

An aerial photograph of a university campus. The campus features several large, modern buildings with white and red roofs, interspersed with lush green lawns and trees. A winding river or canal flows through the campus, with a bridge crossing it. In the foreground, there's a large, curved building with a red roof. The overall scene is bright and sunny, with a clear blue sky.

# WEB ACADEMY

## Tópicos Emergentes

Daniel Augusto Nunes da Silva

# **Apresentação**

# Ementa

- Soluções de apoio ao processo de decisão baseadas em dados. Pensamento Analítico. **Análise de Dados**. Integração de dados e o processo de **ETL**. Processamento Analítico (**OLAP**). **Visualização de dados**. *Dashboards*. Ferramentas ETL e de visualização de dados.



# Objetivos

- **Geral:** Apresentar conceitos e práticas relacionados à **utilização de técnicas de integração e visualização de dados**, fornecendo uma visão geral do emprego de ferramentas de suporte ao processo de tomada de decisão baseado em dados.
- **Específicos:**
  - Discutir a importância e os principais recursos para utilizar o pensamento analítico na solução de problemas;
  - Demonstrar o uso de ETL para auxiliar soluções voltadas ao processamento analítico;
  - Apresentar técnicas, conceitos e ferramentas para criação de visualizações que dão suporte ao processo de tomada de decisão guiado por dados.

# Conteúdo programático

## Introdução

- Dado, informação e conhecimento.
- Tomada de decisão guiada por dados.
- Pensamento Analítico.
- O caminho dos dados à tomada de decisão.

## Tratamento e integração de dados

- Integração de dados e processo de ETL.
- Processamento analítico *online* – OLAP.
- OLTP x OLAP.
- Ferramenta ETL: Talend Open Studio.

## Visualização de dados

- Introdução à visualização de dados.
- Métricas e indicadores de desempenho.
- Ferramenta de visualização de dados: Looker Studio.

# Bibliografia



## **Business Intelligence e Análise de Dados para Gestão do Negócio**

Ramesh Sharda, Dursun Delen e Efraim Turban

4ª Edição – 2019

Editora Bookman

ISBN 978-8582605196



## **Storytelling com Dados**

Cole Nussbaumer Knaflitz

1ª Edição – 2018

Editora Alta Books

ISBN 978-8550804682

# Sites de referência

- **Storytelling com Dados**

- <https://www.storytellingwithdata.com/>

- **Talend Open Studio for Data Integration User Guide**

- <https://help.talend.com/r/en-US/8.0/studio-user-guide-open-studio-for-data-integration>

- **Ajuda do Looker Studio**

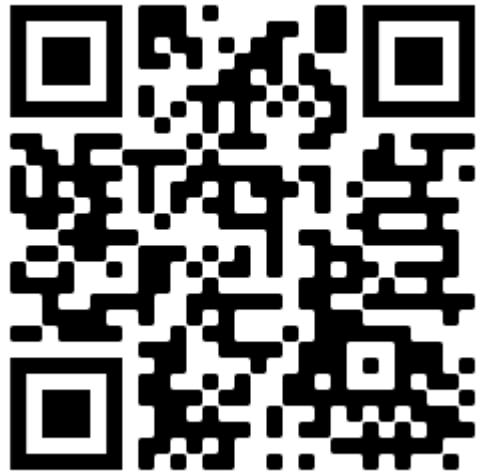
- <https://support.google.com/looker-studio>

# Ferramentas

- **Talend Open Studio for Data Integration**
  - Ferramenta ETL.
  - <https://www.talend.com/lp/open-studio-for-data-integration/>
- **Looker Studio (Google Data Studio)**
  - Ferramenta de visualização de dados.
  - <https://lookerstudio.google.com/>



