



GRIMALDO OLIVEIRA APRESENTA

# Cursos área de Dados

Visão Geral e Ferramentas

# O que você conhecerá

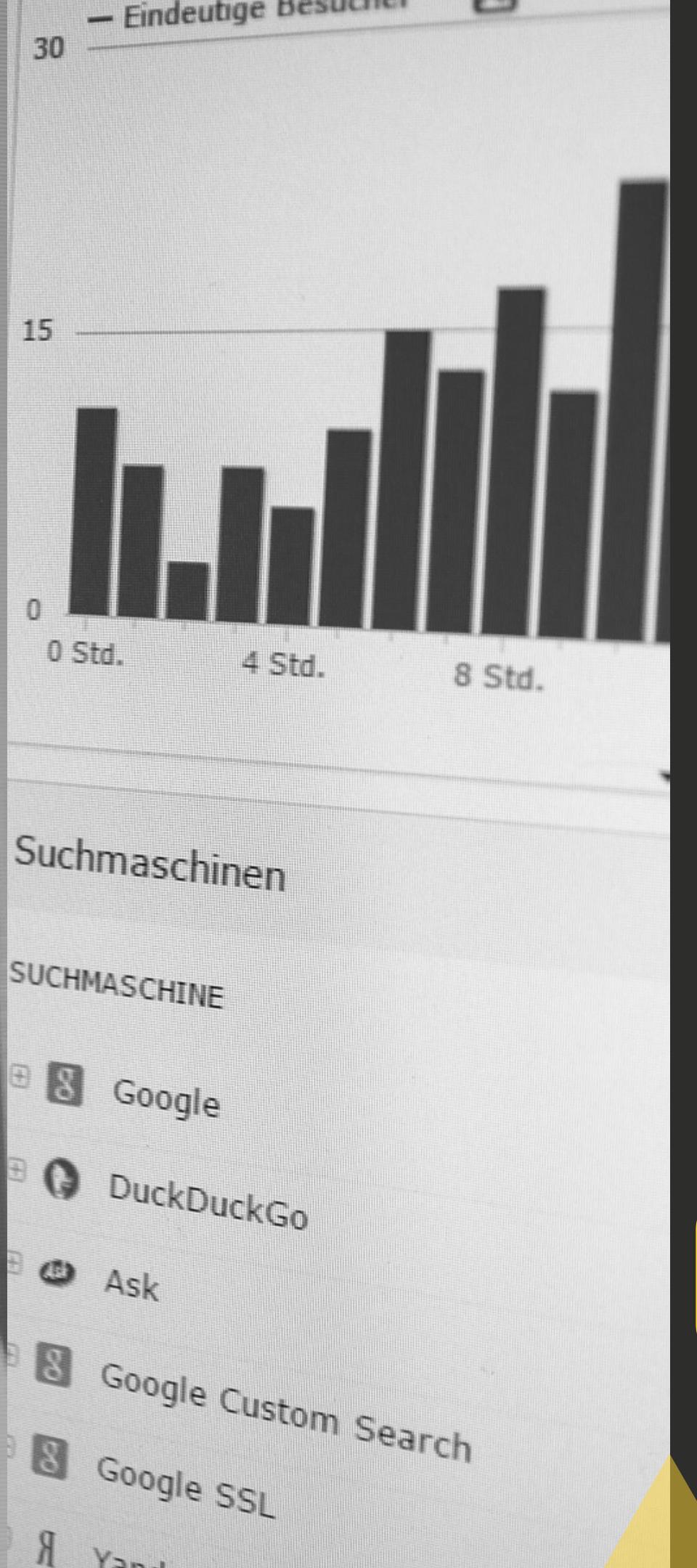
## PRINCIPAIS TÓPICOS



Sistematica sobre as diferentes ferramentas  
Informações sobre os cursos  
Profissionais que se identificam com a área de dados  
Sobre mim

# QUEM SOU EU

Professor das pós-graduações das universidades UNIFACS, CATÓLICA DO SALVADOR e ISL Wyden. Mestre pela Universidade do Estado da Bahia (UNEBA) no Curso de Mestrado Profissional Gestão e Tecnologias Aplicadas à Educação (GESTEC). Possuo Especialização em Análise de Sistemas pela Faculdade Visconde de Cairu e Bacharelado em Estatística pela Universidade Federal da Bahia.



FOLLOW  
FOLLOW  
FOLLOW



**50**  
TO +++++ FOLLOW

DADOS & BI



VALE  
A PENA  
SEGUIR

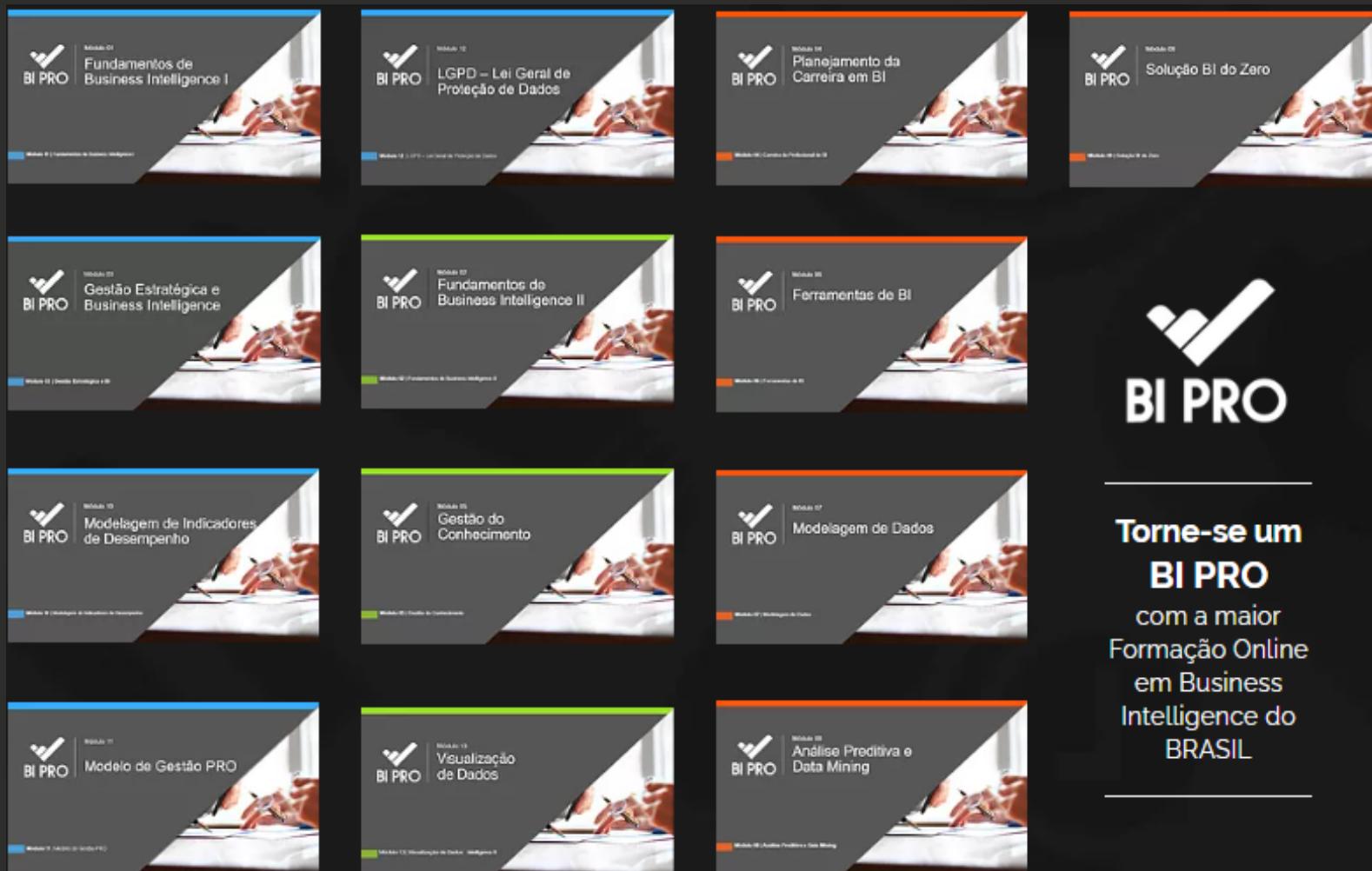
**GRIMALDO  
OLIVEIRA**



## 50 TO FOLLOW EM 2023

Escolhido como um dos 50 principais profissionais em dados do Brasil pela Gama Academy, instituição que indica os profissionais de 5 categorias do mercado de tecnologia destacando a representatividade, influência e trabalhos realizados no Brasil no ano de 2023.

[CLIQUE E CONFIRA LISTA COMPLETA DOS 50 PROFISSIONAIS EM DADOS E BI](#)



# BI PRO

## CURSO DE FORMAÇÃO

Um programa online de formação continuada que desenvolve o profissional com as principais habilidades da área de dados, contendo:

- Mentoria personalizada
- 13 módulos de formação + quizzes + atividades
- Ebook BI COMO DEVE SER
- Mais de 180 cursos técnicos da Udemy inclusos

[CLIQUE E CONFIRA TUDO SOBRE O BI PRO](#)

# BI COMO DEVE SER

O GUIA DEFINITIVO



Diego Elias Oliveira  
Grimaldo Lopes de Oliveira

2ª edição

## EBOOK BI COMO DEVE SER

Um dos livros mais lidos pela comunidade de dados do Brasil, contém todos os passos para a construção de um projeto de BI do ZERO.

Acompanha o livro uma série de vídeoaulas gratuitas sobre modelagem de dados + material sobre Business Intelligence + Material sobre Machine Learning.

[CLIQUE E CONFIRA TUDO SOBRE O LIVRO BI COMO DEVE SER](#)



## CANAL WHATSAPP DIRETO

Nosso Canal no WhatsApp é uma fonte exclusiva de informações e conteúdo. Aqui, você será o primeiro a receber todas as novidades sobre cursos, eventos, workshops e materiais exclusivos. Isso significa que você estará na vanguarda do conhecimento, com acesso a informações privilegiadas antes de qualquer outra pessoa.

[CLIQUE E ENTRE NO CANAL](#)



SOMOS

+ 80.000

ALUNOS

Total de inscrições 

81.254

# LIVROS QUE INDICO LEITURA



- 1- [Storytelling com dados: um guia sobre visualização de dados para profissionais de negócios](#)
- 2- [Estatística prática para cientistas de dados: 50 conceitos essenciais](#)
- 3- [Introdução à Linguagem SQL: Abordagem Prática Para Iniciantes](#)
- 4- [Use a cabeça! SQL](#)
- 5- [SQL Guide](#)
- 6- [Building the Data Warehouse](#)
- 7- [The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data](#)
- 8- [Estatística: O que é, para que serve, como funciona](#)
- 9- [Python Para Análise de Dados: Tratamento de Dados com Pandas, NumPy e IPython](#)



# Antes de começar

## O QUE VOCÊ ENTENDE SOBRE DADOS?

O mundo vive um grande momento nesta área, muitas são as empresas que procuram profissionais para trabalhar com dados, para se manterem em um mercado cada vez mais competitivo. Caso você esteja começando agora, abaixo links de conteúdos sobre como iniciar na área.

<SE TORNE UM PROFISSIONAL DA ÁREA DE DADOS>

<APRENDA A TRABALHAR COM DADOS>

<MINHA ARQUITETURA BI>



PENSE NISSO

O MUNDO ESTÁ NAS MÃOS  
DAQUELES QUE TEM A  
CORAGEM DE SONHAR E  
CORRER O RISCO DE VIVER  
SEUS SONHOS

- PAULO COELHO

# Quais são suas perguntas

## PARA QUE VOCÊ ESTÁ APRENDENDO SOBRE DADOS?

Hoje todas as empresas do mundo, por menor que seja seu porte, produzem e guardam dados. Existem inúmeras formas de entender como o negócio está seguindo, analisando exatamente estes dados.

## QUAL É O PROBLEMA?

Há poucos profissionais realmente qualificados e que sejam capazes de produzir um ecossistema confiável e duradouro na área de dados.

## DE QUE FORMA IREI ADQUIRIR ESTE CONHECIMENTO?

Os cursos permitirão moldar o seu perfil, fazendo com que você identifique-se com que tipo de atividade deseja trabalhar na área de dados.

A black and white photograph showing several white, rounded rectangular blocks arranged horizontally. Each block has a large, bold, black capital letter on it. From left to right, the letters spell out "STUDY". The background is dark and slightly blurred, making the white blocks stand out.

# O QUE VOCÊ APRENDERÁ NOS CURSOS?

## MAPEANDO A JORNADA DO CLIENTE

Aprenderá a levantar os dados que interessam ao negócio, entendendo com documentos e explicações como você registrará tudo que é importante

## MAPAS DE AÇÕES

Verá dentre muitas ferramentas, como você poderá promover o uso de dados na organização. Entenderá que as ações que os dados podem promover facilitarão o crescimento do negócio da sua empresa

## ANÁLISE DA CADEIA DE VALOR

Você conseguirá gerar senso crítico sobre como podem ser aplicados os dados em cada tipo de negócio, aprenderá a ter uma estratégia organizada e com critérios lógicos

## CRIAÇÃO DE STORYBOARD

Aprendendo sobre dados, cada vez mais você poderá construir as histórias sobre eles e implementar esta cultura de forma organizacional



# Vamos falar sobre os treinamentos

A partir de agora você saberá o que compõem os cursos que ministro. Para você ter uma ideia, já estamos com mais de 80 mil alunos em mais de 90 países.

Os cursos foram criados para permitir que do iniciante ao mais avançado, o profissional consiga realizar um treinamento de acordo com seu perfil.

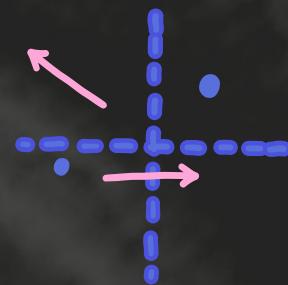
Há mais de 180 cursos sobre dados que ministro, são treinamentos onde o aluno assiste na internet por meio de videoaulas. Respondo aos alunos no máximo em 24 horas. Você irá encontrar os cursos de forma isolada, em combos ou em academias, que reúnem muitos cursos de acordo com seu propósito.

# IDENTIFICANDO O PERFIL DO TREINAMENTO



## BI

cursos destinados aos profissionais que querem aprender sobre Business Intelligence



## ESTATÍSTICA

Curso direcionado aos profissionais que desejam aprender estatística



## DATA SCIENCE

Para quem deseja adentrar na área de estatística, computação e construção de modelos estatísticos



## VISUALIZAÇÃO DE DADOS

Profissionais que desejam aprender sobre construção de gráficos, relatórios e painéis



# SUPER ACADEMIA DADOS

## KPI, BUSINESS INTELLIGENCE, DW

A Super Academia Dados é um projeto que criamos para o que há de mais relevante na área para que você entenda o que é necessário para adentrar na área de dados, montamos um grande roadmap de materiais teóricos e práticos.

Traremos uma visão de entendimento sobre como levantar o negócio do seu cliente de forma eficiente, para que você construa uma solução de dados completa.

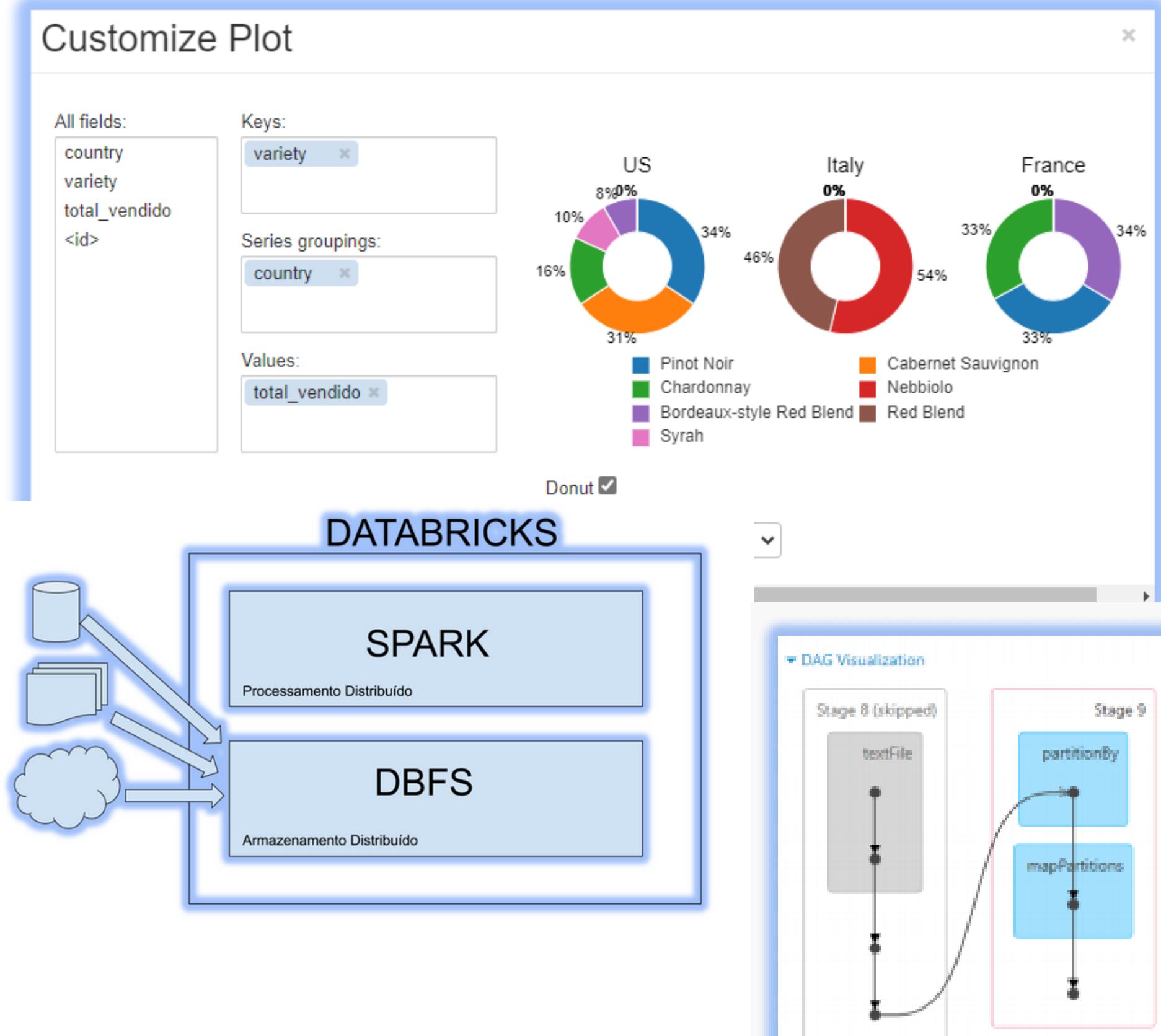




# SUPER ACADEMIA ENGENHARIA DE DADOS

**DATABRICKS, PYSPARK,  
ELASTICSEARCH, NIFI, APACHE  
HOP**

A Academia Engenharia de Dados é um super projeto que reúne as ferramentas mais utilizadas para quem deseja ser um engenheiro de dados. Aqui será possível aprender detalhes sobre temas que estão sendo utilizados nas grandes empresas de todo o mundo. Você aprenderá e discutirá sobre CLUSTER, SPARK, PYTHON, R, SCALA, SQL, NÓS, CONFIGURAÇÃO E INSTALAÇÃO DE SPARK, ELASTICSEARCH, KIBANA, LOGSTASH e muito mais.



In [20]: # Exibindo os dados sobre os profissionais  
profissional.collect()

Out[20]: ["Carlos, oncologista, hospital",  
"Ana, dentista, clinica",  
"Fernanda, enfermeira, hospital",  
"Sandra, pediatra, clinica",  
"Fatima, dentista, clinica",  
"Gilmar, cardiologista, hospital",  
"Fabio, pediatra, clinica",  
"Hilton, enfermeiro, clinica",  
"Thiago, fisioterapeuta, hospital"]

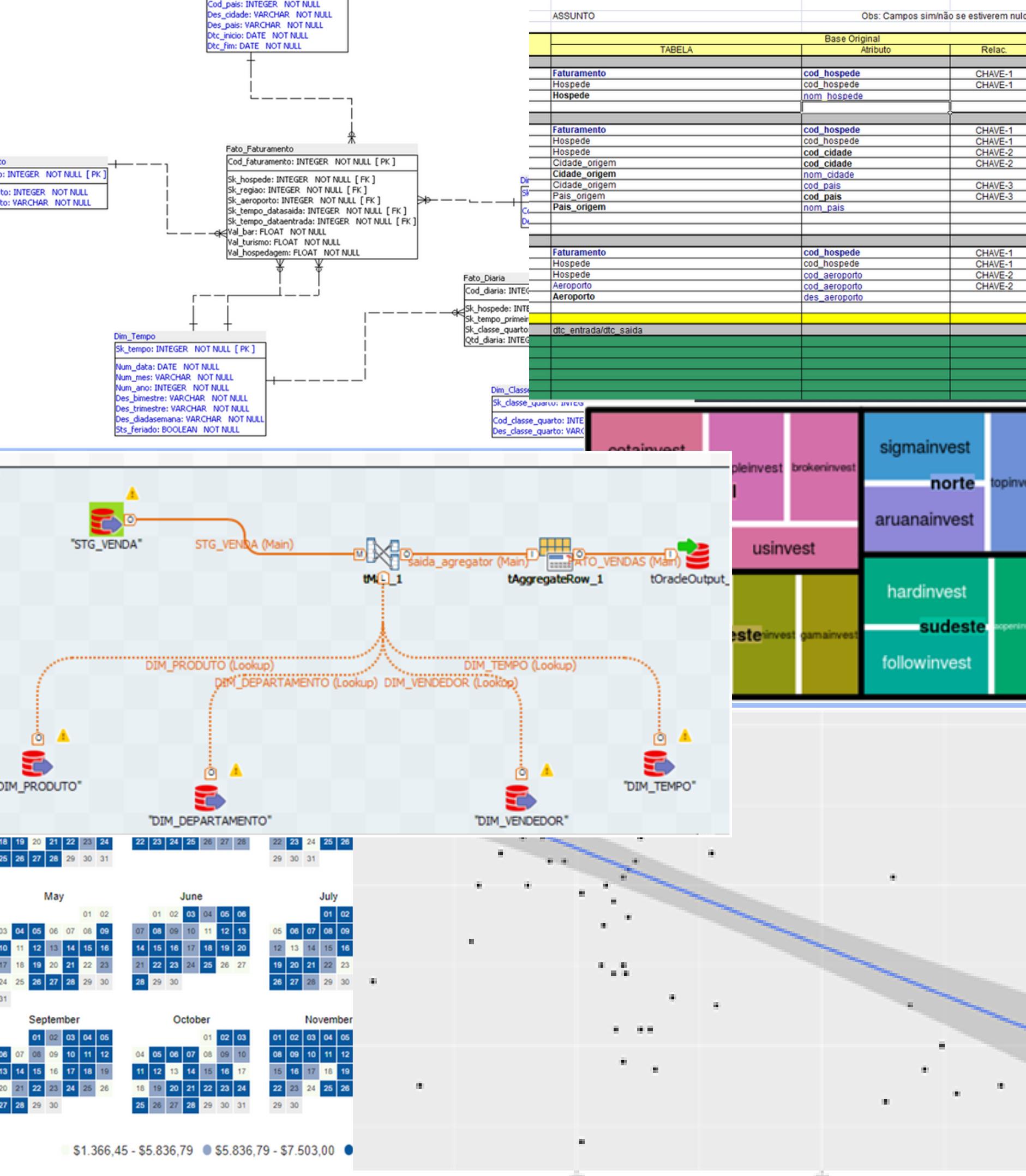


# SUPER ACADEMIA BI

## 16 CURSOS JUNTOS

Desde o entendimento de como coletar dados, realizando o tratamento e higienização das bases de dados, até a preparação de um conjunto de ferramentas para carga dos dados, visualização de dados e uso de Machine Learning fazem parte desde treinamento. Super completo para quem deseja iniciar na área de BI.

