

Power BI para Ciência de Dados

ESPECIALIZAÇÃO EM ANÁLISE E CIÊNCIA DE DADOS



Vinícius Rodrigues Oviedo

Santa Maria, 2024

CONTEXTUALIZANDO O POWER BI...

Microsoft Power BI

- Ferramenta de BI (*Business Intelligence**)
- Análise e Visualização de Dados (interativa) – **relatórios e dashboards**
- Deploy de um produto de dados

***Business intelligence (BI)** é um conjunto de processos tecnológicos para coletar, gerenciar e analisar dados organizacionais para gerar insights que informam estratégias e operações de negócios.

IBM. **O que é business intelligence (BI)?** Disponível em: <https://www.ibm.com/br-pt/topics/business-intelligence>. Acesso em 19 Nov. de 2024.

Microsoft Power BI



CONTEXTO DE NEGÓCIO

- Qual(is) problemas se busca(m) resolver?
- Objetivo do dashboard?
- Informações relevantes para a tomada de decisão?
- Quem vai consumir?



DADOS

- Origem?
- Quem gera?
- Metadados
- Atualização?



PUBLICAÇÃO

- Local
- Online

Microsoft Power BI

Capacidades da ferramenta:

- Tratar dados (ETL)
- Combinar/conectar com várias fontes de dados
- Visualizar dados de maneira interativa e de fácil compartilhamento
- Extrair *insights* e tomar decisões *data-driven*

Ferramentas de BI



Power BI (Microsoft)



Tableau



Looker Studio (Google)



Qlik sense



Streamlit, Dash & plotly, Panel, Taipy

CONHECENDO A INTERFACE...

Primeiro dashboard...

Dashboard Logística

Ano, Mês

2022



R\$ 5.114.737

Receita



R\$ 4.327.324

Lucro



23306

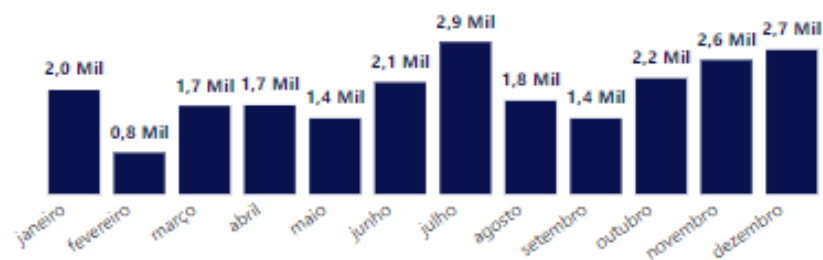
Viagens



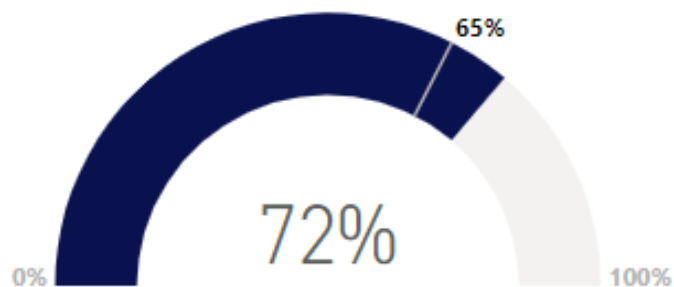
174.139

KM Rodados

Quantidade de Viagens por Mês



Quantidade de Viagens por Mês

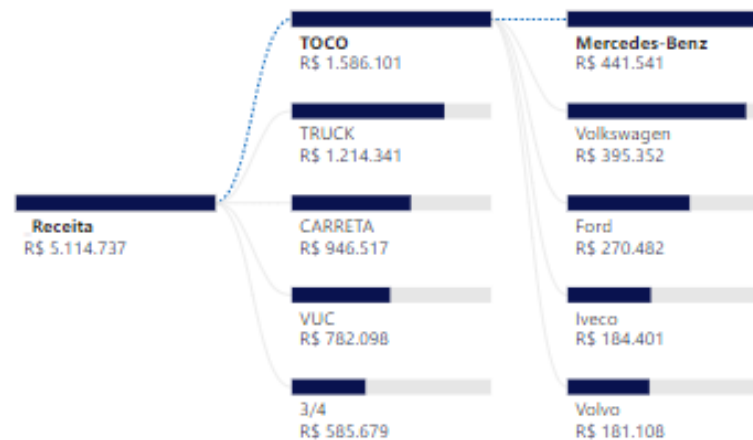


Receita vs Veículo e Marca

Veículo

TOCO

Marca



POWER QUERY...

Power Query

- Funções de transformação (**linguagem m**)
- Novas colunas
- *Extract, Transform and Load* (**ETL**)
- Exploração dos dados (estatística descritiva, distribuição, valores únicos, qualidade do dado)

Power Query

1. Funções de texto
2. Funções de número
3. Funções de data
4. Acrescentar consultas (*CONCAT*)
5. Mesclar consultas (*MERGE*)

Vejam os um exemplo...

