

Qlik Sense

Aprenda a desenvolver suas
próprias aplicações BI

Jair Carrion

Qlik  LEAD WITH DATA™

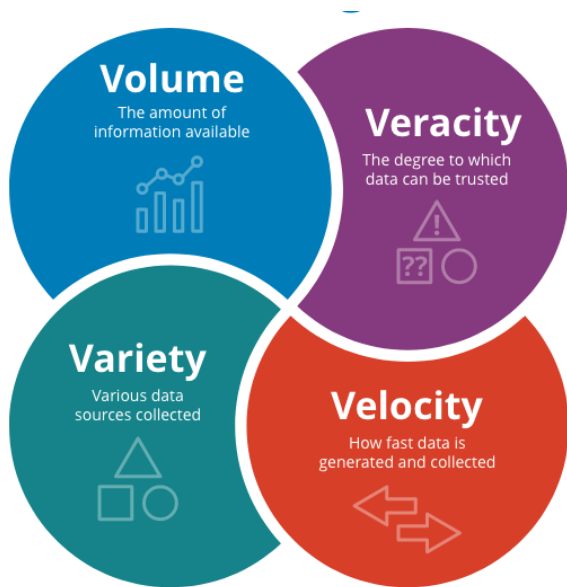


Agenda

- Big Data;
- O que é BI ?;
- BI vs Data Science;
- Diferenças entre DW, DL e DM;
- Diferencias do Qlik Sense;
- Modelo de desenvolvimento (Kanban);
- ETL;
- DER;
- Modelagem dimensional;
- OLTP x OLAP;
- Arquitetura do BI (que vamos desenvolver);
- Desenvolvimento das dashboard.

Big Data

De acordo com um estudo realizado pela IBM, existe uma estimativa de que sejam gerados 2,3 trilhões de Gigabytes por dia. (Estimasse que em 2025 seja 150 trilhões de gigabytes)



4 principais V's do Big Data

- **Volume:** Quantidade de um amontoado de informações disponíveis;
- **Velocidade:** Velocidade em que os dados são gerados e coletados;
- **Variedade:** Os vários tipos de formato de dados na qual estão disponibilizados;
- **Veracidade:** A acurácia com que a informação esteja correta.

— ANALYTICS

APPLICATIONS – ENTERPRISE

SALES

MARKETING - B2B

MARKETING - B2C

CUSTOMER EXPERIENCE / SERVICE

HUMAN CAPITAL

LEGAL

RETECH & COMPLIANCE

FINANCE

AUTOMATION & RPA

SECURITY

PARTNERSHIPS

APPLICATIONS - INDUSTRY

ADVERTISING

EDUCATION

REAL ESTATE

GOVT & INTELLIGENCE

COMMERCE

FINANCE - LENDING

INSURANCE

HEALTHCARE

LIFE SCIENCES

TRANSPORTATION

AGRICULTURE

INDUSTRIAL

OTHER

The chart displays a wide array of open-source tools and frameworks, organized into 14 categories. Each category is represented by a colored header and a row of logos. The categories and their associated tools are as follows:

- FRAMEWORKS**: TensorFlow, PyTorch, Keras, PySpark, etc.
- FORMAT**: Avro, Parquet, ORC, etc.
- QUERY / DATA FLOW**: Spark, Flink, Tez, etc.
- DATA ACCESS**: Uber, etc.
- DATABASES**: PostgreSQL, MySQL, Oracle, etc.
- ORCHESTRATION**: Apache Airflow, Prefect, etc.
- INFRA-STRUCTURE**: Kubernetes, Docker, etc.
- DATA OPS**: Databricks, etc.
- STREAMING & MESSAGING**: Apache Kafka, etc.
- START TOOLS & LANGUAGES**: Jupyter, etc.
- ML OPS & INFRA**: MLflow, etc.
- AI / MACHINE LEARNING / DEEP LEARNING**: TensorFlow, PyTorch, etc.
- SEARCH**: Elasticsearch, etc.
- LOGGING & MONITORING**: Prometheus, etc.
- VISUALIZATION**: Tableau, etc.
- COLLABORATION**: etc.
- SECURITY**: etc.