**+ 9 MIL ALUNOS!**



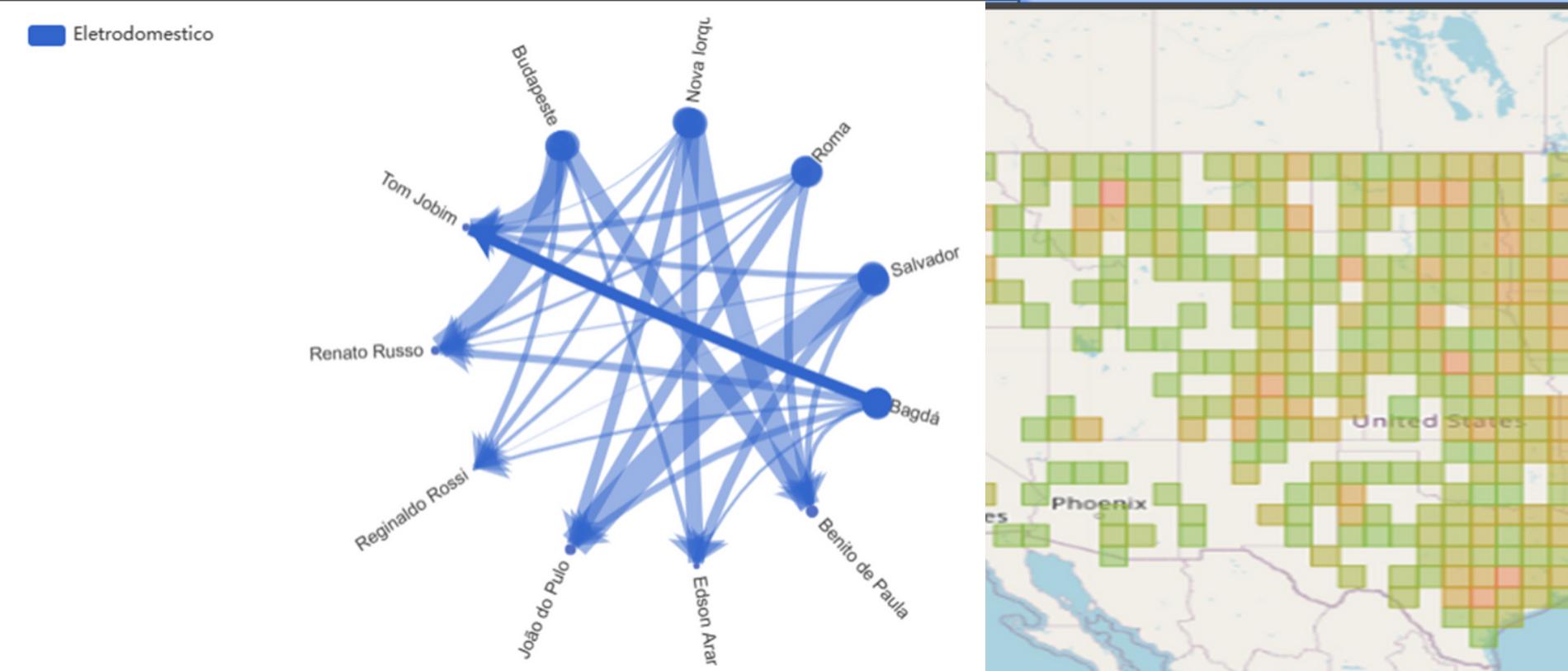
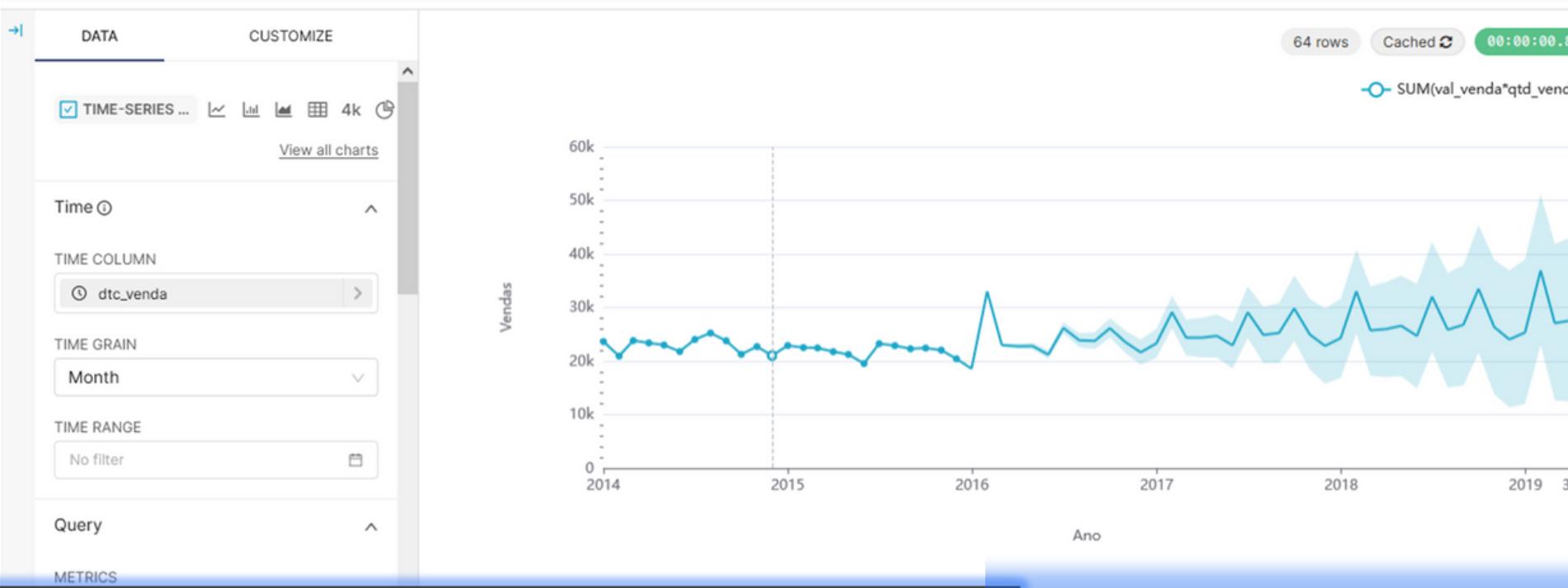
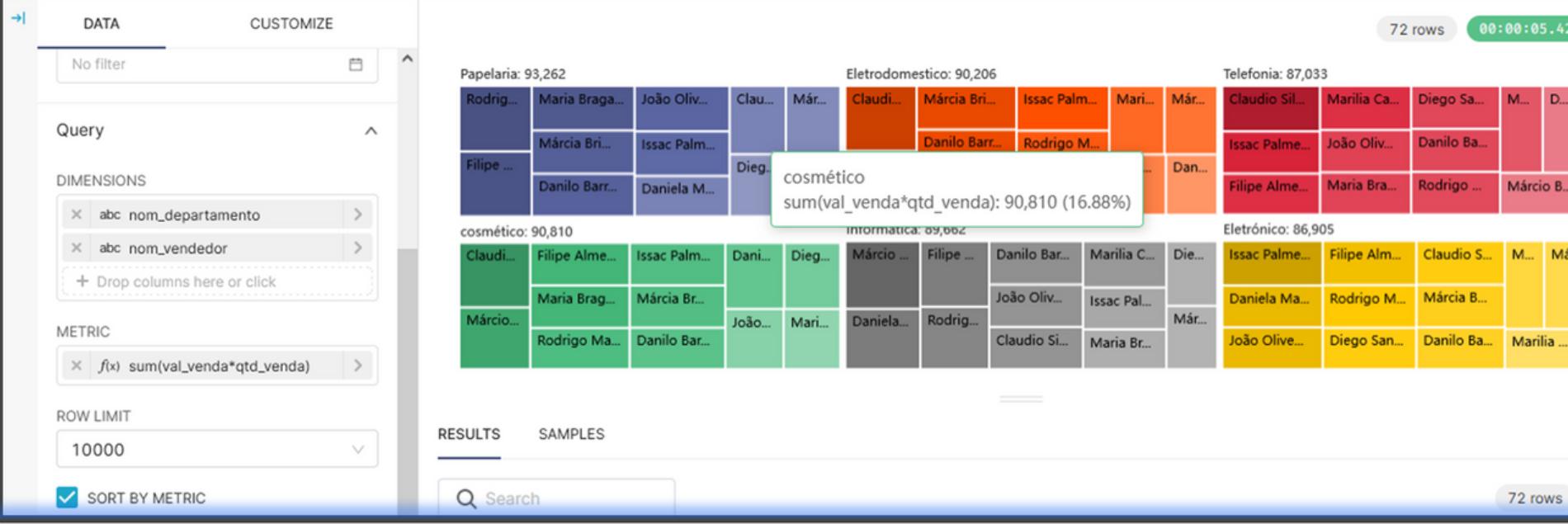


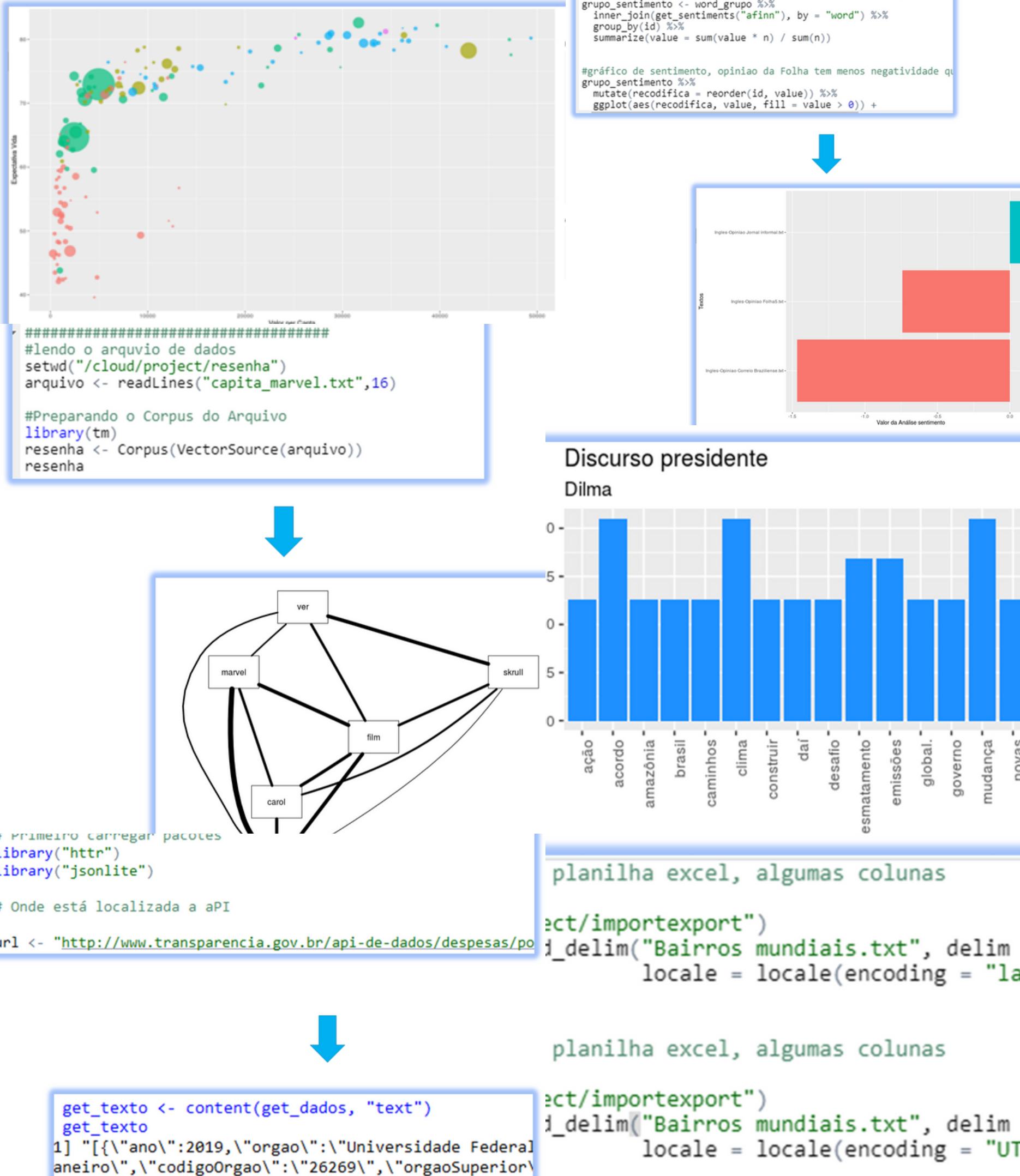
# SUPER ACADEMIA

## Visualização de Dados

### 07 CURSOS JUNTOS

Vamos aprender como construir uma storytelling poderosa, que irá prender a atenção de seu público e transmitir suas ideias de maneira envolvente e memorável. E o melhor de tudo é que vamos explorar várias ferramentas importantes do mercado, como o Power BI, Tableau, Qlik Sense, ODV, Apache Superset, e Metabase. Ou seja, você vai estar preparado para qualquer situação no mundo dos dados!



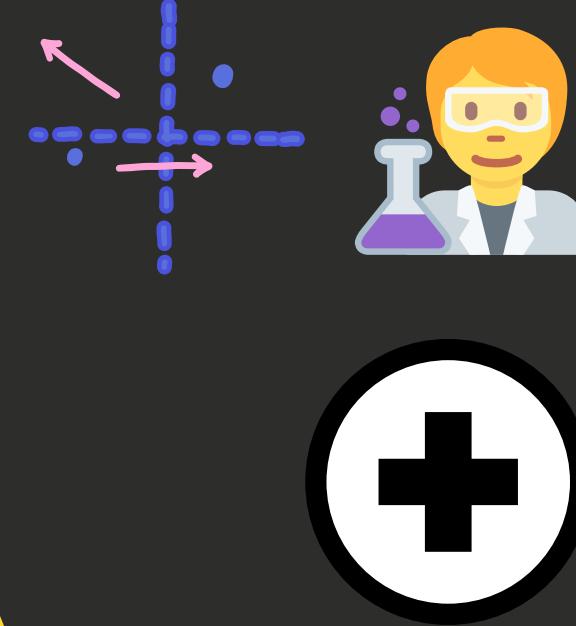


# SUPER ACADEMIA ESTATÍSTICA

## 09 CURSOS JUNTOS

Possui uma das bases mais sólidas em estatística já ministradas. Introdução ao curso de estatística básica, conceitos básicos de estatística, gráficos estatísticos, medidas de tendência central, medidas de dispersão, probabilidade e inferência estatística, teste de hipóteses e além disso uma seção completa de aplicações em Data Science com linguagem R, Orange Data Science.

**+ 6 MIL ALUNOS!**





## Exemplo 2

PLOT 1

```
set.seed(123)
n <- 1000
x1 <- matrix(rnorm(n), ncol = 2)
x2 <- matrix(rnorm(n, mean = 3, sd =
x <- rbind(x1, x2)
smoothScatter(x, xlab = "x1", ylab = "x2")
```

Apresentação Gráfica

Gráfico Plot

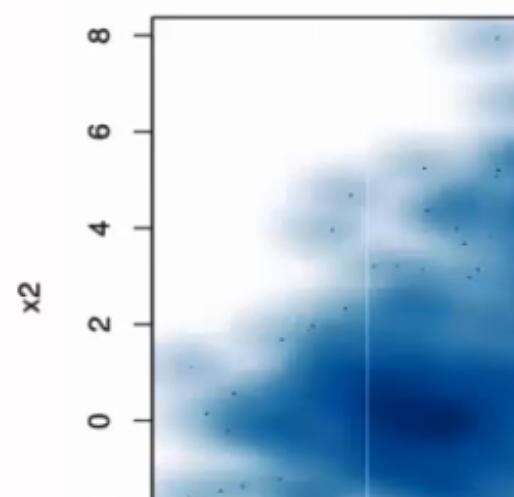
Gráfico Plot e COS

Gráfico Imagem

GRÁFICO EM DIVERSOS FORMATOS:

- Gráfico Plot
- Gráfico Plot e COS
- Gráfico Imagem

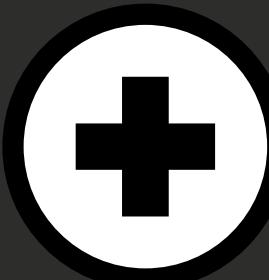
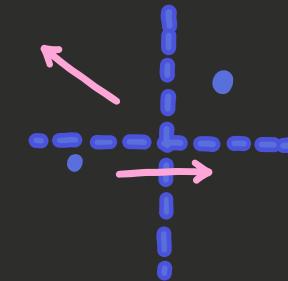
Treinamento R Markdown -  
Prof. Grimaldo Oliveira

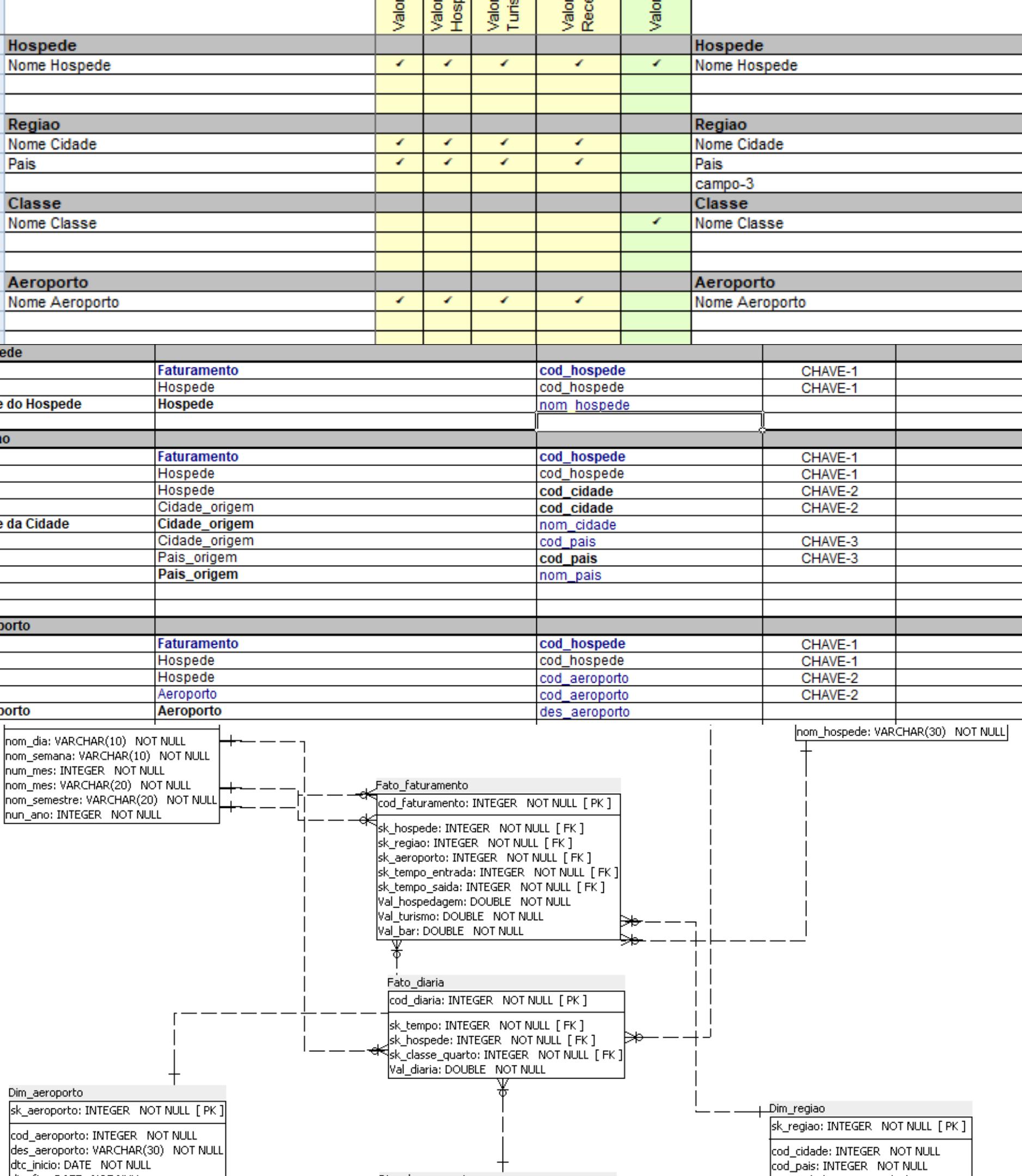


# SUPER ACADEMIA R

## 06 CURSOS JUNTOS

Aprendizado com uma das principais linguagens utilizadas em todo o mundo. Este treinamento compõe desde a introdução dos comandos básicos, passando pela construção de relatórios dinâmicos no R MARKDOWN, pelo uso da linguagem R em painéis de dados com o uso do SHINY, passando pela geração de infográficos no GGPLOT2 e construção de consultas a banco de dados pela biblioteca DPLYR.

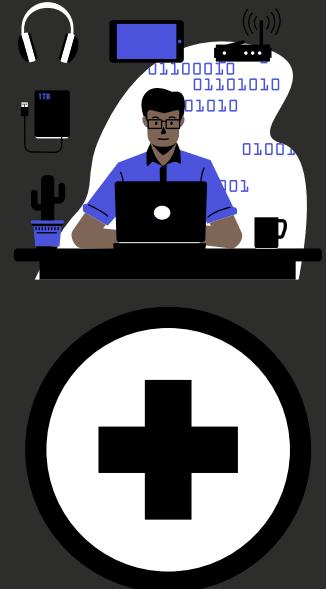


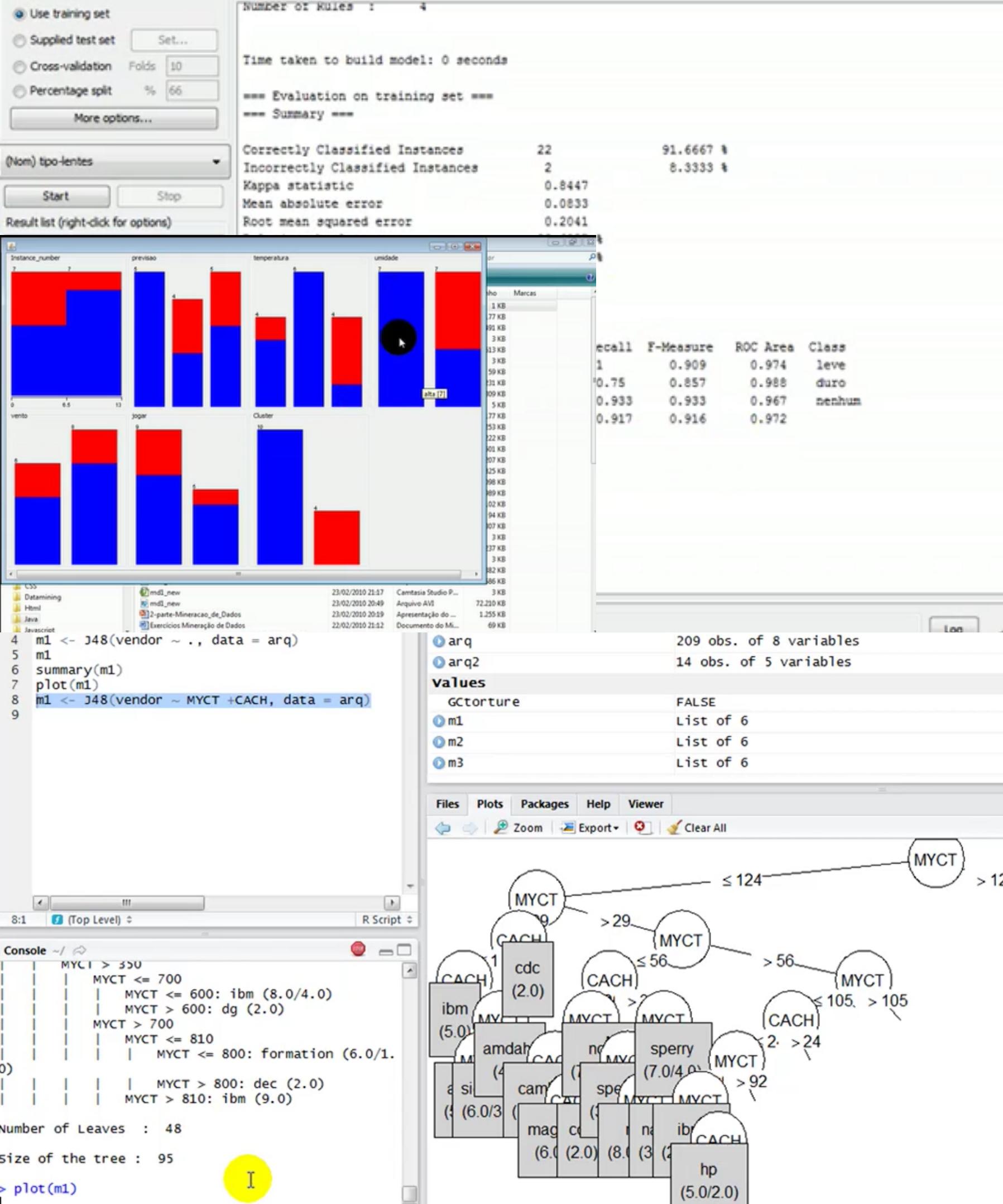


# SUPER ACADEMIA DATA WAREHOUSE

## PREPARANDO DW

Treinamento criado para os profissionais que desejam construir bases de dados multidimensionais. Neste curso será possível aprender a realizar o levantamento de dados, construção do modelo multidimensional e aplicação em uma ferramenta de extração, carga e tratamento dos dados.