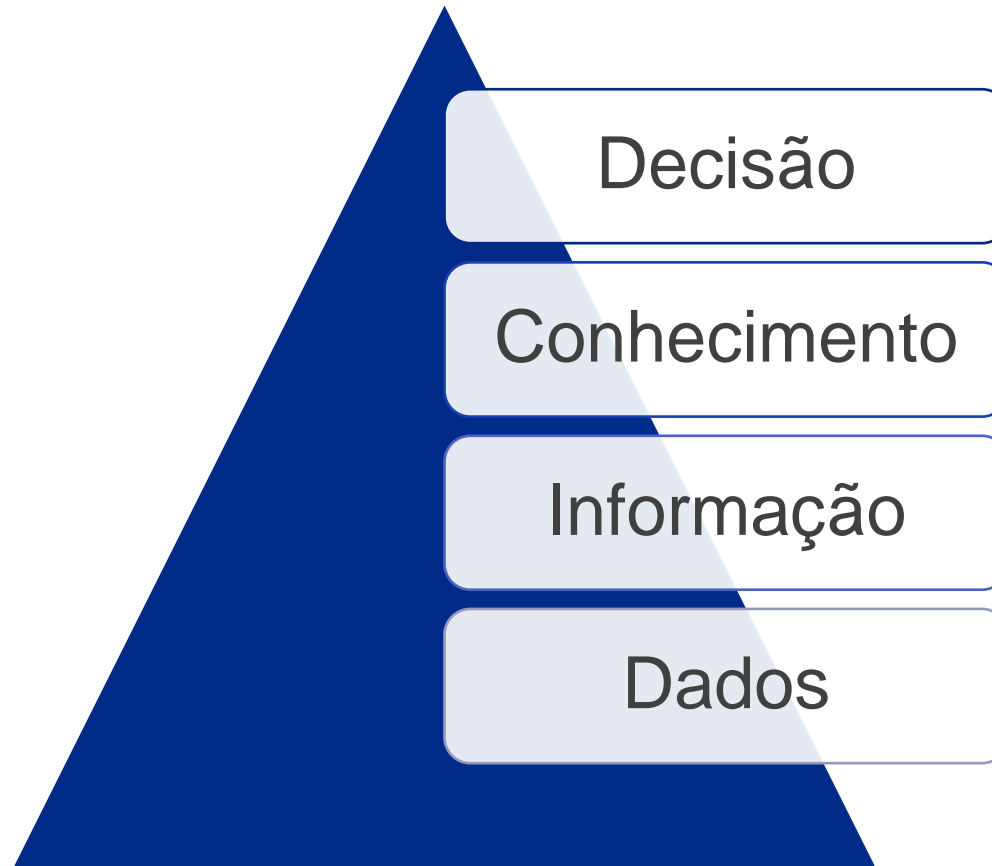
# Contato



<https://linkme.bio/danielnsilva/>

# Introdução

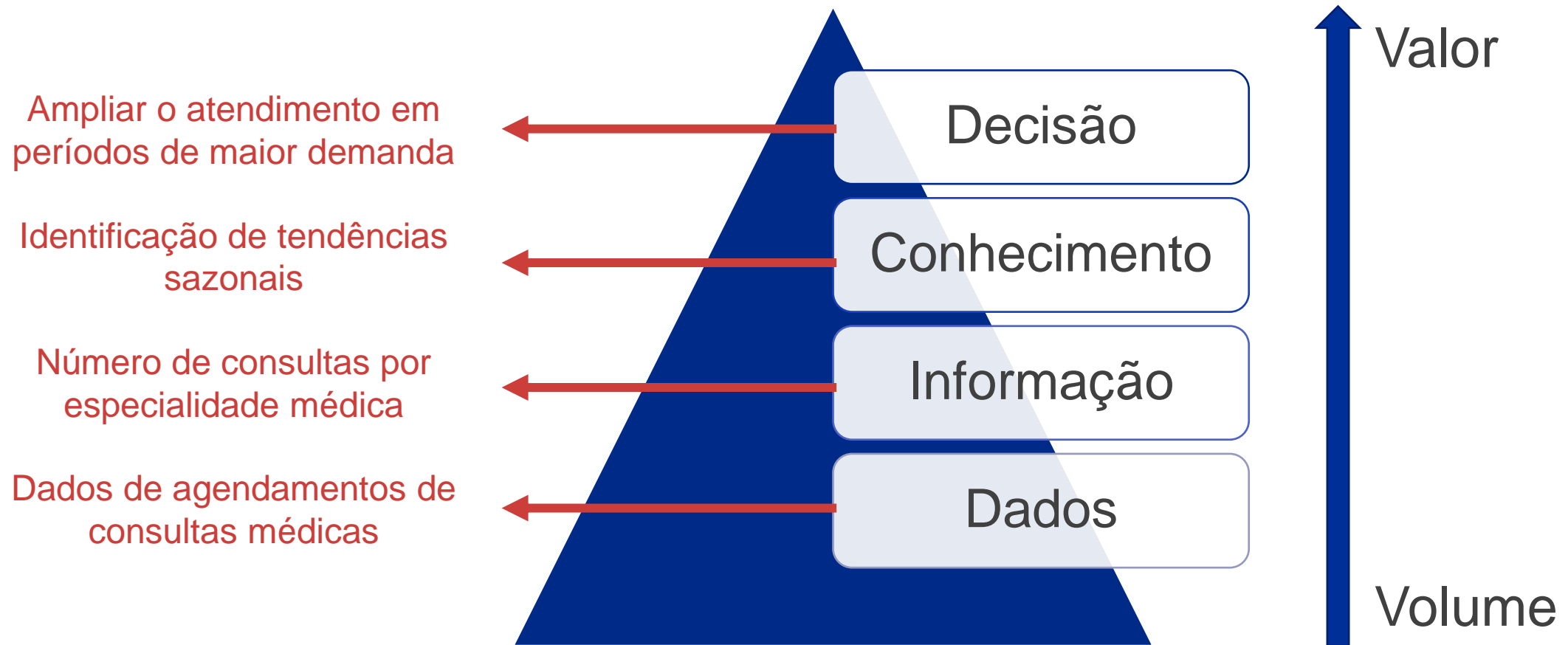
# Dado, informação e conhecimento



# Dado, informação e conhecimento

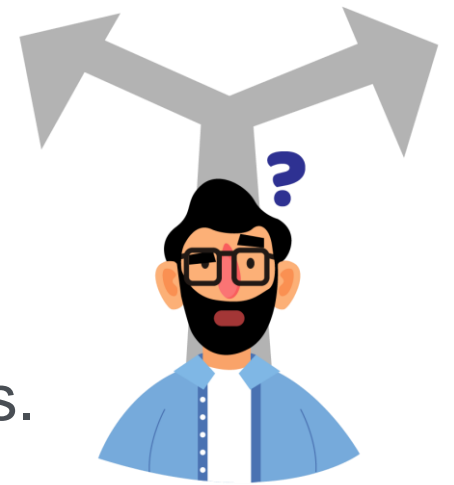


# Dado, informação e conhecimento



# Tomada de decisão guiada por dados

- Tomada de decisão guiada por dados refere-se à prática de **basear as decisões na análise dos dados, em vez de apenas na intuição.**
- Exemplo: um negociante poderá **selecionar anúncios baseados somente em sua longa experiência na área** e em sua intuição, ou pode basear sua escolha na **análise dos dados sobre a forma como os consumidores reagem a diferentes anúncios.**
- Também é possível utilizar uma combinação dessas abordagens.





# Pensamento Analítico

- Pensamento analítico consiste na **habilidade de reunir informações e, de forma cautelosa, buscar a solução de problemas.**
- Quando confrontados com um problema de negócios, devemos **avaliar se e como os dados podem ajudar a resolver o problema.**
- Alguns **conceitos e ferramentas dão suporte ao pensamento analítico**, permitindo que uma análise possa ser feita de forma sistemática.
- O **pensamento analítico** ajudará a **identificar oportunidades para melhorar a tomada de decisões orientada por dados.**





Felipe Santana • 2º

Cientista de Dados Sênior | Machine Learning | Python | Deep Learn...

2 sem •

+ Seguir

# Foco na solução do problema

[https://www.linkedin.com/posts/felipesf\\_datascience-trabalho-carreira-activity-6949350306933579776-Xp-M/](https://www.linkedin.com/posts/felipesf_datascience-trabalho-carreira-activity-6949350306933579776-Xp-M/)

Estes são erros comuns de iniciantes (e até avançados) em Data Science.