Power Query

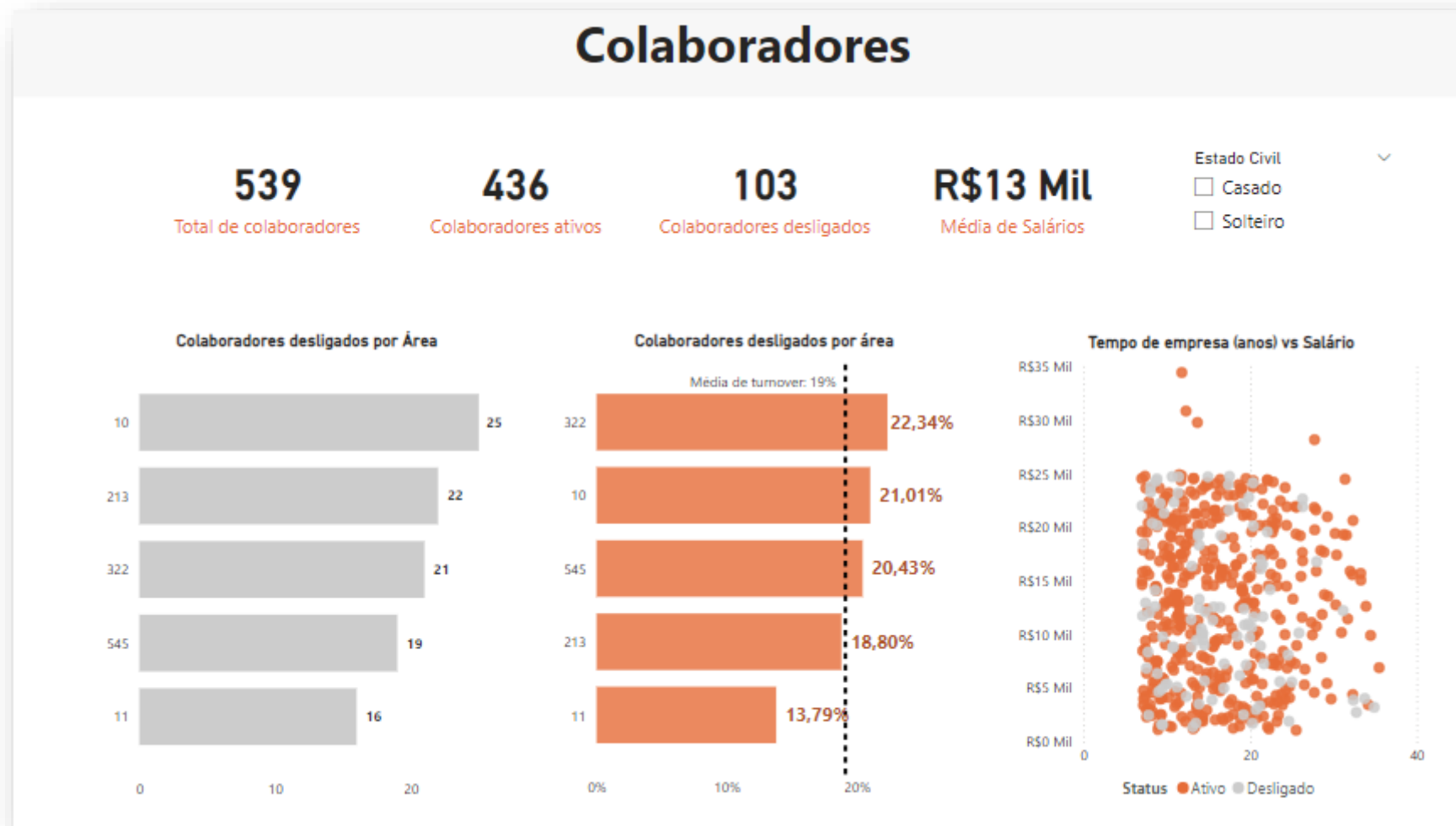
Dataset BaseColaboradores:

1. Total de colaboradores na base?
2. Colaboradores ativos?
3. Colaboradores desligados?
4. Taxa de **turnover***?
5. Média de salários?
6. Colaboradores desligados por área?
7. Turnover por área?
8. Existe relação entre tempo de empresa e salário?
9. Adicionar filtro de estado civil.

***Proporção de colaboradores que deixam a empresa (em relação ao total).**



Power Query



MODELAGEM DE DADOS

Modelagem de dados

Dois principais tipos de modelagem relacional:

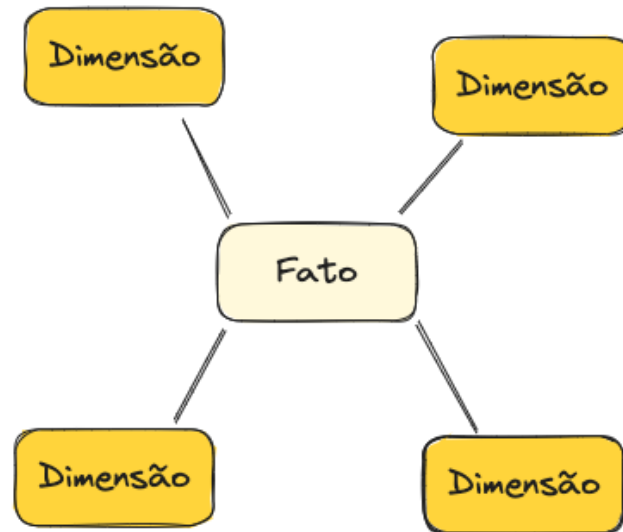
1. Esquema estrela (*star schema*)
2. Esquema floco de neve (*snowflake schema*)

Conceitos:

