

Introdução a Banco de Dados

Trabalho Prático 1

SQL

Outubro 2023

1 Introdução

O objetivo deste trabalho é permitir que o aluno pratique os conceitos vistos em sala de aula e reforce seus conhecimentos sobre a linguagem SQL e álgebra relacional. Um arquivo notebook (*.ipynb) será disponibilizado contendo todos os recursos necessários para a conexão com o banco de dados e para a realização das consultas solicitadas. O aluno deverá elaborá-las a partir de expressões em linguagem SQL.

2 Esquema Relacional

O arquivo *bancoDadosViagens.sql*, utilizado pelo notebook, contém o banco de dados que iremos utilizar durante o trabalho e poderá ser acessado ao executar as 3 primeiras células do arquivo notebook (TP1.ipynb), conforme instruções presentes no tópico 4. Este banco possui um esquema elaborado exclusivamente para esta tarefa, que foi baseado no banco de dados abertos de "Viagens realizadas a serviço" disponibilizado pelo site Portal da Transparência - Dados Abertos (<https://portaldatransparencia.gov.br/download-dados/viagens>). Os dados são referentes às viagens realizadas por servidores do Governo Federal durante o ano de 2022. Contém dados das viagens, passagens, trechos viajados e pagamentos. O arquivo *dicionario.pdf* possui uma explicação detalhada das tabelas e atributos do banco. A figura 1 mostra o diagrama ER do banco de dados.

Atenção: alguns registros foram modificados para melhor aproveitamento dos dados no TP. Portanto, algumas informações podem não corresponder à realidade completa.