Adianto que já cometi alguns, principalmente o terceiro.

1. Entender de forma superficial o requisito de negócio.
2. Menosprezar a etapa de análise exploratória de dados e ir direto para experimentação com algoritmos.
3. Subestimar técnicas simples até quebrar a cara e voltar com o rabo entre as pernas. Esse é classico, que atire a primera pedra quem nunca.

4. Se apegar a ferramentas e tecnologias ao invés de focar na solução do problema.

5. Tomar decisão puramente técnica sem levar em consideração a experiência do usuário.
6. Não investigar de forma minuciosa os resultados do modelo e alinhar com as métricas de negócio.
7. Não saber se comunicar e apresentar de forma clara e sucinta o seu trabalho.

# O caminho dos dados à tomada de decisão

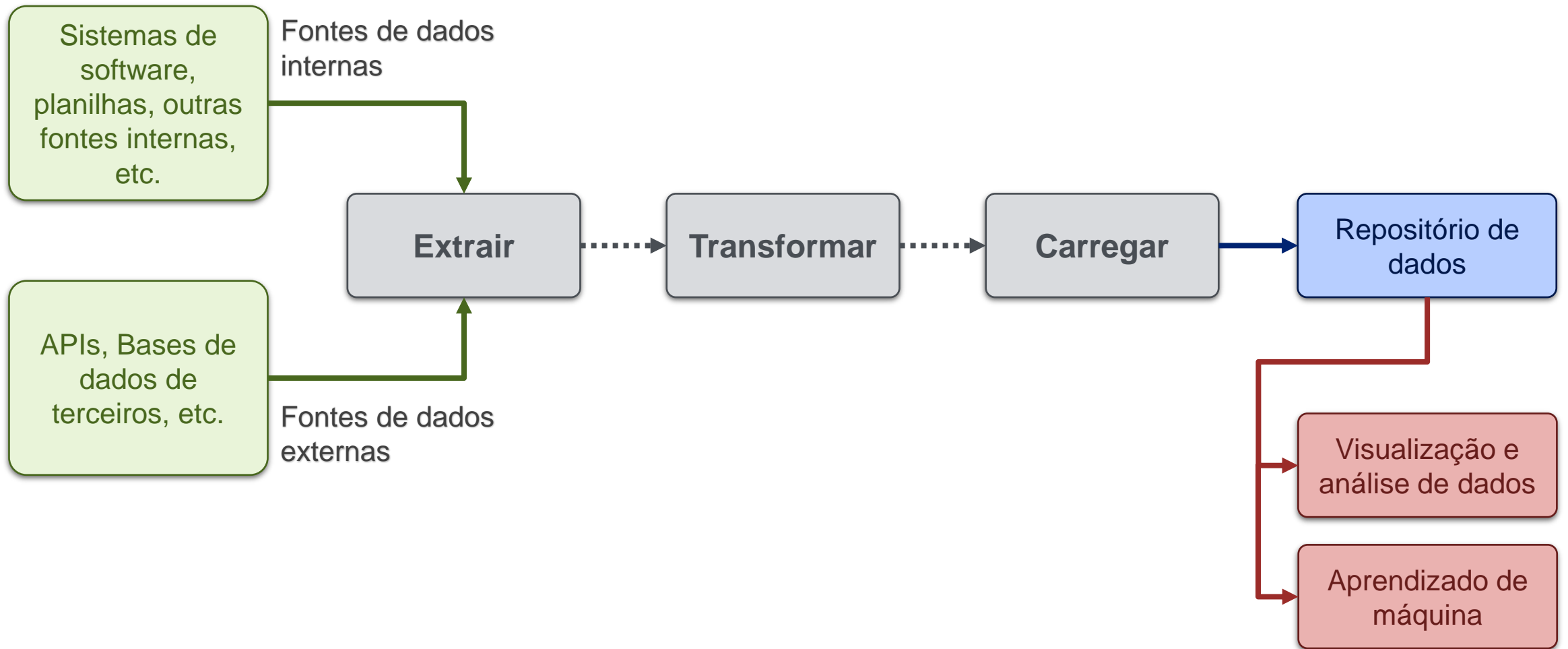
- Todo processo de tomada de decisão inicia com as etapas de **compreensão do problema** e **preparação dos dados**, independente das ferramentas adotadas e das abordagens (*dashboards*, *apredendizado de máquina*, *otimização de processos*, etc.).
- O objetivo é **responder questões de negócios** com técnicas de análise de dados.
- E quanto ao **SGCM**?
  - Qual é o percentual de pacientes ausentes (no-show)?
  - Quantos atendimentos por dia?
  - Qual especialidade tem mais atendimentos?, etc...