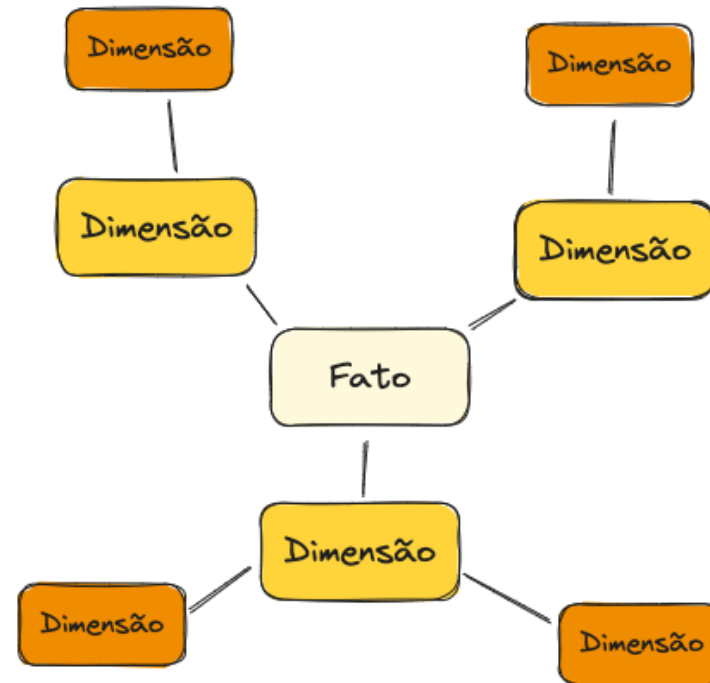
- Tabela fato
- Tabela dimensão

Modelagem de dados

Star Schema



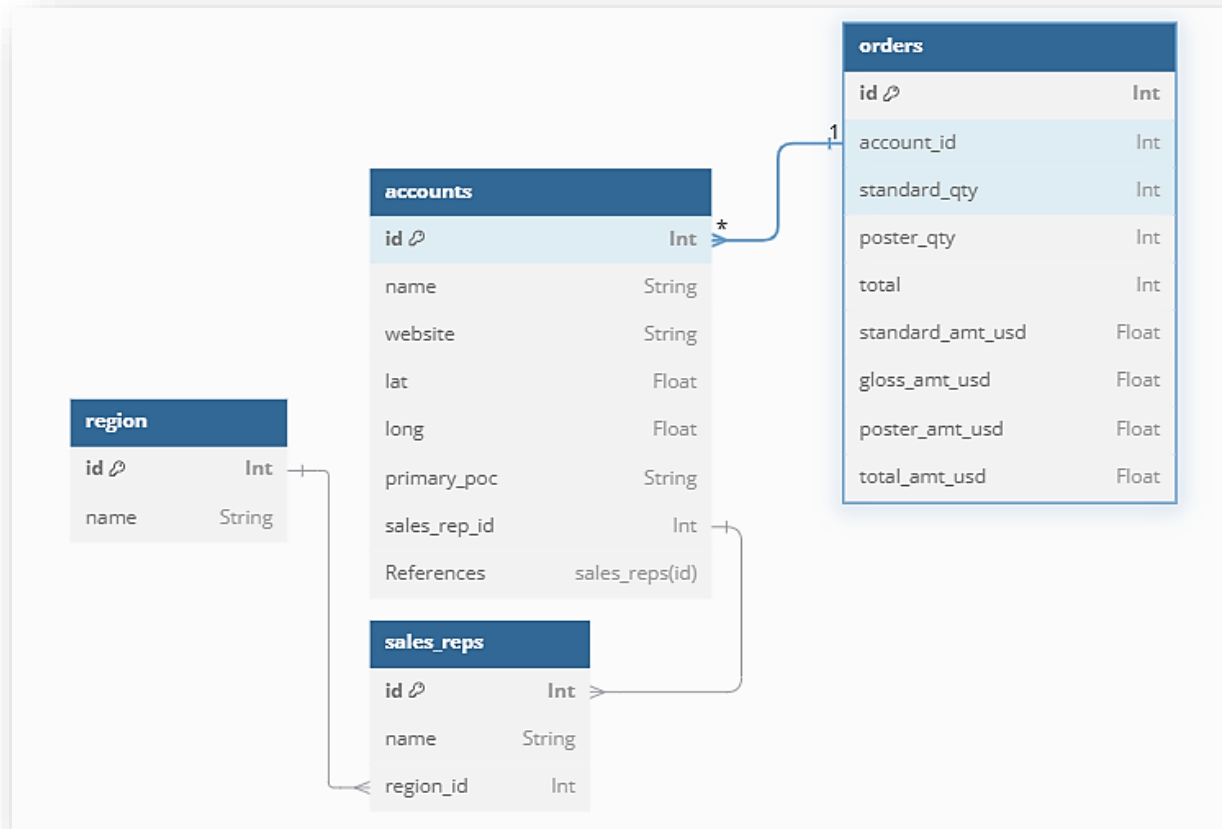
Snowflake



Fonte: [Linkedin](#).

Vejam os um exemplo...

Caso prático no PBI



Empresa Parch & Posey (Diagrama Entidade-Relacionamento).

