

Trabalho Prático 1

Introdução a Banco de Dados

Data de Entrega: 2 de junho de 2025

1 Introdução

Este trabalho tem por objetivo a sedimentação dos conhecimentos vistos em sala relativos a bancos de dados relacionais, especialmente álgebra relacional e a linguagem SQL, por meio de contato prático com SGBDs relacionais. Conforme detalhado adiante, o trabalho consiste na elaboração de consultas SQL a partir de expressões em álgebra relacional e linguagem natural das informações que se pretende recuperar. Este trabalho utilizará o módulo VPL (*Virtual Programming Lab*) do Moodle e o SGBD SQLite.

2 Esquema Relacional

O banco de dados a ser utilizado contempla dados históricos reais da Fórmula 1 entre 1950 e 2018, com um esquema elaborado especialmente para este trabalho, conforme exibido na Figura 1. Arquivos contendo o banco de dados (`database.sql`) e uma descrição detalhada das tabelas e campos (`dicionario.pdf`) podem ser baixados no Moodle.

3 Especificação das Consultas

Para submissão, deverão ser elaboradas 10 (dez) consultas na linguagem SQL, sendo 5 (cinco) selecionadas da Seção 3.1 e outras 5 (cinco) da Seção 3.2.

3.1 Especificações em Álgebra Relacional

Neste bloco, você deverá elaborar consultas SQL equivalentes a expressões em álgebra relacional. Selecione as 5 (cinco) consultas da lista a seguir correspondentes aos 5 primeiros dígitos de seu número de matrícula, prefixados com **A**. Caso haja algum dígito repetido dentre esses, considere a consulta posterior¹ mais próxima ainda não selecionada. Por exemplo, para o número de matrícula 2015021992, as consultas seriam **A2** (correspondente ao 1º dígito), **A0** (2º), **A1** (3º), **A5** (4º) e **A3** (5º, considerando que as consultas **A0**, **A1**, e **A2** já teriam sido selecionadas).

A0. $\pi_{\text{NOME_PAIS}, \text{NOME_CIRCUITO}}(\text{PAIS} \bowtie \text{CIRCUITO} \bowtie \sigma_{\text{ANO_CORRIDA}=1987}(\text{CORRIDA}))$

A1. $\pi_{\text{NOME_CONSTRUTOR}}(\text{CONSTRUTOR} \bowtie \text{RESULTADO} \bowtie \sigma_{(\text{DESC_STATUS}='Velas' \vee \text{DESC_STATUS}='Turbo')}(\text{STATUS}))$

¹Considere o número 0 como posterior ao 9.

