apenas dois. Portanto, utilizamos uma técnica de agrupamento de dados: Todos os indivíduos que tiveram entre 7 e 9 horas de sono foram classificados como "sono saudável", diferente dos demais. O resultado final pode ser visualizado na Figura 8.

A partir desses dados, fizemos uma regressão linear simples para cada modelo, onde cada resultado era predito a partir de apenas a variável booleana de sono saudável recém-calculada. Mais uma vez, a intenção com esses modelos não era atingir uma grande acurácia na regressão - afinal, é impossível regredir bem um valor contínuo a partir de somente uma entrada booleana -, mas principalmente obter intervalos de confiança para as correlações entre a variável independente e o resultado de cada teste. Dessa forma, conseguimos os resultados na Tabela 4.

A partir dela, podemos observar que os resultados são relativamente significantes para o teste de Percepção: considerando o ínfimo de seu intervalo de confiança de 95 %, o coeficiente angular da variável de sono ainda está consideravelmente distante de zero. Intuitivamente, faz sentido que esse seja o teste com maior correlação, dado que ele exige do usuário um processamento rápido das informações apresentadas e uma reação instintiva

Teste	Coef. Angular (95% de conf.)	R2	MAE
Atenção	0.5789 ± 11.91	0.002	18.739
Coordenação	9.8939 ± 10.19	0.062	14.653
Memória	9.5329 ± 8.19	0.087	11.22
Percepção	19.5885 ± 14.33	0.116	19.376
Raciocínio	4.0118 ± 5.86	0.032	8.835

Tabela 4: Coeficientes angulares para a variável de "sono saudável" na regressão da pontuação de cada um dos testes.

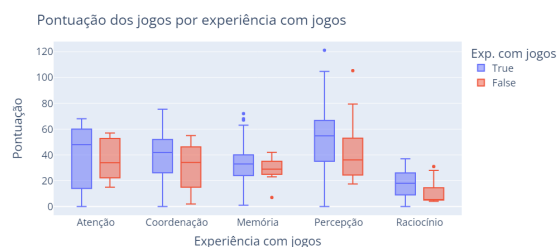


Figura 9: Box plots das pontuações com relação à coluna de experiência transformada.

acurada. Por outro lado, é intrigante que os resultados teste de atenção tenham apresentado uma interação com o sono tão próxima de zero. A nossa suspeita é de que o fato de a população dos testes ter sido majoritariamente jovem tenha algum tipo de impacto que faça com o que o sono não seja tão relevante para prever capacidade de atenção.

3.4. Item d

O item (d) se refere a correlação entre desempenho e experiências em jogos eletrônicos. A técnica de agrupamento utilizada no item anterior foi reutilizada aqui, colocando em um grupo aqueles indivíduos que diziam ter experiência zero ou um com jogos eletrônicos - segundo os textos do questionário original, indivíduos que não tinham experiência com nenhum jogo ou que eram familiares com um ou outro jogo -; e colocando em outro grupo os demais. O resultado desse agrupamento se encontra na Figura 9, e o resultado da aplicação da OLS na Tabela 5.

A partir desses resultados, vemos que todos os intervalos de confiança para os coeficientes angulares englobam o valor zero, o que indica que existe uma probabilidade considerável de que nossos testes não estejam atrelados ao nível de experiência em jogos eletrônicos. Essa é uma propriedade

Teste	Coef. Angular (95% de conf.)	R2	MAE
Atenção	4.1288 ± 14.25	0.006	18.57
Coordenação	9.5447 ± 12.37	0.04	14.929
Memória	5.7386 ± 10.17	0.022	11.29
Percepção	11.0995 ± 18.05	0.026	20.827
Raciocínio	6.125 ± 6.96	0.052	8.742

Tabela 5: Coeficientes angulares para a variável de "experiência com jogos eletrônicos" na regressão da pontuação de cada um dos testes.

muito desejável, já que, embora os testes não deixem de ser, de certa forma, outros jogos eletrônicos, eles foram feitos com o objetivo de serem imparciais quanto à experiência do jogador e avaliar apenas a sua capacidade cognitiva em determinada área.

É importante notar que alguns deles, apesar de conterem o zero em seu intervalo de confiança, estão extremamente enviesados para cima, como é o caso dos testes de Coordenação e Percepção. Isso significa que, embora exista a possibilidade de que seus resultados não estejam atrelados à experiência do jogador, existe uma probabilidade muito maior de que, na verdade, esteja. Esse resultado era esperado, porém, uma vez que ambos os testes exigem habilidades que jogos eletrônicos geralmente já exigem de seus jogadores, como raciocínio rápido e movimentos precisos. Isso faz com que eles treinem suas capacidades cognitivas nessas áreas naturalmente e, portanto, recebam maiores notas. Dessa forma, é possível concluir que todos os resultados aqui estão de acordo com o desejado.

