|  |  |
| --- | --- |
|  |  |

**TESTE TÉCNICO**

**BUSINESS**

**INTELLIGENCE**

**Breves explicações sobre o processo de solução dos problemas**

**Glauber César Silva e Silva**

**Criar dashboard automatizado para acompanhamento de métricas de Apps música da Apple Store**

**1ª VERSÃO - MVP UTILIZANDO POWER BI PARA DATAVIZ E POWER QUERY PARA ETL:**

1. Conexão com arquivo csv no Google Drive
2. Criação do Modelo Dimensional:

|  |  |
| --- | --- |
| - d\_genre: gênero primário do aplicativo e id do gênero primário  - d\_app: dados categóricos referente ao aplicativo  - f\_apple\_store: demais dados quantitativos |  |

1. Criação dos gráficos e tabelas solicitadas:

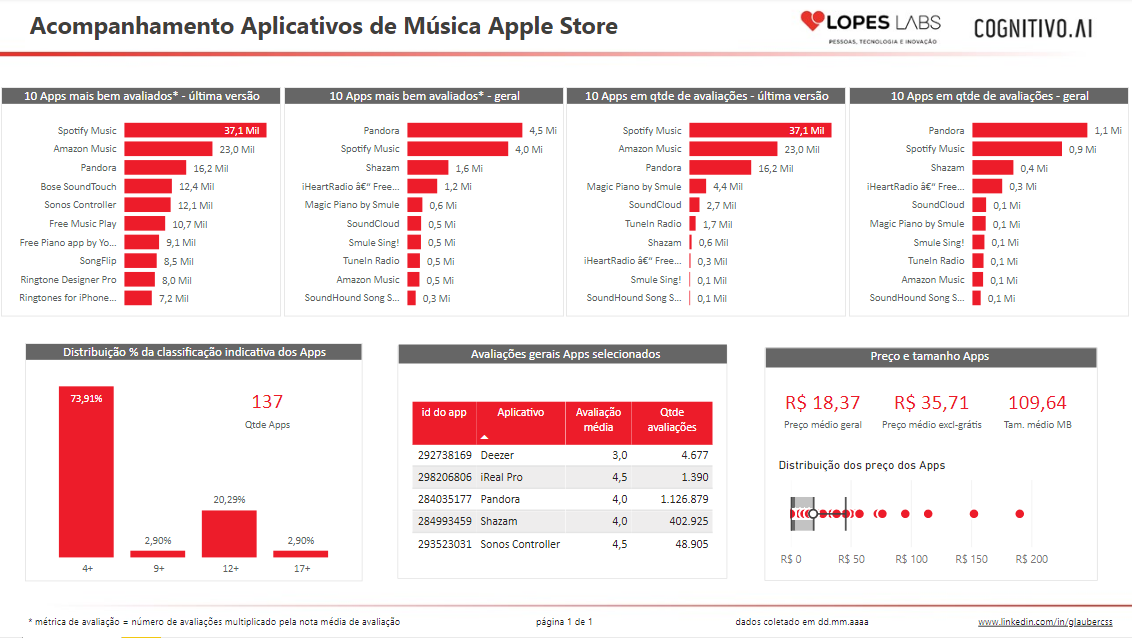
- Foi criada métrica para os gráficos “ranking dos 10 aplicativos mais bem avaliados” obtida pela multiplicação do número de avaliações pela nota média das avaliações. Do ponto de vista de negócio, penso ser importante ponderar a nota média com o número de avaliações afim de evitar distorções como aplicativo com apenas 1 avaliação e nota máxima figurar entre os mais ranqueados.