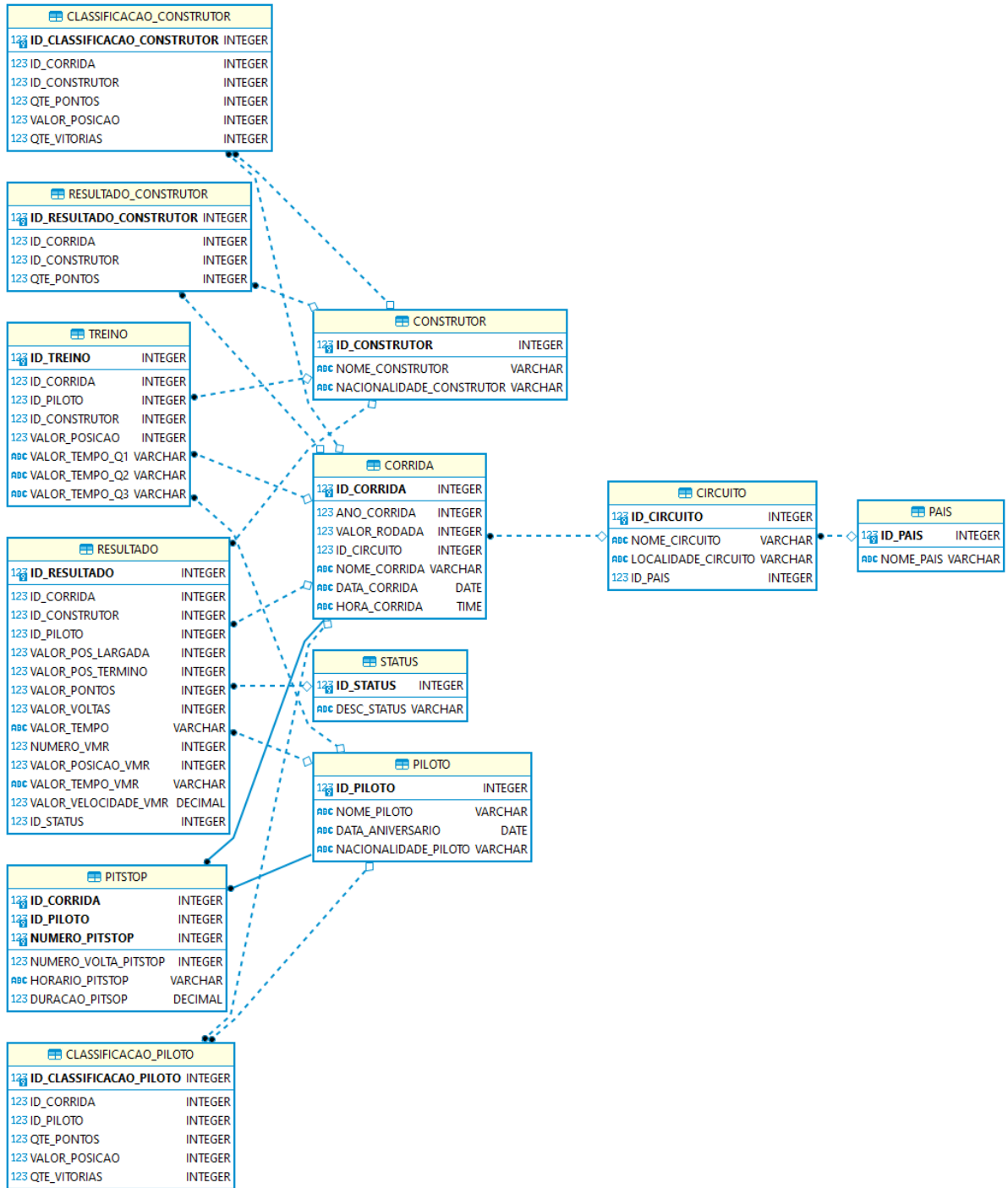


Figura 1: Esquema do banco de dados.

- A2. $\pi_{\text{NOME_PILOTO}, \text{NACIONALIDADE_PILOTO}}(\text{PILOTO} \bowtie \sigma_{(\text{VALOR_POS_TERMINO}=1 \wedge \text{VALOR_POS_LARGADA}=1 \wedge \text{VALOR_POSICAO_VMR}=1)}(\text{RESULTADO}))$
- A3. $\pi_{\text{NOME_PAIS}, \text{NOME_CIRCUIITO}}(\text{PAIS} \bowtie \text{CIRCUIITO})$
- A4. $\pi_{\text{NACIONALIDADE_PILOTO}}(\text{PILOTO}) \cap \pi_{\text{NACIONALIDADE_CONSTRUTOR}}(\text{CONSTRUTOR})$
- A5. $\pi_{\text{NOME_CIRCUIITO}, \text{ANO_CORRIDA}, \text{NOME_PILOTO}}(\text{CIRCUIITO} \bowtie \sigma_{\text{VALOR_POS_TERMINO}=1}(\text{RESULTADO})) \bowtie$