3.5. Item e

O item e se refere à correlação entre o local de nascimento do usuário e seu desempenho nos testes. O local de nascimento na base de dados é descrito da seguinte forma: se o usuário for estrangeiro, ele é dado pelo país onde o usuário nasceu. Se o usuário tiver nacionalidade brasileira, porém, tem-se também a informação de a qual estado ele pertence. Com essa análise, esperávamos ver como os níveis de educação de cada região impactam os resultados dos testes.

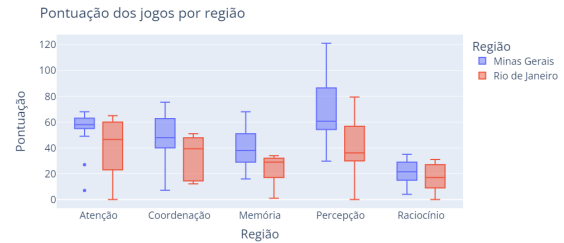


Figura 10: Box plots das pontuações com relação à região do usuário.

Teste	Coef. Angular (95% de conf.)	R2	MAE
Atenção	15.0889 ± 14.58	0.148	12.84
Coordenação	12.3311 ± 15.84	0.09	15.198
Memória	16.0111 ± 11.39	0.243	10.938
Percepção	28.3057 ± 19.42	0.257	19.868
Raciocínio	3.1222 ± 8.22	0.023	8.091

Tabela 6: Coeficientes angulares para a variável de "mora em Minas Gerais" na regressão da pontuação de cada um dos testes.

Novamente o baixo número de respostas por categorias fez com que tivéssemos que repensar a leitura dos dados. Nesse caso, optamos por selecionar apenas as duas regiões com maior número de respostas, Minas Gerais e Rio de Janeiro, com 19 e 11 respostas, respectivamente. O resultado do agrupamento está na Figura 10, e da aplicação da OLS na Tabela 6.

É possível perceber que há certa tendência para resultados maiores se o usuário é de Minas de Gerais em comparação a usuários do Rio de Janeiro. É difícil dizer a real razão, mas acreditamos que esse fator esteja relacionado com o fato de os níveis de educação do estado de Minas Gerais serem no geral mais elevados que os do estado do Rio de Janeiro - INEP (2023), e que isso de alguma forma impacte os resultados dos testes cognitivos.

3.6. Item f

Por fim, utilizamos todas as informações obtidas nas análises anteriores para tentar prever com algum grau de precisão o resultado do teste de Atenção do usuário a partir das outras variáveis disponíveis utilizando um modelo estatístico mais complexo. O teste de atenção foi escolhido arbitrariamente, e resultados similares podem ser compu-

	Coef.	[0.025	0.975]	std
Intercept	2.9564	2.472	3.440	0.247
Coordenação	0.0173	0.007	0.027	0.005
Memória	-0.0046	-0.017	0.008	0.006
Percepção	0.0082	0.001	0.016	0.004
Raciocínio	-0.0096	-0.028	0.009	0.009

Tabela 7: Fórmula: Atenção Coordenação + Memória + Percepção + Raciocínio. AIC: 343.2357

	Coef.	[0.025	0.975]	std
Intercept	2.9984	2.486	3.511	0.262
minas[T.True]	0.1430	-0.346	0.632	0.249
safesleep[T.True]	-0.1169	-0.567	0.333	0.230
Coordenação	0.0171	0.007	0.027	0.005
Memória	-0.0050	-0.019	0.009	0.007
Percepção	0.0076	-0.001	0.016	0.004
Raciocínio	-0.0094	-0.028	0.010	0.010

Tabela 8: Fórmula: Atenção minas[T.True] + safesleep[T.True] + Coordenação + Memória + Percepção + Raciocínio. AIC: 341.4859

tados para os outros testes cognitivos, mas foram aqui resumidos para manter o escopo de número de páginas desejado.

Como dito anteriormente, modelamos as notas do teste por meio de um GLM com a função Gamma como família e a função log como link. Essa opção foi escolhida pois o resultado de um teste é não-negativo, O que fazia com que uma simples regressão para uma gaussiana fizesse previsões claramente incabíveis. Testamos diversos conjuntos como variáveis independentes, e avaliamos a medida de ajuste AIC como indicadora de um modelo capaz de balancear corretamente qualidade de predição e complexidade. Ao fim, chegamos nas tabelas 7, 8 e 9.

Logo, conclui-se que o melhor modelo é o que considera as informações de todas as colunas.