# Tratamento e integração de dados

# Integração de dados e processo de ETL

- **Integração de dados** consiste em reunir dados de diferentes origens para dar suporte ao processo de tomada de decisão, seja por meio da **análise e visualização de dados**, ou outras finalidades, como a construção de modelos de **aprendizado de máquina**.
- Uma das tecnologias que permitem a integração de dados é o processo de ETL:
  - **Extração**: leitura dos dados a partir de diferentes fontes;
  - **Transformação**: conversão dos dados extraídos para um formato novo;
  - **Carga**: colocar os dados em um novo espaço de armazenamento, para ser utilizado em outras etapas do processo de tomada de decisão guiada por dados.

# O processo de ETL





# Processamento analítico *online* – OLAP

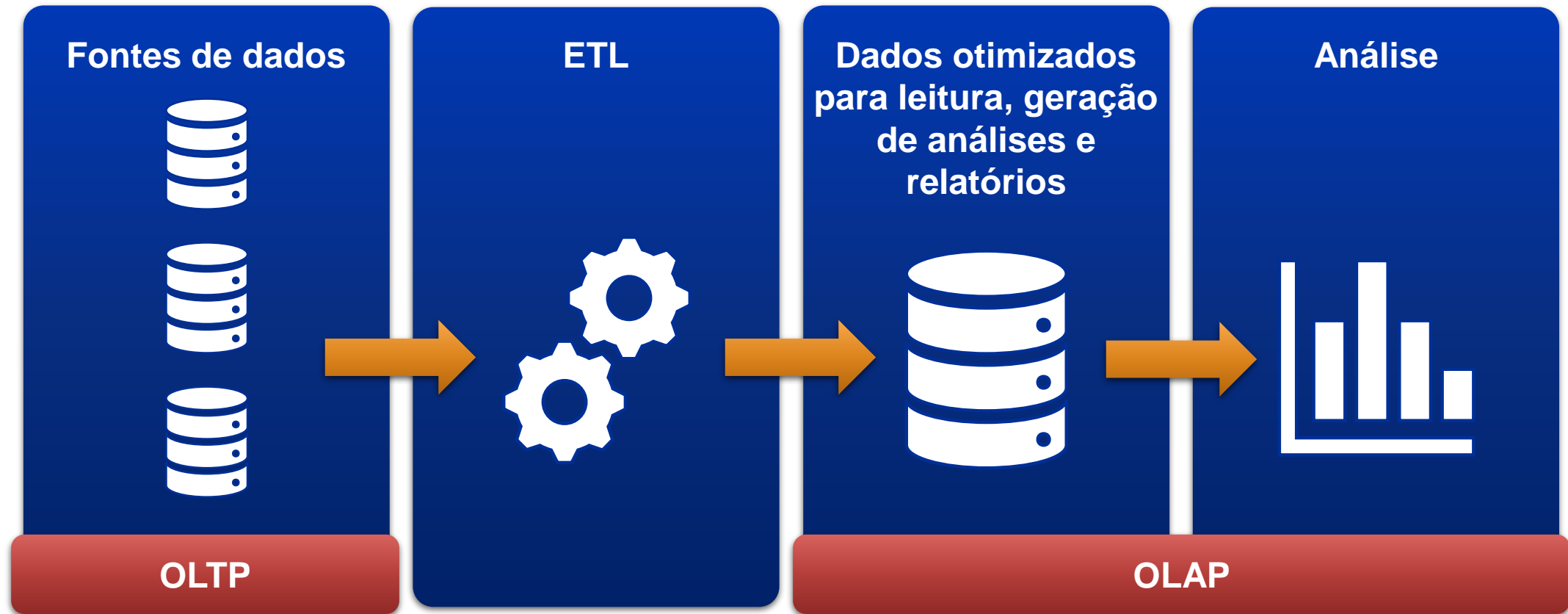
- O termo faz referência a uma **variedade de atividades voltadas à análise de dados**, normalmente executadas por usuário finais, o que pode incluir:
  - Geração de consultas;
  - Solicitação de relatórios e gráficos de rotina e *ad hoc*;
  - Realização de análises estatísticas;
  - Construção de apresentações visuais.



