

EDA

Análise Exploratória de Dados

→ Uma etapa crucial no processo de entender e utilizar dados. Pense nela como uma investigação inicial em que analisamos os dados de diferentes formas para entender suas principais características, identificar padrões e detectar possíveis problemas.

Algumas etapas do processo:

Compreensão dos Dados → Procura de Padrões → Detecção de Anomalias → Construção de Hipóteses

Questionamentos que podem surgir durante o entendimento do problema que são respondidas pelo processo de análise exploratória de dados (AED ou EDA em inglês):

- **Quais as principais características do conjunto de dados?**

→ *O conjunto é composto por 2557 linhas e 19 colunas em seu estado original.*

→ Suas Colunas têm diferentes tipos de dados, sendo eles dados dos tipos, string, float e int.

- int: Número inteiro
- string: Dado textual
- float: Número com ponto flutuante

- Colunas:

gender: gênero do indivíduo

age: idade do indivíduo

working_professional_or_student: informa se o indivíduo é estudante ou trabalha

profession: profissão do indivíduo, informa com o que trabalha ou se está desempregado.

academic_pressure: reserva a pressão acadêmica do indivíduo

work_pressure: reserva a pressão que o indivíduo sofre no trabalho

study_satisfaction: informa o quão satisfeito está o indivíduo com o que ele estuda

job_satisfaction: reserva o quanto o indivíduo se sente satisfeito quanto ao cargo que exerce profissionalmente

sleep_duration: informa o período de sono do indivíduo

dietary_habits: reserva informações a respeito de como o indivíduo se alimenta

degree: nível de especialização técnica

have_you_ever_had_suicidal_thoughts_?: reserva se o indivíduo teve ou tem pensamentos suicidas

work/study_hours: informa quanto tempo o indivíduo passa estudando ou trabalhando (ou os dois)

financial_stress: reserva a informação do quanto o indivíduo se estressa com o financeiro de sua vida cotidiana

family_history_of_mental_illness: informa se o indivíduo tem histórico depressivo na família

depression: informa se o indivíduo tem depressão diagnosticada ou não

general_pressure: pressão geral no cotidiano do indivíduo

general_satisfaction: informa o quanto o indivíduo está satisfeito num contexto geral

- **Há valores ausentes ou sem propósito no conjunto?**

→ *O conjunto de dados conta com algumas colunas que não vão de encontro com nenhuma feature/ferramenta de resolução do problema, tem campos nulos, e inconsistência de padrão no tipo de valores armazenados, isso é ligeiramente resolvido com uma normalização e padronização desses mesmos dados*

- **Existem outliers nos dados?**

→ *Outliers são dados que excedem a média, são muito acima ou abaixo do que se enxergar se comparando aos seus semelhantes e nesse conjunto de dados é possível enxergar que sim existem dados desse tipo, mas eles não interferem nos resultados que o conjunto poderá fornecer pois quando os dados são normalizados e padronizados a existência de outliers se torna nula.*

- **Analisando o conjunto, quantas pessoas têm pensamentos suicidas?**

→ *Segundo as análises que foram realizadas no conjunto, quase 18% do público tem depressão diagnosticada, sendo a maior parte desse público jovem.*

- **Quantas das pessoas que têm pensamentos suicidas têm depressão ?**

→ *Pessoas que já tiveram pensamentos suicidas tem uma porcentagem de 28% do seu público com depressão confirmada.*

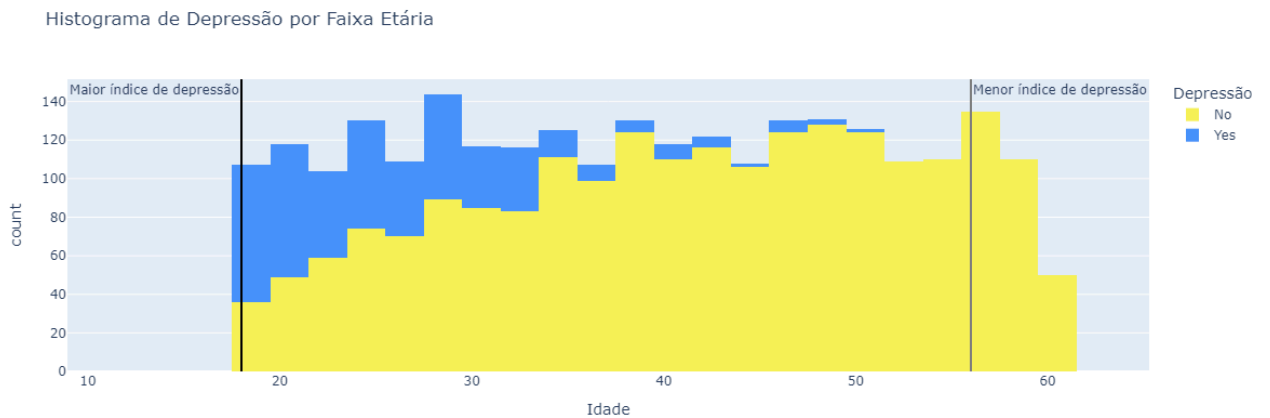
- **Pessoas com depressão são influenciadas por histórico depressivo familiar?**

→ *Por mais que pareça que o questionamento terá um resultado que afirme o mesmo, a relação de depressão com o histórico familiar do indivíduo é quase nula, sendo menor que 1%*

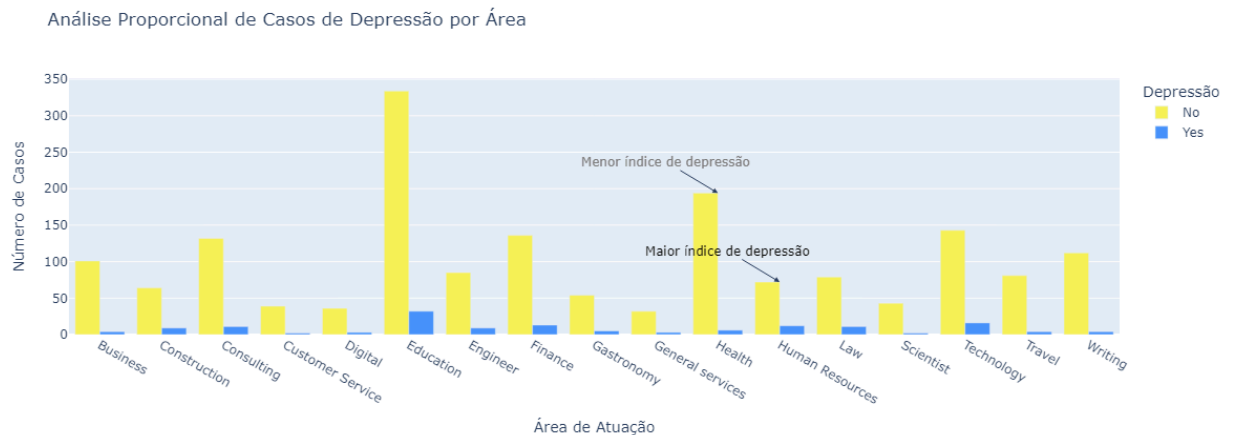
- As pessoas costumam ter depressão mais por estresse financeiro ou por pressão externa?

→ Analisando as duas relações a que se sobressai é a pressão geral, tendo 24%, enquanto o estresse financeiro tem relação 8% menor com a depressão, sendo assim, 16%

- Relação de depressão por faixa etária:



- Relação de casos de depressão por área:



- Índice de depressão por gênero e situação

Índice de depressão por Gênero e Situação