DATA RESOURCES

The banner displays a wide array of startup logos, organized into nine distinct categories. The categories and their respective logos are as follows:

- DATA MARKETPLACES & DISCOVERY:** Includes logos for DataCamp, Kaggle, and others.
- FINANCIAL & ECONOMIC DATA:** Includes logos for Bloomberg, Thomson Reuters, Dow Jones, and others.
- AIR / SPACE / SEA:** Includes logos for SpaceX, Virgin Galactic, and others.
- PEOPLE / ENTITIES:** Includes logos for Zillow, Indeed, and others.
- LOCATION INTELLIGENCE:** Includes logos for Foursquare, Mapbox, and others.
- OTHER:** Includes logos for various miscellaneous startups.
- DATA SERVICES:** Includes logos for DataRobot, H2O, and others.
- INCUBATORS & SCHOOLS:** Includes logos for Y Combinator, Techstars, and others.
- RESEARCH:** Includes logos for OpenAI, Google Research, Facebook Research, and others.

O que é BI (Business Intelligence)

Analytics e business intelligence (ABI) é um conjunto de ferramentas, técnicas, infraestrutura e melhores práticas, possibilitando a análise de informação para suporte na tomada de decisões.

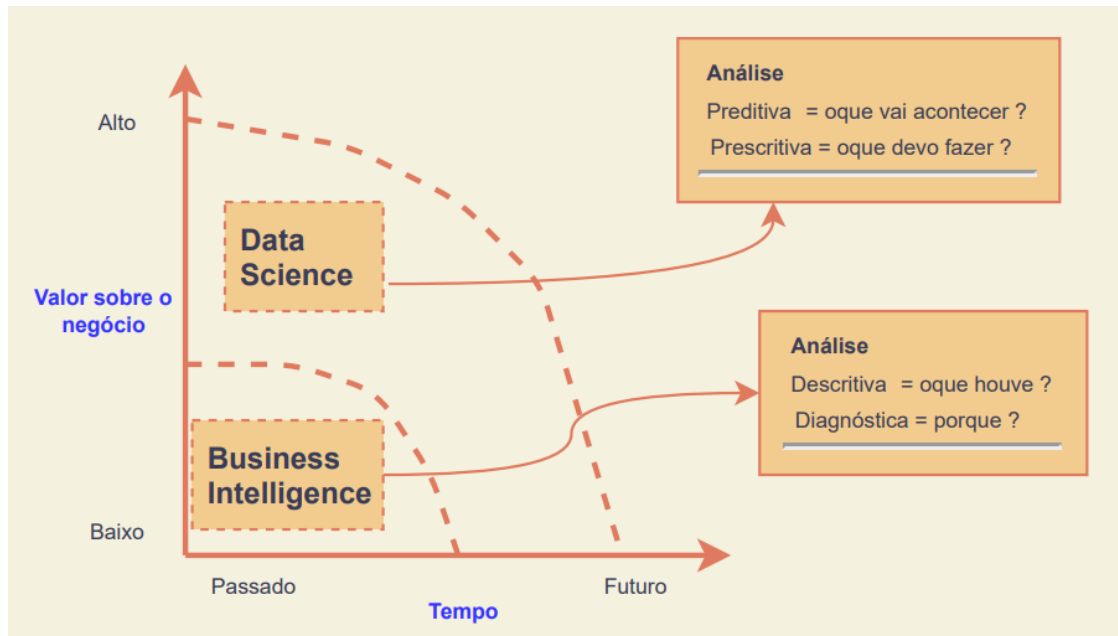


Fonte: GARTNER

<https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi>

Business Intelligence x Data Science

BI e Data Science possui diferenças ?



DW x DL x DM

Quais as diferenças desses conceitos ?

DW (Data Warehouse)



Estamos preocupados com o armazenamento e a recuperação dos dados. (Schema definido na gravação).

Usamos o ETL

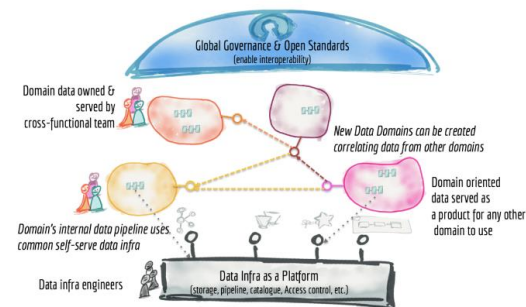
DL (Data Lake)



A preocupação maior é com o armazenamento em si. (Schema definido na leitura).

Usamos o ELT

DM (Data Mesh)



Democratização dos dados.