Múltiplas tabelas



ordens



contas



regiões



representantes

Visuais

Dados:

Referentes à Parch and Posey, uma empresa fictícia de venda de papel. Essa empresa tem 50 representantes de vendas em quatro regiões dos Estados Unidos. Constam 5 tabelas associadas:

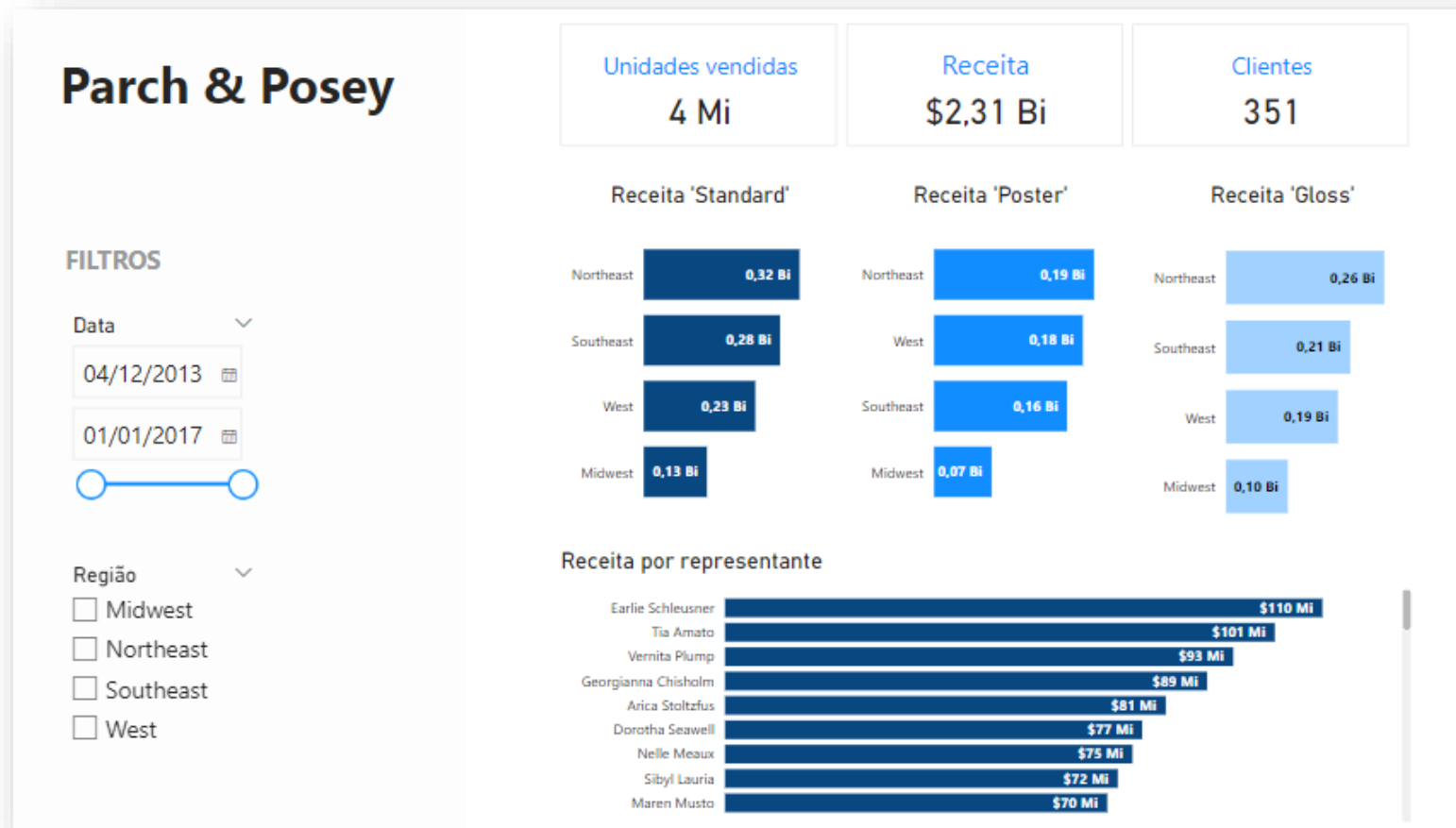
- **orders:** todos os pedidos feitos de 2014-12-03 a 2017-01-01
- **accounts:** todas as contas de Parch & Posey
- **regions:** as 4 regiões nos Estados Unidos onde a empresa opera
- **sales_reps:** informações sobre os representantes de vendas
- **web_events:** todos os dados de eventos da web para as contas de Parch & Posey

Visuais

Questões:

1. Qual foi a quantidade total vendida no período?
2. Qual foi a receita em Novembro de 2016?
3. Qual tipo de papel mais gerou receita na região `West` (oeste)?
4. Qual representante de vendas vendeu região `Midwest` (centro-oeste)?

Caso prático no PBI



Dashboard de referência.

VISUAIS...

