

Criando um BD analítico com o Pentaho Data Integration





Quem sou eu?

- ❖ Magno Macedo de Oliveira
 - Estudante do BSI (**B**acharelado em **S**istemas de **I**nformação)
 - Participante do LEDS (**L**aboratório de **E**ducação e **D**esenvolvimento de **S**oluções)

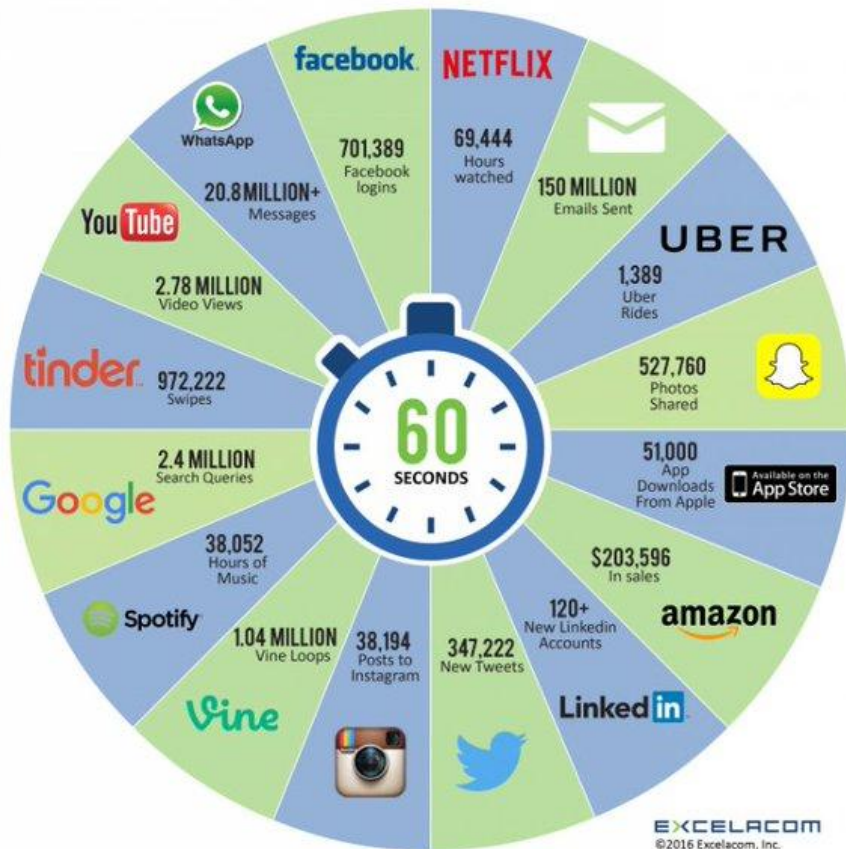




**Qual a importância da
informação no mundo
atual?**

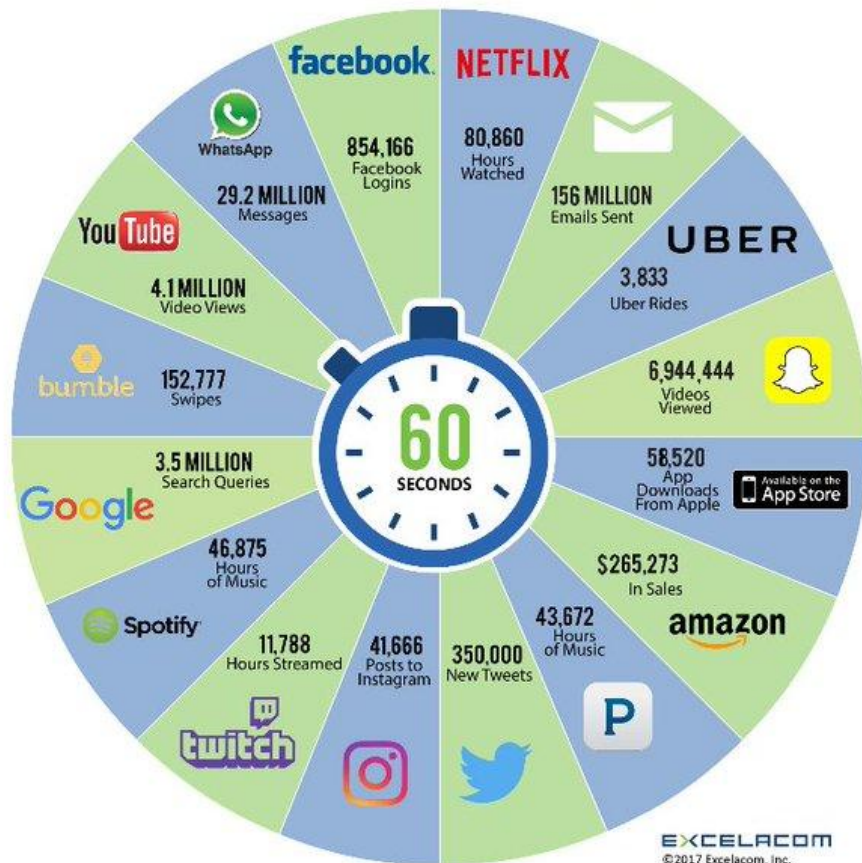


2016 What happens in an INTERNET MINUTE?



EXCELACOM
©2016 Excelacom, Inc.

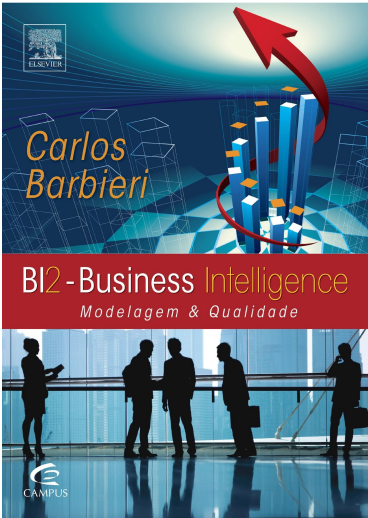
2017 What happens in an INTERNET MINUTE?



EXCELACOM
©2017 Excelacom, Inc.

-

(Barbieri, Carlos - **BI2: Business Intelligence: Modelagem & Qualidade**, Ed. Elsevier, 2011)





OLTP vs OLAP

- ❖ Essas siglas são bastante utilizadas no universo do BI;
 - OLTP:
 - On-Line Transaction Processing;
 - Termo usado para tratar sistemas transacionais.
 - OLAP:
 - On-Line Analytical Processing;
 - Trata de bancos de dados que possuem foco em analisar grandes volumes de informações;
 - Dá suporte para as funções de análises de negócio organizacional.