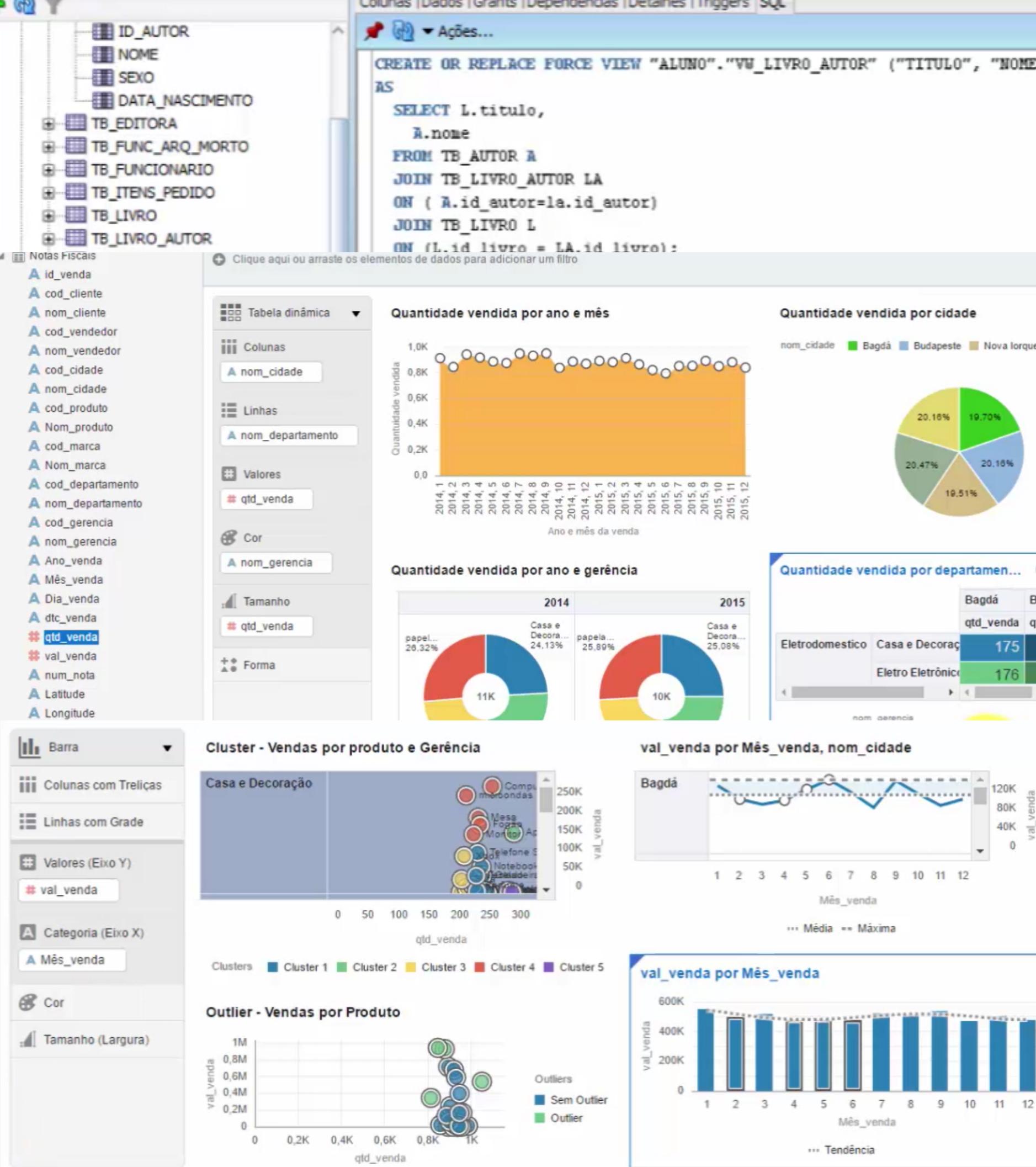




# SUPER ACADEMIA DATA SCIENCE

## TRABALHANDO COM MACHINE LEARNING

Reunidos neste curso as principais ferramentas do mercado na área de Data Science: R, Python, ORANGE e WEKA. Com um portfólio variado com diversos exercícios e tudo muito prático, trabalhando diretamente com o uso de estatística e scripts.



# SUPER ACADEMIA ORACLE

## TRABALHANDO COM BANCO ORACLE

Para os profissionais que desejam ser especialista em banco de dados ORACLE, este é o melhor treinamento. Aqui será possível aprender a trabalhar com SQL e PLSQL com diversos exercícios, além da construção, conta com um portfólio de ferramentas de BI, a famosa OBIEE e o ODV ambas ferramentas gráficas.



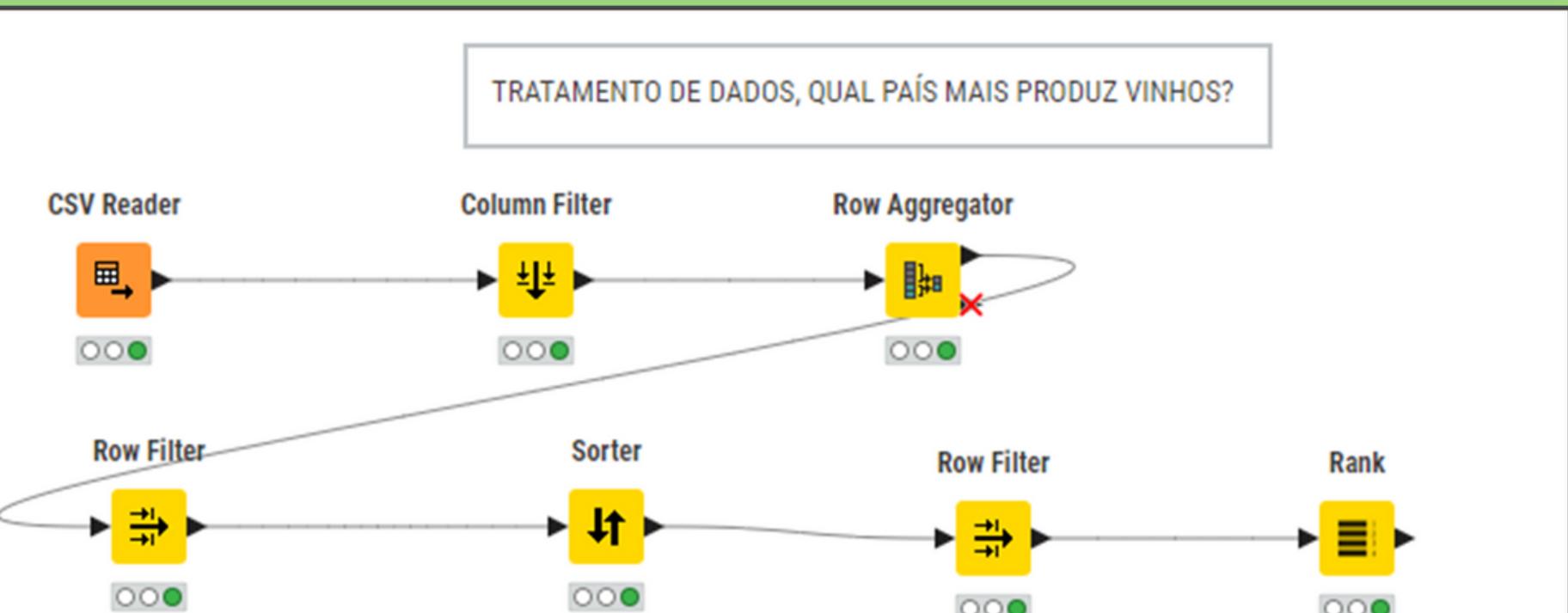
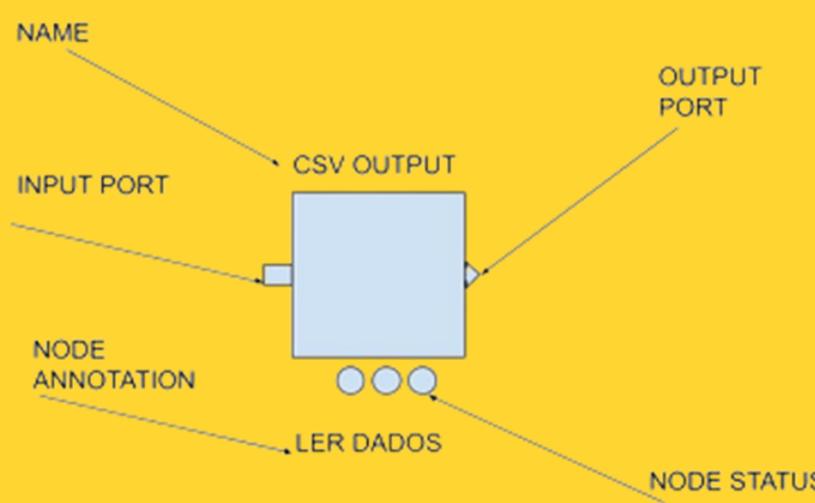
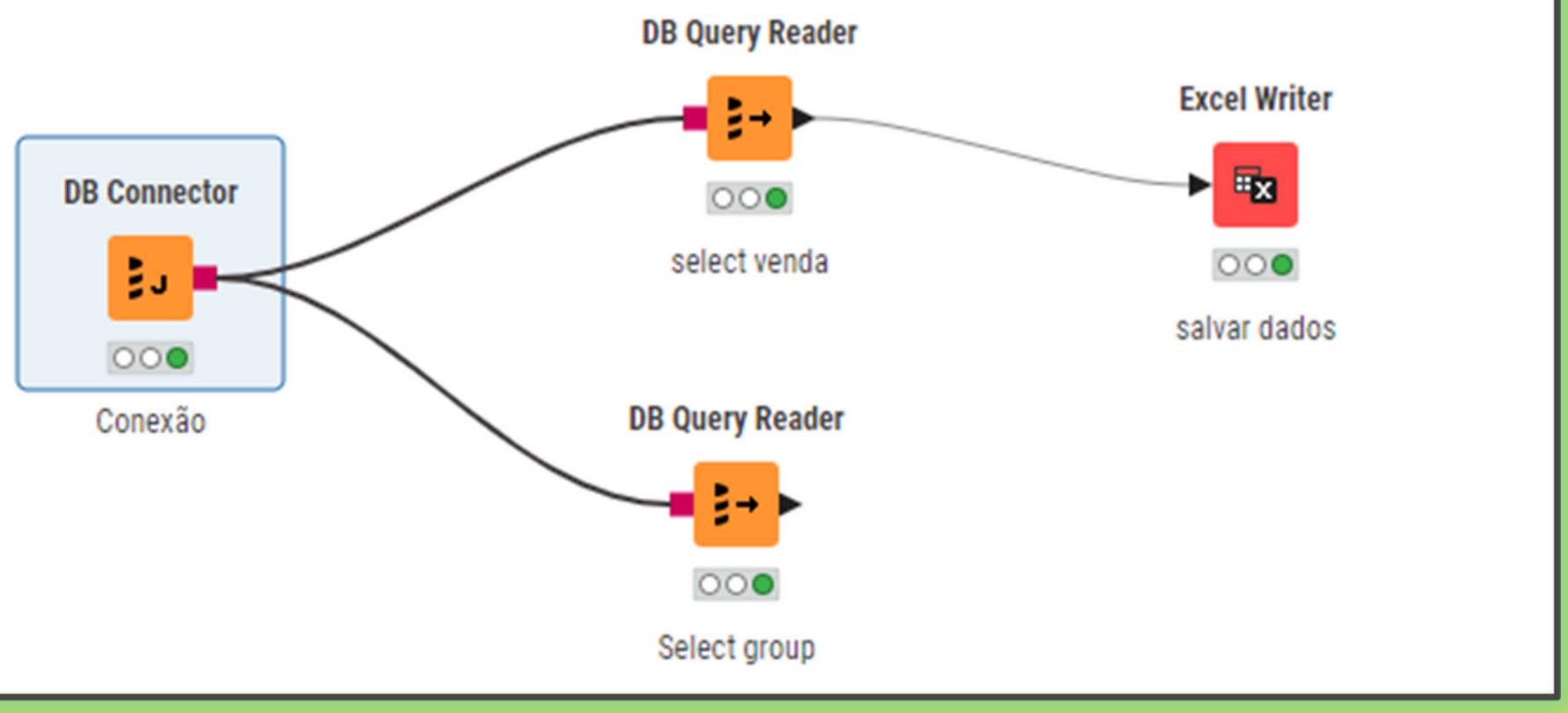


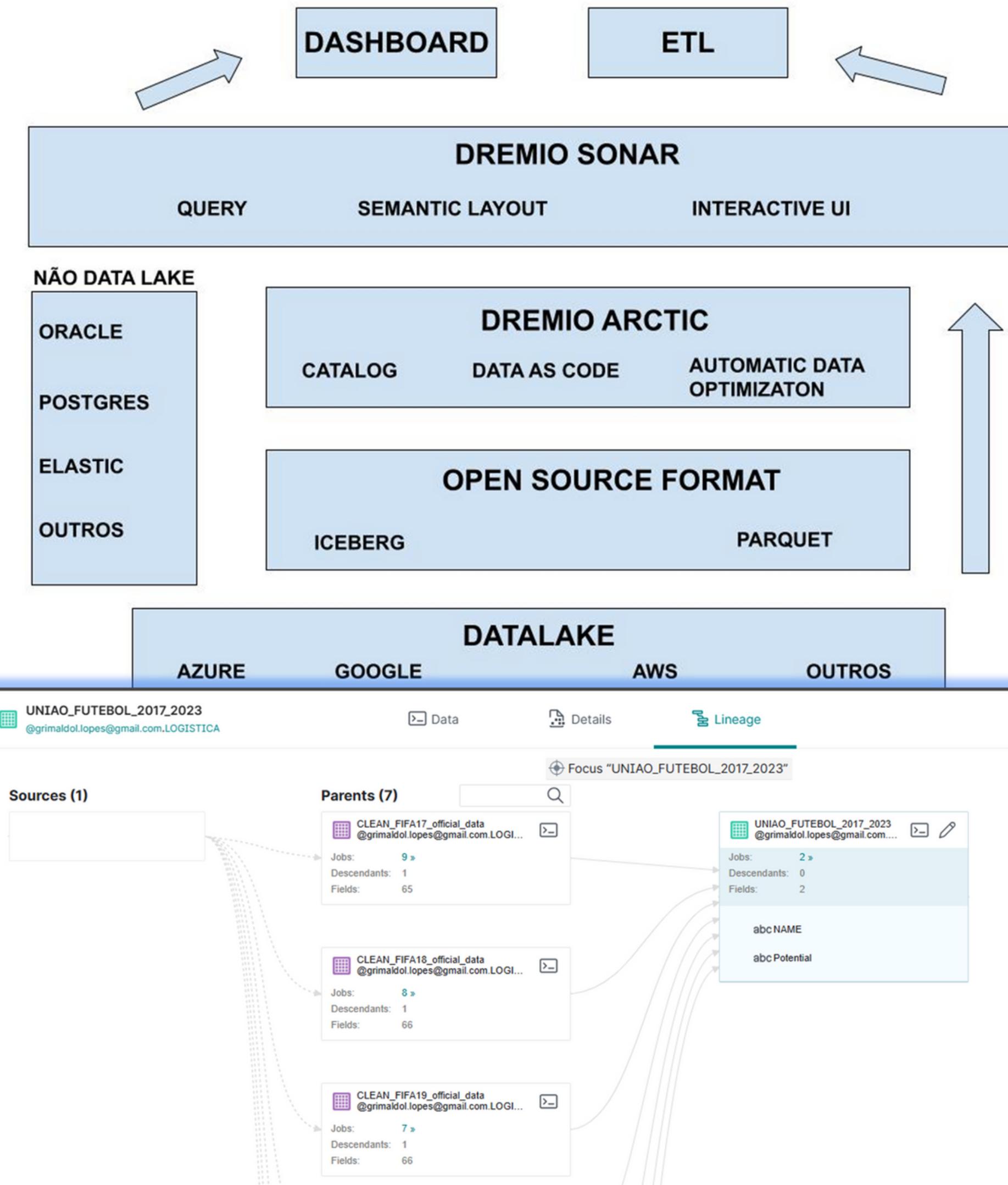
# KNIME DATA ENGINEER

## INTEGRE E OTIMIZE PIPELINES DE DADOS

O KNIME oferece uma plataforma intuitiva e poderosa para a integração de dados, permitindo que profissionais construam fluxos de trabalho eficazes. Sua capacidade de lidar com diversas fontes, transformar dados e garantir qualidade coloca os usuários em posição estratégica para enfrentar desafios complexos na engenharia de dados.

Desenvolverá habilidades cruciais para explorar fontes variadas, realizar transformações eficientes e garantir a confiabilidade dos dados. Ao final, você será capaz de enfrentar projetos em dados, impulsionando sua carreira na engenharia de dados.





# DREMIO DATA LAKE

## CONSTRUÇÃO E GOVERNANÇA EM DATA LAKE

Dremio é uma plataforma inovadora de análise de dados que permite que as organizações desbloqueiem todo o potencial de suas fontes de dados. Uma das principais vantagens do Dremio é sua abordagem focada na governança de dados.

Com o Dremio, a exploração de dados se torna uma tarefa intuitiva e interativa. Você entenderá como criar integrar com as ferramentas populares de visualização de dados, como Tableau, Power BI, em conjunto com o Dremio.

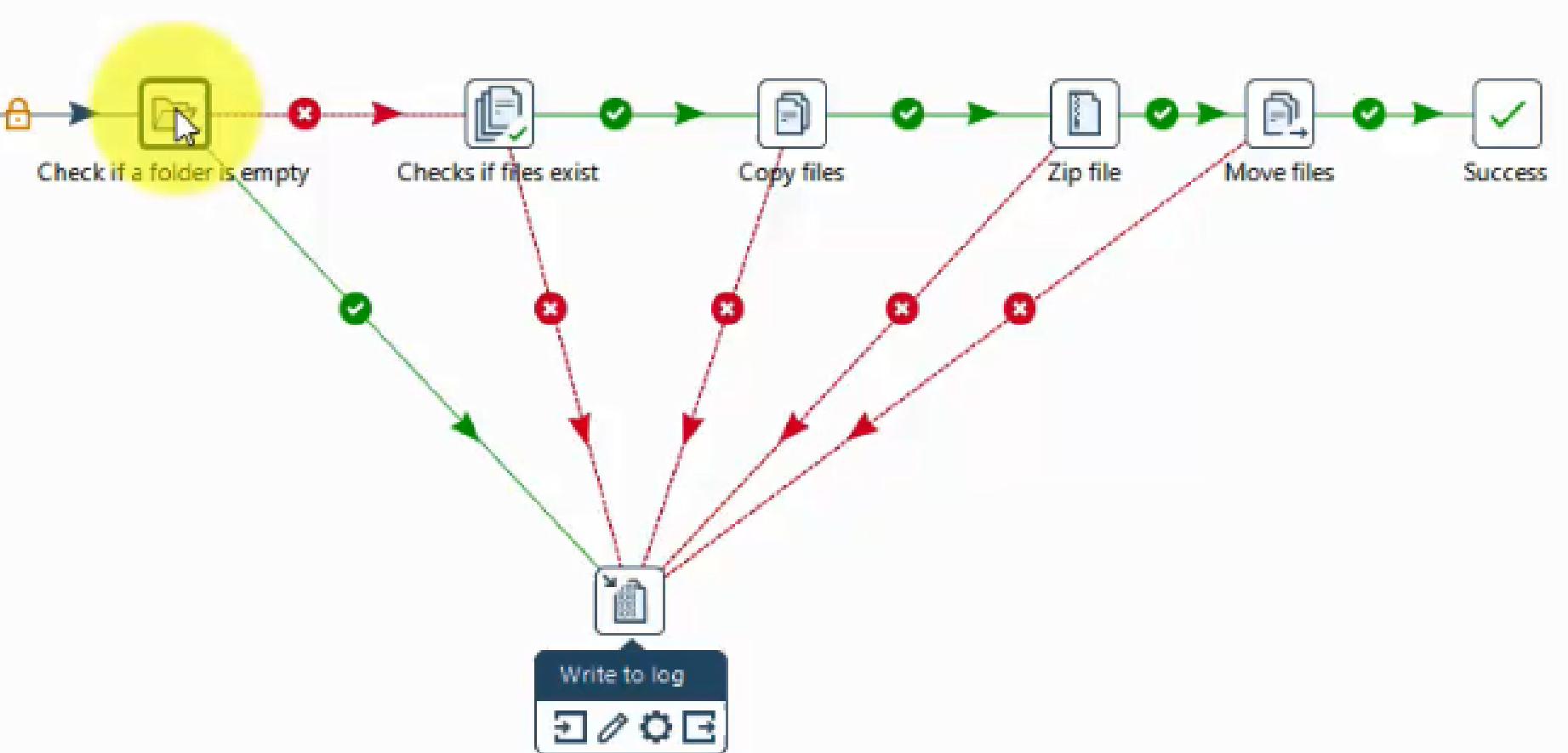
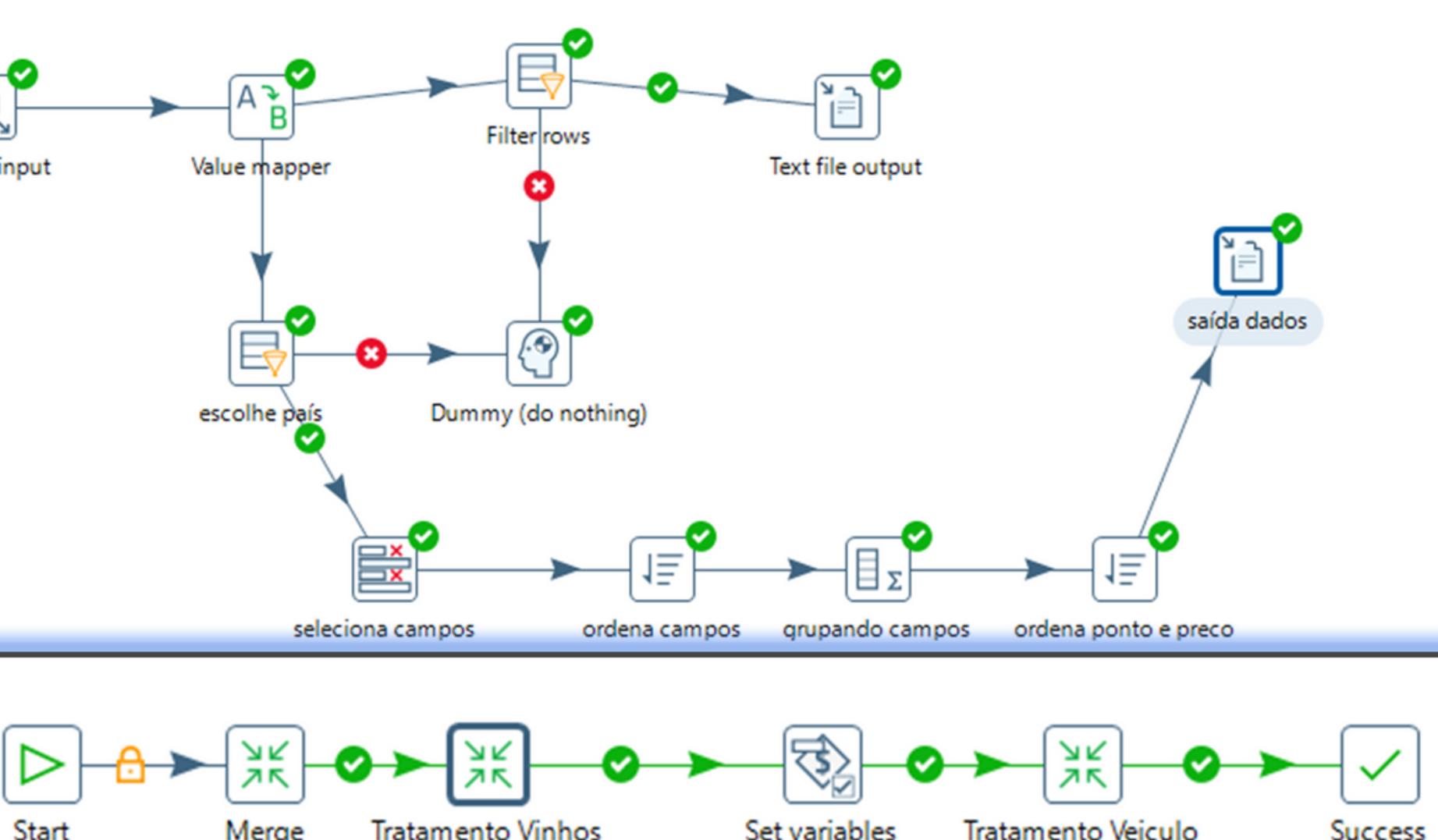
A trilha da linhagem de dados é fundamental para garantir a transparência e confiabilidade das informações. Vamos mostrar como o Dremio rastreia a linhagem de dados, permitindo que você compreenda a origem e o impacto de cada informação presente em seus relatórios e análises.





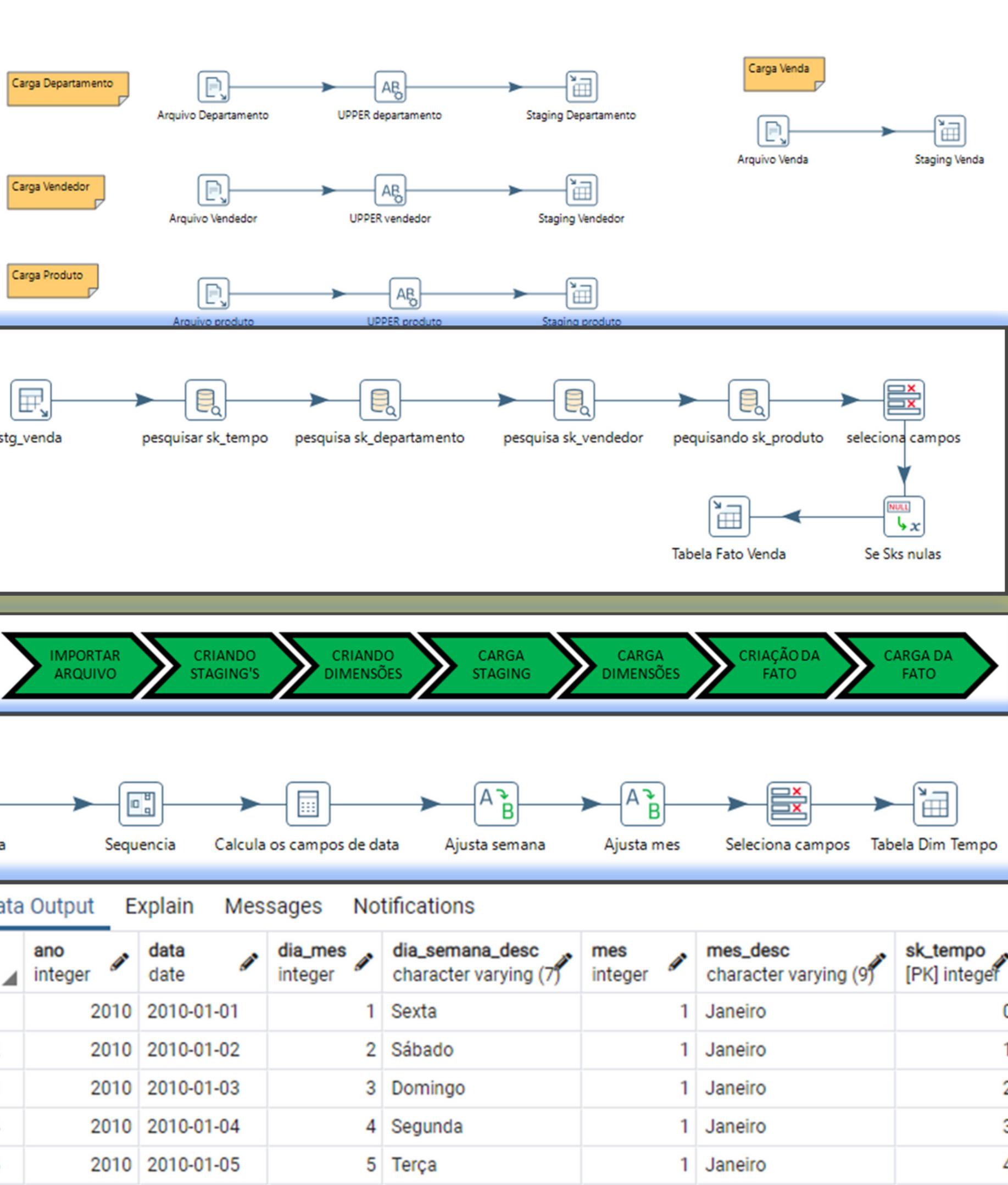
# PENTAHO PDI

## INTEGRAÇÃO E INGESTÃO DE DADOS



Com o aumento exponencial da quantidade de dados gerados diariamente pelas empresas, a necessidade de profissionais capacitados em análise e processamento de dados tem crescido de forma significativa. O Pentaho PDI é uma ferramenta poderosa que permite a integração de diferentes fontes de dados e o tratamento de informações para transformá-las em insights úteis para as empresas.