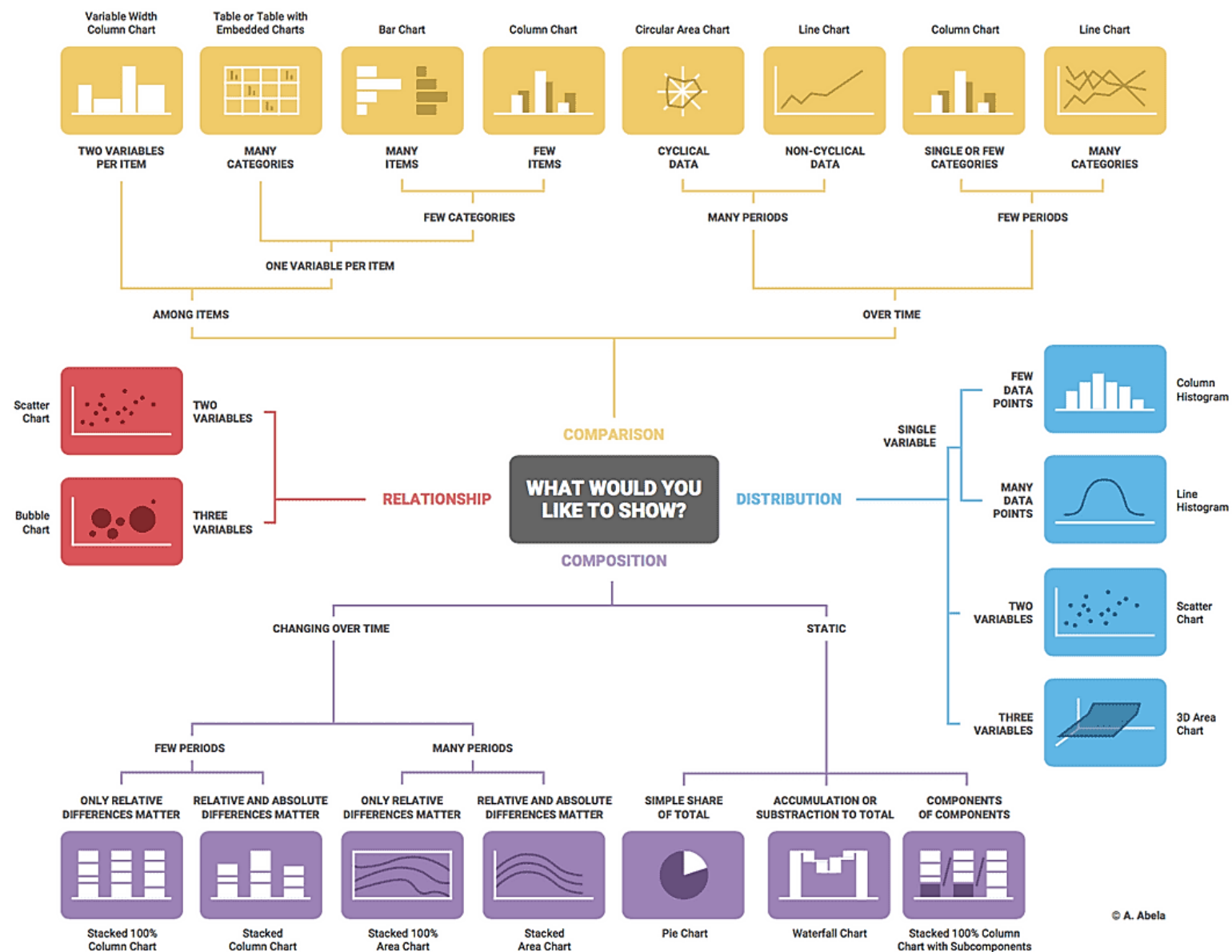
Visuais

O PBI possui:

1. Visuais *built-in*
2. De terceiros ([comunidade](#))

Qual gráfico escolher? ([from data to viz](#))

1. Aquele que melhor transmite a informação
2. Fácil comunicação e entendimento (por parte da audiência)
3. Contexto
4. Baixa carga cognitiva



© A. Abela

Fonte: (MORALES, 2020).

Vejamos alguns exemplos...

Caso no PBI: Vendas em uma *Coffee Store*

Objetivo:

Extrair insights para uma Coffee Store, levando em conta dados históricos de vendas.

Contexto:

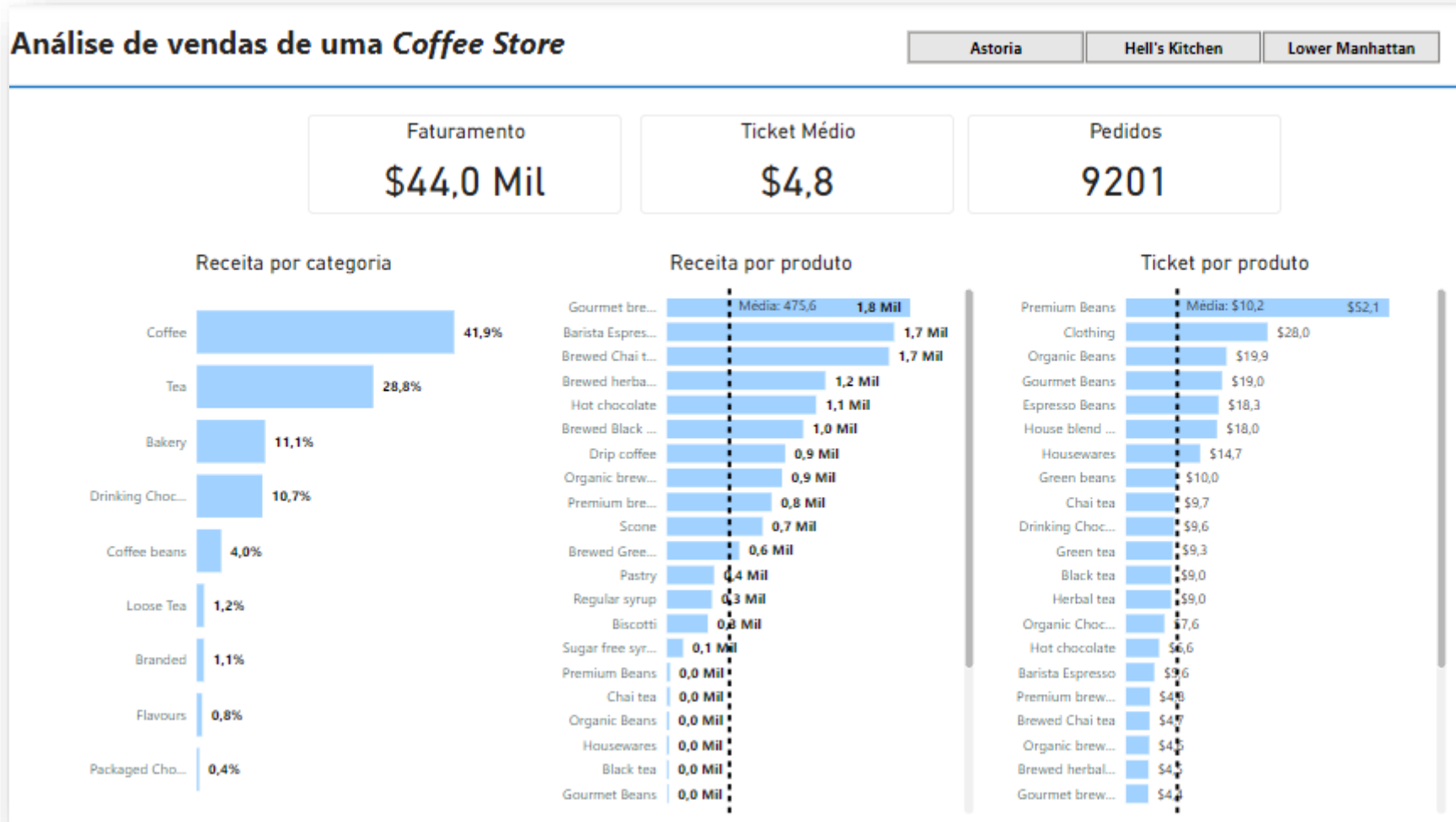
- Período: 7 dias
- Filiais: 3 (Hell's Kitchen, Lower Manhattan, Astoria)
- Categorias de produto: 9
- Tipos de produtos: 26
- Produtos distintos: 71

Perguntas de negócio:

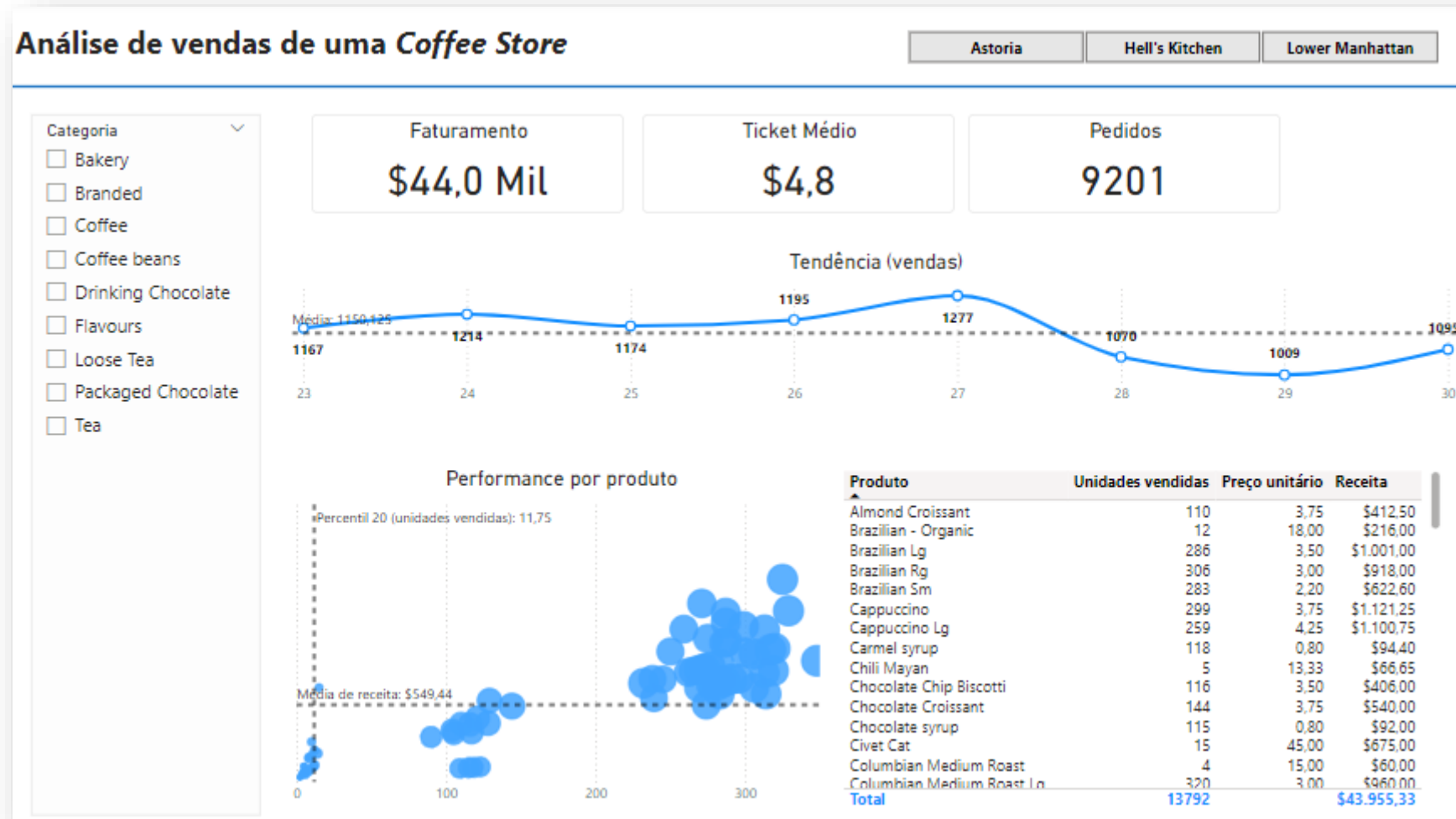
1. Qual foi o faturamento nos últimos 7 dias?
2. Qual o ticket médio?
3. Quais categorias trazem maior parte dos resultados?
4. Ranking de produtos por receita.
5. Quais são os produtos de maior ticket? eles têm vendido bem? devem ser descontinuados?
6. Qual a tendência de vendas nos últimos 7 dias?
7. Quais produtos tem performado melhor? existem segmentos distintos?