Diferenciais do Qlik Sense



Visualização self-service



Dashboards interativos



Pesquisa e conversacional



Alertas e ação



Relatórios



Mobilidade

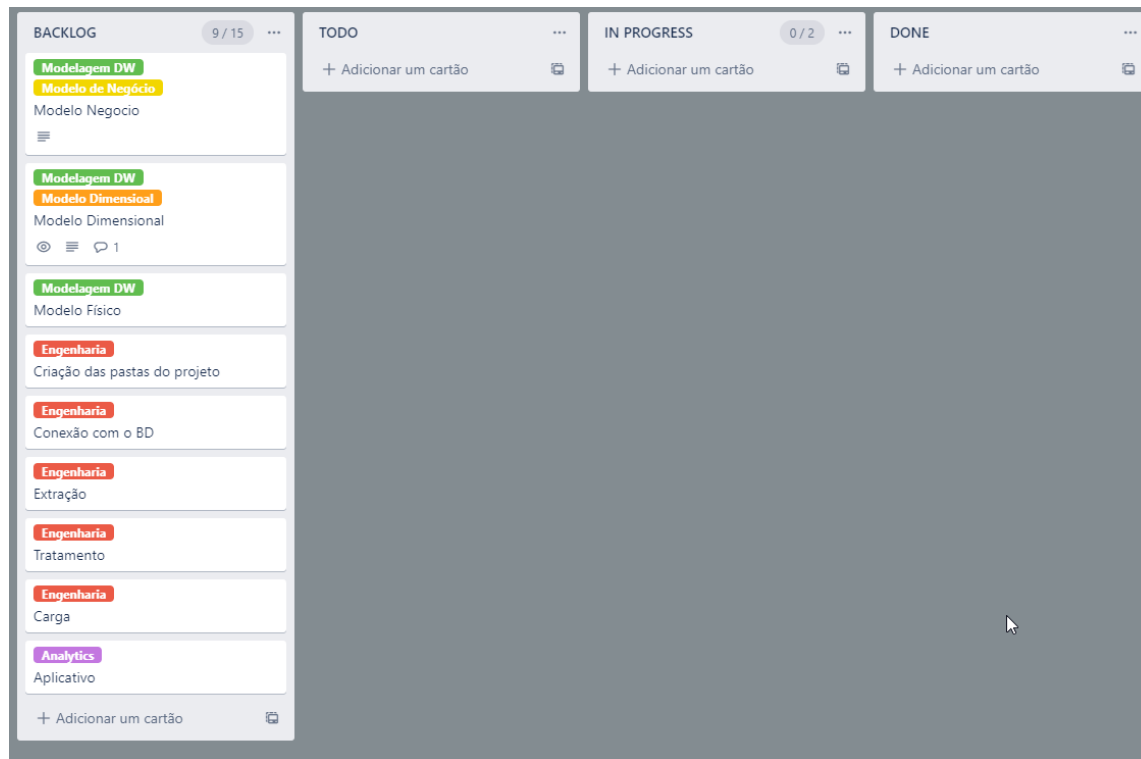


Personalizada e incorporada



Analytics avançado

Modelo de desenvolvimento (Kanban)