- Itens adicionadas:

. preço médio do aplicativo em Reais desconsiderando os aplicativos gratuitos ou *fremium* – como é o caso do Spotify -, que puxam a média para baixo;

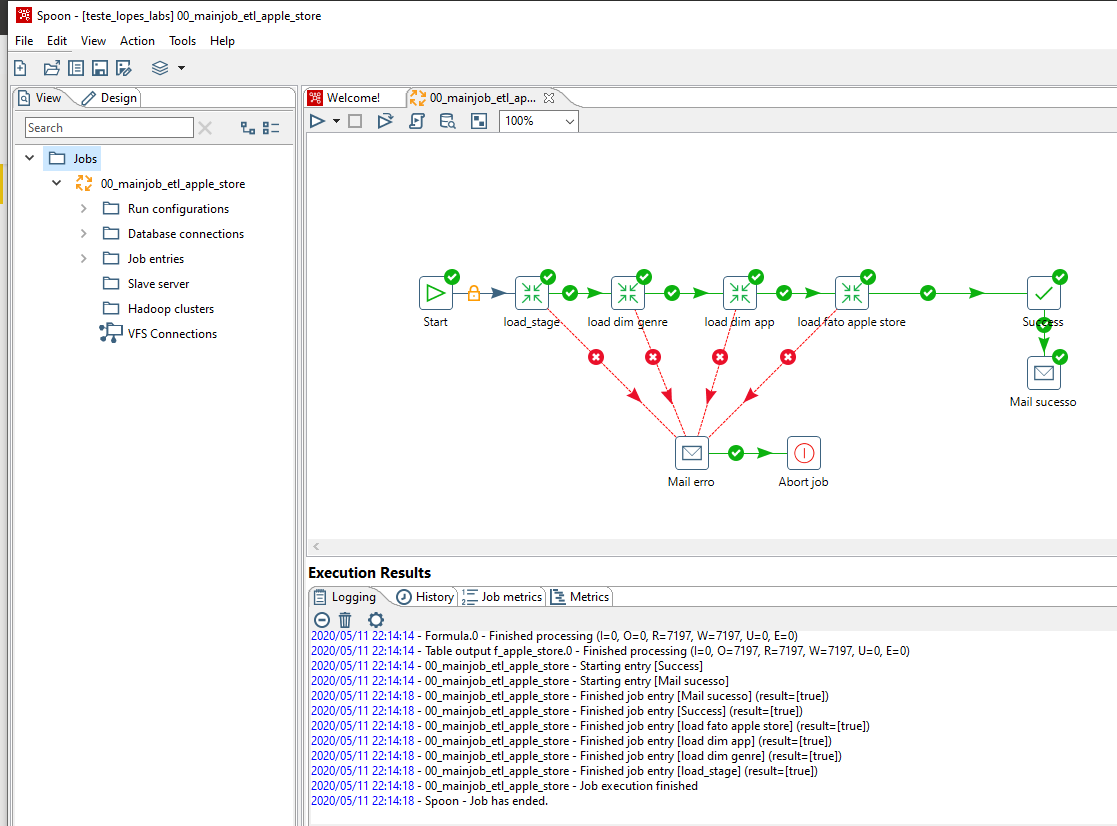
. quantidade de aplicativos de música;

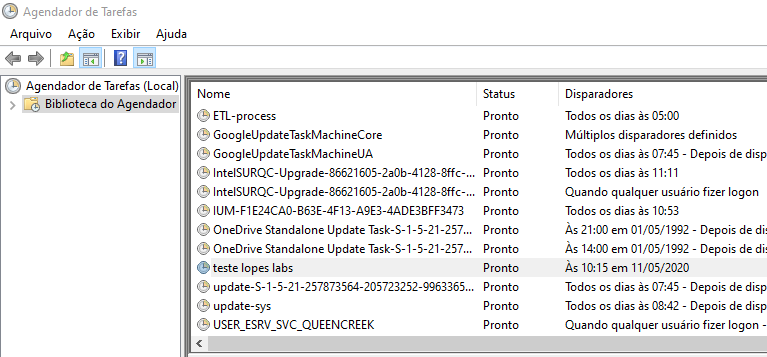
. Gráfico BoxPlot com a distribuição de preço dos aplicativos



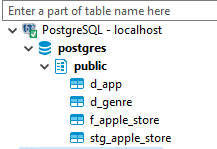
**2ª VERSÃO – UTILIZAÇÃO DO PENTAHO DATA INTEGRATION PARA ETL:**

- O processo de ETL para criação do modelo de dados e preparação foi realizado via a ferramenta *Pentaho Data Integration* inclusive com agendamento diário para execução automática do processo e economizar tempo dada a restrição para realização desta tarefa. O log do processo é enviado por e-mail para acompanhamento.