A ferramenta é conhecida em todo o mundo como uma das melhores OPEN SOURCE que existe, trabalhando com criação de fluxos de dados, passando por transformações e limpezas de dados, além de automatização de processos de ETL.

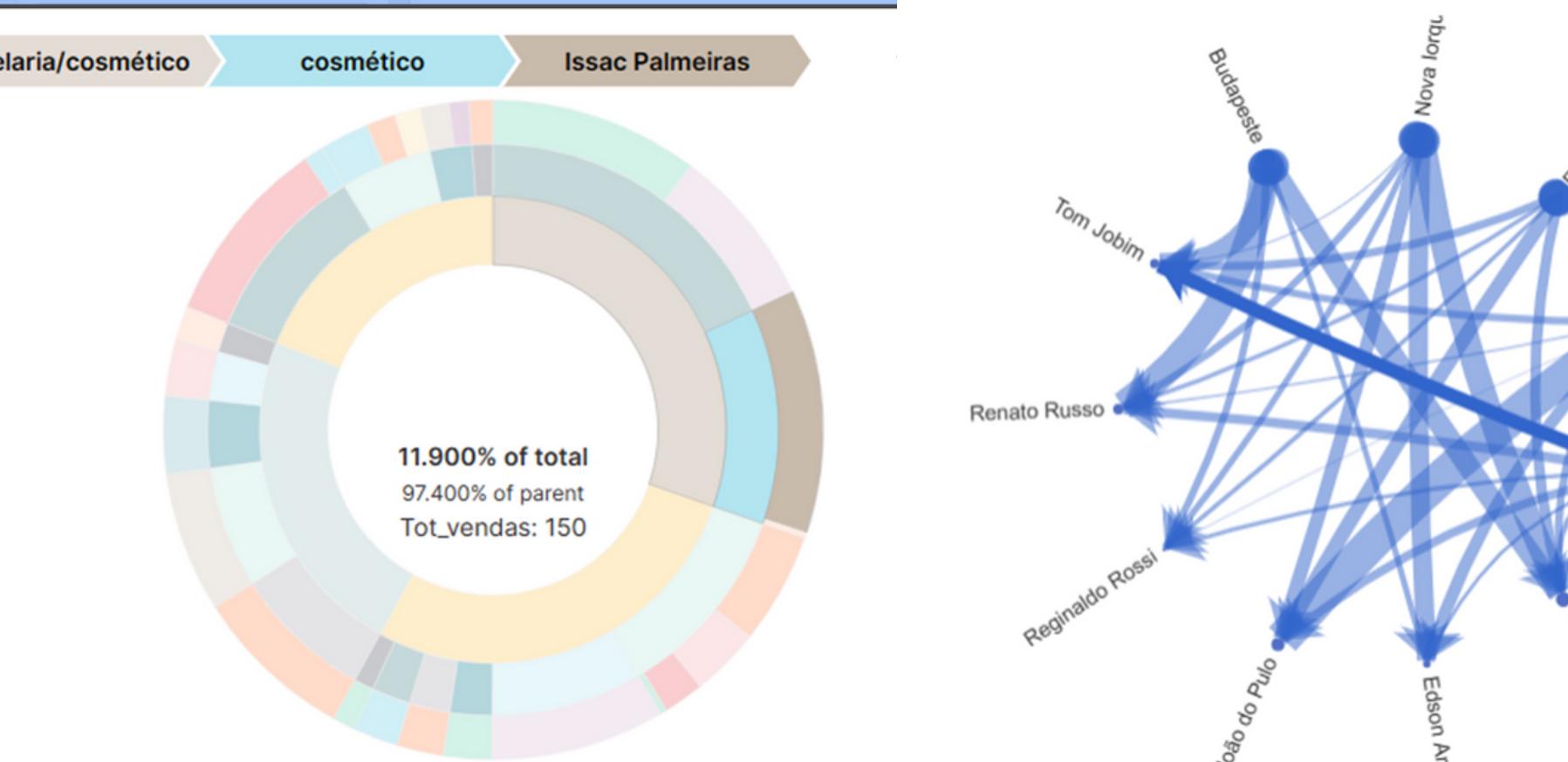
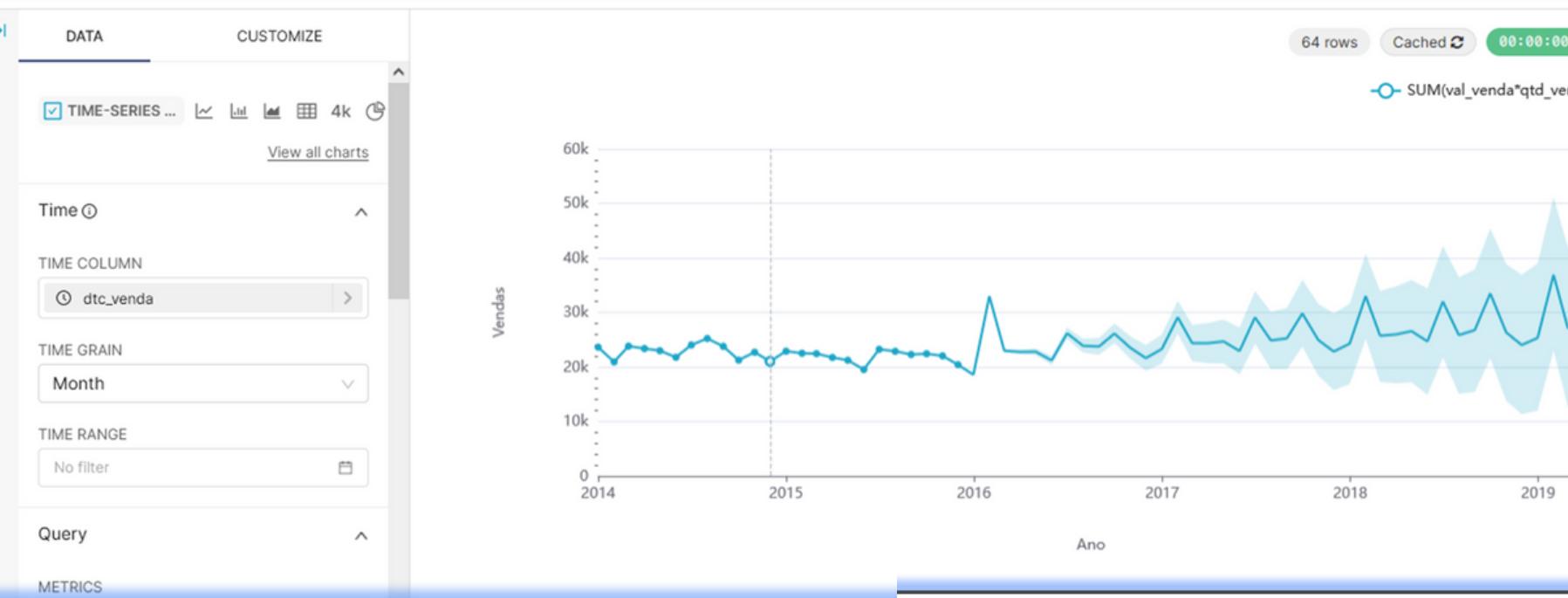


# PENTAHO PDI

## PARA PROJETOS DE DATA WAREHOUSE

No ambiente empresarial moderno, a quantidade de dados gerados cresce exponencialmente a cada dia. Para que essas informações se tornem úteis, é necessário organizá-las de maneira estruturada e coerente, permitindo análises complexas e tomadas de decisões fundamentadas. É nesse ponto que o conceito de Data Warehouse desempenha um papel crucial.

Um Data Warehouse é um repositório centralizado de dados que combina informações provenientes de diferentes fontes e as organiza de maneira consistente para análise e geração de relatórios. Ao aprender a construir e gerenciar um Data Warehouse, você estará apto a auxiliar as empresas na transformação de dados brutos em conhecimentos estratégicos.



# PRESET APACHE SUPERSET

O Preset - Apache Superset é uma ferramenta poderosa para análise e visualização de dados, que permite aos profissionais de dados criar dashboards e explorar dados de forma eficiente e intuitiva. Aprender a utilizar essa ferramenta pode trazer vários benefícios para a carreira de um profissional de dados.

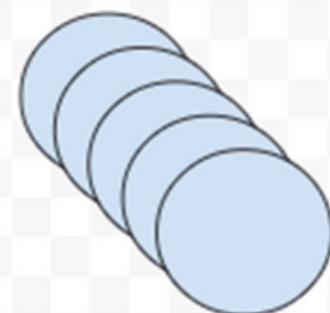
APACHE Superset é open source, mas o Preset é uma versão que embarca a ferramenta da APACHE, de forma a facilitar seu entendimento e manuseio. Ela disponibiliza gratuitamente acesso e permite que você desenvolva seus projetos na área de análise de dados sem problemas.

1. Análise de dados mais rápida e eficiente
2. Criação de dashboards interativos e fáceis de compartilhar
3. Integração com fontes de dados diferentes, incluindo bancos de dados relacionais e NoSQL

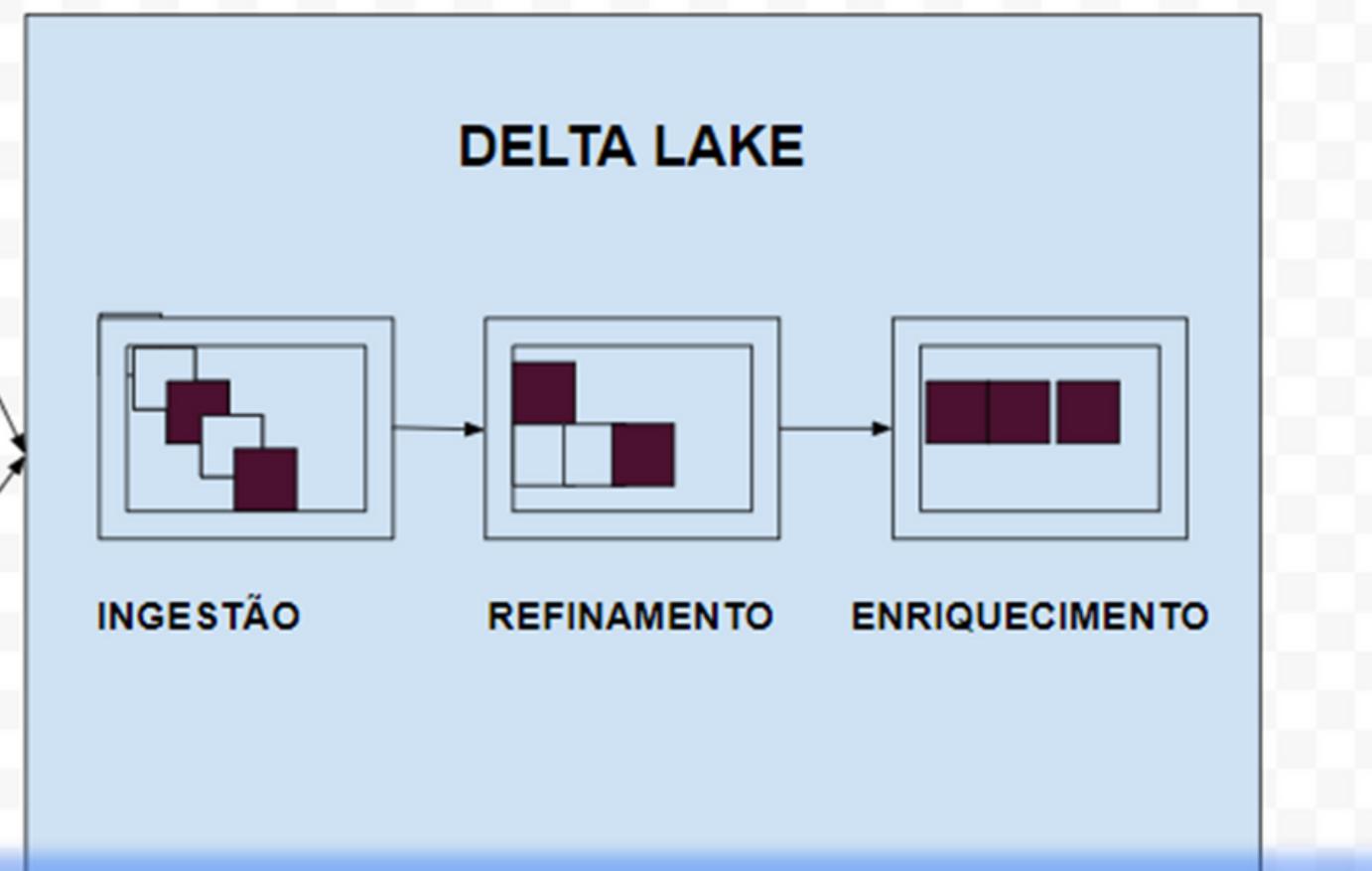
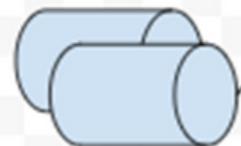


	userId	userName	operation	operationParameters
51:11.000+0000	6174905436003004	grimaldo_lopes@hotmail.com	DELETE	▶ {"predicate": "((CAST(spark_catalog.default.co)"}
49:32.000+0000	6174905436003004	grimaldo_lopes@hotmail.com	UPDATE	▶ {"predicate": "(cast(id#6147 as int) = 4)"}
48:14.000+0000	6174905436003004	grimaldo_lopes@hotmail.com	MERGE	▶ {"predicate": "(CAST(spark_catalog.default.compa"}
45:59.000+0000	6174905436003004	grimaldo_lopes@hotmail.com	CREATE OR REPLACE TABLE	▶ {"isManaged": "true", "description": null, "partiti"

STREAMING



BATCH



ROM delta. `/user/hive/warehouse/compras` VERSION AS OF 3

bs

	date_order	customer	product	unit	price
	2021-03-22	Ana	AR-CONDICIONADO	6	121.6
	2021-04-23	Sandra	Geladeira	8	313

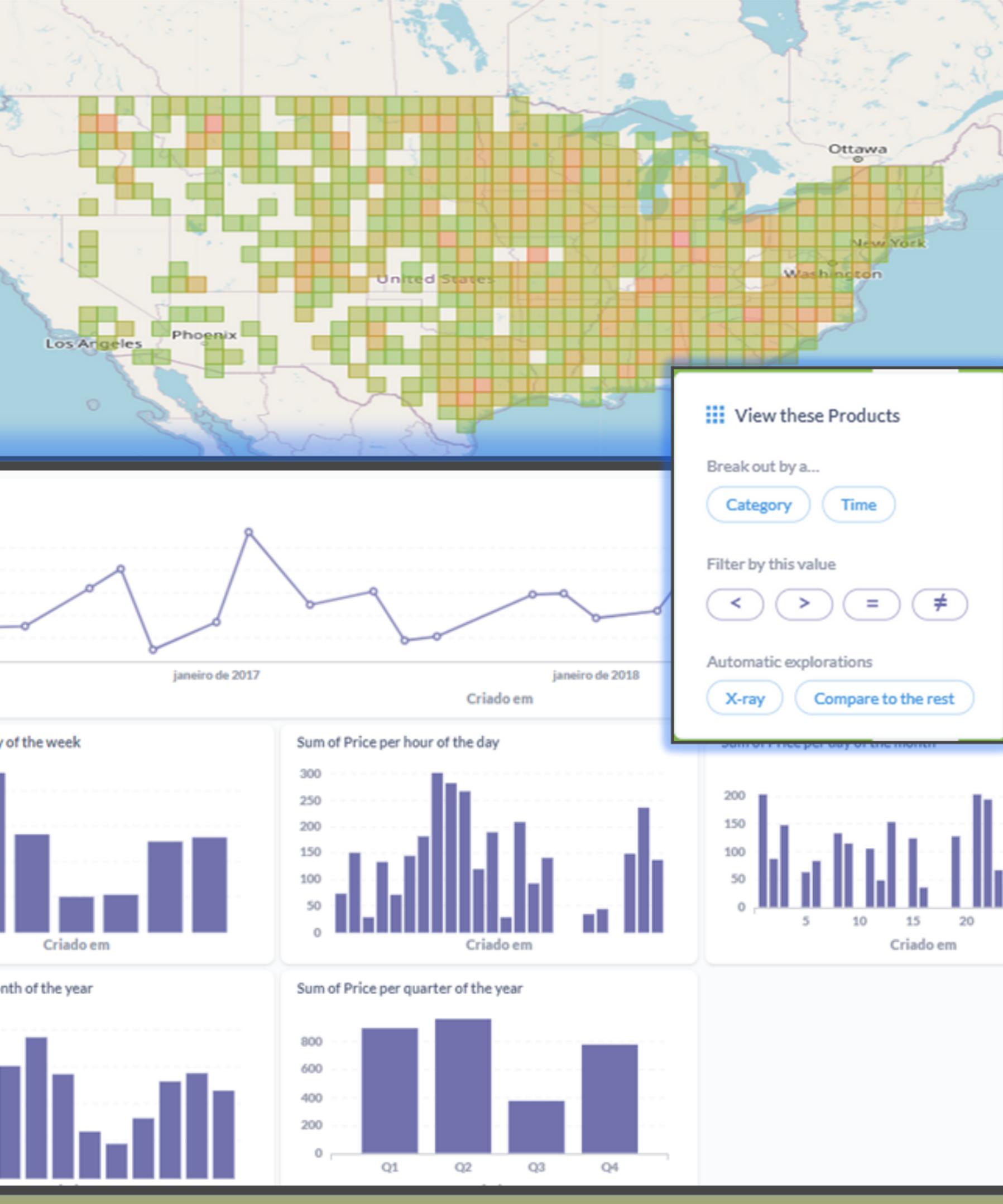
# DATABRICKS

## DELTA LAKE

O Databricks Delta Lake é um projeto de código aberto que permite construir uma arquitetura chamada Lakehouse sobre o Data Lake no Databricks. O Delta Lake fornece processamento em batch, streaming em lote, além de controle de transações sobre os dados, como um banco de dados.



- **FACILIDADE:** gerenciamento de dados temporais que simplifica seu pipeline de dados, facilitando a auditoria.
- **AUDITORIA:** Organizações que trabalham com sistemas de dados tradicionais para tecnologias de Big Data sempre tem necessidade de auditar os dados, fundamental em termos de conformidade de dados
- **REPRODUZIR EXPERIMENTOS:** analistas ou cientistas de dados projetam práticas recomendadas criando várias cópias dos dados, levando a um aumento dos custos de armazenamento. Tudo isso para simular a história dos dados, no Delta Lake você utiliza qualquer versão dos dados, pois eles ficam gravados e são historizados dentro do DBFS.



# METABASE

## BI + SQL + DASHBOARD

No METABASE é possível criar gráficos, explorá-los com várias visões simplesmente trocando as funcionalidades com o clique no mouse, construção de consultas SQL nativas, e publicação e exportação dos resultados. Veja o que mais você aprenderá com o METABASE:

- Criar análises de forma simples que permite que qualquer profissional utilize-a para tomar decisões com dados
- Permite a construção de consultas SQL de forma nativa, com ambiente próprio para isso
- Permite que você construa painéis com gráficos e sua storytelling de dados diretamente nas chamadas Collections
- Possui uma interface intuitiva e simples utilizando objetos em menus para a construção de suas análises
- É uma ferramenta que utiliza os dados para responder as suas principais perguntas



## FONTE DE DADOS



Ferramentas Back End  
Fontes Externas  
Extração  
Transformação  
Carga  
Atualização



Meta Dados



Serv OLAP



Ferramentas Front End  
Análise  
Data Mining  
Ger. Relatórios



BDs Operacionais



Serv OLAP

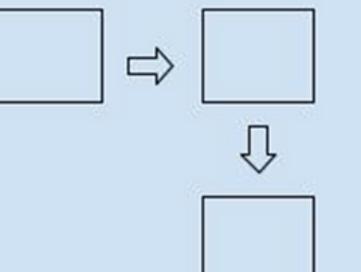


## JAVA

## APACHE HOP

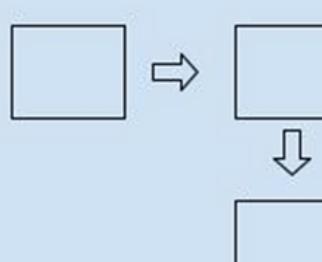
## PIPELINE

## TRANSFORMAÇÕES



## WORKFLOW

## FLUXO DE TRABALHO

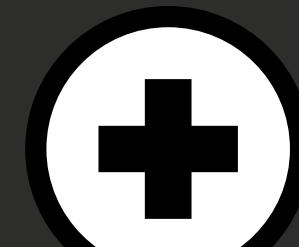


# PROJETOS DW com APACHE HOP

## ETL PARA DATA WAREHOUSE

O APACHE HOP é conhecido no mercado pela sua grande versatilidade e facilidade na construção de pipelines de dados sem o uso de código. Para demonstrar o quanto é poderoso, trago um curso inédito e que vai ajudar você a entender como construir um projeto de Data Warehouse do zero, com os principais componentes para realizar as principais tarefas, aqui você vai compreender:

- Como funciona a modelagem de dados para Data Warehouse
- Como se cria um ecossistema para construção de um Business Intelligence
- Entendendo o que é uma dimensão tempo e como construí-la
- O que é uma staging area
- O que são dimensões Slow Change Dimension1 e Slow Change Dimension2
- O que é uma tabela Fato e como devemos construí-la





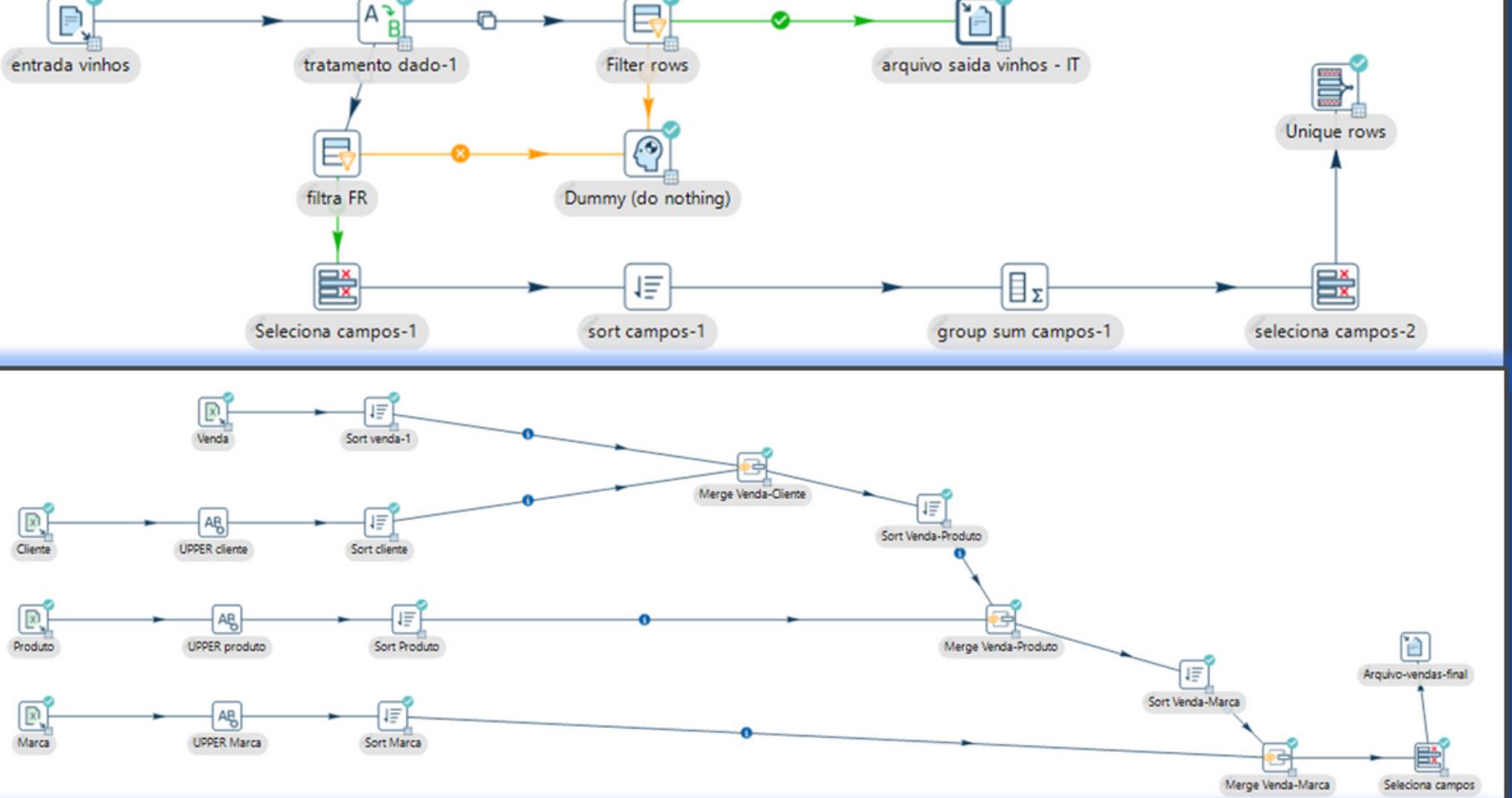
# APACHE HOP

## INTEGRAÇÃO E INGESTÃO DE DADOS

o APACHE HOP é a abreviação de Hop Orchestration Platform, é uma plataforma de orquestração de dados e engenharia de dados que visa facilitar todos os aspectos da orquestração de dados e metadados, por padrão o HOP vem com cerca de 400 plugins ou componentes.