- $$\sigma_{\text{NACIONALIDADE_PILOTO} = \text{'Brasileira'}}(\text{PILOTO}) \bowtie$$
- $$\sigma_{(\text{ANO_CORRIDA} \geq 1971 \wedge \text{ANO_CORRIDA} \leq 1980)}(\text{CORRIDA}))$$
- A6. $\pi_{\text{NOME_PILOTO}}(\text{PILOTO} \bowtie \sigma_{\text{ANO_CORRIDA}=1991}(\text{CORRIDA}) \bowtie$
 $\sigma_{\text{QTE_VITORIAS} > 4}(\text{CLASSIFICACAO_PILOTO}))$
- A7. $\pi_{\text{ID_CORRIDA}, \text{DURACAO_PITSOP}}(\sigma_{\text{DURACAO_PITSTOP} < 14.0}(\text{PITSTOP}))$
- A8. $\pi_{\text{ANO_CORRIDA}, \text{NOME_PILOTO}}(\text{PILOTO} \bowtie \sigma_{\text{VALOR_POS_TERMINO}=1}(\text{RESULTADO}) \bowtie$
 $\sigma_{\text{ID_CIRCUITO}=5}(\text{CORRIDA}))$
- A9. $\pi_{\text{ANO_CORRIDA}, \text{NOME_CONSTRUTOR}, \text{QTE_PONTOS}}(\text{CONSTRUTOR} \bowtie$
 $\sigma_{\text{QTE_PONTOS} > 10}(\text{RESULTADO_CONSTRUTOR}) \bowtie$
 $\sigma_{\text{ID_CIRCUITO}=10 \wedge \text{ANO_CORRIDA} \geq 2016}(\text{CORRIDA}))$

3.2 Especificações em Linguagem Natural

Neste bloco, você deverá elaborar consultas SQL equivalentes a expressões em linguagem natural. Selecione as 5 (cinco) consultas da lista a seguir correspondentes aos 5 últimos dígitos de seu número de matrícula, prefixados com B. Caso haja algum dígito repetido dentre esses, considere a consulta posterior² mais próxima ainda não selecionada. Por exemplo, para o número de matrícula 2015021992, as consultas seriam B2 (correspondente ao 6º dígito), B1 (7º), B9 (8º), B0 (9º, considerando que a consulta B9 já teria sido selecionada) e B3 (10º, considerando que a consulta B2 já teria sido selecionada).

- B0. Listar (1) o nome do circuito e (2) nome do piloto vencedor de cada corrida de 2017;
- B1. *Pole position* é quando o piloto larga da 1ª posição. Listar (1) nome, (2) nacionalidade e (3) número de *pole positions* do piloto que mais vezes largou em 1º em 2015;
- B2. Liste (1) nome da equipe e (2) a duração do pit stop mais curto registrado na base;
- B3. Liste (1) o nome dos pilotos brasileiros que nunca pisaram no pódio, i.e. nunca terminaram uma corrida entre as 3 primeiras posições;
- B4. Liste (1) as nacionalidades e (2) o número de abandonos das equipes que já tiveram pelo menos 10 abandonos de prova por culpa da embreagem;
- B5. Liste (1) o nome dos países que já receberam mais de uma prova em uma mesma temporada;
- B6. Liste (1) o ano e (2) o nome do circuito de todas as corridas que já aconteceram em um mês de dezembro.
- B7. Liste (1) ano, (2) nome e (3) quantidade de pontos do piloto campeão de cada ano desde 2001.
- B8. Liste (1) os anos em que a última prova do ano foi disputada em um circuito do Brasil
- B9. Liste (1) os nomes dos circuitos nos quais o piloto Lewis Hamilton já venceu corridas e (2) quantas vitórias ele tem em cada um deles.

²Considere o número 0 como posterior ao 9.