- As tabelas foram criadas em banco de dados PostgreSQL localhost.

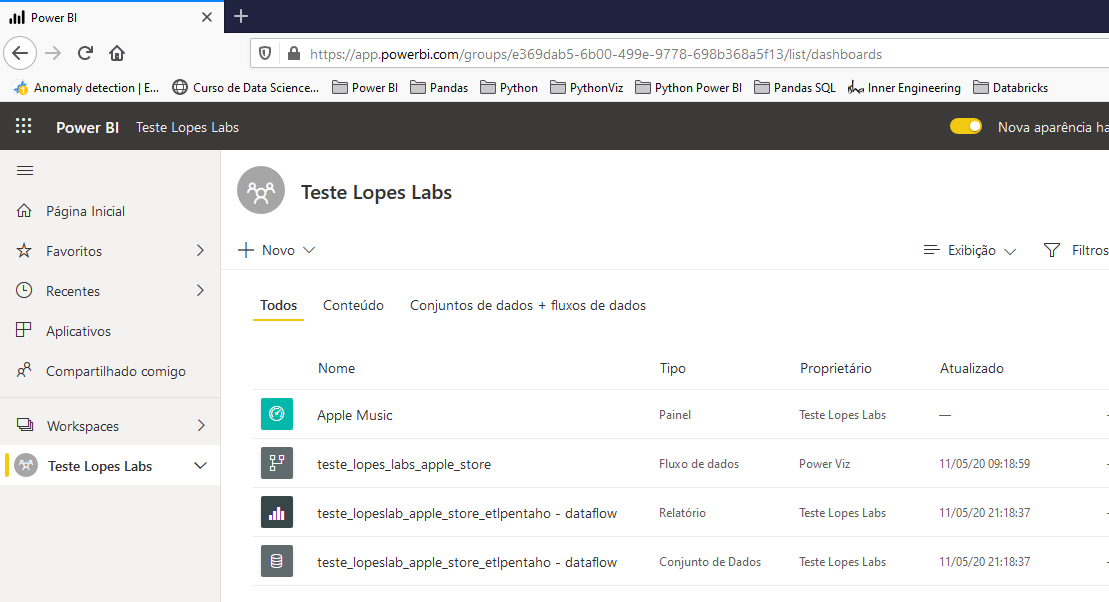


- Os dashboards executados na etapa 1 foram copiados (sem esforço extra) e são idênticos.