São criados fluxos de trabalho (Workflow) e pipelines em um ambiente de desenvolvimento visual chamado Hop Gui.

Com o APACHE HOP é possível combinar, enriquecer, limpar e de muitas outras maneiras manipular dados.

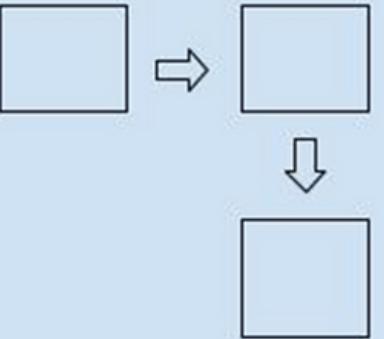


JAVA

APACHE HOP

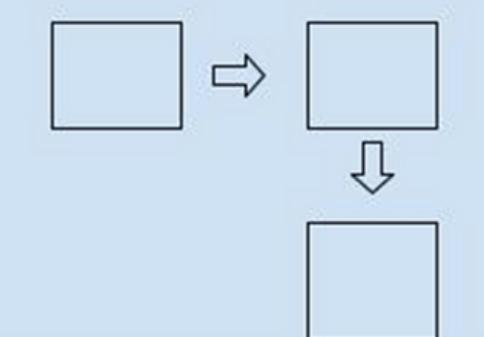
PIPELINE

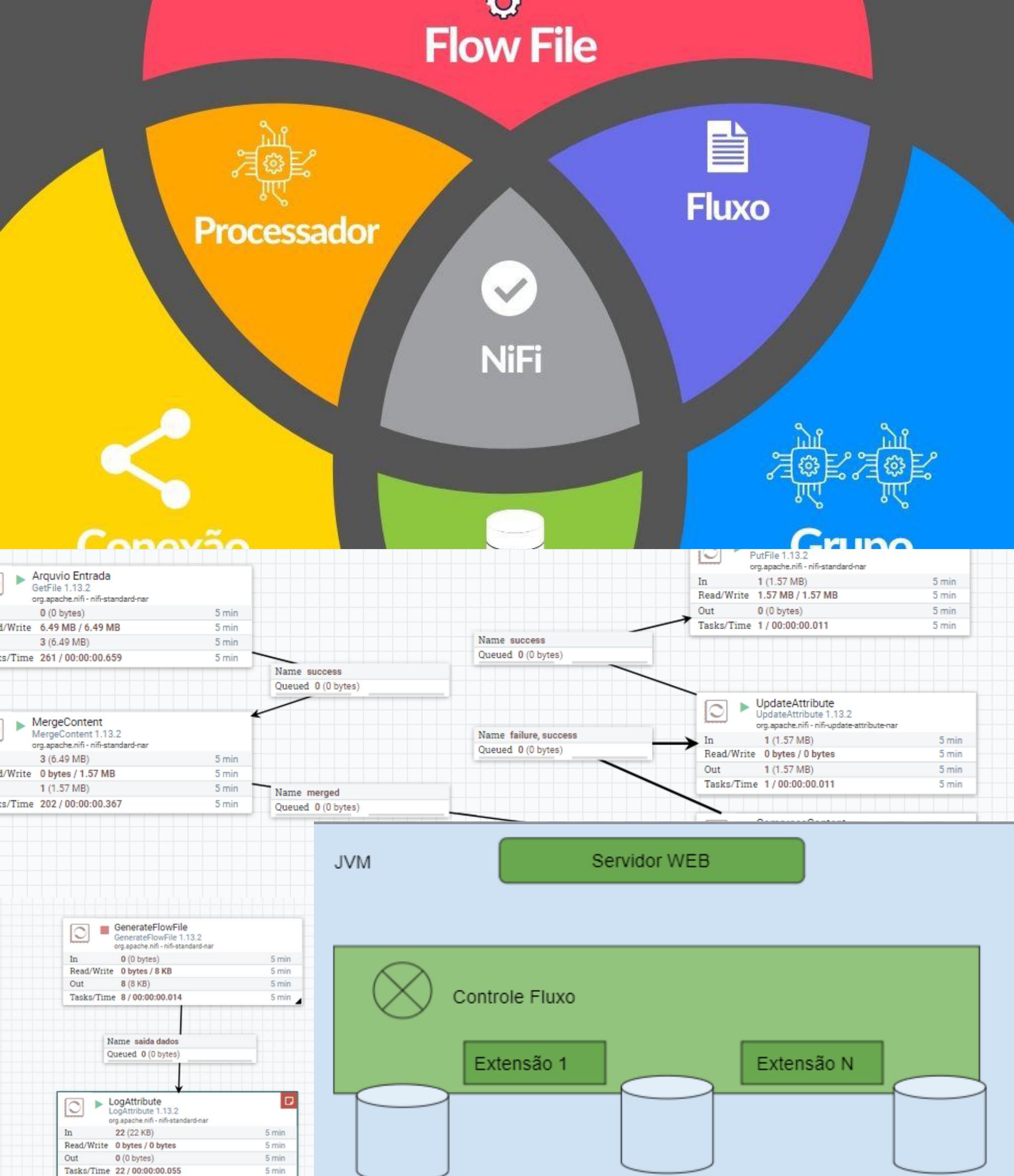
TRANSFORMAÇÕES



WORKFLOW

FLUXO DE TRABALHO





# NiFi

## CRIANDO FLUXO DE DADOS

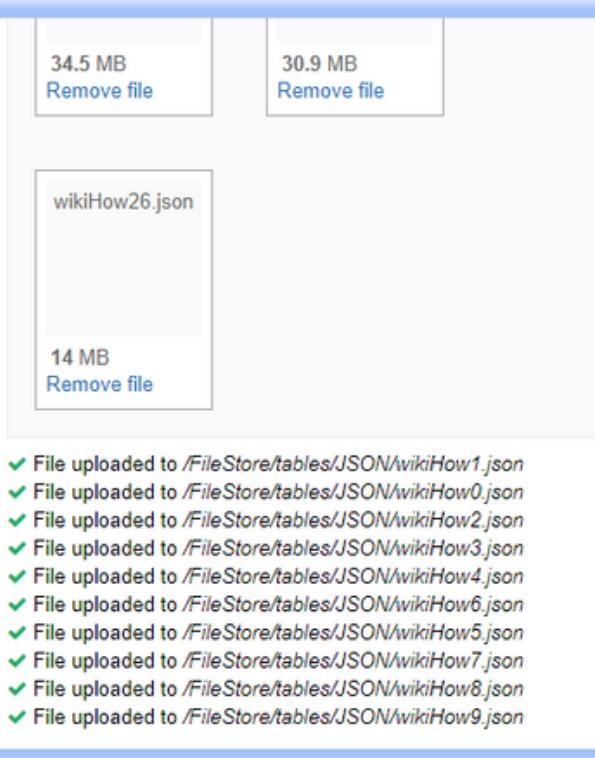
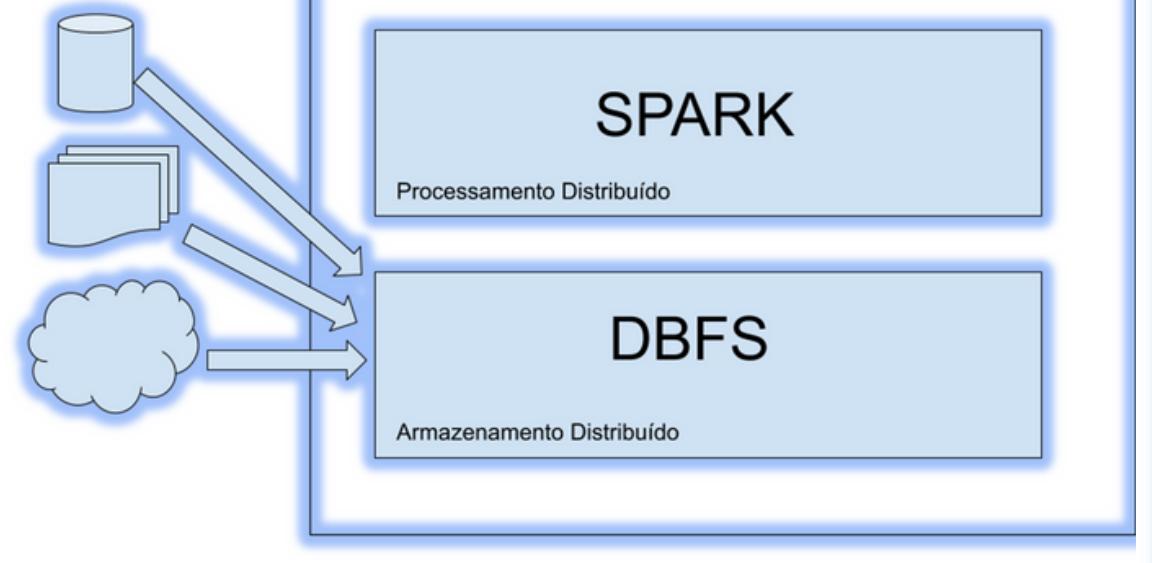
Apache Nifi é uma plataforma de ingestão de dados, criada para processar e distribuir dados entre diferentes sistemas.

Sua principal funcionalidade é o gerenciamento e a automatização do fluxo de dados entre os sistemas.

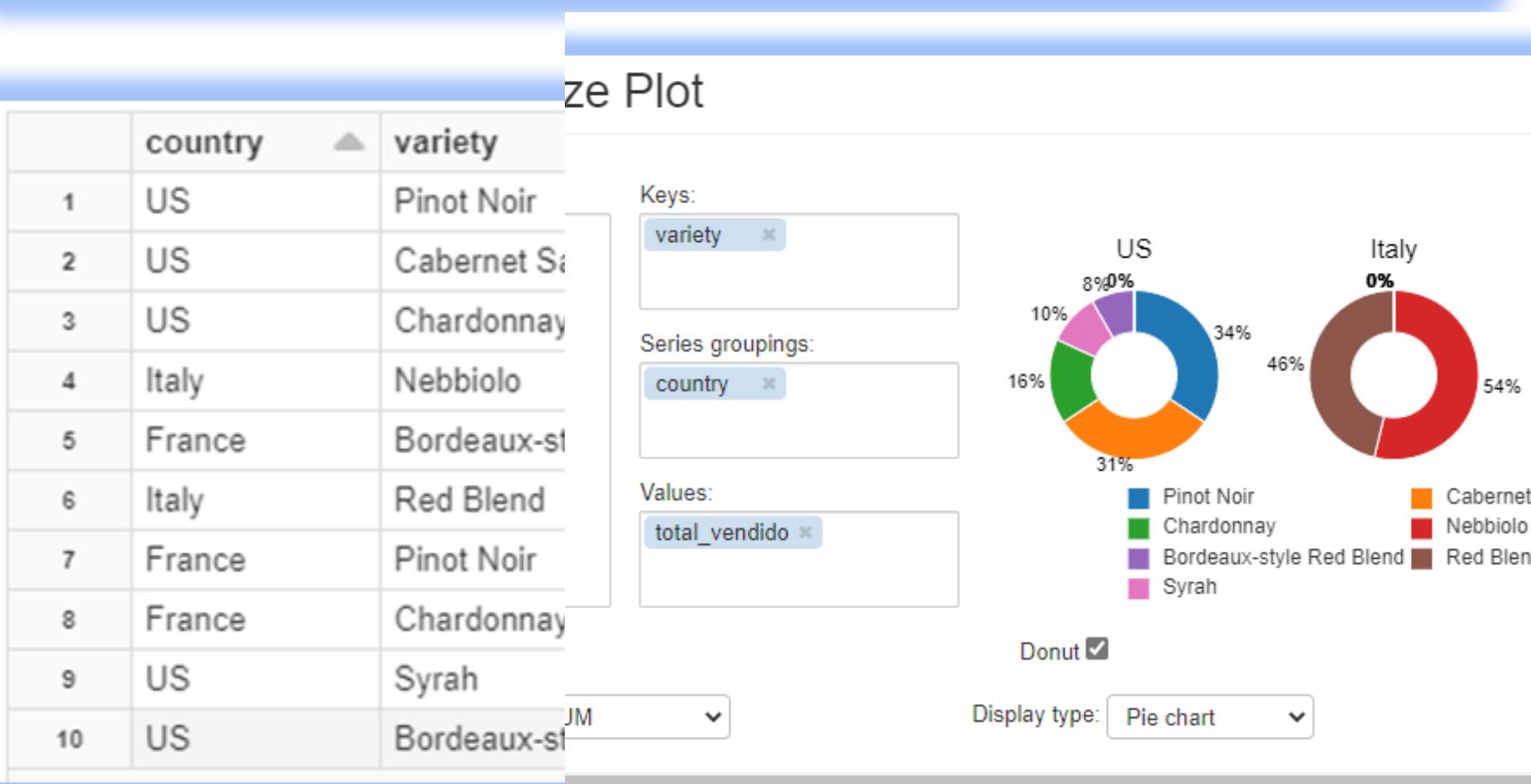
Para execução destes fluxos de dados é fornecida uma interface WEB . É a programação baseada em fluxos.

É possível realizar: coleta de dados, transmissão de dados, armazenamento de dados, dentre outras tarefas.

# DATABRICKS



```
1 %sql
2 select country, sum(price) as total_vendido from dados_vinho
3 where price > 0
4 group by country
5 order by total_vendido desc
6 limit 10
7
```



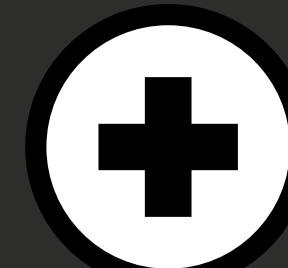
# DATABRICKS

## UNIFICANDO SEUS DADOS

Uma poderosa plataforma de colaboração entre os profissionais da área de dados. É um ecossistema fácil de usar para aqueles que desejam executar consultas em diversos conjuntos de dados.

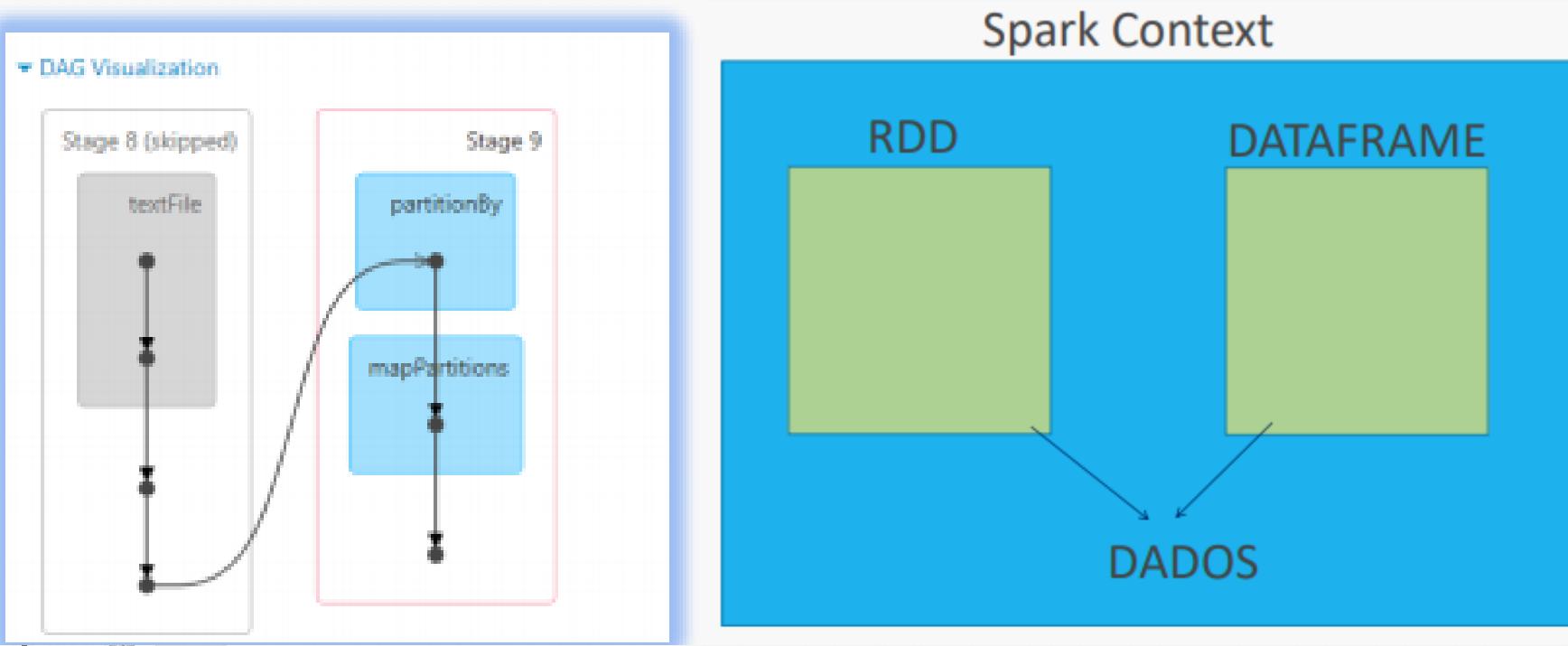
Permite que os analistas e cientistas de dados construam seus scripts nas linguagens Python, SQL, R, Scala, fazendo consulta aos dados. Separados por células em um mesmo notebook.

Com tantos gráficos sendo criados, você pode organizar tudo em dashboards dentro do Databricks.



## UTILIZANDO OPERAÇÕES DE JOIN, LEFT e RIGHT

```
In [19]: # Leitura dos arquivos de profissionais e salários  
profissional=spark.sparkContext.textFile("/Treinamento/Dados/Folhapagamento/profissionais.txt")  
salario=spark.sparkContext.textFile("/Treinamento/Dados/Folhapagamento/salario.txt")  
  
In [20]: # Exibindo os dados sobre os profissionais  
profissional.collect()  
  
Out[20]: ['Carlos,oncologista,hospital',  
'Ana,dentista,clinica',  
'Fernanda,enfermeira,hospital',  
'Sandra,pediatra,clinica',  
'Fatima,dentista,clinica',  
'Dilmar,cardiologista,hospital',  
'Fabio,pediatra,clinica',  
'Hilton,enfermeiro,clinica',  
'Daiane,dentista,clinica',  
'Paulo,farmacêutico,clinica',  
'Gilberto,pediatra,hospital']
```

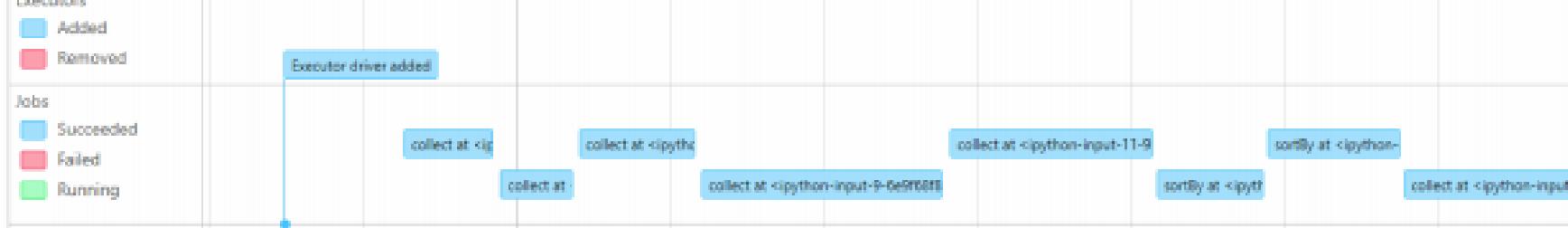


## Spark Jobs

User: IEUser  
Total Uptime: 6.1 min  
Scheduling Mode: FIFO  
Completed Jobs: 8

### Event Timeline

Enable zooming



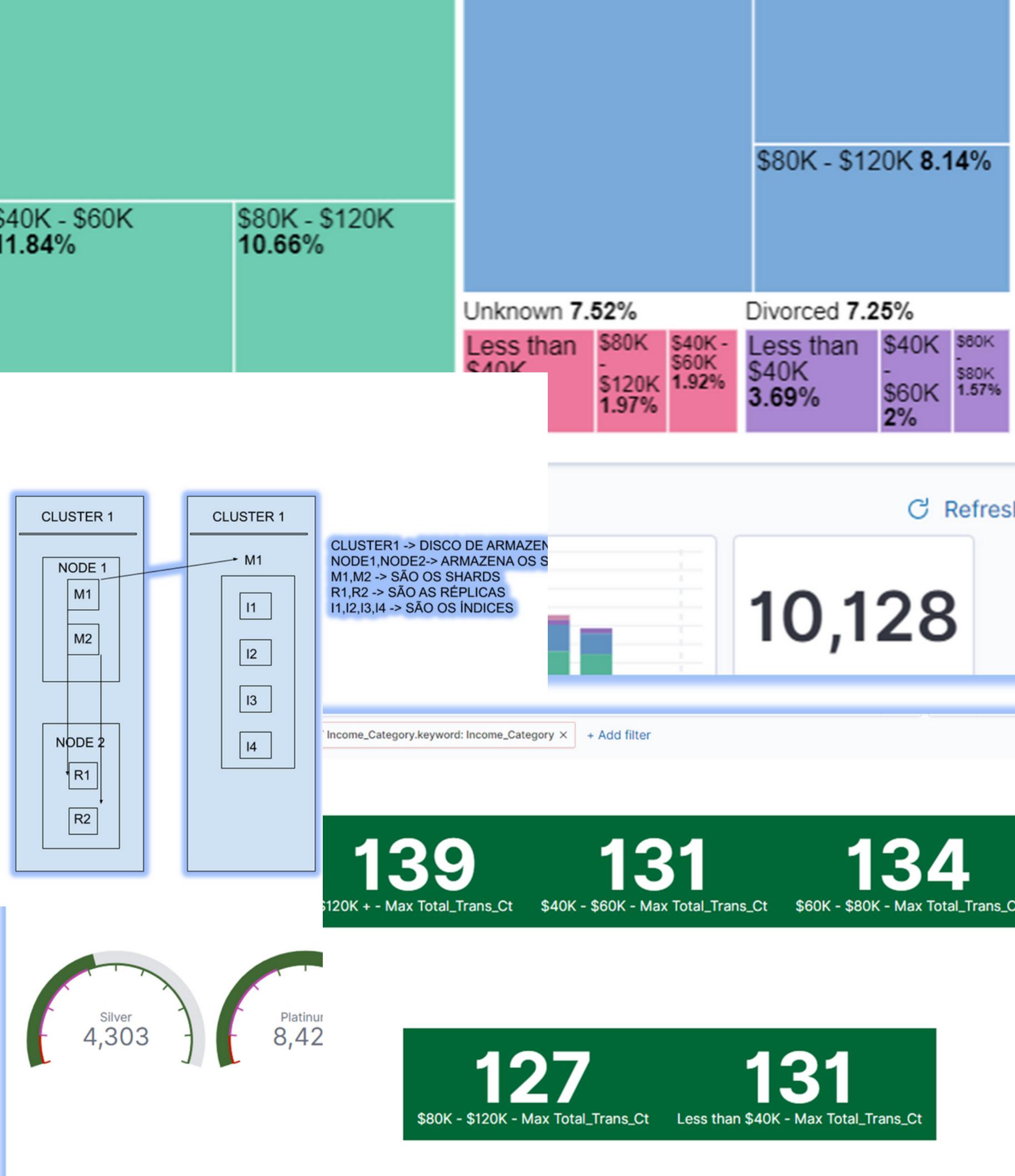
# PYSPARK

## SPARK E PYTHON PARA ANALISAR DADOS

O PySpark é uma API Python para Apache SPARK que é denominado como o mecanismo de processamento analítico para aplicações de processamento de dados distribuídos em larga escala e aprendizado de máquina, ou seja, para grandes volumes de dados.

Geralmente as aplicações criadas e executadas no PySpark são 100x mais rápidas que outras em sistemas de dados conhecidos. É um contexto novo de analisar e gerenciar dados já utilizado em dezenas de empresas pelo mundo.