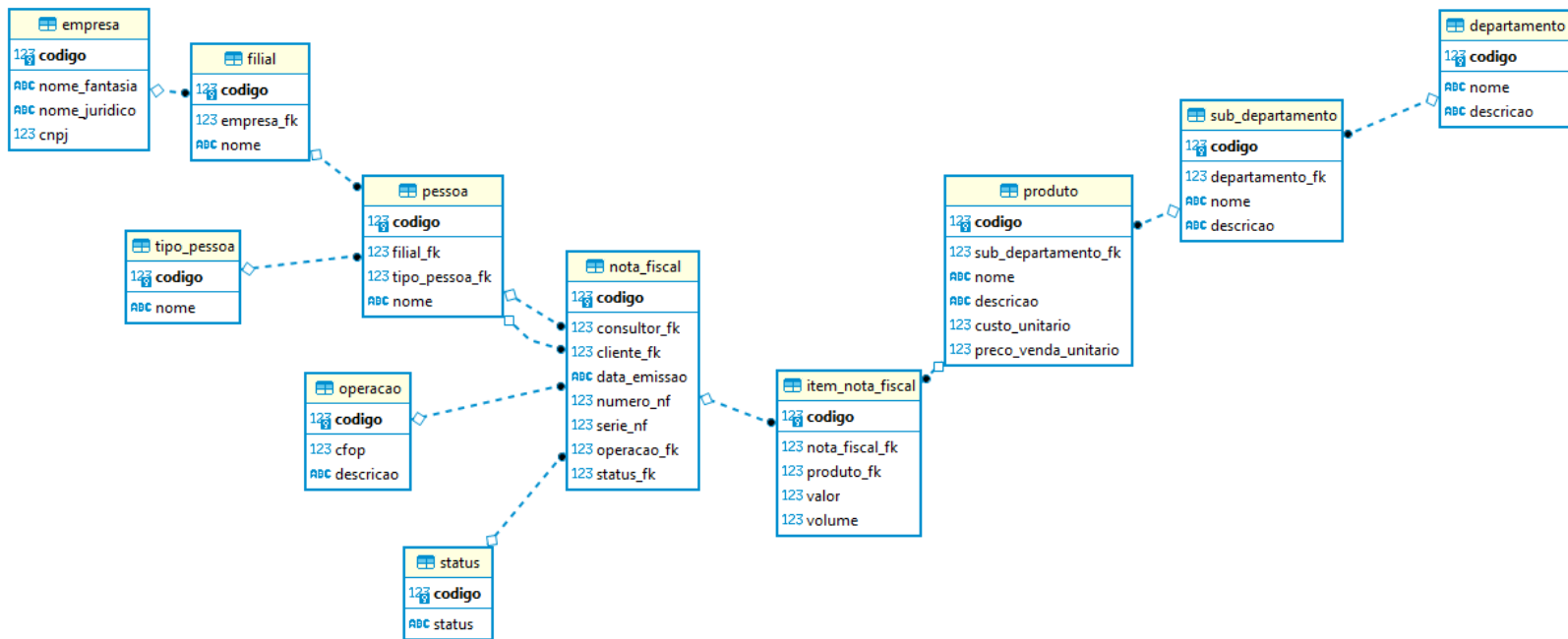
ETL

Extração (Extract), Tratamento (Transform) e Carga (Load)

- Extração;
 - Entender e realizar a leitura dos dados na fonte, copiando os dados necessarios para futura manipulação – tratamento.
- Tratamento;
 - Após a coleta dos dados a partir da ferramenta de ETL, poderá ser necessario realizar varios tipos de transformações, ou seja, limpeza nos dados (corrigir valores nulos, arrumar formatação) combinar varias fontes de dados, corrigir valores duplicados, etc.
- Carga;
 - Na etapa final do processo de ETL, é o processo de modelagem fisica e ingestão dos dados na fonte de destino.

DER (Diagrama Entidade Relacionamento)

Como dados são armazenados dentro de um ERP ?

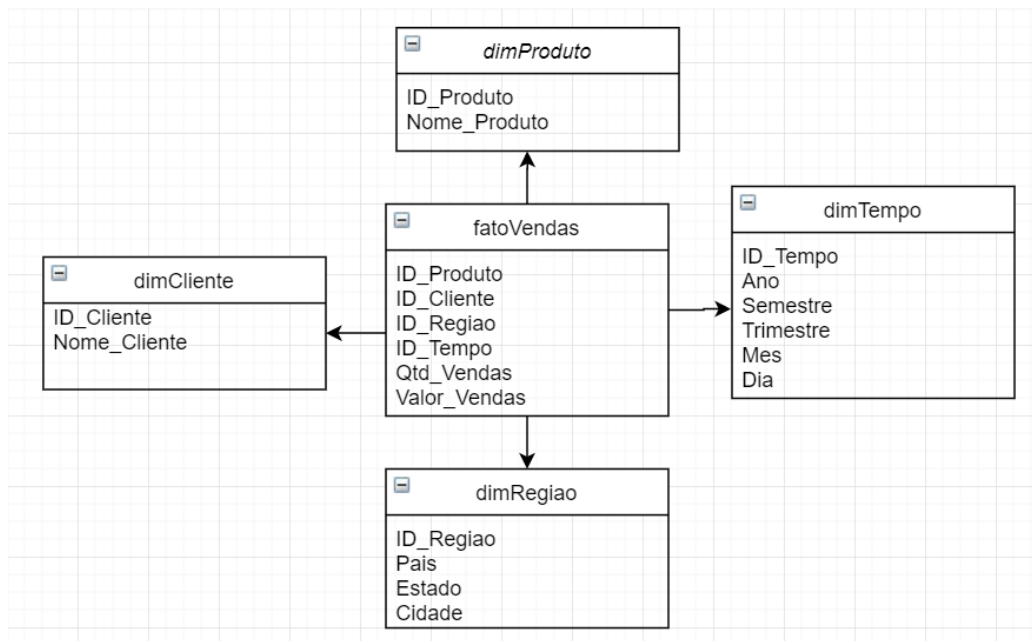


Exemplo DER: Modelo de Banco de Dados que estaremos utilizando no projeto

Modelagem Dimensional

Modelo Estrela

- Dimensões;
 - Nome Cliente;
 - Cidade;
 - Endereço;
 - Nome do produto;
 - Etc..
- Fatos;
 - Numero da Nota Fiscal;
 - Valor (comercializado);
 - Quantidade;



OLTP x OLAP

OLTP (Online Transaction Processing)

É um sistema de transações, onde o principal objetivo é na gravação.