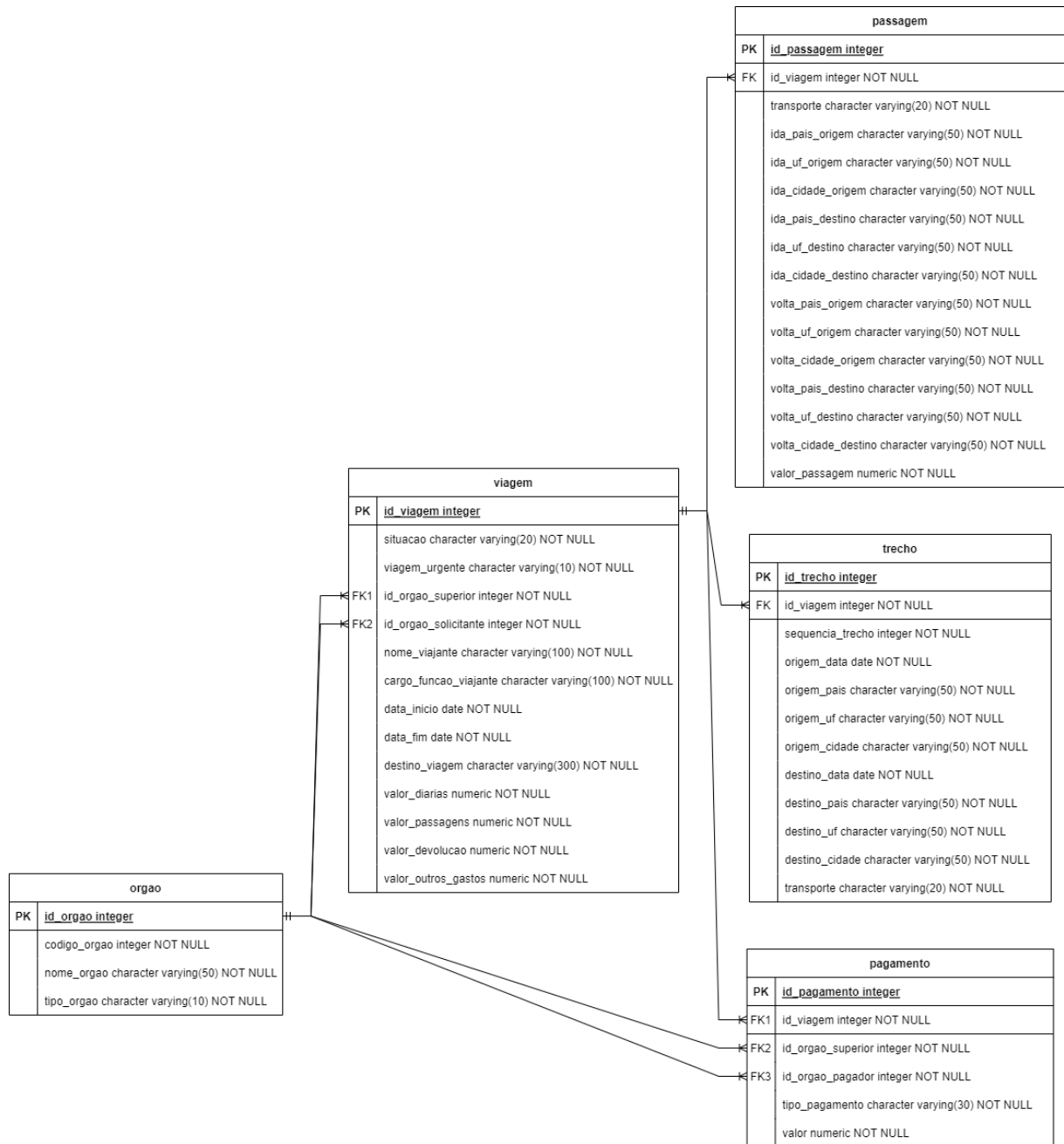


Figura 1: Esquema Relacional do Banco de Dados

3 Especificações das Consultas

3.1 Álgebra Relacional

Você deverá elaborar as consultas SQL equivalentes às expressões em álgebra relacional a seguir.

1. $\pi_{id_viagem, origem_cidade, destino_cidade} \sigma_{transporte=Rodoviario} (trecho)$

2. $\pi_{id_pagamento, valor} \sigma_{valor > 1000}$ (pagamento)
3. $\pi_{orgao.nome_orgao, pagamento.valor} \sigma_{pagamento.id_orgao_superior = orgao.id_orgao \wedge pagamento.valor < 200}$ (pagamento \times orgao)
4. $\pi_{id_viagem, sequencia_trecho, destino_viagem} \sigma_{sequencia_trecho = 5}$ (viagem \bowtie trecho)
5. $\pi_{id_pagamento, valor} \sigma_{viagem_urgente = SIM}$ (pagamento \bowtie viagem)

3.2 Especificações em Linguagem Natural

Você deverá elaborar as consultas SQL equivalentes às expressões em linguagem natural a seguir.

1. Liste os campos id_viagem, situacao, nome_viajante, data_inicio, data_fim e destino_viagem para todas as viagens com data de inicio entre 2022-01-09 e 2022-01-16 (incluso).
2. Liste todos os códigos e nomes dos órgãos que tenham a palavra "Nacional" no nome.
3. Liste os campos nome_viajante e cargo_funcao_viajante dos servidores que tenham nome que termina com "SANTOS" (pesquisar a palavra em caixa alta)
4. Liste os 50 primeiros nomes dos servidores e valores das passagens que tiveram "São Paulo" como o estado de origem da ida da passagem. O resultado deverá ser ordenado primeiramente pelo valor e depois pelo nome, os dois em ordem crescente.
5. Liste o nome do órgão superior e a quantidade de viagens urgentes realizadas que cada órgão superior pagou. Em ordem crescente de número de viagens.
6. Liste o campo id_viagem e o total do campo valor do pagamento de viagens solicitadas pelo órgão de código 52000.
7. Liste os ids das viagens que tenham mais de 3 passagens (4 ou mais) e o valor da passagem mais barata de cada uma.
8. Liste todos os ids dos trechos das viagens que não tem "Brasil" como país de origem e que o viajante possui a palavra "MINISTRO" no nome da função/cargo.
9. Liste, em ordem alfabética, os nomes de todos os órgãos superiores que custearam as despesas das viagens que tiveram gastos com diárias maiores que 0 e menores que 1000.
10. Liste os ids das viagens urgentes que têm algum trecho em que o transporte não foi "Aéreo" nem "Rodoviário", o nome do viajante e o valor total dos gastos da viagem indicados na tabela pagamento (com o nome "valor_total_pago_viagem"). O resultado deverá considerar apenas as viagens que tiveram gastos com "PASSAGEM" e deverá ser exibido em ordem alfabética em relação ao nome do viajante.