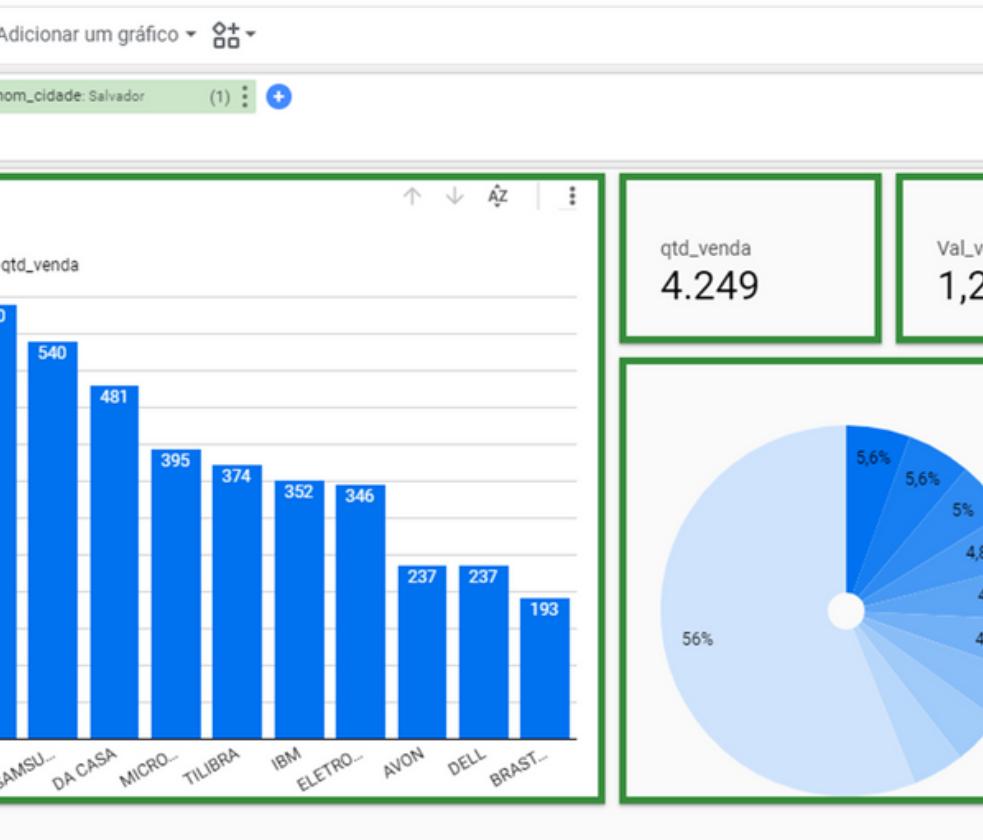


# ELASTICSEARCH

## BUSCA E ANÁLISE DE DADOS

Com o Elasticsearch você pode trabalhar com uma variedade de fontes de dados, incluindo logs, arquivos de dados em texto, e aplicações web. Ele oferece um mecanismo de pesquisa de texto completo distribuído em uma interface web e documentos JSON. Aqui neste curso nós iremos trabalhar com o banco search-engine ElasticSearch , a ferramenta de ingestão de dados conhecida como Logstash e a ferramenta de visualização e construção de painéis chamada Kibana. Você poderá construir dashboards incríveis com um banco de dados ultra rápido.



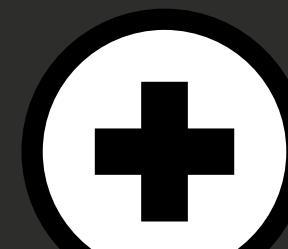


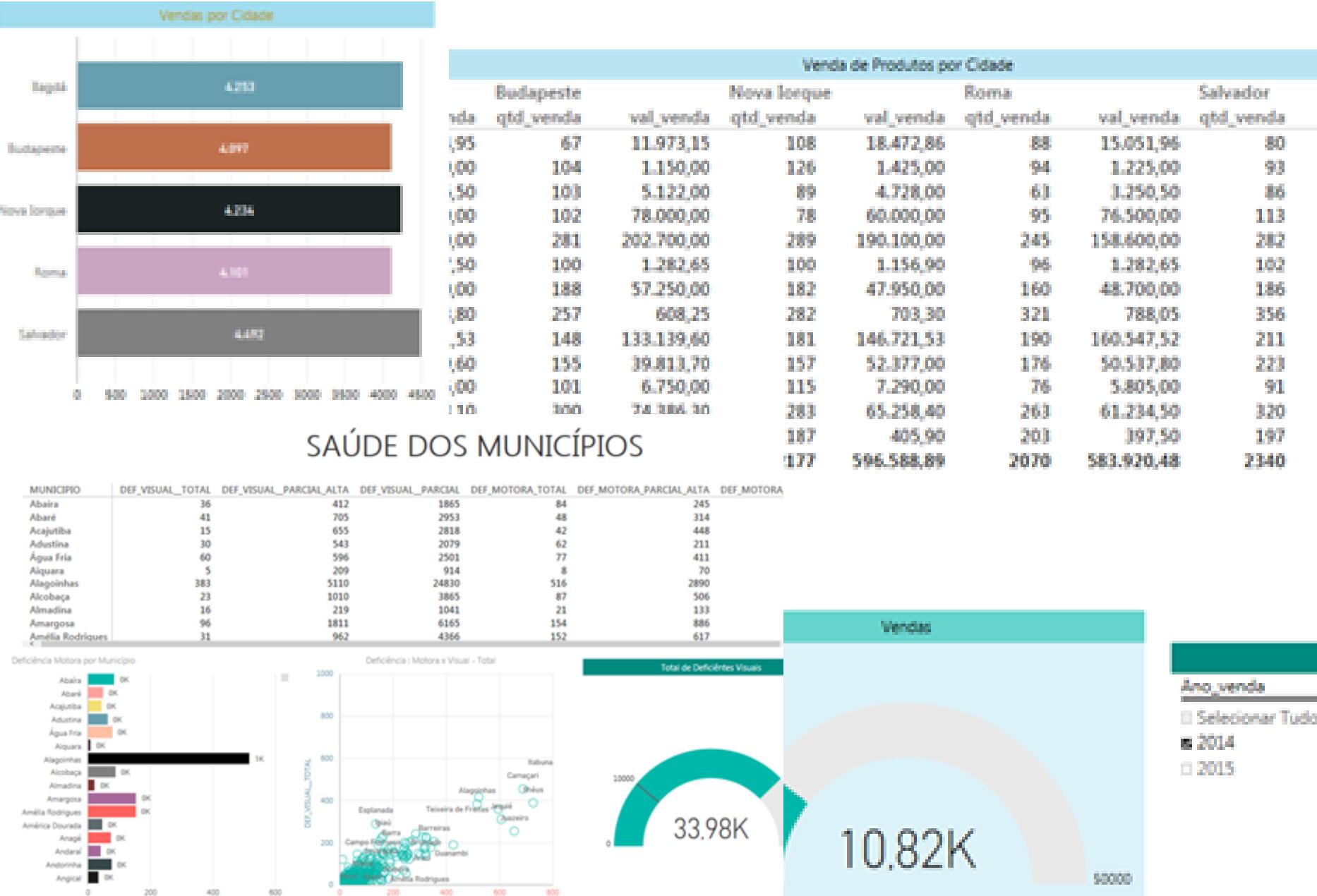
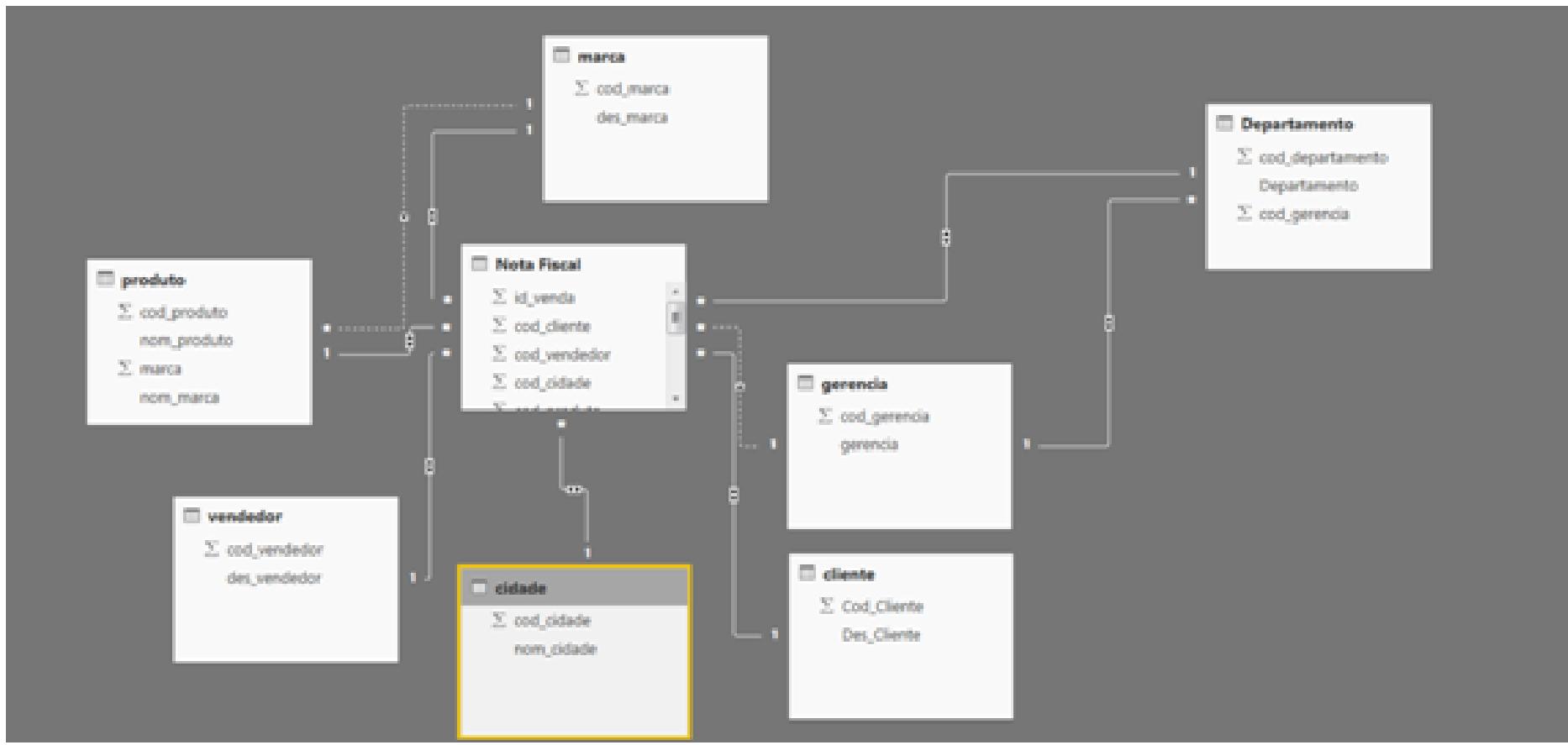
# GOOGLE DATA STUDIO

# VISUALIZAÇÃO DE DADOS

É a ferramenta de dashboard do Google totalmente personalizável e muito fácil de usar. Com ela você tem liberdade para construir seus relatórios e análises. É possível reunir múltiplas fontes de dados para a construção de seus relatórios, gráficos e dashboards . É possível compartilhar seu relatório, criar filtros personalizados, incluir imagens, estilo e temas de cores.

É uma das mais utilizadas em todo o mundo. Neste treinamento você conta com um projeto a ser construído.



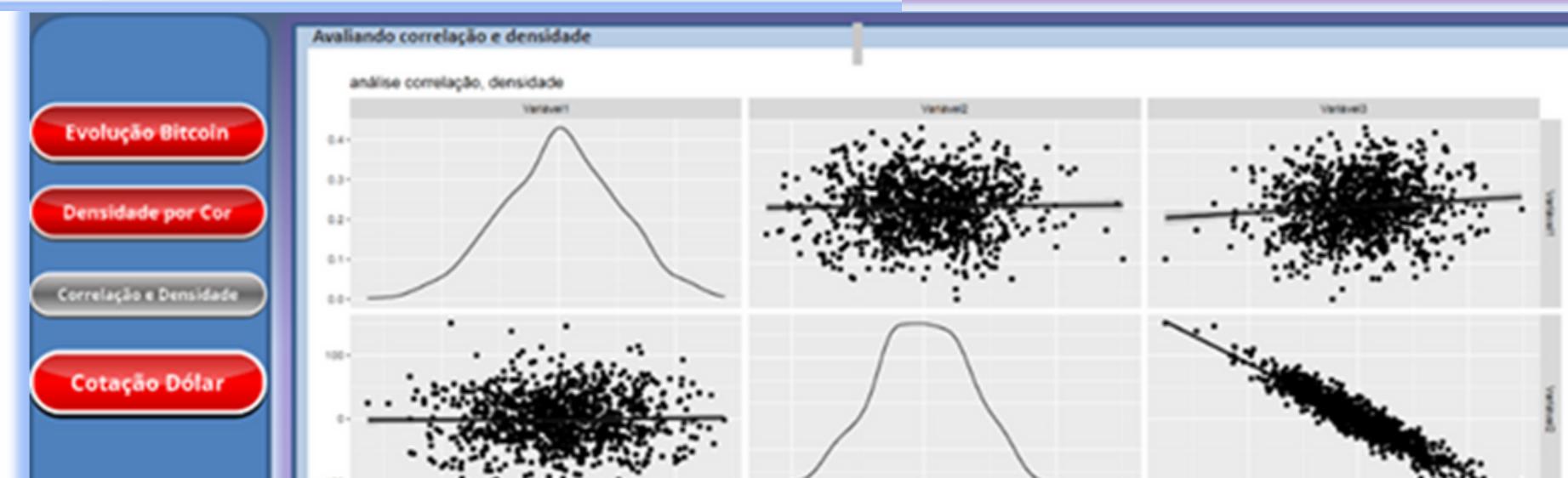
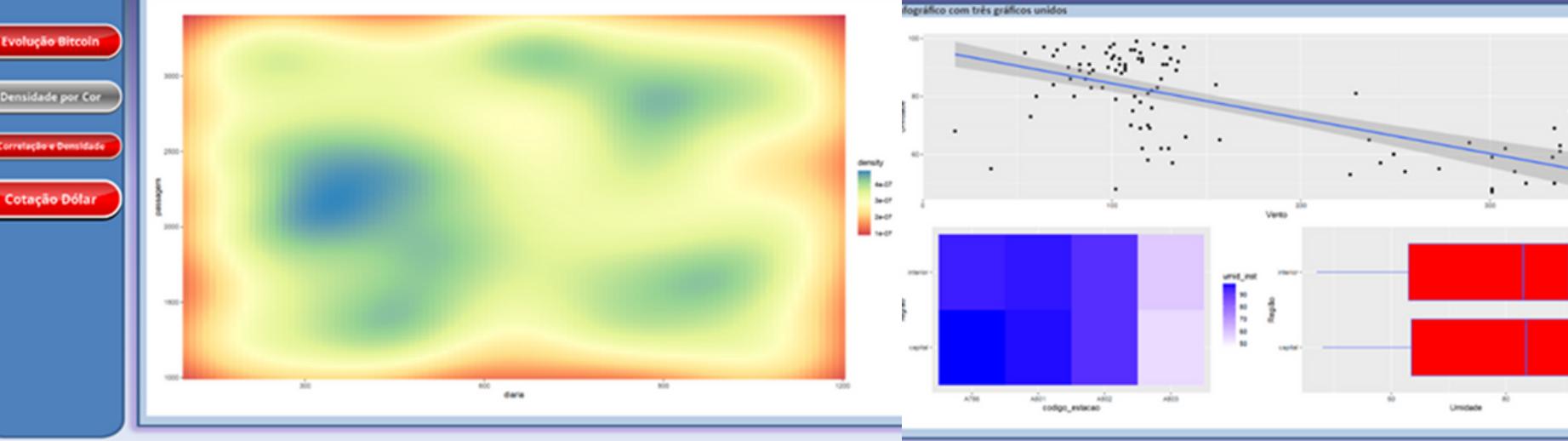
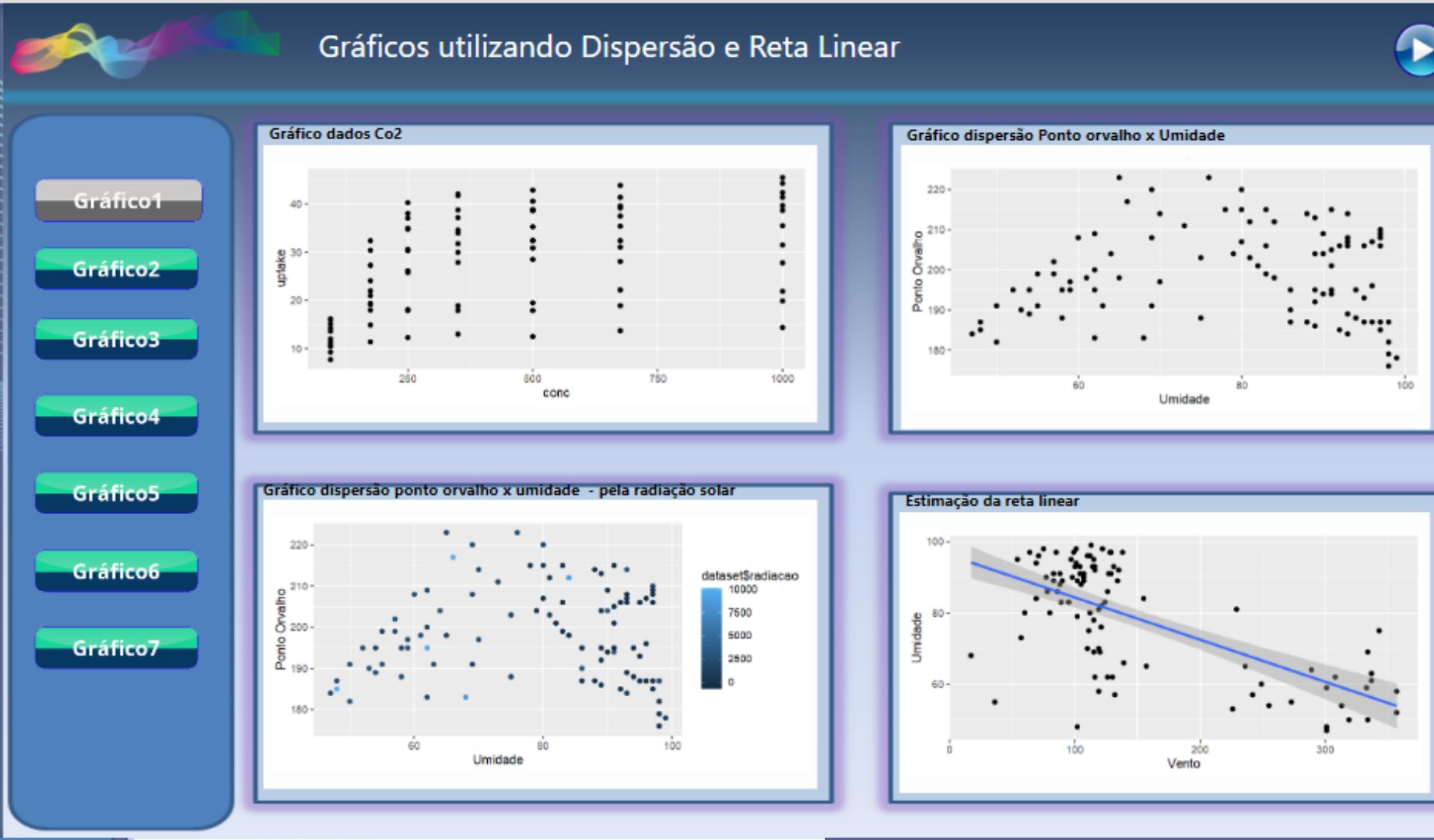


# POWER BI

## GRÁFICOS E DASHBOARDS

Uma das mais conhecidas ferramentas do mercado, neste curso diversas formas de uso do Power Query, uma tecnologia para conectar dados, além da construção de gráficos e painéis. Há uma bateria de exercícios para você testar o seu conhecimento, além de uma apostila para você ler e utilizar quando quiser.





# POWER BI: 100% gráficos criados em R

## GRÁFICOS E DASHBOARDS

Este treinamento é um grande avanço para os profissionais e alunos que trabalham com POWER BI e a linguagem R, pois permite que aprendizados adquiridos em ambas as ferramentas possam ser incorporados juntos, integrando os conhecimentos.

Será criado um projeto completo de dados, com 100% dos gráficos na linguagem R dentro do POWER BI. É um dos primeiros passos para você adentrar no mundo de analytics e Machine Learning.



The screenshot shows the Power BI desktop application. At the top, there's a ribbon with various tabs like 'Gerenciar Relações', 'Nova Medida', 'Nova Coluna', 'Nova Tabela', etc. Below the ribbon, there are two data tables. The left table has columns 'DATA', 'VALOR VENDA', and 'NEXTDAY'. The right table has columns 'MES', 'VALOR VENDA', and 'NEXTMONTH'. In the center, there's a small preview of a chart. At the bottom, the DAX formula bar is visible with the text 'Retorna uma tabela com novas colunas especificadas pelas expressões DAX.' and several DAX functions listed: ADDCOLUMNS, ADDROWSITEMS, ALL, ALLEXCEPT, ALLNOBLANKROW, ALLSELECTED, CALCULATETABLE, CALENDAR, CALENDARAUTO, CROSSJOIN, DATATABLE, and others.

This screenshot shows the Power BI DAX formula editor. A context menu is open over a selected item named 'Métodos DAX'. The menu includes options like 'Nome' (Name), 'Modo de armazenamento' (Storage mode), and 'Última atualização' (Last update). The background shows a list of DAX functions and measures, such as 'ACUMULADO...', 'DATESMTD', 'DATESQTD', 'DATESYTD', 'endofmonth', 'ENTRE DATAS', 'Firstdate', 'Lastdate', 'NEXTDAY', 'NEXTMONTH', 'NEXTQUARTER', 'NEXTYEAR', 'PREVIOUSDAY', 'PREVIOUSMONTH', 'PREVIOUSQUARTER', 'PREVIOUSYEAR', 'PROXIMO MES', 'startofmonth', 'VENDAS DO A...', 'VENDAS DO ...', and 'VENDAS TRIM...'. On the right side of the slide, there's a decorative graphic featuring a monitor with charts, a person working at a desk, and a large circular icon with a plus sign.