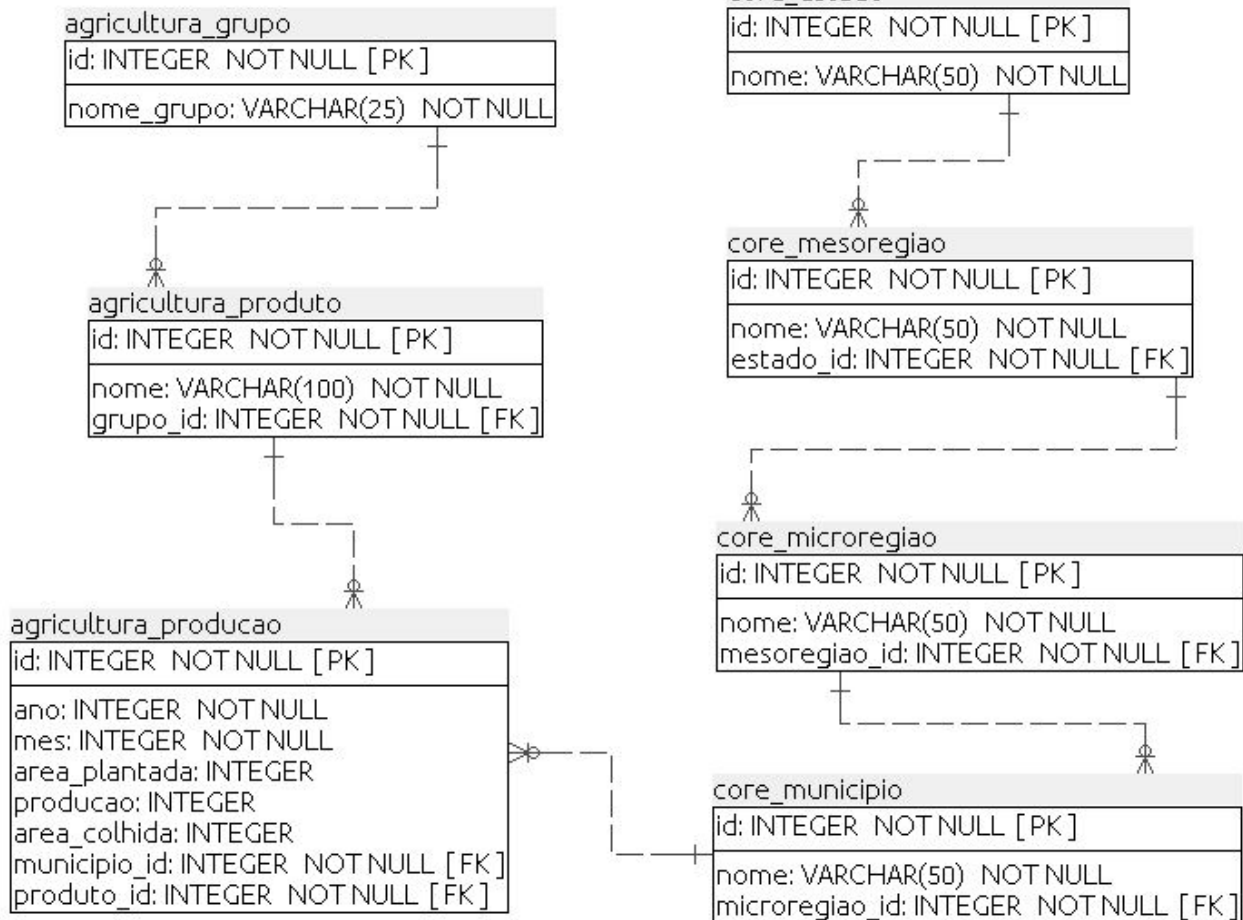
OLTP vs OLAP

	OLTP	OLAP
Foco	Foco no nível operacional da organização. Visa a execução operacional do negócio.	Foco no nível estratégico da organização. Visa a análise empresarial e tomada de decisão.
Performance	Alta velocidade na manipulação de dados operacionais, porém ineficiente para geração de análises gerenciais complexas.	Otimização para a leitura e geração de análises e relatórios gerenciais.
Frequência de atualização	A atualização dos dados é feita no momento da transação. Frequência alta.	A atualização é feita no processo de carga dos dados. Frequência baixa.

Logo, um complementa o outro.