4 Instruções para Submissão

Para submissão, as consultas elaboradas deverão ser incluídas em um único arquivo `submission.sql`, conforme template fornecido na tarefa criada para este trabalho prático no Moodle. A seguir, apresentamos um exemplo desse arquivo, considerando o número de matrícula 2015021992 utilizado nos exemplos da Seção 3.

```
-- 2015021992
-- A2
SELECT ... ;
-- A0
SELECT ... ;
-- A1
SELECT ... ;
-- A5
SELECT ... ;
-- A3
SELECT ... ;
-- B2
SELECT ... ;
-- B1
SELECT ... ;
-- B9
SELECT ... ;
-- B0
SELECT ... ;
-- B3
SELECT ... ;
```

Note que a primeira linha contém um comentário (denotado por `--` no início da linha) incluindo o número de matrícula do aluno. Em seguida, são listadas as 10 consultas elaboradas conforme a Seção 3, cada uma precedida por seu código correspondente (A2, A0, etc.) e encerrada com ponto-e-vírgula (;).³ O arquivo `submission.sql` completo deverá ser submetido via Moodle VPL para avaliação automática.

IMPORTANTE: Antes de submeter as consultas para avaliação automática no VPL, é aconselhável que você as teste localmente. Para tanto, você deverá criar e popular localmente o banco de dados a partir do script `database.sql` fornecido. Como preparação para as próximas atividades da disciplina, em particular o TP2, recomendamos que utilize o Jupyter Notebook⁴ para desenvolvimento, conforme visto nas aulas.

5 Avaliação

Este trabalho utilizará um modelo de avaliação automática via Moodle VPL. O avaliador segue as seguintes premissas básicas:

³Isso é fundamental, pois o arquivo será executado como uma lista única de comandos SQL.

⁴<https://jupyter.org/>

1. Serão avaliadas somente as 10 consultas selecionadas para cada aluno conforme seu número de matrícula (vide Seção 3), devidamente identificadas conforme o formato de submissão descrito (vide Seção 4).
2. As consultas elaboradas deverão executar sem erro no SQLite (sqlite3).
3. Cada consulta deverá retornar exatamente as colunas definidas e na ordem definida na especificação, conforme determinadas pela operação de projeção na Seção 3.1 ou explicitamente listadas na Seção 3.2.

Respeitadas as premissas, cada consulta será avaliada automaticamente quanto à correspondência do conjunto de resultados retornados R em relação ao conjunto de resultados esperados G . Para medida de correspondência, o avaliador utiliza o índice de Jaccard:⁵

$$J(R, G) = \frac{|R \cap G|}{|R \cup G|}. \quad (1)$$

Note que $J(R, G) \in [0, 1]$, o que permite atribuir pontuação parcial a cada uma das 10 consultas submetidas. A nota final do trabalho prático será a soma das notas atribuídas às 10 consultas, totalizando o máximo de 5 pontos (0,5 ponto por consulta).

5.1 Solicitação de Avaliação

Não há limite quanto ao número de submissões permitidas ao longo da duração do trabalho. Entretanto, somente a versão mais recente do arquivo `submission.sql` será avaliada. Para solicitar uma avaliação, o aluno deverá clicar no botão de avaliar disponível no editor do VPL, conforme Figura 2, ou na tela de “Visualizar envios”, conforme Figura 3.

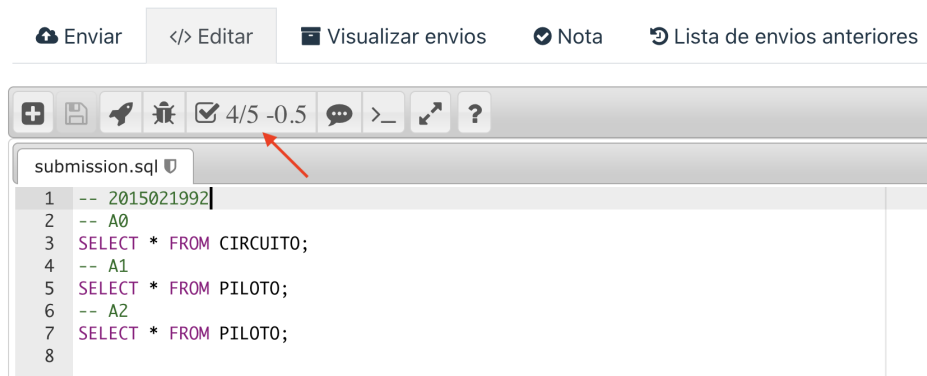


Figura 2: Editor do VPL e botão avaliar.

