**Entrega AI & CHATBOT (Resumo):**

**Arquivo de Componentes:**

* Crie um arquivo componentes.txt com os nomes completos e RM dos alunos participantes.

Modelo de Dados Lógico:

* **Utilize o Data Modeler para desenvolver um modelo de dados lógico.** 
  + Este deve incluir:
    - Entidades.
    - Atributos (chave primária, chave estrangeira).
    - Relacionamentos.
    - Considere as classes desenvolvidas em backend Java e anexe o diagrama de classes. Este modelo será usado no CRUD em Python.
* **Script DDL para Estrutura do BD:**
  + Crie um script DDL para a estrutura do banco de dados, incluindo PKs (Primary Keys), FKs (Foreign Keys), NOT NULL, etc.
  + Use nomes significativos para tabelas, colunas e restrições.
* **Script DML para Popular Tabelas:**
  + Desenvolva um script DML para inserir dados nas tabelas, com pelo menos 10 linhas em cada tabela.
* **Script DQL/DRS para Consultas:**
  + Elabore pelo menos 4 consultas DQL/DRS, incluindo uma explicação do que se deseja recuperar do banco de dados. As consultas devem ser:
    - Uma consulta simples (SELECT/FROM/WHERE/ORDER BY).
    - Uma consulta com uma ou mais junções de tabelas (SELECT/FROM/WHERE/ORDER BY).
    - Uma consulta com função de grupo e agrupamento.
    - Uma consulta com função de grupo, agrupamento com filtro (HAVING) e junção de tabelas.
* **Critérios de Avaliação:**
  + Modelo de Dados Relacional (até 20 pontos).
  + SCRIPT DDL (até 20 pontos).
  + SCRIPT DML (20 pontos).
  + SCRIPT DQL (até 40 pontos).

**Entrega AI & CHATBOT (completo):**

* O uso de dados e inteligência artificial na área da saúde está revolucionando o diagnóstico, tratamento e gestão de pacientes, melhorando a precisão e eficiência dos cuidados médicos. Essas tecnologias têm o potencial de salvar vidas, otimizar recursos e impulsionar avanços médicos significativos.
* Nesta entrega sua equipe será desafiada a coletar e analisar dados de saúde relacionados às temáticas selecionadas (escolha 1 das “Metas da ODS 3” – veja slides anteriores).

**Qual a fonte de dados?**

* Vocês podem buscar dados de Saúde sobre doenças diversas no Kaggle ou qualquer outra fonte confiável. Não serão aceitos dados sintéticos (gerados aleatoriamente). Caso queiram trabalhar com uma base pronta, disponibilizaremos uma base de ataques cardíacos no Teams. Isto é um exemplo de doença crônica, não transmissível.

**Entrega:**

* Enviar via portal um Jupyter Notebook único para o grupo contendo (1) nome e RM de todos os alunos do grupo; (2) fonte da base de dados; (3) todas os códigos executados e sem erros; (4) comentários markdown sobre os principais blocos de código e conclusões
* Arquivo .py do servidor Flask adaptado para disponibilizar os modelos treinados no Jupyter Notebook.

**Uma vez com os dados em mãos, vocês devem:**

* + - * + **(20 pts) Entrega 1:** Fazer uma análise estatística exploratória da base de dados, resumindo as variáveis e o contexto geral da base selecionada;
        + **(20 pts) Entrega 2:** Construir 5 gráficos distintos utilizando as variáveis da base. Vocês podem fazer gráficos selecionando parte dos dados de acordo com filtros que julgarem pertinentes. Atenção:

(1) vocês devem variar a seleção de tipos de gráficos e variáveis investigadas. A repetição da mesma variável ou do tipo de gráfico em todas as figuras acarretará desconto de nota;

(2) todo gráfico deve conter título e legenda. Vocês estão livres para usarem quaisquer configurações a mais que desejarem;

(3) cada gráfico gerado precisa comunicar uma conclusão relevante/interessante. Caso seu gráfico seja complexo, não deixe de resumi-lo com algumas linhas em comentários markdown;

* + - * + **(40 pts) Entrega 3:** Construir 2 modelos preditivos distintos, selecionando-os entre os tipos vistos em sala: regressão, classificação ou clusterização. A avaliação partirá dos seguintes critérios:

PARA MODELOS SUPERVISIONADOS: (1) seleção de variável relevante à Saúde como label; (2) seleção apropriada do tipo de modelo com o label selecionado (regressão X classificação); (3) aplicação de métricas pertinentes ao modelo implementado.

PARA MODELOS NÃO SUPERVISIONADOS: (1) seleção de número de clusteres aderente ao problema; (2) interpretação dos clusteres formados.

PARA TODOS OS MODELOS: (1) implementação correta dos conceitos e passos de machine learning; (2) código executado sem erros; (3) conclusões analisando os resultados dos modelos e sua aplicabilidade real.

* + - * + **(20 pts) Entrega 4:** Adaptar o script do servidor Flask fornecido para que os modelos de IA criados na entrega 3 possam ser integrados a sua aplicação;