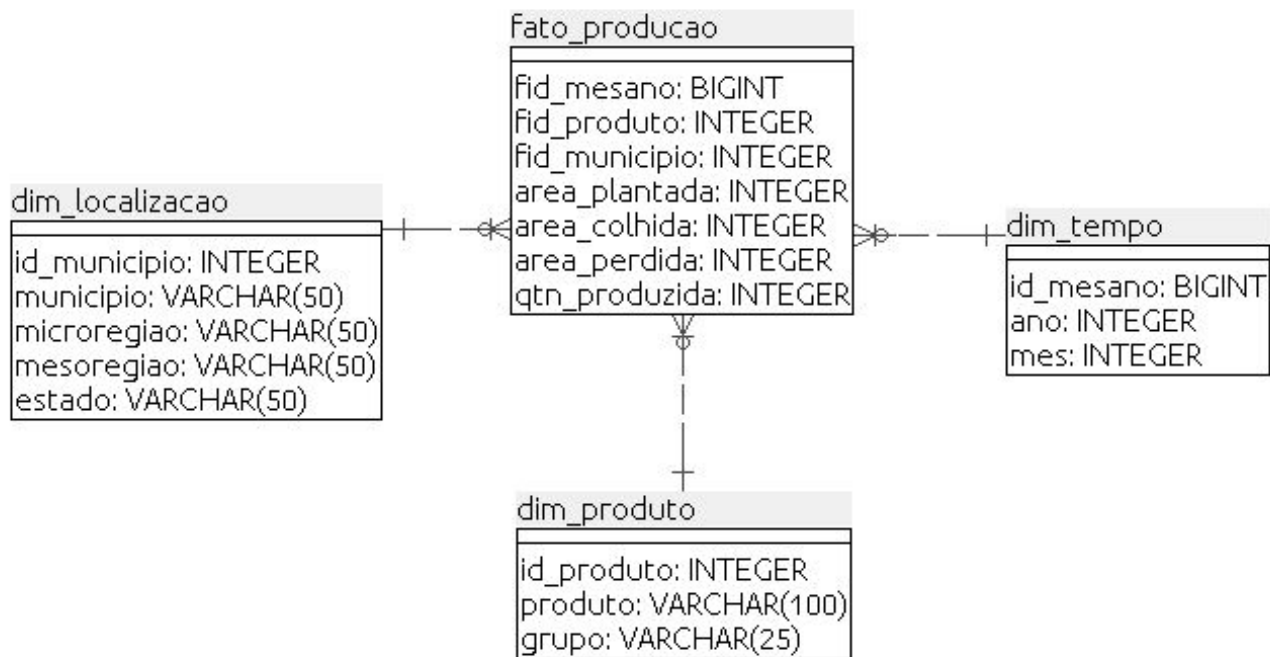


OLTP





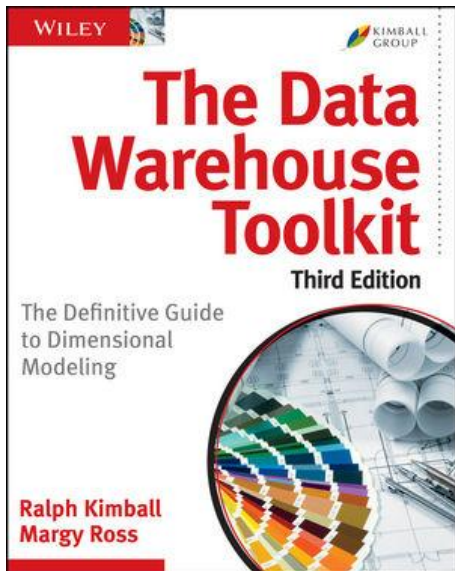
OLAP



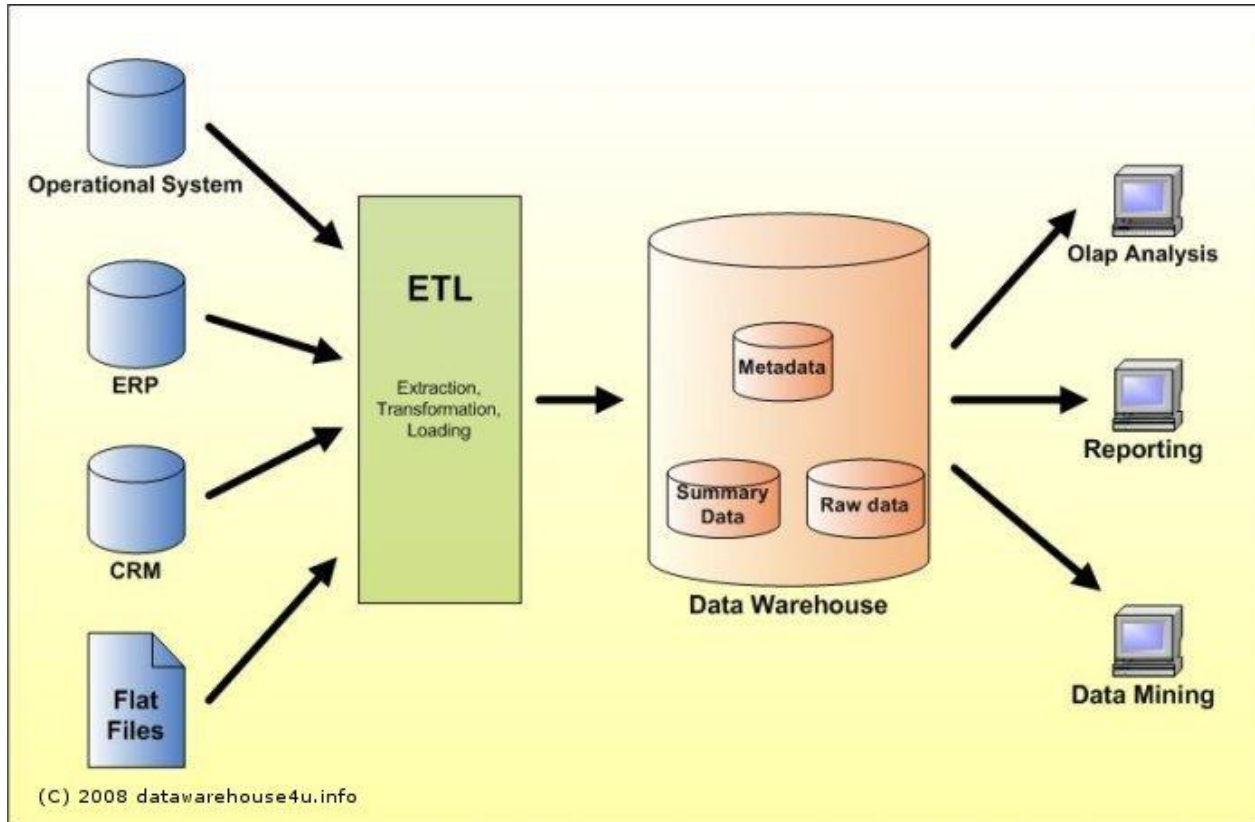
Data Warehouse

- ❖ Para Ralph Kimball, criador do conceito, Data Warehouse é...
- ❖ “...Uma cópia de dados de transação (dados transacionais dos sistemas de origem), especificamente estruturado para consulta e análise”.

(Kimball, Ralph - **The Data Warehouse Toolkit**, Ed. John Wiley & Sons, 2013)



Data Warehouse





Pentaho

- ❖ Desenvolvido em 2004 pela Pentaho Corporation, essa suíte, desenvolvida em Java, é um software de código aberto voltado para BI.
 - Dentro do pentaho encontra-se soluções voltadas a:
 - ETL (**E**xtraction, **T**ransformation and **L**oad)
 - Reporting
 - OLAP
 - Data-mining