Após a execução da avaliação, a nota atribuída, bem como as mensagens de erro/sucesso poderão ser conferidas no painel da direita, no caso do editor do VPL (Figura 4), ou diretamente na tela de “Visualizar envios” (Figura 5). Caso deseje, o aluno poderá utilizar o *feedback* provido por essa avaliação para melhorar seu trabalho, realizar nova submissão e nova avaliação. **O aluno poderá solicitar até 5 (cinco) avaliações sem desconto na nota.** A partir da 6ª avaliação será descontado 0,5 (meio) ponto da nota final a cada nova solicitação. **A nota final será a nota da última avaliação realizada.**

⁵https://en.wikipedia.org/wiki/Jaccard_index



Figura 3: Botão “Avaliar” da tela “Visualizar envios”.

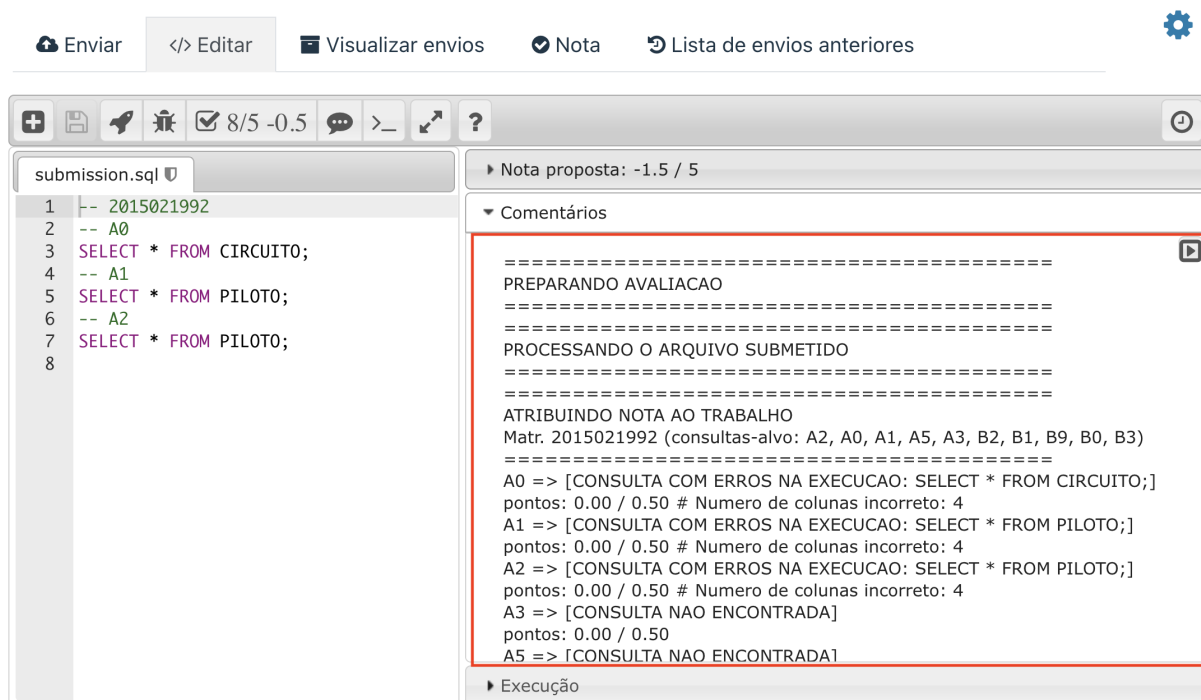


Figura 4: Resultado da avaliação na tela do editor VPL.

