

Visualização de dados com PowerBI e Linguagem R

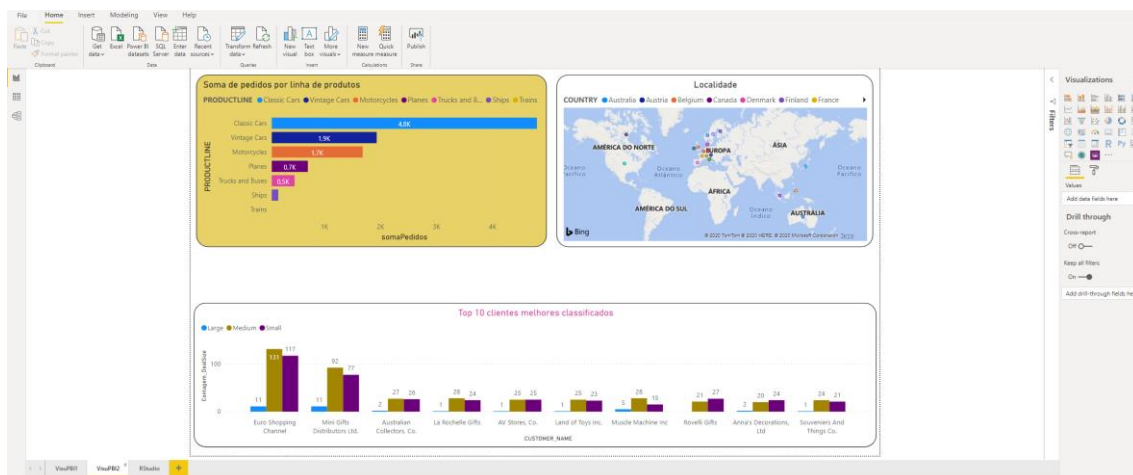
O Power BI vem com inúmeros recursos internos, e talvez dependendo do contexto é preciso buscar mais. Então, é hora de fazer o PowerBI trabalhar junto com outras ferramentas como a Linguagem R e Python.

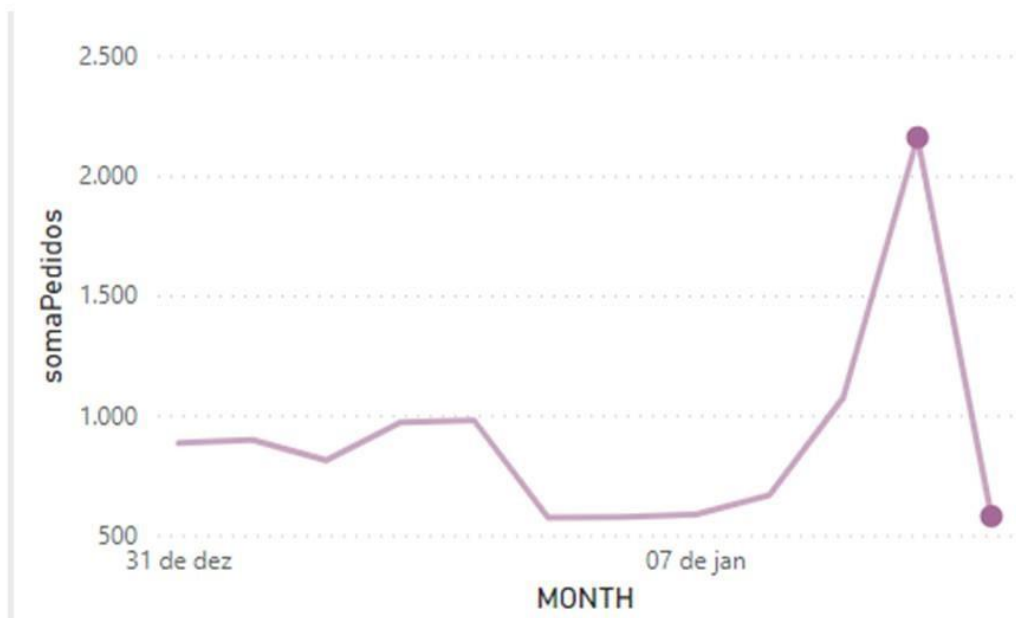
A Linguagem R é uma poderosa linguagem estatística que funciona muito bem junto com modelos de Machine Learning para produzir modelos preditivos no contexto de Data Science e também com o PowerBI para funcionalidades de BI(Business Intelligence) .