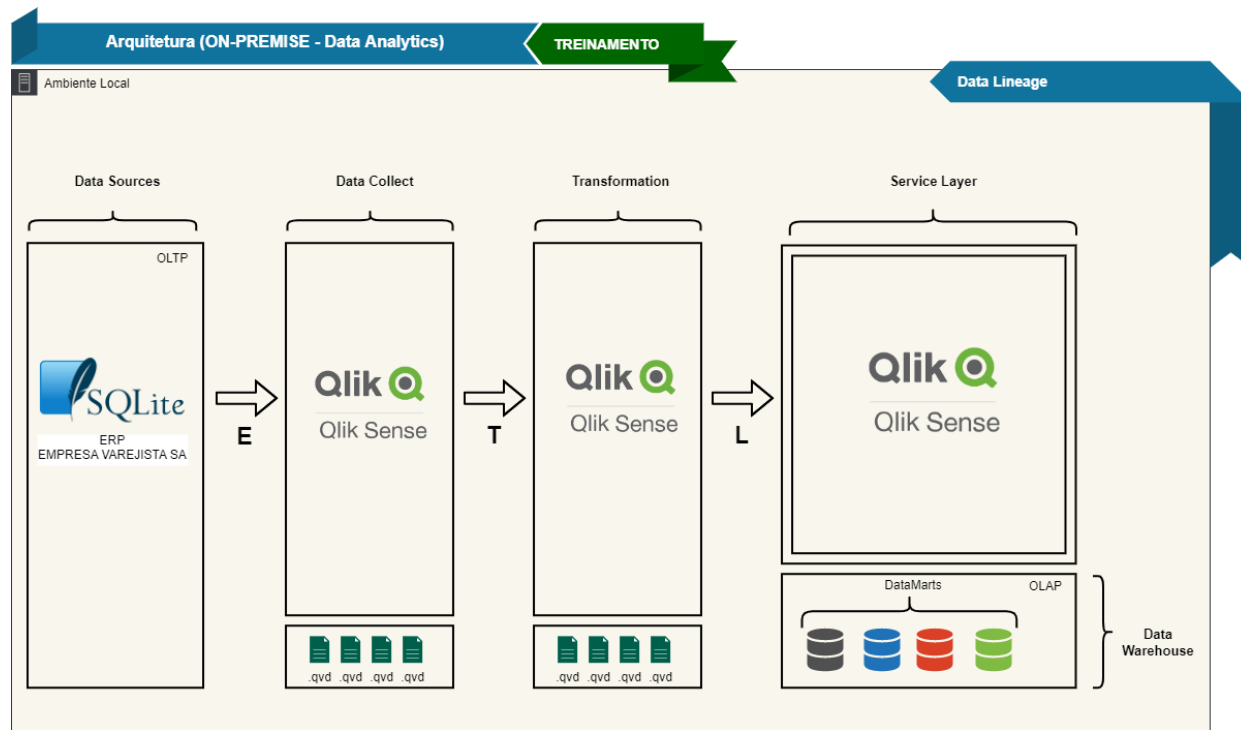
OLAP (Online Analytical Processing)

É um sistema de processamento analítico, que por sua vez, têm como objetivo na velocidade de consultas analíticas multidimensionais.



Arquitetura do BI

Modelo de arquitetura a ser desenvolvido no projeto



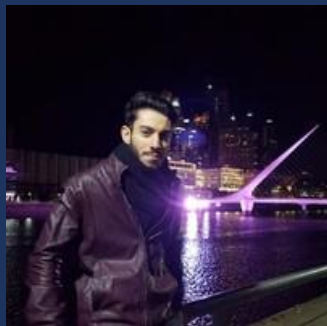
Dashboards

Sempre tentando responder os 5W2H

- O que ?;
- Quem ?;
- Quando ?;
- Onde ?;
- Por quê ?;
- Como ?;
- Quanto ?.



Apresentador



Jair Carrion

- Bacharel: Ciência da Computação;
- Pós-graduação: Análise de Dados;
- +4 anos de experiência com a plataforma Qlik;
 - Desenvolvimento;
 - Consultoria;
 - Treinamento;
 - Pré-vendas;
- Certificado QlikSense Business Analyst.





Obrigado !





www.qlik.com