4 Tutorial - Configurações do banco de dados

Você pode utilizar a plataforma Google Colab (<https://colab.research.google.com/>) ou a plataforma Anaconda para usar o sqlite, construir o banco e executar suas consultas SQLs.

4.1 Google Colab

Na interface do Google Colab, vá em Arquivo - Abrir notebook e selecione o arquivo do TP1. Depois é só seguir as instruções no notebook.

4.2 Anaconda

1. O download da plataforma pode ser realizado através do site <https://www.anaconda.com/download/>.
2. Execute o arquivo de instalação e siga as instruções em <https://docs.anaconda.com/free/anaconda/install/>.
3. Acesse o Moodle e baixe o arquivo TP1.ipynb.
4. Abra o Anaconda Navigator e clique em 'Launch' no aplicativo Jupyter Notebook.

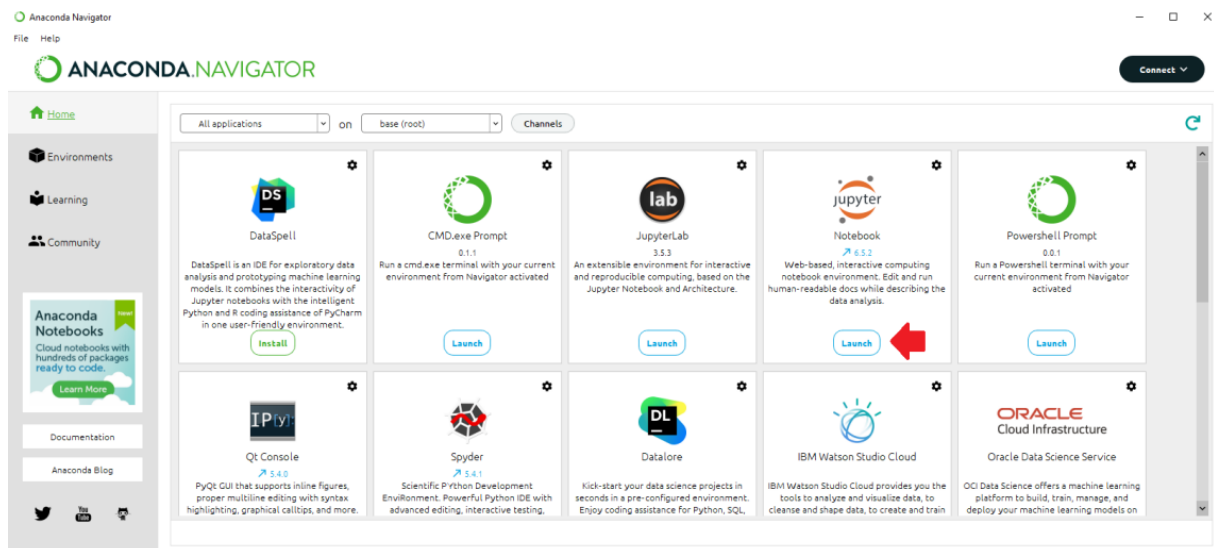


Figura 2: Acessando o Jupyter Notebook

5. Faça o upload do arquivo TP1.ipynb
6. Clique sobre o arquivo e veja que uma nova aba do navegador foi aberta apresentando o conteúdo do arquivo TP1.ipynb.