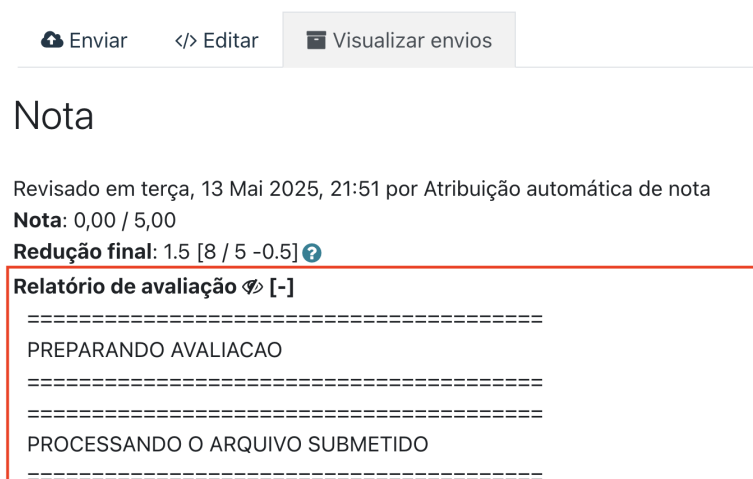


Figura 5: Resultado da avaliação na tela “Visualizar envios”.

5.2 Exemplo de Avaliação

A título de exemplo somente, o trecho a seguir mostra a saída produzida pelo avaliador para um arquivo de teste em que somente 4 consultas foram submetidas. A saída foi

editada para omitir o texto das consultas, substituídas pelo texto “[OMITIDO]”. As saídas produzidas são explicadas na sequência.

```
=====
PREPARANDO AVALIACAO
=====
=====
PROCESSANDO O ARQUIVO SUBMETIDO
=====
=====
ATRIBUINDO NOTA AO TRABALHO
Matr. 2015021992 (consultas-alvo: A2, A0, A1, A5, A3, B2, B1, B9, B0, B3)
=====
A0 => [CONSULTA EXECUTADA COM SUCESSO: [OMITIDO]]
pontos: 0.00 / 0.50
A1 => [CONSULTA EXECUTADA COM SUCESSO: [OMITIDO]]
pontos: 0.50 / 0.50
A2 => [CONSULTA EXECUTADA COM SUCESSO: [OMITIDO]]
pontos: 0.50 / 0.50
A3 => [CONSULTA EXECUTADA COM SUCESSO: [OMITIDO]]
pontos: 0.40 / 0.50
A5 => [CONSULTA COM ERROS NA EXECUCAO: [OMITIDO]]
pontos: 0.00 / 0.50 # Numero de colunas incorreto: 1
B0 => [CONSULTA NAO ENCONTRADA]
pontos: 0.00 / 0.50
B1 => [CONSULTA NAO ENCONTRADA]
pontos: 0.00 / 0.50
B2 => [CONSULTA NAO ENCONTRADA]
pontos: 0.00 / 0.50
B3 => [CONSULTA NAO ENCONTRADA]
pontos: 0.00 / 0.50
B9 => [CONSULTA NAO ENCONTRADA]
pontos: 0.00 / 0.50
TOTAL (antes dos descontos): 1.40 / 5.00
```

A consulta “A0” não levantou erros de execução e tinha as colunas corretas, por isso o avaliador informa sua execução com sucesso. No entanto, como nenhuma linha produzida coincidiu com o esperado, a nota atribuída foi 0 (zero) de 0,5 ponto possível.

As consultas “A1” e “A2” não continham erros e retornaram o resultado esperado, por isso receberam nota total de 0,5 ponto.

A consulta “A3” não continha erros e especificava as colunas corretamente, no entanto o coeficiente de similaridade de Jaccard entre o conjunto das linhas retornadas e esperadas foi 0,4 e a nota atribuída foi proporcional.

A consulta “A5” especificava incorretamente a quantidade de colunas frente ao solicitado e por isso recebeu avaliação 0 (zero).

As consultas “B1”, “B2”, “B3” e “B9” não estavam presentes no arquivo.

A nota exibida na última linha é a soma das notas individuais, sem contar eventuais descontos por solicitações adicionais de avaliação. No exemplo, a submissão obteve a nota não descontada de 1,40 de um total de 5 pontos possíveis.