# OLAP *versus* OLTP

- Durante muito tempo o foco era o **processamento de transações (OLTP)**, normalmente baseados em sistemas que utilizam bases de dados relacionais.
- **OLTP** é voltado para o processamento de transações repetitivas em grandes quantidades (leitura, inserção, modificação e exclusão).
- **OLAP** foca em relacionamentos complexos e na busca de padrões e tendências (diretamente relacionado com o suporte à decisão);
- **OLTP** geralmente envolve **normalização de dados**, o que pode afetar o desempenho nas operações de leitura, que é o foco do **OLAP**.

# OLAP *versus* OLTP



# Talend Open Studio for Data Integration

- **Ferramenta ETL para integração de dados**, baseada na IDE Eclipse.
- A ferramenta gera um aplicação Java, mas a maior parte do recursos exige apenas operações de arrastar e soltar.
- Suporte a múltiplas fontes de dados: BDs relacionais, Serviços de nuvem, APIs, Big Data, etc.
- Permite que a aplicação seja compilada e executada de forma independente.
  - <https://www.datalytyx.com/scheduling-talend-open-studio-jobs-in-windows-without-talend-administration-center-tac/>
- Tutorial: <https://www.javatpoint.com/talend>

# Visualização de dados

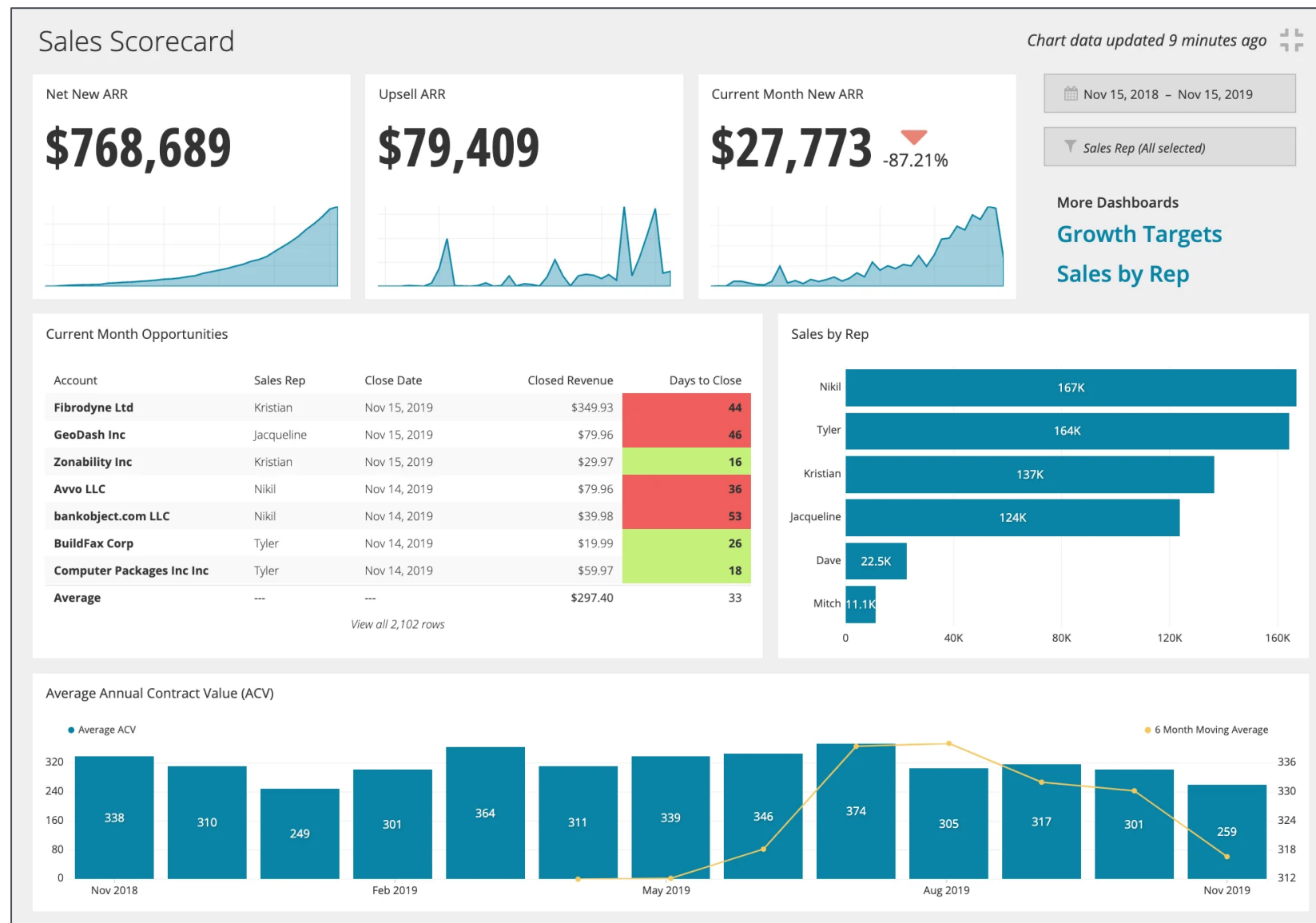
# Introdução à visualização de dados

- Visualização de dados: **qualquer representação gráfica**, animada ou estática, utilizada para apresentação de dados.
- Fornece capacidade para **identificar rapidamente as tendências importantes nos dados**.
- Método tradicional: visualização por planilhas.
- Métodos mais modernos: **dashboards** e **indicadores**.





# Visualização de dados: *dashboard*



# Métricas e indicadores de desempenho

- Dashboards são formados principalmente por métricas e indicadores de desempenho (**KPI** – *Key Performance Indicator*).