Pentaho

File View Tools Help

Home ▾

admin ▾

Browse Files

Create New

Manage Data Sources

Documentation

Recents

 2005 Q1 Produc... ☆

 SampleTransfor... ☆

 Chart Pick List ☆

 SQL Rule - Mana... ☆

Pentaho Business Analytics

Learn

Contribute

Meet the Family

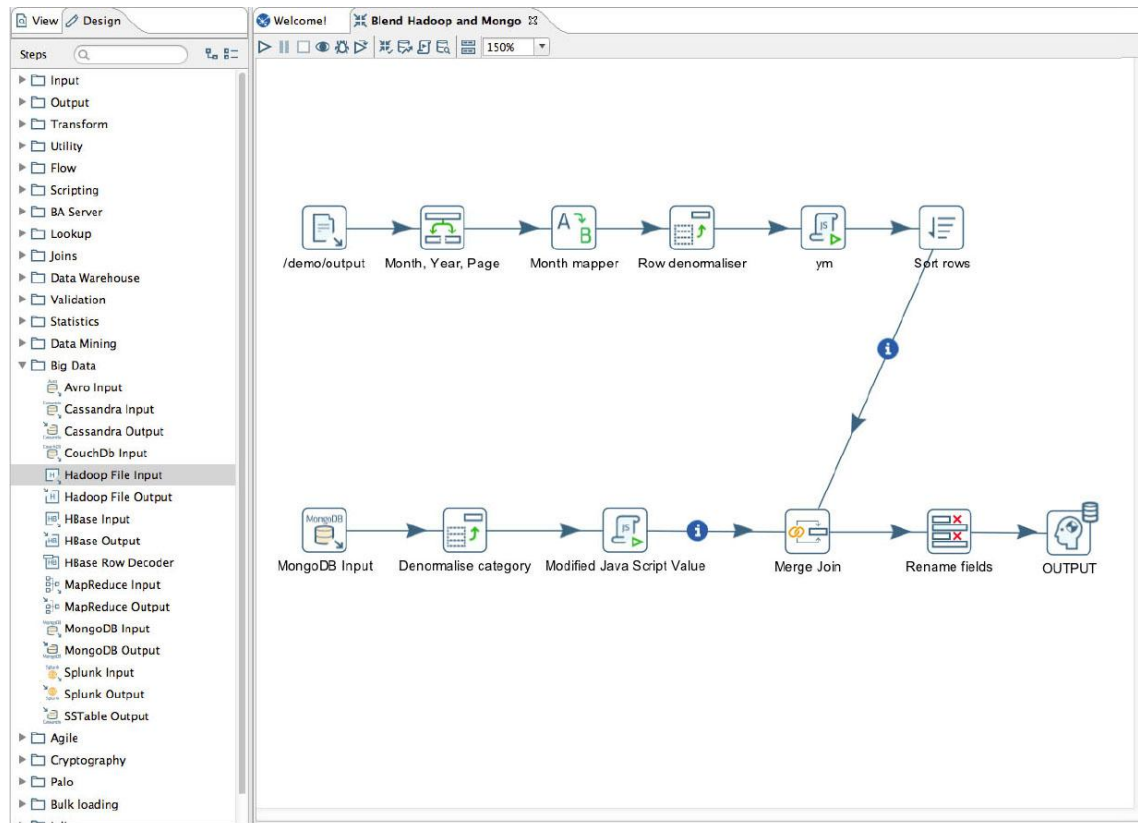
Upgrade to Enterprise Edition

The Pentaho Business Analytics Platform delivers an open, unified, end-to-end solution including data integration, visualization and consumption of data.

Either by providing ad-hoc tools to analyze and visualize data or by serving pre-created content, different content types including analysis, reports, dashboards, data mining and even community created plugins can be seamlessly connected.



Pentaho





Pentaho

Schema Workbench

File Edit View Tools Windows Help

Schema - SugarCRM Sales Performance (SugarCRM-Sales-Performance2.xml)

Schema

- SalesAnalysis
 - Table: sales_fact
 - Start Period
 - Close Period
 - Sale State
 - Sales Person
 - Account Categorization
 - Account Location
 - Total Sale Amount
 - Number of Sales
 - Time To Close (Days)
 - Sum Close Probability
 - CM Avg Sale Amount
 - CM Avg Time To Close (Days)
 - CM Avg Close Probability
 - Sale State
 - Account Categorization
 - Account Location
 - Sales Person
 - Start Period

Dimension Usage for 'SalesAnalysis' Cube

Attribute	Value
name	Start Period
foreignKey	sales_date_of_start_id
source	probability
level	sales_account_id
usagePrefix	sales_date_of_close_id
caption	sales_date_of_start_id
	sales_location_id
	sales_person_id
	sales_state_id
	time_to_close

Database - sugarcrm (MySQL)



Pentaho

Pentaho Report Designer - Buyer Report - C:\temp\prd\report-designer\samples\Operational Reports\Buyer Product Analysis.prpt

File Edit View Insert Format Data Extras Window Help

Buyer Report X Product Sales X

147% 16 B I U #133484

Operations

Page Header

Report Header

Buyer's Product Status Report

Group Header

LabelPRODUCTLINE

Group Header

LabelPRODUCTVENDOR

Details Header

History SKU Name

No Data

No Data Available

Details

PRODUCTCODE PRODUCTNAME

PRODUCTDESCRIPTION

Report Footer

End of Report

Structure Data

Master Report: Buyer Report

- Page Header
 - horizontal-line
 - message-field
 - Ab label
- Report Header
 - horizontal-line
 - image
 - Ab label
- Group: Line
- Report Footer
- Page Footer
 - message-field

Style Attributes

Name	Inh	Value	Expr
size & position			
x		0.0	
width		526.987	
y		5.378	
height		24.0	
invisible-consumes-s...		<input checked="" type="checkbox"/>	
visible		<input checked="" type="checkbox"/> true	
dynamic-height		<input checked="" type="checkbox"/> true	
font			
family		<input checked="" type="checkbox"/> Arial	
font-size		16	

Messages

Description

- Expression with name 'Product' misses a value for the mandatory property 'Fail On Error'.
- Element " " is smaller than the actual height of a single line of text. The element will be empty.

218.5 of 272.0 MB

Mão na massa!





Dúvidas?





#FicaDica

❖ Contatos:

- facebook.com/magno.macedo2099
- github.com/magnoDev
- gitbook.com/@magnodev
- **Pentaho para novos aventureiros**

❖ LEDS:

- github.com/leds
- gitbook.com/@leds

OBRIGADO