Caso no PBI: Vendas em uma *Coffee Store*



Caso no PBI: Vendas em uma *Coffee Store*



DICA DE FERRAMENTA

Dica de ferramenta

Existem dois tipos:

1. Padrão (mostra valores)
2. Outro visual

Vejam os um exemplo...

Dica de ferramenta

Daset marketing_campaigns:

Carregar os dados no PBI, e obter as seguintes colunas:

- Total gasto
- Compras
- Dias como cliente
- Campanhas aceitas

i) Gerar um gráfico com dica de ferramenta com outro visual (próximo slide)

ii) Gerar um gráfico com dica de ferramenta padrão (próximo slide)

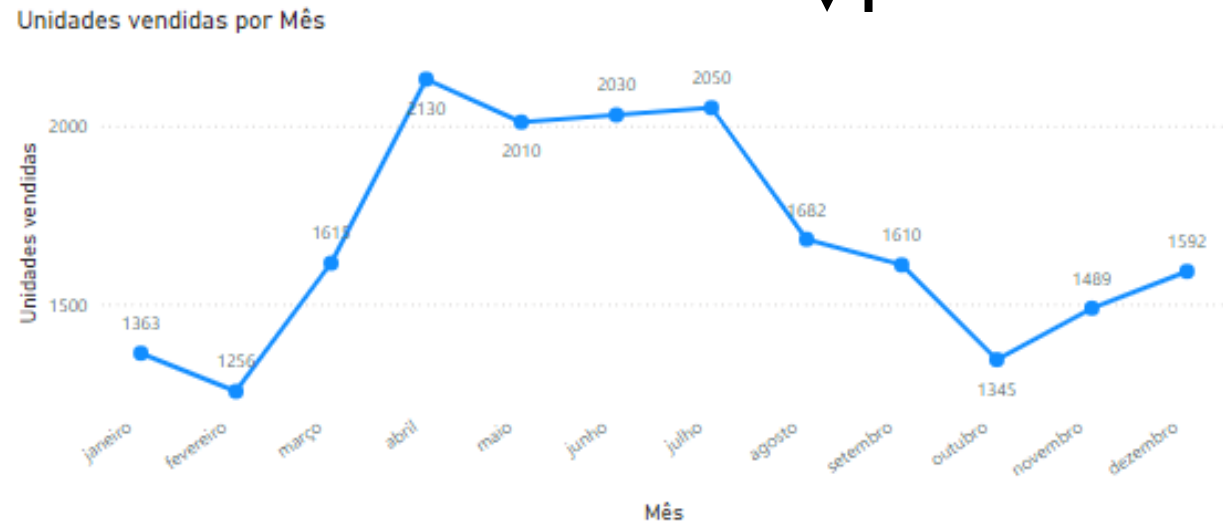
Dica de ferramenta



DRILL UP & DRILL DOWN

Drill up / Drill down - dataset Bikes

↓ ↑ Hierarquias



DATA ANALYSIS EXPRESSIONS (DAX)

DAX

Pode ser utilizado para gerar:

1. Novas colunas
2. Medidas
3. Novas tabelas (análogo à tabela dinâmica do Excel)
4. Medidas rápidas (lista de funções DAX *built-in*)

Em termos de performance:

Medida > Coluna

DAX

Funções:

1. Básicas
2. Iterativas
3. Cumulativas,
4. De texto,
5. Lógicas ([Switch](#))...

SWITCH(*tabela[coluna], valor1, resultado1, valor2, resultado2, ...*)

DAX

Tabela Calendário:

1. Criada manualmente ou via DAX
2. Melhora a interação com filtros
3. Acrescenta granularidade nos dados
4. Permite os *drills*
5. **Boa prática**



Observação: precisa estar relacionada com as tabelas do modelo de dados.

DAX

Dataset segmented_customers:

1. Qual total de clientes? e por cluster? (criar medida, cartão e usar filtro de cluster)
2. Classifique os clusters como 'top' ou 'ordinary'. (aqueles que gastam acima da média do cluster são `top`, caso contrário, são `ordinary`)

EXEMPLO NO PRÓXIMO SLIDE...