- A atualização do Power BI Desktop é feita manualmente

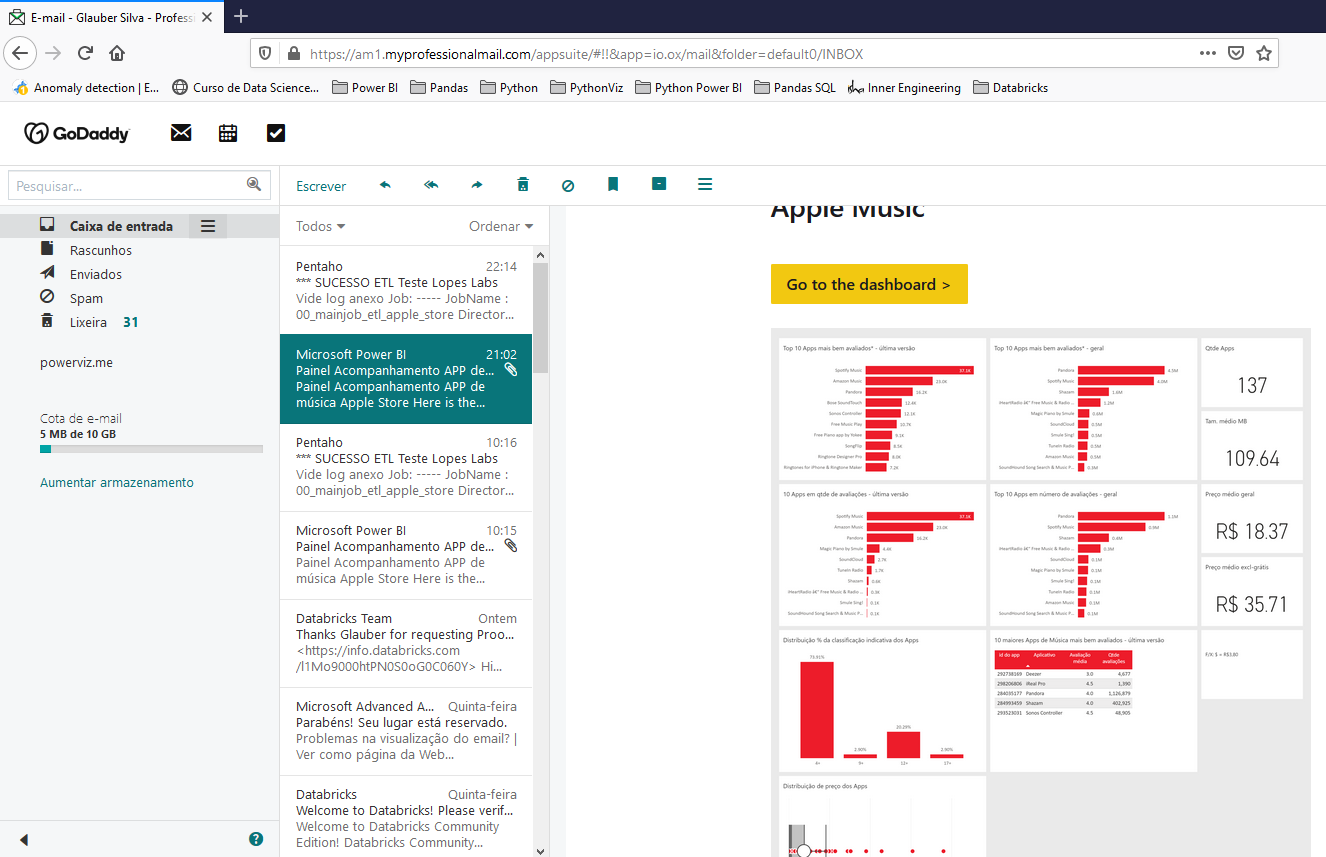
**3ª VERSÃO – UTILIZAÇÃO DO DATAFLOW DO POWER BI SERVICES PARA AUTOMATIZAÇÃO DA ATUALIZAÇÃO DOS DADOS DO DASHBOARD E ENVIO DE E-MAIL DIARIAMENTE:**

- Utilização do dataflows do Power BI para executar a carga diária automática dos novos dados e horário posterior à atualização do modelo de dados via Pentaho Data Integration.



- O banco de dados PostgreSQL *on premisses* é acessado diretamente pelo Power BI Services via *gateway* instalado na máquina.

- Foi gerado um painel contendo as informações do dashboard e agendado para envio diário automático para o e-mail cadastrado. Desta forma todo o processo o dashboard também será enviado automaticamente para a área de negócios em horário pré-determinado.



O dashboard pode ser acessado online via <http://www.bit.ly/testelopeslab>

**Avaliação de Conhecimento de SQL**

As resoluções estão no arquivo exercicios.sql, que contém o processo passo-a-passo e observações.

**Arquivos Anexados**

Os arquivos estão disponíveis no arquivo .Zip anexo:

- pasta dashboard:

teste\_lopeslab\_apple\_store\_etl\_powerquery.pbix

teste\_lopeslab\_apple\_store\_etl\_pentaho.pbix

teste\_lopeslab\_apple\_store\_etlpentaho - dataflow.pbix

- pasta etl:

|  |  |
| --- | --- |
| 00\_run\_etl\_teste\_lopes\_labs.bat  00\_mainjob\_etl\_apple\_store.kjb  01\_stg\_apple\_store.ktr | 02\_dim\_genre.ktr  03\_dim\_app.ktr  04\_fato\_apple\_store.ktr |

- pasta sql:

Exercicios.sql