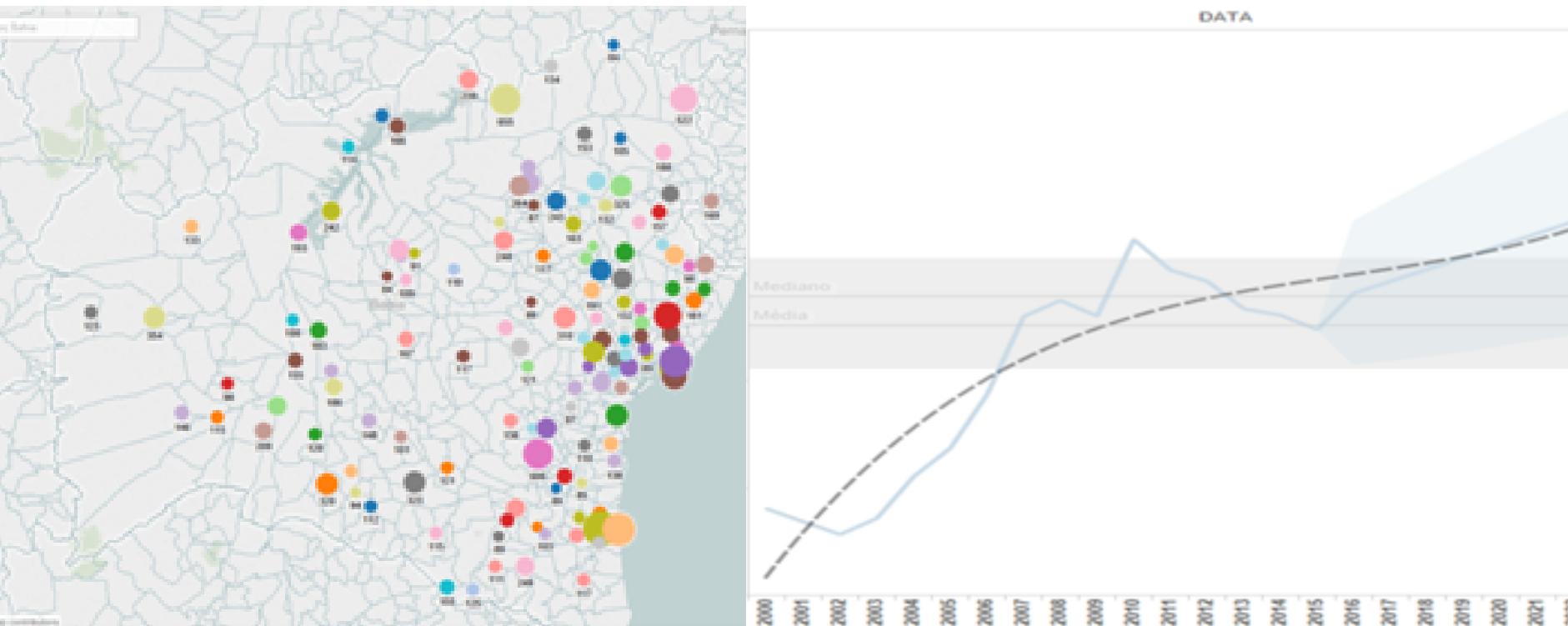
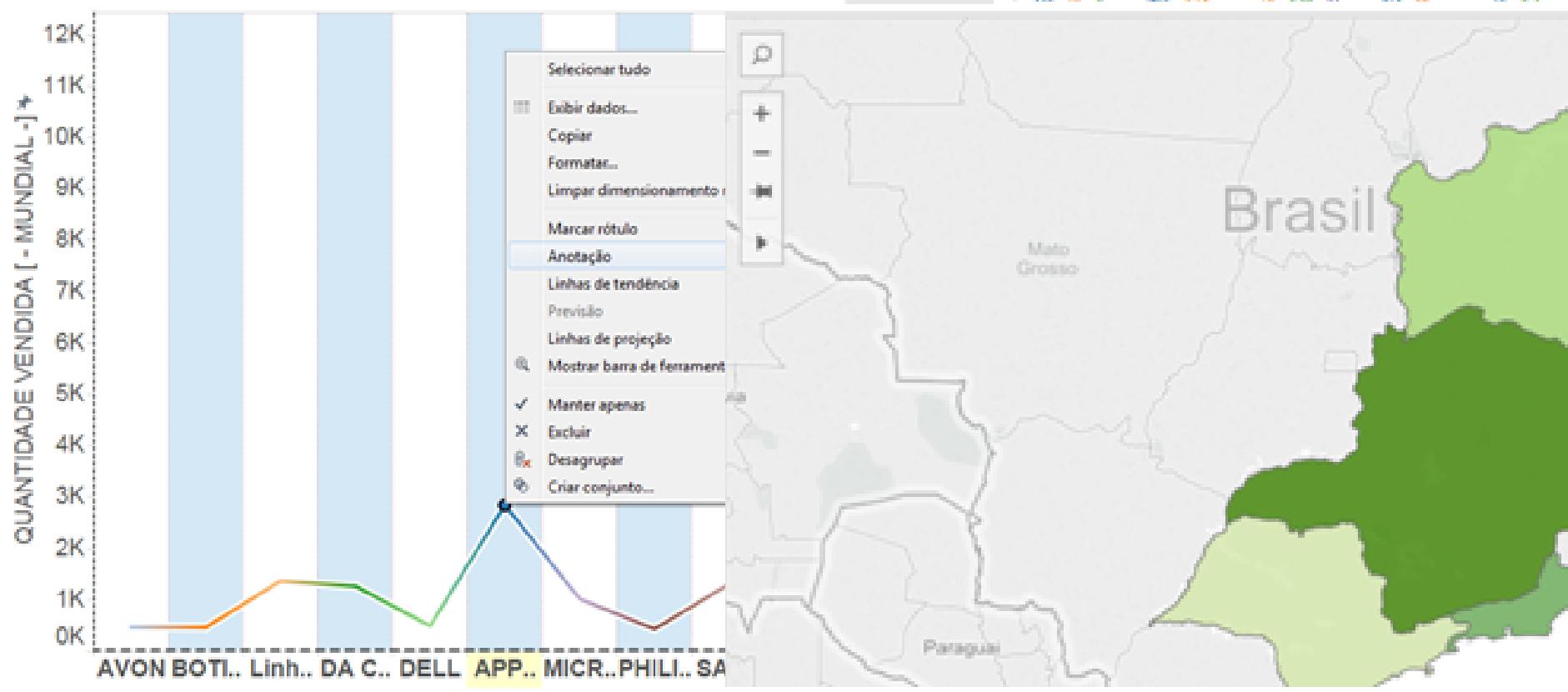
# POWER BI 65

## DAX

### GRÁFICOS E DASHBOARDS

O POWER BI possui uma linguagem que permite construir medidas e colunas por meio de funções, operadores e constantes que podem ser usados em uma fórmula ou expressão, para calcular e retornar um ou mais valores. Neste treinamento você terá a chance de conhecer diversas situações e como aplicá-las.

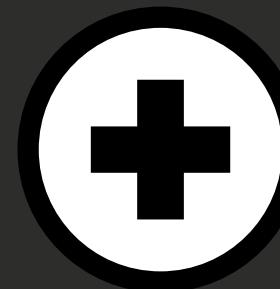
Ano de Dtc Venda	Trimestre de Dtc Venda	APPLE					AVON					
		Benito de Paula	Edson Arantes	João do Pulo	Reginaldo Rossi	Renato Russo	Tom Jobim	Benito de Paula	Edson Arantes	João do Pulo	Reginaldo Rossi	Renato Russo
2014	T1	8 8 10 19	26 34 15	31 2 5 5	29 11 8 4 21	15 18 31 3 4	4 5 16 21	29 23 24	6 40 12 8	19 15 20	10 2 8 9	31 30 10 16
		✓ Manter apenas	X Excluir									
		1657 itens selecionados - SOMA(Qtd Venda): 39.114										
2014	T4	16 31 16 25	19 25 12 12	10 20 16 5 9	15 24 6 4 9	6 4 19 15 ...	21 17 25 ...	14 30 15 ...	40 4 24 9	10 13 10 22	8 3 10 15 24	4 31 11 5
		10 9 2 15 9	49 24 13 ...	8 26 8 13 ...	15 6 11 5	9 27						

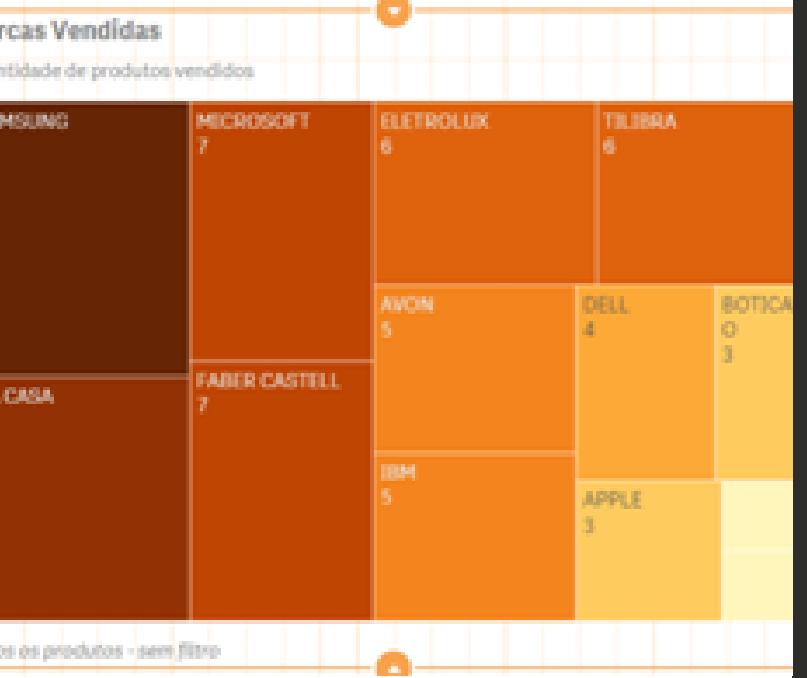
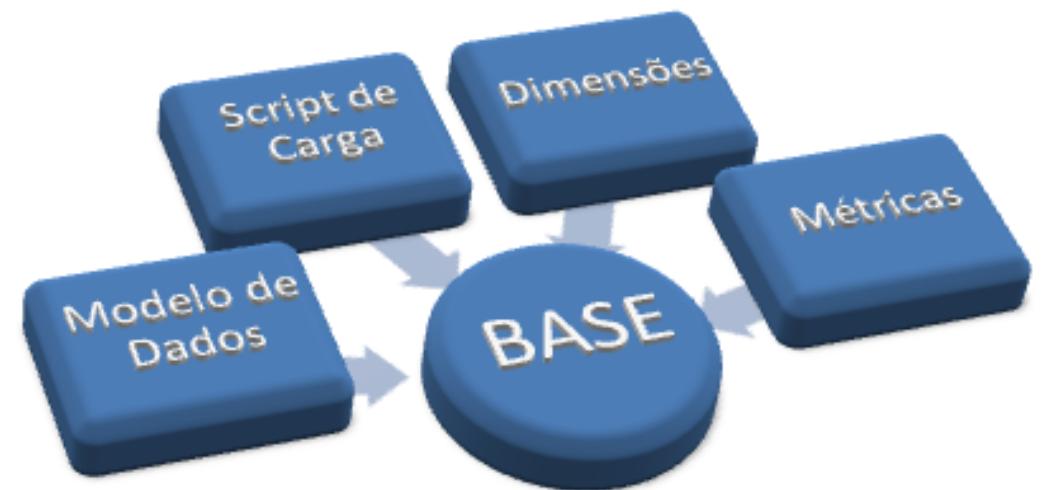


# TABLEAU

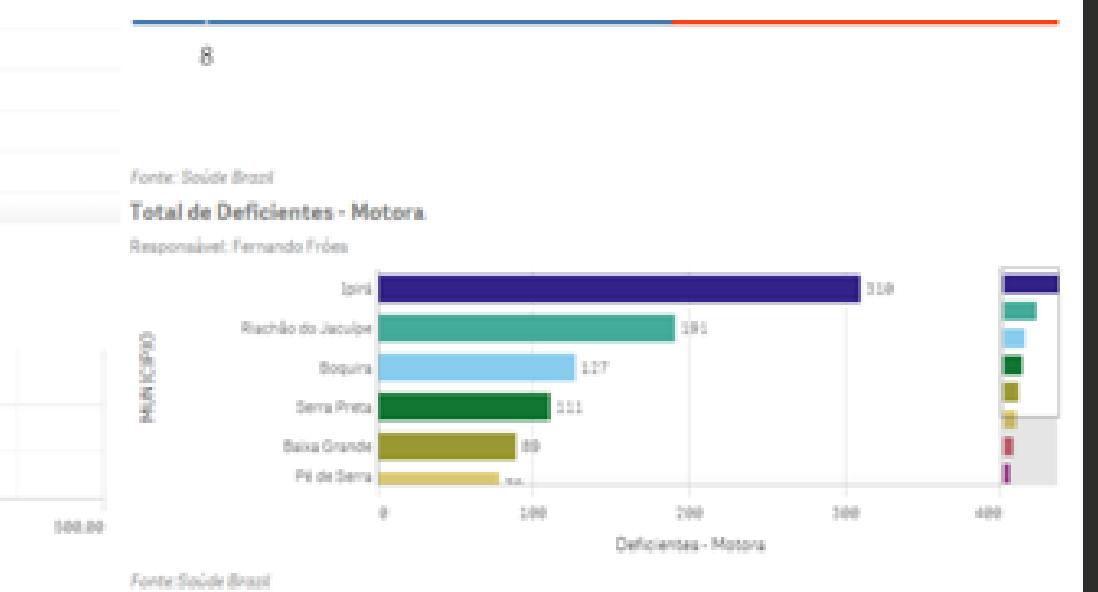
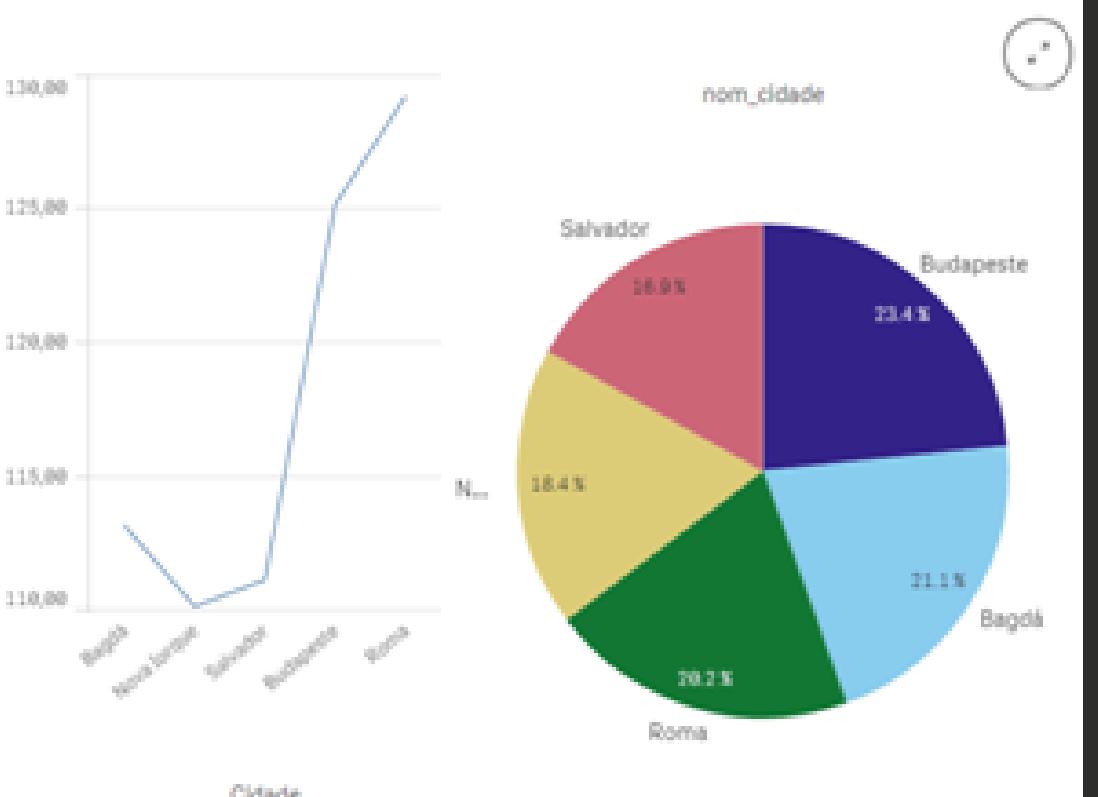
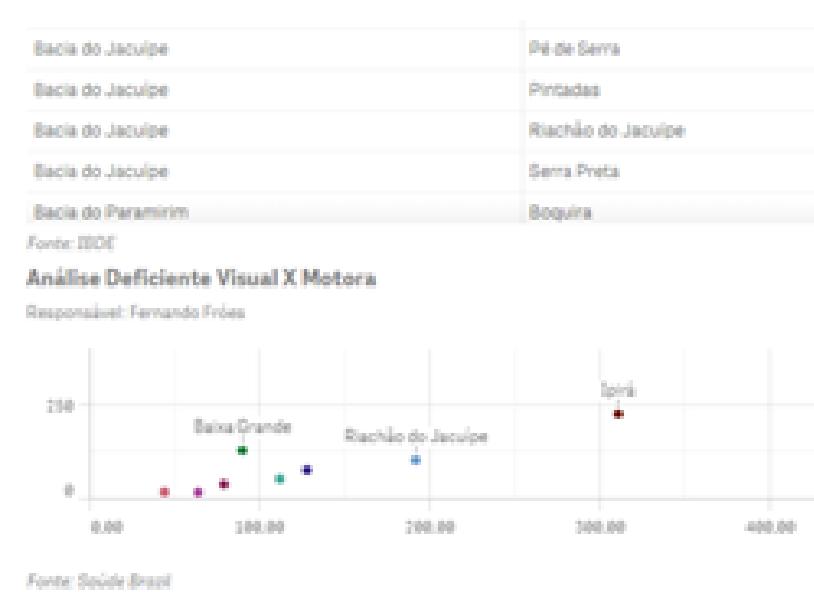
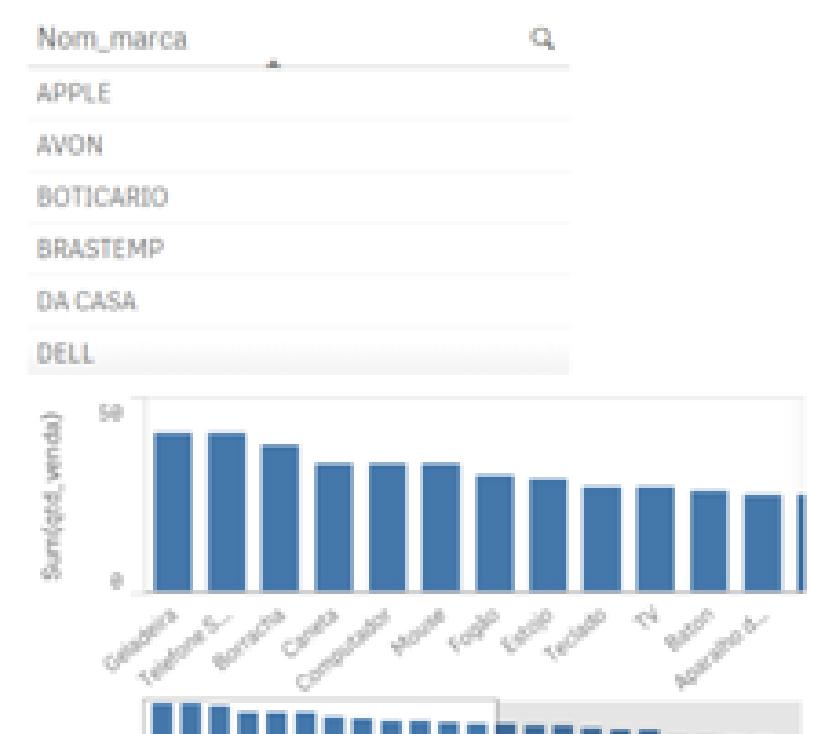
## ELEGÂNCIA GRÁFICA

É uma ferramenta de última geração denominada Data Discovery, que trabalha com consulta de bancos de dados relacionais, cubos OLAP, bancos de dados da nuvem e planilhas, gerando vários tipos de gráficos. O TABLEAU é mundialmente conhecido, já presente nas grandes corporações. Neste treinamento abordamos desde o carregamento dos dados até a construção da chamada história com os dados. E tem uma super apostila também.





Apuração das vendas



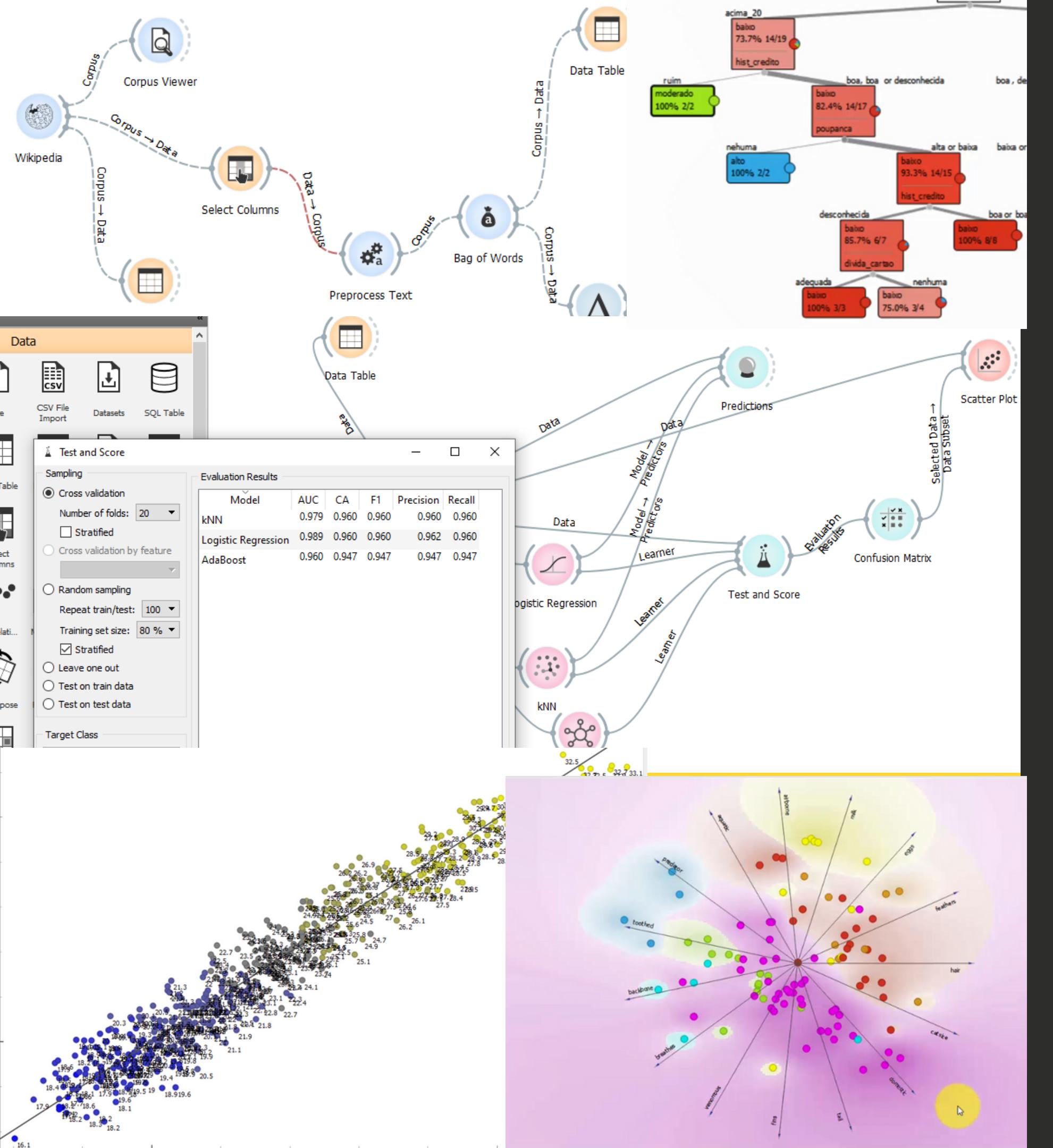
# QLIK SENSE

## ÚNICA COM MÉTODO ASSOCIATIVO DOS DADOS

Você pode compartilhar análises relevantes, incluindo aplicativos e histórias, através de um hub centralizado. É possível unificar fontes de dados para uma visão completa da informação, sem lacunas, e facilita a descoberta de informações ocultas. Utiliza interfaces simples de arrastar e soltar para criar visualizações de dados flexíveis e interativas.

E tem uma super apostila também.