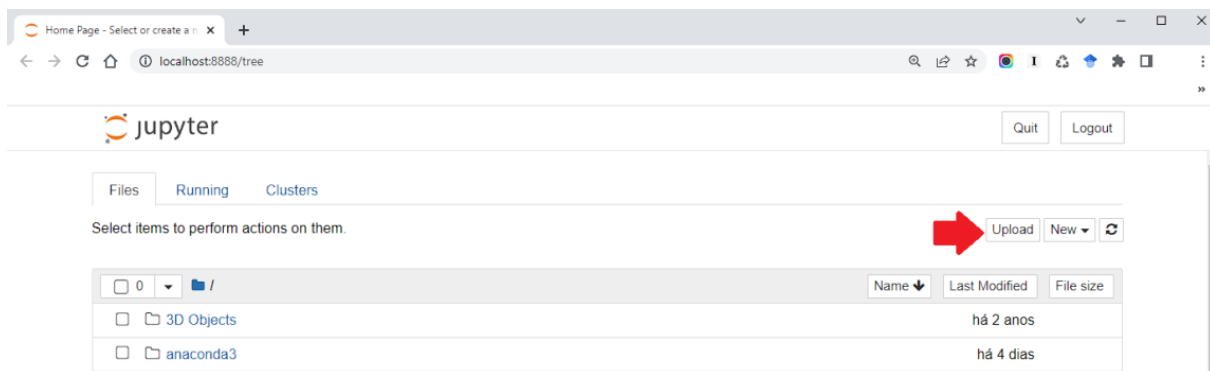


Figura 3: Upload do arquivo TP1.ipynb

7. Em caso de erro de importação, pode ser necessário baixar algum dos pacotes como o `sqlite3`, `pandas`, `csv` ou `gdown`. Para isso, abra o prompt de comando e digite o comando: **`pip install <nome_do_pacote>`**

```

Prompt de Comando
Microsoft Windows [versão 10.0.19045.2846]
(c) Microsoft Corporation. Todos os direitos reservados.

C:\Users\...>pip install gdown
Collecting gdown
  Downloading gdown-4.7.1-py3-none-any.whl (15 kB)
Collecting tqdm
  Downloading tqdm-4.65.0-py3-none-any.whl (77 kB)
Collecting requests[socks]
  Downloading requests-2.29.0-py3-none-any.whl (62 kB)
Collecting six
  Downloading six-1.16.0-py2.py3-none-any.whl (11 kB)
Collecting beautifulsoup4
  Downloading beautifulsoup4-4.12.2-py3-none-any.whl (142 kB)
Collecting filelock
  Downloading filelock-3.12.0-py3-none-any.whl (10 kB)
Collecting colorama; platform_system == "Windows"
  Downloading colorama-0.4.6-py2.py3-none-any.whl (25 kB)
Collecting urllib3<1.27,>=1.21.1
  Downloading urllib3-1.26.15-py2.py3-none-any.whl (140 kB)
Collecting certifi>=2017.4.17
  Downloading certifi-2022.12.7-py3-none-any.whl (155 kB)
Collecting charset-normalizer<4,>=2
  Downloading charset-normalizer-3.1.0-cp38-cp38-win_amd64.whl (96 kB)

```

Figura 4: Baixando o pacote gdown

8. Siga as instruções do arquivo TP1.ipynb.

5 Entrega

Você deverá entregar o arquivo .zip gerado pelo notebook contendo um arquivo .csv com o SQL para cada questão e o notebook com suas modificações. No final, seu arquivo

.csv deverá ter 16 linhas, uma para cada questão, e ainda o header do arquivo (*Query Number;Row Count;Query*).

6 Dúvidas

Em caso de dúvidas utilize o fórum criado para discussão sobre o tema no Moodle ou envie um e-mail para a monitora fernanda.ramalho@dcc.ufmg.br