

## Processo Seletivo 2024

Guilherme Braga Pinto – gbp1999@gmail.com

### Respostas:

1. No dataset existem alguns valores faltantes. Antes de começar a manipular os dados, trate essas informações e descreva sucintamente as alterações feitas.

Limpeza feita em código. Defini o *schema* dos dados, importei, verifiquei a quantidade de nulos, preenchi os faltantes, determinei se as categorias de uso de drogas não tinham um valor que não podia estar ali (sempre CL0, CL1, CL2, CL3, CL4, CL5 e CL6), limpei o nome dos países e salvei o dataframe.

2. Qual é a distribuição da idade dos indivíduos na amostra? Existem diferenças significativas nas faixas etárias predominantes de consumo entre os grupos de usuários de diferentes substâncias?

- Legalh - relação forte, pouco uso. Uso expressivo na faixa: 18-24 anos (CL3).
- Nicotine - relação forte, muito uso. Uso expressivo na faixa: 18-24/25-34 (CL6). Nenhum uso expressivo na faixa: 25-34 (CL0).
- Amphet - relação forte, pouco uso. Uso maior na faixa: 18-24 (CL3).
- Amyl - relação forte, pouco uso. Uso maior na faixa: 25-34 e 18-24 (CL2).
- Cannabis - relação forte, muito uso. Uso maior na faixa: 18-24 (CL6 e CL5).
- Coke - relação forte, pouco uso. Uso maior na faixa: 18-24/25-34 (CL3 e CL2 respectivamente).
- Ecstasy - relação forte, pouco uso. Uso maior na faixa: 18-24 (CL3 e CL4).
- LSD - relação forte, pouco uso. Uso maior na faixa: 18-24 (CL3).

- Mushrooms - relação forte, pouco uso. Uso maior na faixa: 18-24 (CL3).
- VSA - relação forte, pouco uso. Uso maior na faixa: 18-24 (CL2).
- **Faixa etária de interesse: 18-24 anos aparenta ser uma faixa etária de risco.**

3. Há uma relação entre o nível educacional e o consumo de substâncias?

### **Respondendo a questão 3:**

Sim, há relação relevante para algumas drogas.

### **Respondendo a questão 10:**

- Uso de Legalh aparenta ter muitos grupos de níveis educacionais maiores que nunca tiveram contato com a droga. A relação é negativa (maior nível educacional, menor uso). Categoria de interesse: "dropouts" com nível CL3 (o resto das categorias com maior uso não apresentam uso considerável).
- Uso de Nicotina aparenta ter muitos grupos de níveis educacionais maiores que tiveram contato com a droga. A relação é positiva (maior nível educacional, maior uso). Categoria de interesse: pessoas que terminaram o ensino superior que nunca tiveram contato com a droga (o resto das categorias com maior uso apresentam uso considerável).
- Uso de Amphet aparenta ter muitos grupos de níveis educacionais maiores que nunca tiveram contato com a droga. A relação é negativa (maior nível educacional, menor uso). Categoria de interesse: "dropouts" no topo das pessoas que tiveram contato com a droga com categoria CL3. Categoria de interesse: pessoas que terminaram o ensino superior que nunca tiveram contato com a droga (o resto das categorias com maior uso apresentam uso considerável).
- Uso de cannabis aparenta estar fortemente ligado com o nível educacional "Some college or university, no certificate or degree". Ou seja, pessoas que entram na universidade e saem, além de pessoas que terminaram a universidade ("University degree"), em categoria CL6. A relação é positiva (maior nível educacional, maior uso). Categoria de interesse: pessoas que terminaram o ensino superior que nunca tiveram contato com a droga (o resto das categorias com maior uso apresentam uso considerável).