Vale lembrar que BI tem foco em analisar o que aconteceu no passado e Data Science o foco é em prever o futuro e apesar do Microsoft PowerBI ser uma ferramenta de BI, pode ajudar muito no contexto de gerar relatórios que ajudam de forma expressiva o processo de que, olhando para o passado fornece insumos para que Data Science possa prever o futuro.

Neste artigo mostrarei a conexão do PowerBI com o RStudio que é uma IDE para se trabalhar com a linguagem R.

Nas imagens abaixo é mostrado dashboard analítico construído no PowerBI de uma dataset retirado do repositório de dados Kaggle em formato csv chamado "Sales data sample".





The screenshot displays the Microsoft Power BI Desktop interface. The top ribbon shows various tabs including Home, Model, and Fields. The main workspace shows a data model with the following tables and their fields:

- Location**: CITY, COUNTRY, POSTAL_CODE, STATE, TERRITORY
- Order**: MONTH, ORDER_DATE, ORDER_NUMBER, QUANTITY_ORDERED, STATUS, YEAR, Cancel, InvalidPeriods
- Product**: PRICE_BASK, PRODUCTCODE, PRODUCTLINE, CHD_ID, Cat, Product, price
- TableInfo**: ID_CUSTOMER, ORDER_NUMBER, POSTAL_CODE, PRODUCT_CODE
- Customer**: ADDRESS, CONTACT_NAME, CUSTOMER_NAME, DEALSIZE, ID_CUSTOMER, PHONE
- sales_data_sample**: ADDRESSLINE1, ADDRESSLINE2, CITY, CONTACTFIRSTNAME, CONTACTLASTNAME, COUNTRY, CUSTOMERNAME, DEALSIZE, MONTH_ID, MGRP, ORDERDATE, ORDERNUMBER, ORDERNUMBER

The relationships between the tables are as follows:

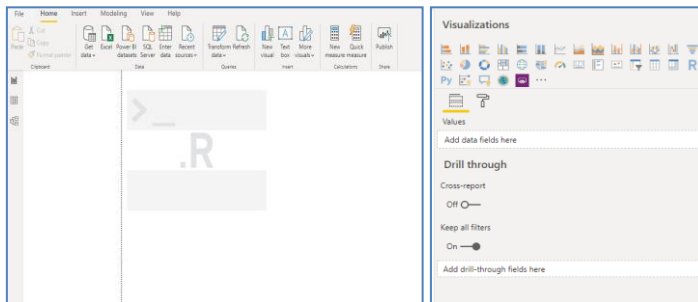
- Location** is related to **Order** (1:M relationship on CITY).
- Location** is related to **Product** (1:M relationship on POSTAL_CODE).
- Order** is related to **Product** (1:M relationship on ORDER_NUMBER).
- Order** is related to **Customer** (1:M relationship on ORDER_NUMBER).
- Order** is related to **sales_data_sample** (1:M relationship on ORDER_NUMBER).

The **Fields** pane on the right shows the available fields for each table, and the **Properties** pane on the left shows the properties of the selected field.

The screenshot shows a data science workflow in RStudio. The top toolbar includes functions like 'Get', 'Export', 'SQL', 'Enter', 'Recent', 'Transform', 'Refresh', 'New', 'Test', 'More', 'Quick', and 'Publish'. The main window displays a box plot titled 'Quantidade de pedidos por estado' (Quantity of orders by state). The y-axis is 'QUANTITY_ORDERED' and the x-axis is 'STATE'. The plot shows the distribution of orders for various states including SC, GA, CT, WA, NH, NJ, NY, OR, IL, CO, NM, and VT. A legend on the right lists the states. The bottom pane shows R script code for installing packages, loading data, and creating a faceted box plot.