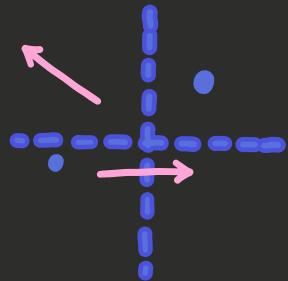


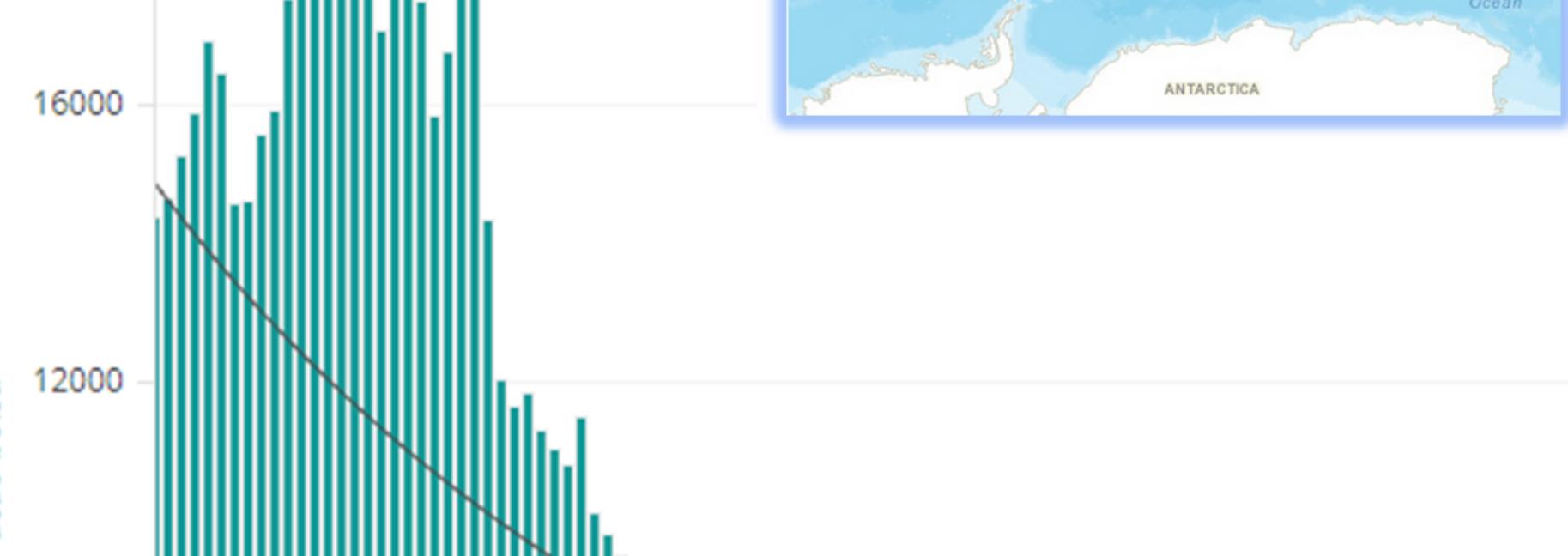
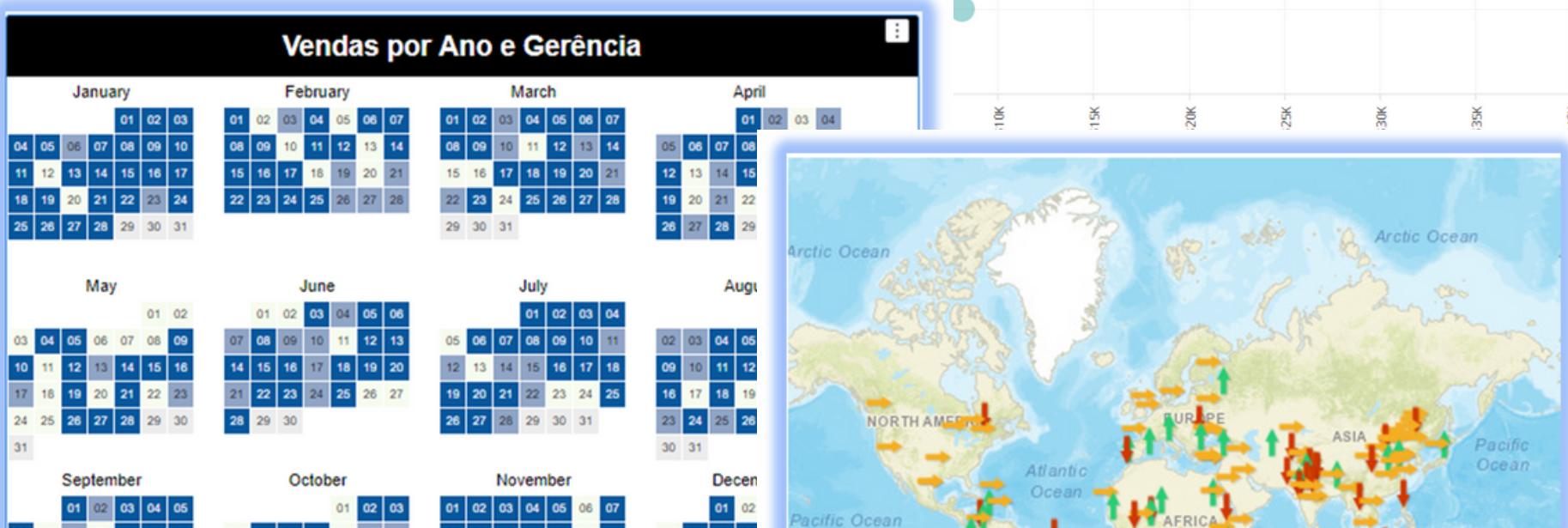
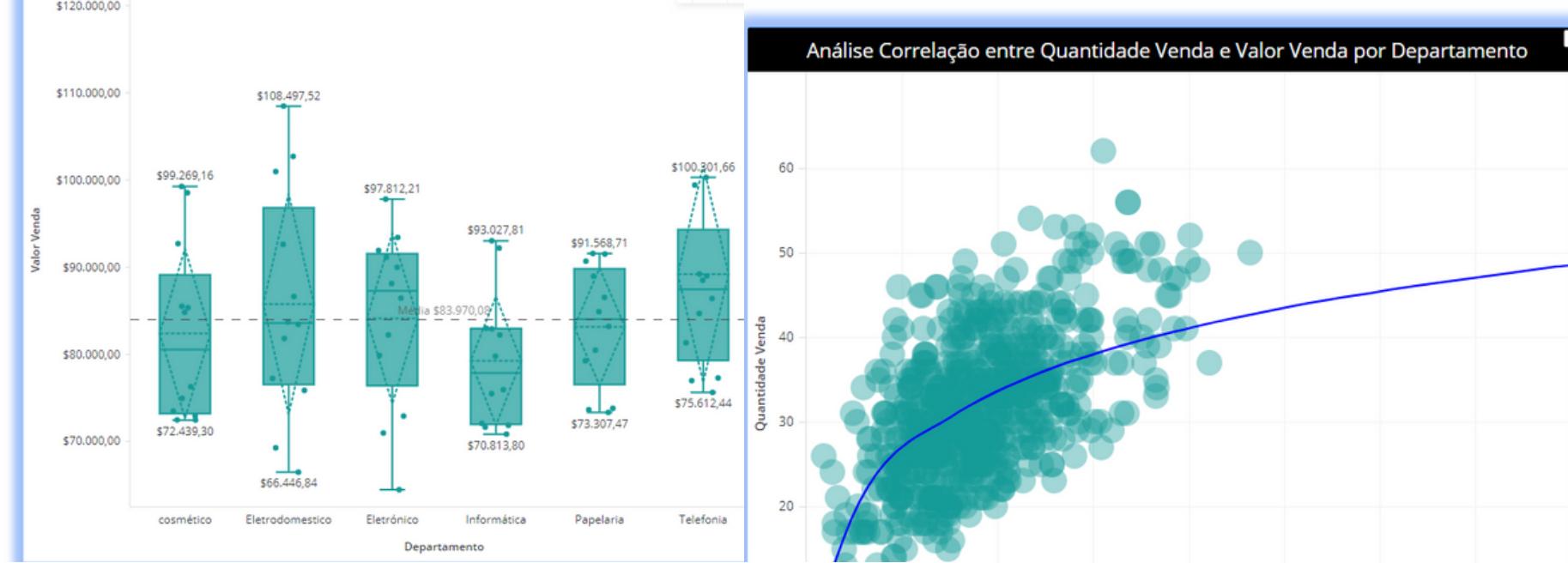


# ORANGE DATA SCIENCE

## MACHINE LEARNING GRAFICAMENTE

O ORANGE é uma das poucas ferramentas de mercado totalmente construída em python, que o analista de dados pode trabalhar de forma totalmente visual, com um amplo aspecto de atendimento a diversas demandas na área de mineração de dados por meio do uso de algoritmos de Machine Learning. Neste treinamento a prática é 90% de todo o curso.





# MICROSTRATEGY

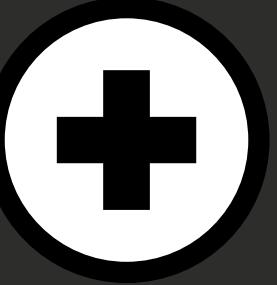
## STORYTELLING DOS SEUS DADOS

A MicroStrategy é uma das principais ferramentas do mercado a construir este caminho juntamente com você. Tem características ímpares que outras ferramentas não possui. Embarcada com uma inteligência que analisa os dados carregados e disponibiliza para você diversas opções de análise, por meio da chamada linguagem natural. Você pode utilizá-la ou criar perguntas que sejam respondidas, tudo isso em um clique. Este treinamento permitirá que você construa grandes gráficos e dashboards.



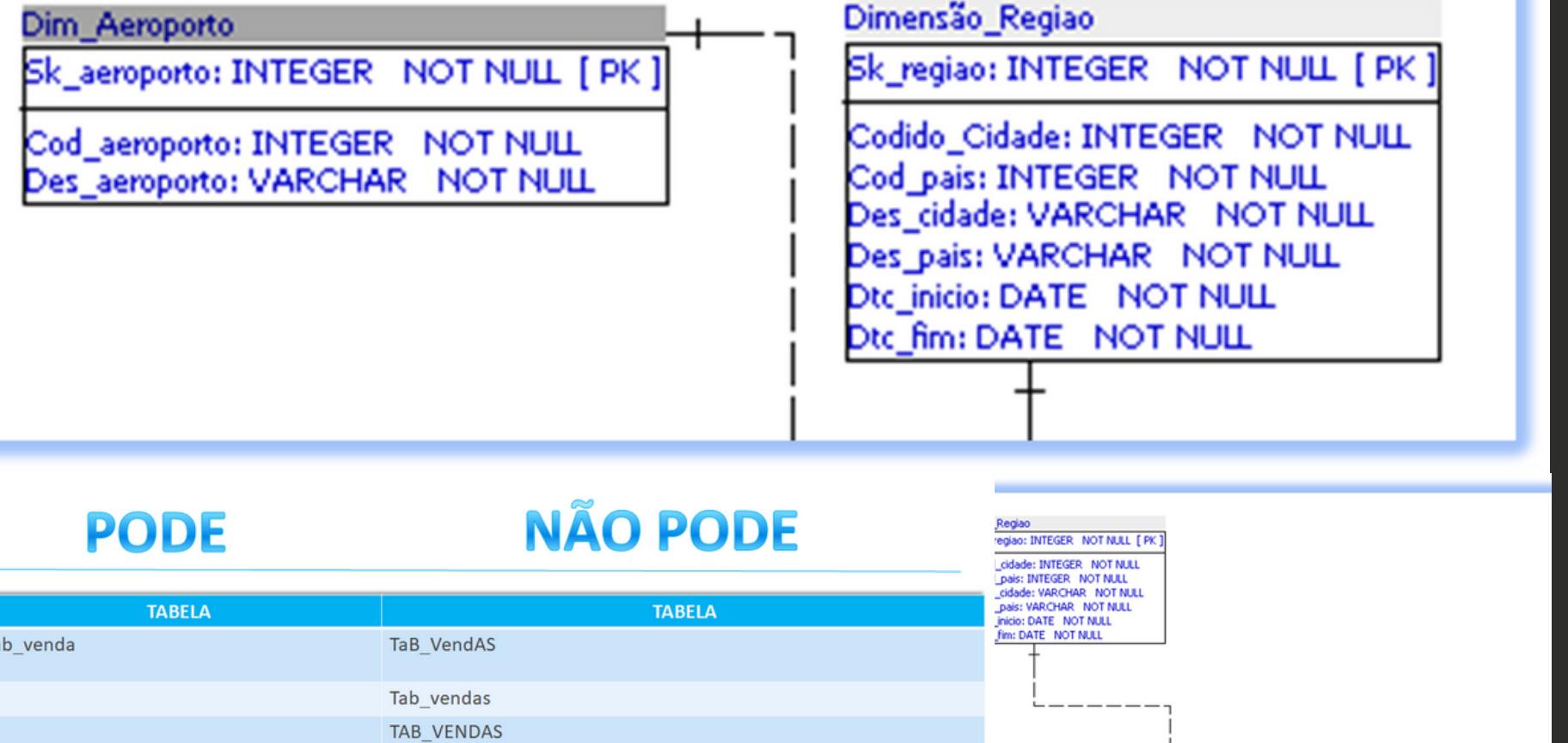


# PADRÕES EM BANCO DE DADOS

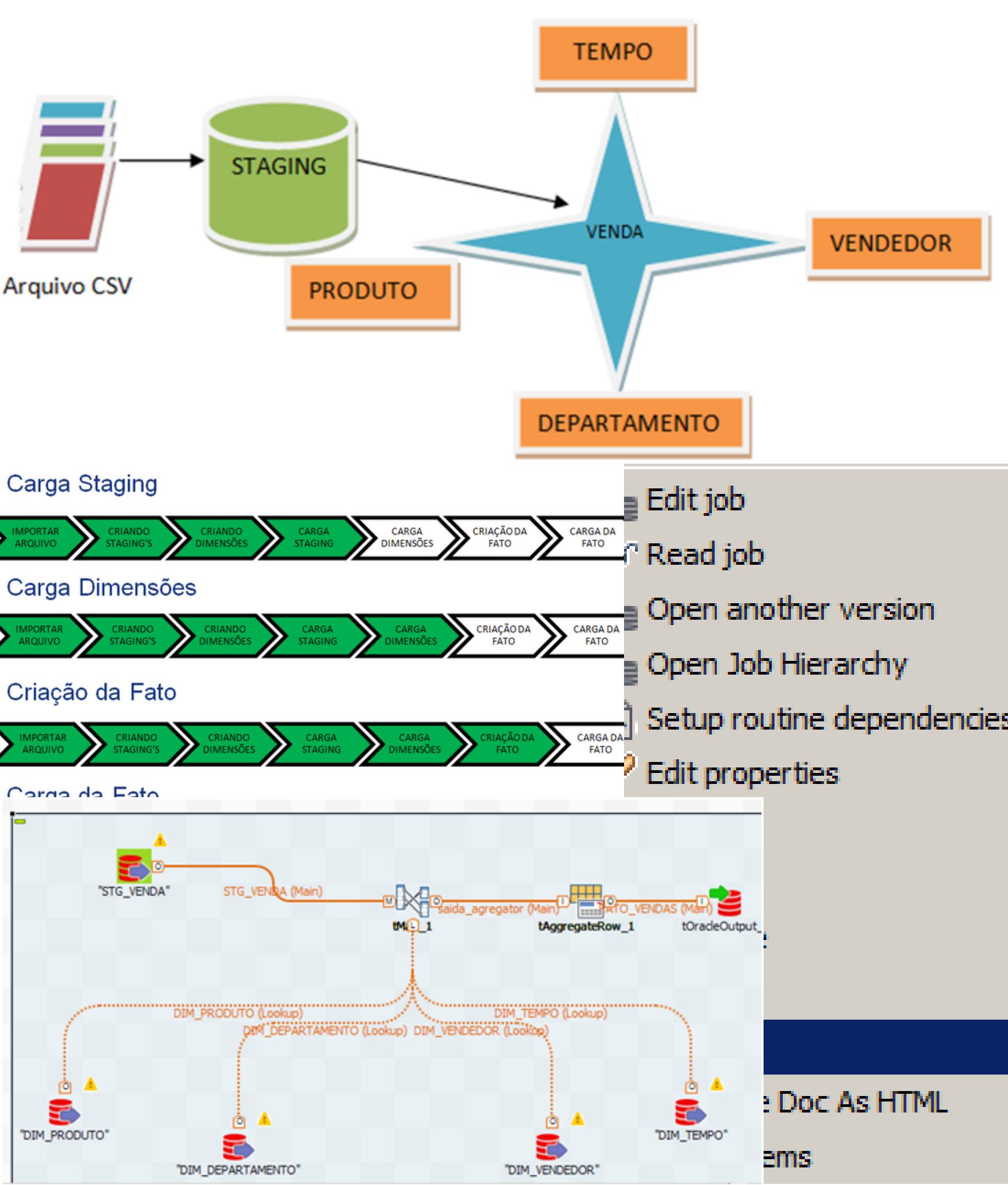


## PADRONIZANDO SUAS BASES

Muitos projetos na área de dados estão sendo criados em grande escala nas empresas de todo o mundo, o que não percebemos é que toda esta estrutura, em grande parte, diz respeito a guarda dos dados em bases de diversos formatos, seja em bancos transacionais, NoSql, tabulares ou multidimensionais. Neste treinamento será possível estruturar: nomenclatura das tabelas, padrão de nomenclaturas dos atributos, garantia de que novos objetos ao serem construídos sigam os padrões da organização, dentre outros assuntos.



ELA_ORI	CONSTRAITS_ORIGEM	COLUNA_ORIGEM	TIPO	NULIDADE	ESCALA	PREC
CEDOR	SYS_C0009585	COD_FORNECEDOR	NUMBER	Y	0	12
CEDOR	SYS_C0009768	DES_FORNECEDOR	VARCHAR	N	0	30
CEDOR	SYS_C0009354	DES_ENDERECHO	VARCHAR	N	0	15
CEDOR	SYS_C0009932	NUM_GNDI	VARCHAR	N	0	12



# ETL TALEND

## EXTRAÇÃO DE DADOS

O Talend Data Integration é uma solução aberta e escalável de integração de dados e qualidade de dados para integrar, limpar e perfilar todos e qualquer tipo de dado. Possui mais de 900 componentes pré-construídos. Trabalha com armazenamento em nuvem, integração de dados, gerenciamento de dados , gerenciamento de dados mestre, qualidade de dados, preparação de dados e software e serviços de integração empresarial. A ferramenta é open source e uma das mais completas do mundo. Na área de Data Warehouse (DW), ela é utilizada nas mais diversas empresas dos ramos governamentais, financeiro, dentre outros.



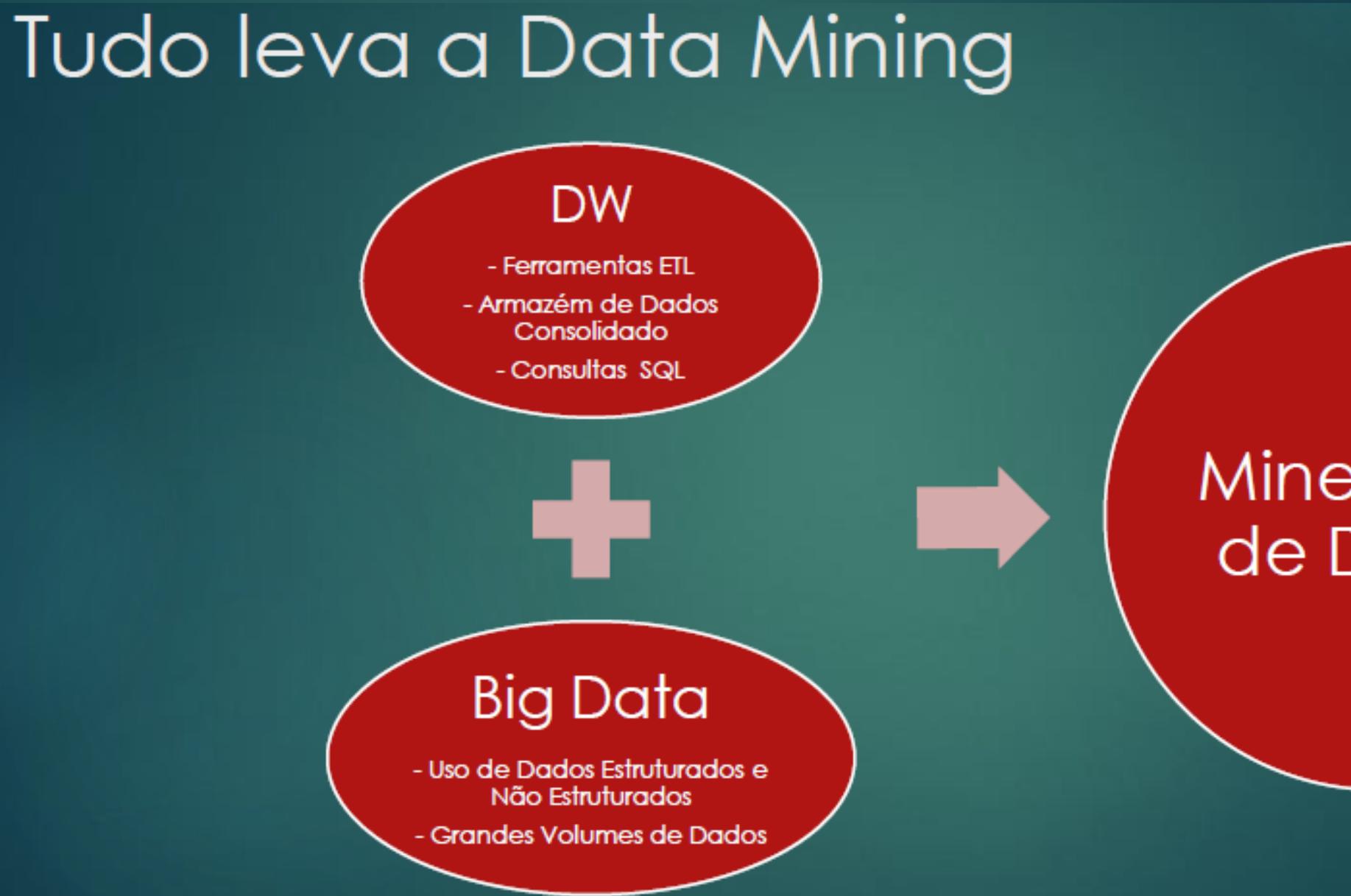
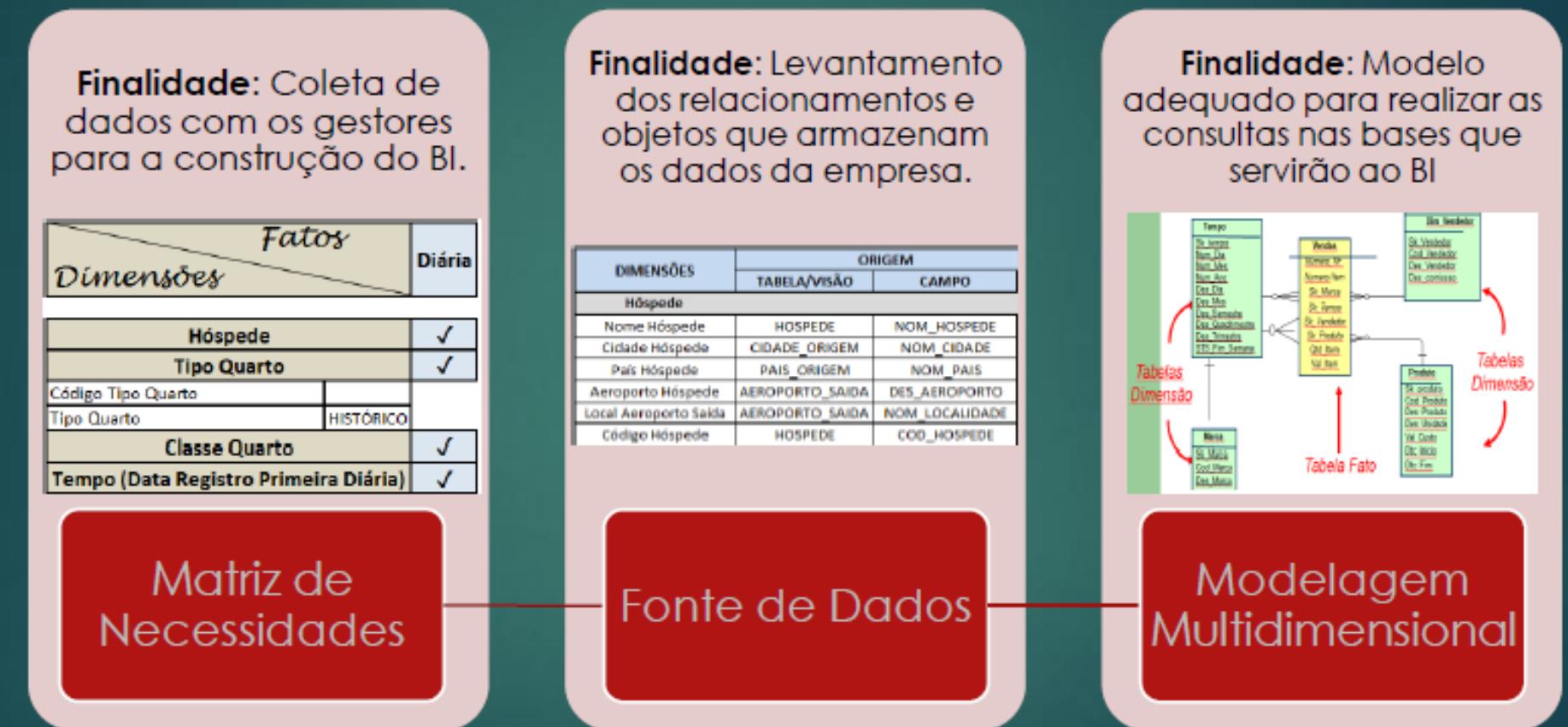


# BUSINESS INTELLIGENCE

## CONCEITOS E APLICAÇÕES

Há um portfólio de cursos que você poderá aprender sobre Business Intelligence, destaco 2 deles:

- Resumo BI: um resumo completo e objetivo sobre os fundamentos de Business Intelligence (BI), o que é BI, Data Warehouse, OLTP, OLAP, metadados, Modelagem Dimensional, Data Mining, Big Data.
- Business Intelligence - as etapas para um BI Perfeito: trabalhar com dois documentos essenciais (e inéditos) para o início do projeto de DW: Matriz de Necessidade e Fonte de Dados.



# Os dados são o novo petróleo



FRASE QUE AGITOU O MUNDO DOS  
NEGÓCIOS FOI CRIADA POR CLIVE HUMBY,  
UM MATEMÁTICO LONDRINO  
ESPECIALIZADO EM CIÊNCIA DE DADOS

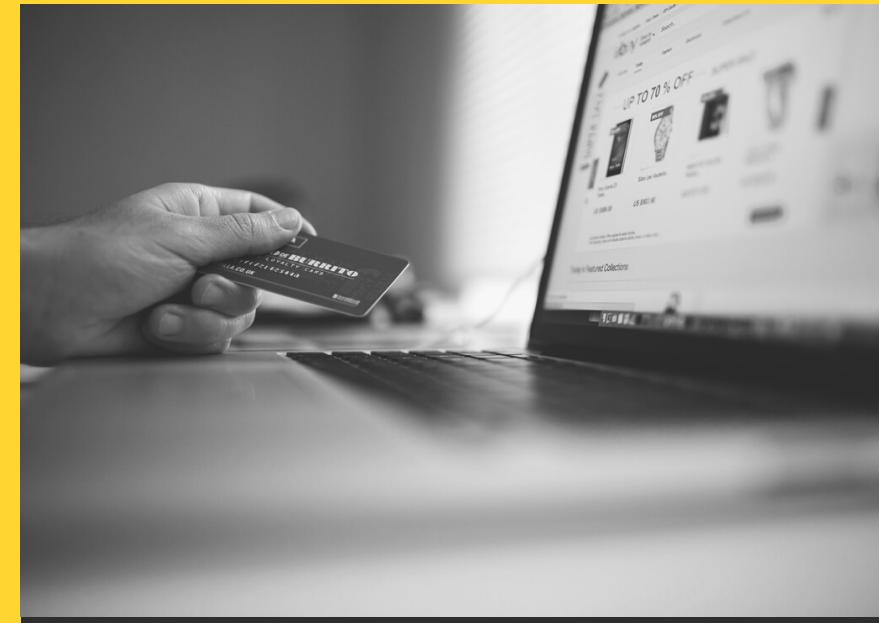
# ONDE ENCONTRO OS CURSOS?



## DIRETAMENTE

Você pode acessar diretamente todos os meus cursos.

Link: [TODOS OS CURSOS](#)



## SITE DE CUPONS

Aqui mantendo regularmente promoções de todos os cursos que são constantemente atualizados, inclusive com cursos gratuitos.

Link: [CUPONS DESCONTO](#)



## CONVERSANDO

Você pode escrever para mim, podemos conversar sobre os cursos que deve realizar.

E-mail: [grimaldo\\_lopes@hotmail.com](mailto:grimaldo_lopes@hotmail.com)



FAÇA A SUA HISTÓRIA NA  
ÁREA DE DADOS