- Uso de Ecstasy aparenta ter muitos grupos de níveis educacionais maiores que nunca tiveram contato com a droga. A relação é negativa (maior nível educacional, menor uso). Categoria de interesse: "dropouts" no topo das pessoas que tiveram contato com a droga com categoria CL3/CL2/CL4 (o resto das categorias com maior uso não apresentam uso considerável).
  - Uso de LSD aparenta ter muitos grupos de níveis educacionais maiores que nunca tiveram contato com a droga. A relação é negativa (maior nível educacional, menor uso). Categoria de interesse: "dropouts" no topo das pessoas que tiveram contato com a droga com categoria CL3 (o resto das categorias com maior uso não apresentam uso considerável).
  - Uso de Mushrooms aparenta ter muitos grupos de níveis educacionais maiores que nunca tiveram contato com a droga. A relação é negativa (maior nível educacional, menor uso). Categoria de interesse: "dropouts" no topo das pessoas que tiveram contato com a droga com categoria CL3/CL2/CL4 (o resto das categorias com maior uso não apresentam uso considerável).
4. Como o gênero influencia no consumo de drogas alucinógenas (LSD, Ecstasy, Ketamine, Cannabis e Mushrooms)? Explique.

Agrupando quem teve **algum contato com a droga**, podemos afirmar que nessa base de estudo mais homens usaram alucinógenos do que mulheres. O mesmo pode ser dito ao agrupar o uso dado sexo para pessoas que efetivamente **abusaram** da substância.

Não é muito interessante a noção de que isso explica o uso de uma substância dado gênero, *minha resposta é mais uma análise exploratória dos dados que se tem em mãos*. Para responder de fato como uma coisa influencia outra, um estudo mais aprofundado deve ser conduzido.

5. Qual é a proporção de participantes que se auto-classificam como impulsivos (score superior a zero)? Existe uma correlação entre a impulsividade e o consumo de substâncias?

Usuários mais impulsivos tiveram aumento sensível de abuso de Nicotina e Cannabis ao se realizar uma análise por alto dos padrões de consumo mais comuns para as drogas (análise de moda).

**6. Classifique as variáveis entre qualitativas (ordinal ou nominal), ou quantitativas (discreta, contínuas).**

|-- ID: Quantitativa discreta  
|-- Age: Qualitativa ordinal  
|-- Gender: Qualitativa nominal  
|-- Education: Qualitativa ordinal  
|-- Country: Qualitativa nominal  
|-- Ethnicity: Qualitativa nominal  
|-- Income (USD): Quantitativa contínua  
|-- Nscore: Quantitativa contínua  
|-- Escore: Quantitativa contínua  
|-- Oscore: Quantitativa contínua  
|-- AScore: Quantitativa contínua  
|-- Cscore: Quantitativa contínua  
|-- Impulsive: Quantitativa contínua  
|-- SS: Quantitativa contínua  
|-- Alcohol: Qualitativa ordinal  
|-- Amphet: Qualitativa ordinal  
|-- Amyl: Qualitativa ordinal  
|-- Benzos: Qualitativa ordinal  
|-- Caff: Qualitativa ordinal  
|-- Cannabis: Qualitativa ordinal  
|-- Choc: Qualitativa ordinal  
|-- Coke: Qualitativa ordinal  
|-- Crack: Qualitativa ordinal  
|-- Ecstasy: Qualitativa ordinal  
|-- Heroin: Qualitativa ordinal  
|-- Ketamine: Qualitativa ordinal  
|-- Legalh: Qualitativa ordinal  
|-- LSD: Qualitativa ordinal  
|-- Meth: Qualitativa ordinal  
|-- Mushrooms: Qualitativa ordinal  
|-- Nicotine: Qualitativa ordinal  
|-- Semer: Qualitativa ordinal

|-- VSA: Qualitativa ordinal

- 7.** Qual é a proporção de consumo de substâncias legais versus ilícitas na amostra (considere a definição de legalidade segundo a legislação brasileira)?

Proporção de 2.7 vezes o abuso de substância de drogas legais para as ilegais. Código para mais detalhes.

- 8.** Quais fatores predizem a probabilidade de um indivíduo consumir crack (Crack)?

Ao se analisar as features mais importantes de uma árvore de decisão que determina o uso de crack, podemos determinar que o uso de Crack está altamente correlacionado com o uso de Metanfetamina, Cocaína e Heroína.

Legenda/explicação:

- 9.** Qual é a média das pontuações Nscore, Escore, Oscore, AScore, Cscore? Calcule a correlação entre elas.

As médias tendem ao seu ponto central.

Maiores correlações (positivas) ocorrem entre CScore e AScore (0.24), além de CScore e OScore (0.30).

Maiores correlações (negativas) ocorrem entre Nscore e CScore (-0.39) e NScore e Escore (-0.43).