4. Conclusão

Tiramos uma série de conclusões dos resultados descritos. Do item (a), percebemos a importância de uma análise exploratória feita com cuidado. Certas hipóteses só passam a ser cogitáveis após uma observação atenciosa dos dados e, para isso, é preciso aproveitar-se de técnicas e tecnologias moder-

	Coef.	[0.025	0.975]	std
Intercept	3.0868	2.483	3.691	0.308
minas[T.True]	0.1539	-0.365	0.673	0.265
safesleep[T.True]	-0.1272	-0.596	0.342	0.239
exp[T.True]	-0.2545	-0.738	0.229	0.246
Coordenação	0.0199	0.009	0.031	0.006
Memória	-0.0056	-0.020	0.009	0.007
Percepção	0.0076	-0.001	0.016	0.004
Raciocínio	-0.0087	-0.029	0.011	0.010

Tabela 9: Fórmula: Atenção minas[T.True] + safesleep[T.True] + exp[T.True] + Coordenação + Memória + Percepção + Raciocínio. AIC: 334.5404

nas para criar visualizações verdadeiramente interpretáveis. Além disso, a partir dessa análise também conseguimos um norte para um futuro desenvolvimento dos testes rumo a um sistema de pontuação mais justo e adequado, possivelmente adaptando os testes para que todos eles retornem pontuações avaliadas em casas decimais.

Do item (b), extraímos como até mesmo os métodos estatísticos mais simples, como uma regressão linear ordinária de uma variável para outra, nos permitem tirar conclusões sobre os dados que outros métodos muito comuns e bem estabelecidos deixam passar, como foi o caso com a matriz de correlação.

Dos itens (c), (e) e (d), observamos como o agrupamento de valores pode evidenciar padrões escondidos devido a baixa quantidade de amostras.

Do item (f), observamos a importância de utilizar estatísticas de bondade de ajuste mais avançadas para avaliarmos modelos não apenas por sua capacidade de prever corretamente os dados, mas também pela complexidade desnecessária introduzida por colunas adicionais.

Ainda, é importante observar algumas limitações claras desses trabalhos: o grupo de amostragem e o número de amostragens. Sobre a primeira, deve-se destacar que os usuários do Mind-Track nesse período recente de quando os dados foram coletados é composto majoritariamente por minha rede de amigos. Dessa forma, a população que forma sua base de dados é de indivíduos jovens (19.85 ± 2.96 anos de idade), onde mais de metade deles (66.12%) são estudantes da FGV. Ou seja, es-

ses resultados não podem ser extrapolados para a população em geral.

Quanto ao número de amostragens, o período de coleta de dados foi extremamente limitado, o que permitiu a conclusão de apenas 64 testes até o início desse trabalho. Mais do que isso, para certas categorias, como pessoas com nível 0 de experiência com jogos - isto é, nenhuma experiência, conseguimos apenas 3 pontos de dados. Dessa forma, uma série de conclusões aqui desenvolvidas foi limitada pelo alto grau de incerteza que esses números causaram.

Em futuros desenvolvimentos desse trabalho, há o interesse de avaliar como as mudanças nos sistemas de pontuações refletirão na distribuição dos testes e ampliar o número de amostras e avaliar como elas afetam os resultados aqui obtidos - como na conclusão de que o nível de experiência com jogos eletrônicos não afeta o resultado dos testes, por exemplo.

Agradecimentos

Meus agradecimentos também a todos aqueles que participaram dos testes e fizeram com que ambos os meus trabalhos fossem possíveis, especialmente a aqueles que participaram desde o início, enquanto os testes ainda estavam em desenvolvimento. Em especial, meus agradecimentos aos dois usuários mais dedicados: Daniel Falqueto, detentor dos recordes nos testes de Memória e Raciocínio, e João Paulo Silva, recordista no teste de Percepção.

Por fim, meus maiores agradecimentos a Tiago Barradas e Maisa Fraiz, membros do grupo da Disciplina de Visualização de Dados e que desenvolveram junto a mim todo o aplicativo do MindTrack.

Referências

- INEP, 2023. Notas estatísticas do censo escolar 2020. URL: [\url{https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_escolar_2020.pdf}](https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_escolar_2020.pdf).
- Luís Laguardia, Tiago Barradas, Maisa Fraiz, 2023. Mind-track publish paper. URL: [\url{https://github.com/fgv-vis-2023/final-project-human-benchmark/blob/main/FinalPaper.pdf}](https://github.com/fgv-vis-2023/final-project-human-benchmark/blob/main/FinalPaper.pdf).

- SciELO, 2023. Journal of nursing education and practice. URL: [\url{https://www.scielo.br/j/rlae/a/js8RQDyNF5pghv9cvC8Y8Gz/?lang=en&format=pdf}](https://www.scielo.br/j/rlae/a/js8RQDyNF5pghv9cvC8Y8Gz/?lang=en&format=pdf).
- Wikipedia contributors, 2023a. Intelligence quotient — Wikipedia. URL: [\url{https://en.wikipedia.org/wiki/Intelligence_quotient}](https://en.wikipedia.org/wiki/Intelligence_quotient).
- Wikipedia contributors, 2023b. Kolmogorov-smirnov test — Wikipedia. URL: [\url{https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test}](https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test).