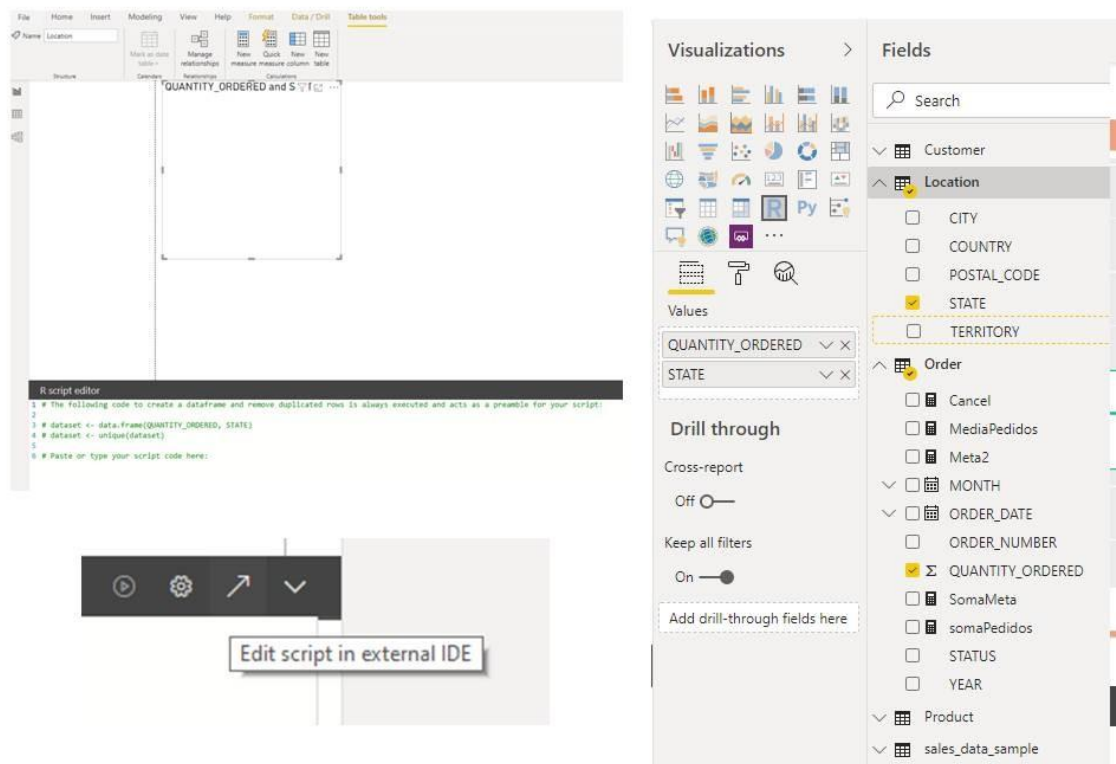
E como Fazer a conexão com o RStudio?

Para criar visualizações no R Selecione o ícone Visual do R no painel Visualização.

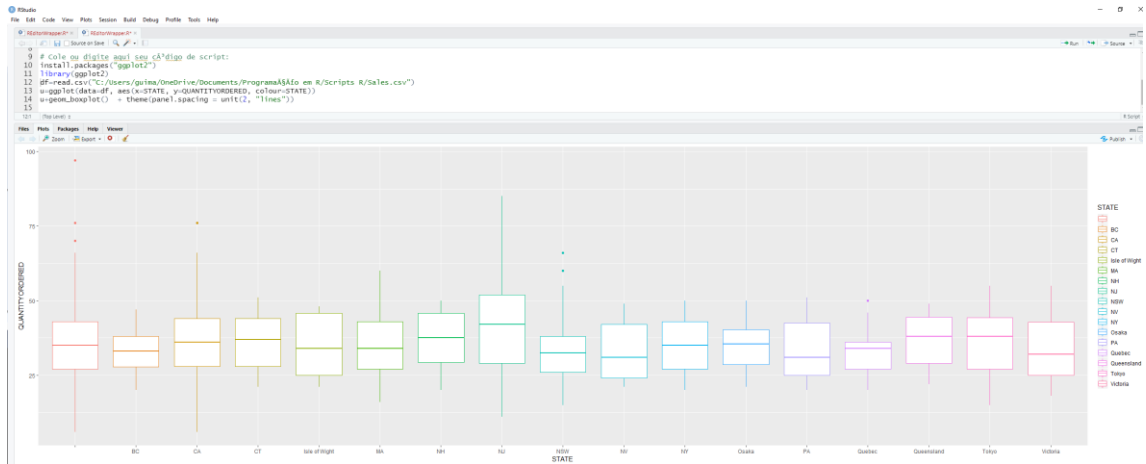


Arraste os campos do painel Campos que você deseja utilizar em seu script R, e aparece automaticamente no R script editor o espaço para se criar o código em linguagem R .

Clicar na seta "Edit script in external IDE" e é direcionado para a IDE do RStudio.



Apos carregar o dataset e instalar os pacotes necessários, basta rodar o escript e a visualização aparece na janela do RStudio, que no caso seria um gráfico de caixa (Boxplot) que mostra informações dos quartis, percentis e outliers das colunas: *"quantidade de ordens por estado"*.



Agora basta copiar o script construído no RStudio, colar no R script editor do PowerBI, executar o código e a visualização será gerada.

