



REQUISITOS DO PROJETO 01

Dados do componente curricular:

- IMD1151 - Ciências de Dados (60h)
- Semestre 2024.1
- Horário: 24T12
- Sala: A305
- Docentes: Daniel Sabino, Heitor Florencio

Objetivo do projeto:

Explorar e analisar uma base de dados estruturada de sua escolha, que contenha uma combinação de dados categóricos e numéricos. Esta é uma oportunidade para aplicarem os conhecimentos em um contexto real e explorarem a diversidade de aplicações da ciência de dados.

Etapas do projeto:

1. Escolha da Base de Dados: escolher uma base de dados estruturada que seja relevante para seus interesses ou área de estudo. Isso pode incluir conjuntos de dados disponíveis publicamente, dados de pesquisa própria ou dados fornecidos pela instituição em que trabalham.
2. Exploração Inicial dos Dados: realizar uma exploração inicial da base de dados para entender sua estrutura e qualidade. Examinem o número de observações, número de variáveis, tipos de variáveis (categóricas, numéricas), valores ausentes, etc.
3. Levantamento de conhecimento dos especialistas da área de negócio: realizar um levantamento de informações importantes sobre a base de dados para definir regras de negócio.
4. Pré-processamento de Dados: realizar o tratamento dos dados, se necessário, para garantir que estejam em um formato adequado para análise. Isso pode incluir o tratamento de valores ausentes, codificação de variáveis categóricas e normalização ou padronização de variáveis numéricas.



5. Análise Exploratória de Dados (AED): realizar uma análise exploratória dos dados para identificar padrões, tendências e insights interessantes. Utilizem estatísticas descritivas, visualizações gráficas, correlações entre variáveis e outras técnicas relevantes.
6. Interpretação e Conclusões: interpretar os resultados da análise e extrair insights significativos sobre o conjunto de dados. Sejam capazes de responder a perguntas importantes relacionadas aos dados e fornecer conclusões fundamentadas.

Requisitos do dataset:

- Pode ser um conjunto de dados de uma base pública, de dados de pesquisa própria ou dados fornecidos pela instituição em que trabalham.
- Deve conter obrigatoriamente dados qualitativos e quantitativos.
- O tamanho mínimo do dataset depende de cada caso e será analisado na etapa de validação do dataset.

Entrega:

Cada grupo deve entregar:

- O notebook no Google Colab contendo todo o processo de análise de dados.
- O relatório da análise de dados.
- O slide da apresentação do projeto.

Outras informações:

- O grupo deve ser composto por, no máximo, 3 membros.
- Todos os membros do grupo devem participar dos encontros de orientação e da apresentação final.
- Haverá uma fase de validação do dataset.

Mais informações, entrar em contato com os docentes da disciplina.