DAX

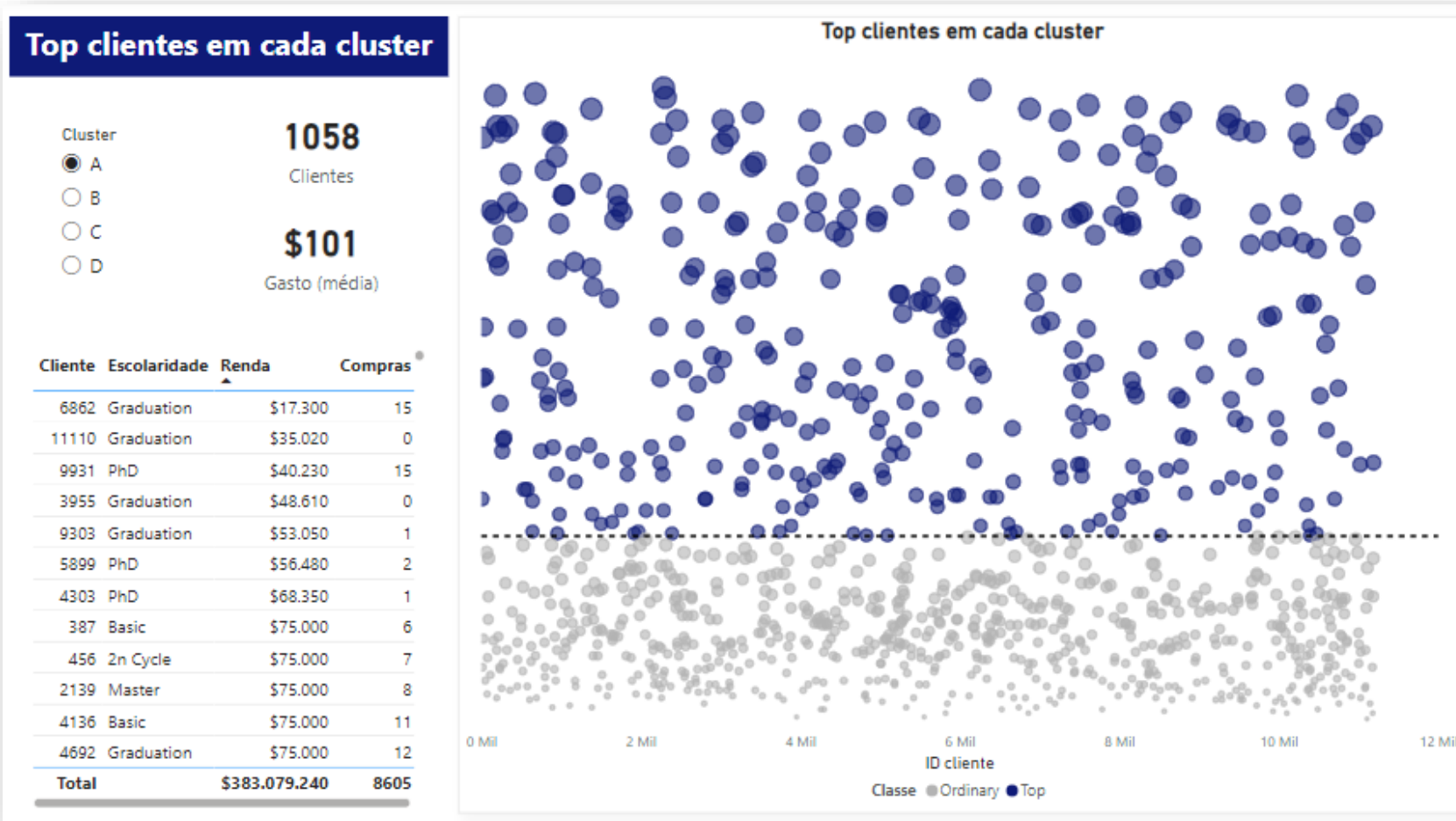


Tabela Calendário no Power BI

Pode ser criada manualmente ou via DAX.

Vantagens:

- Permite trabalhar melhor com filtros;
- Maior granularidade nos dados;
- Boa prática (modelagem de dados).

Criando com DAX:

1. Vá para a guia "Modelagem".
2. Selecione "Nova Tabela".
3. Criar Tabela com a DAX



```
CalendarTable =  
VAR StartDate = MIN(SuaTabela[Data]) // Data mínima da coluna  
VAR EndDate = MAX(SuaTabela[Data]) // Data máxima da coluna  
RETURN  
    ADDCOLUMNS(  
        CALENDAR(StartDate, EndDate),  
        "Year", YEAR([Date]),  
        "Month", FORMAT([Date], "MMMM"),  
        "Month Number", MONTH([Date]),  
        "Quarter", "Q" & QUARTER([Date]),  
        "Day", DAY([Date]),  
        "Day of Week", FORMAT([Date], "dddd"),  
        "Weekday Number", WEEKDAY([Date], 2)  
    )
```

Medidas rápidas...

INTERAÇÃO ENTRE VISUAIS & BOTÕES

Interações & Botões

Interação entre visuais:

- Editável caso se deseje limitar o cruzamento de filtros a alguns visuais.

Botões:

1. Navegação entre páginas
2. Resetar filtros
3. Uso como filtro

LAYOUT & COMPARTILHAMENTO