- 10.** Analise a relação entre o nível de educação (Education) e o consumo de diferentes substâncias ilícitas (como LSD, Amphet, Cannabis, etc.). Identifique se há uma correlação significativa entre essas variáveis e, em caso afirmativo, explore a natureza dessa correlação (positiva/negativa).

Condensei com a questão 3.

- 11.** Treine uma árvore de decisão para prever se um indivíduo consome uma

determinada substância (por exemplo, álcool, anfetaminas, cannabis) com base em suas características demográficas e pontuações de personalidade. Utilize a acurácia para avaliar os seus resultados.

Código no meu notebook, acurácia de 0.85.

Legenda/explicação:

- 12.** Explore a correlação entre a idade (variável Age) e a experimentação de diferentes substâncias ilícitas. Verifique se há uma tendência de aumento ou diminuição do consumo conforme a idade avança.

A tendência é que as ocorrências do uso de substâncias ilícitas diminui com a idade. Mais explicações no print dado no meu notebook.

Legenda/explicação:

- 13.** Quais são as 3 drogas mais utilizadas para cada país presente na amostra? E quais são as 3 menos utilizadas?

-----

USA

Drogas mais frequentes - Alcohol - CL5

Drogas mais frequentes - Caff - CL6

Drogas mais frequentes - Cannabis - CL6

Drogas mais frequentes - Choc - CL6

Drogas mais frequentes - Nicotine - CL6

-----

Republic of Ireland

Drogas mais frequentes - Alcohol - CL5

Drogas mais frequentes - Caff - CL6

Drogas mais frequentes - Cannabis - CL6

Drogas mais frequentes - Choc - CL5

Drogas mais frequentes - Nicotine - CL6

-----

Republic of Ireland

Drogas mais frequentes - Alcohol - CL5  
Drogas mais frequentes - Caff - CL6  
Drogas mais frequentes - Cannabis - CL6  
Drogas mais frequentes - Choc - CL5  
Drogas mais frequentes - Nicotine - CL6

-----

#### New Zealand

Drogas mais frequentes - Caff - CL6  
Drogas mais frequentes - Cannabis - CL6  
Drogas mais frequentes - Choc - CL6

-----

#### Australia

Drogas mais frequentes - Alcohol - CL5  
Drogas mais frequentes - Caff - CL6  
Drogas mais frequentes - Cannabis - CL6  
Drogas mais frequentes - Choc - CL6  
Drogas mais frequentes - Nicotine - CL6

-----

#### UK

Drogas mais frequentes - Alcohol - CL5  
Drogas mais frequentes - Caff - CL6  
Drogas mais frequentes - Choc - CL6

-----

#### USA

Drogas menos frequentes - Alcohol - CL1  
Drogas menos frequentes - Benzos - CL1  
Drogas menos frequentes - Caff - CL0  
Drogas menos frequentes - Cannabis - CL1  
Drogas menos frequentes - Legalh - CL1  
Drogas menos frequentes - Meth - CL1  
Drogas menos frequentes - Nicotine - CL1

-----

#### Republic of Ireland

Drogas menos frequentes - Alcohol - CL0  
Drogas menos frequentes - Amphet - CL1  
Drogas menos frequentes - Nicotine - CL0  
Drogas menos frequentes - Semer - CL0

-----

## Canada

Drogas menos frequentes - Alcohol - CL0

Drogas menos frequentes - Cannabis - CL1

Drogas menos frequentes - Semer - CL0

-----

## New Zealand

Drogas menos frequentes - Amyl - CL0

Drogas menos frequentes - Crack - CL0

Drogas menos frequentes - Ecstasy - CL0

Drogas menos frequentes - LSD - CL0

Drogas menos frequentes - Semer - CL0

-----

## Australia

Drogas menos frequentes - Alcohol - CL1

Drogas menos frequentes - Amphet - CL1

Drogas menos frequentes - Benzos - CL1

Drogas menos frequentes - Caff - CL0

Drogas menos frequentes - Cannabis - CL0

Drogas menos frequentes - Coke - CL1

Drogas menos frequentes - Legalh - CL1

Drogas menos frequentes - Nicotine - CL1

-----

## UK

Drogas menos frequentes - Alcohol - CL0

Drogas menos frequentes - Caff - CL1

Drogas menos frequentes - Semer - CL0