- Uma **métrica é qualquer dado que permita mensurar algum aspecto do negócio**: quantidade de clientes, atendimentos, vendas, etc.
- **KPIs fazem referência aos resultados** de ações e acompanhamento de metas.
- Exemplo:
  - **Métrica**: quantidade de clientes.
  - **KPI**: percentual de clientes satisfeitos.

# Looker Studio

- Até Outubro/2022 era Google Data Studio.
- Permite **visualizar dados com gráficos e tabelas** configuráveis.
- **Suporte a conexão com várias fontes de dados:** BigQuery, **MySQL**, PostgreSQL, Planilhas, CSV, dentre outros.
- Recursos de compartilhamento e colaboração.
- Tutorial para criar um relatório:
  - [https://support.google.com/datastudio/topic/6289358?hl=pt-BR&ref\\_topic=9170843](https://support.google.com/datastudio/topic/6289358?hl=pt-BR&ref_topic=9170843)

**Fim!**



# Referências

- SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. **Business Intelligence e Análise de Dados para Gestão do Negócio**. 4. ed. Porto Alegre: Bookman, 2019. 614 p.
- KNAFLIC, Cole N. **Storytelling com Dados**. Rio de Janeiro: Alta Books, 2018. 256 p.
- FAWCETT, Tom; PROVOST, Foster. **Data Science para Negócios**. 1. ed. Rio de Janeiro: Alta Books, 2016. 408 p.
- KNAFLIC, Cole N. **Storytelling com Dados**. [S. l.], 2022. Disponível em: <https://www.storytellingwithdata.com/>.
- TALEND. **Talend Open Studio for Data Integration User Guide**. [S. l.], 2022. Disponível em: <https://help.talend.com/r/en-US/8.0/studio-user-guide-open-studio-for-data-integration>.
- GOOGLE. **Ajuda do Looker Studio**. [S. l.], 2022. Disponível em: <https://support.google.com/looker-studio?hl=pt-BR>.