

BI e Data Warehouse

Fábio Ávila
Fernando Fonseca
Ana Carolina Salgado



ÁVILA

Ávila Soluções

Business
Intelligence



Power BI
SQL Server
Data Discovery/Analytics

Software sob
Medida

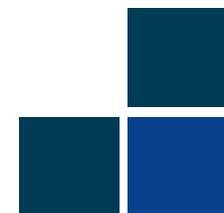


.NET / SQL Server
Intranets SharePoint
Desktop | Web | Celular

DBA
SQL Server
Oracle



Desempenho
Segurança
Alta Disponibilidade



Alguns Produtos



Monitoramento
DBA



Projetos de BI

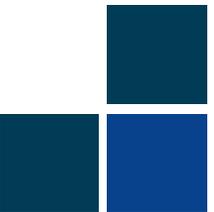


Alguns Clientes



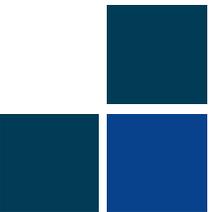
Data Warehouse (Armazém de Dados)

- “Coleção de dados **orientada a assunto, integrada, não-volátil** e **variante no tempo**, utilizada para tomada de decisões”.
 - W. H. Inmon
- “Uma cópia de dados transacionais estruturada especificamente para **consulta** e **análise**”
 - R. Kimball
- “Repositório estruturado e corporativo de dados **orientados a assunto, variantes no tempo** e **históricos**, usados para recuperação de informações e **suporte à decisão**. O DW armazena dados atômicos e sumarizados”.
 - Oracle



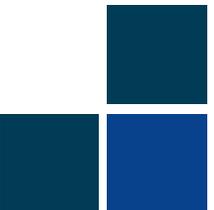
Data Mart

- O que é
 - Subconjunto de um Data Warehouse
 - Específico a um assunto particular ou atividade de negócios
- Por que construir Data Marts
 - Consultas mais rápidas e menor número de usuários
 - Tempo de implantação minimizado
- Data Marts Integrados



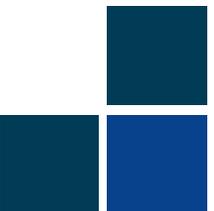
Business Intelligence

- Refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócios.
- Conjunto de técnicas e ferramentas para auxiliar na transformação de dados brutos em informações significativas e uteis a fim de analisar o negócio.
- As tecnologias BI são capazes de suportar uma grande quantidade de dados desestruturados para ajudar a identificar, desenvolver e até mesmo criar uma nova oportunidade de estratégia de negócios.
- O objetivo do BI é permitir uma fácil interpretação do grande volume de dados. Identificando novas oportunidades e implementando uma estratégia efetiva baseada nos dados, também pode promover negócios com vantagem competitiva no mercado e estabilidade a longo prazo.

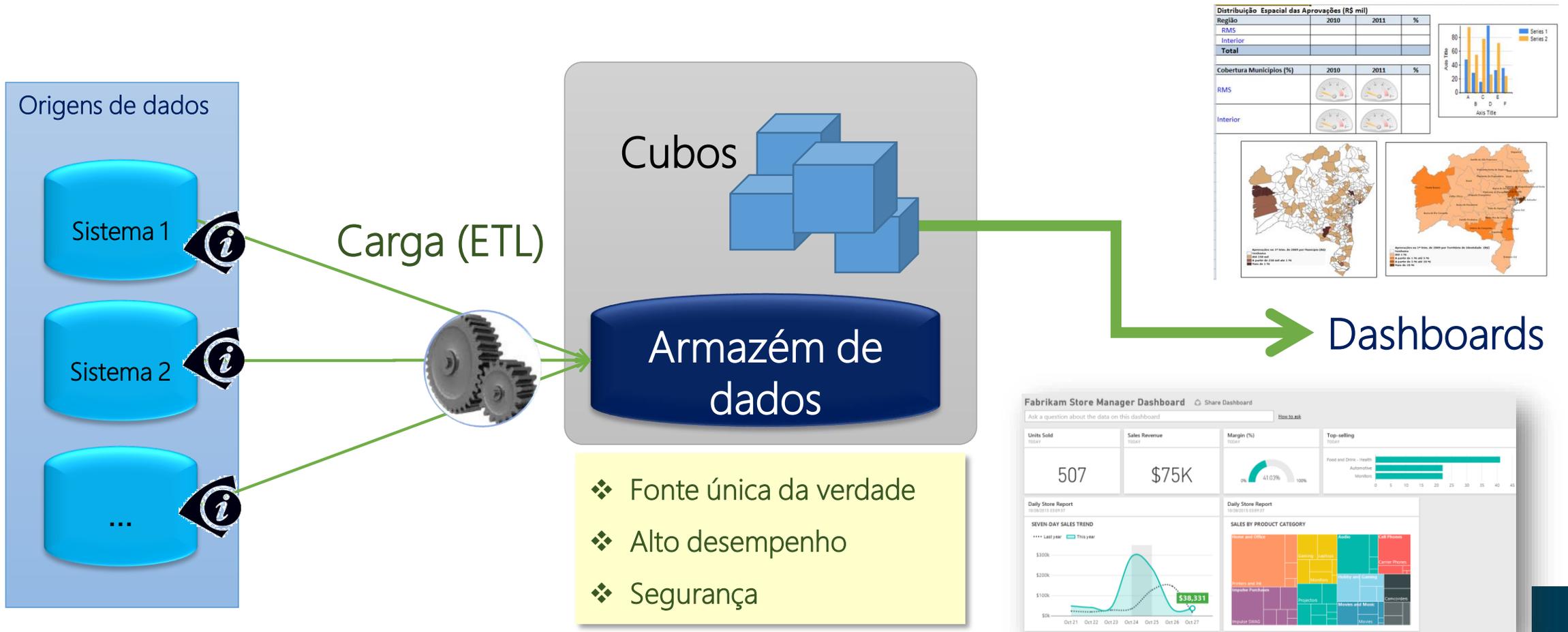


OnLine Transaction Processing (OLTP) vs. OnLine Analytical Processing (OLAP)

OLTP	OLAP
Transações numerosas e curtas	Análise gerencial
Dados detalhados e distribuídos	Dados consolidados
Consultas pré-definidas; pouco volume de dados	Consultas "ad-hoc"; alto volume de dados
Alta concorrência	Baixa concorrência
Transações/s mais uniforme	Transações/s menos uniforme
Mais normalização	Mais redundância
BD reflete o estado atual	Estado do momento da carga

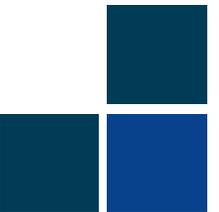


Arquitetura Típica de uma Solução de DW/BI

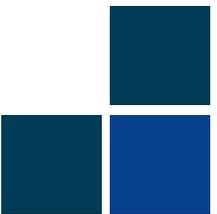


Benefícios de um Projeto de DW/BI

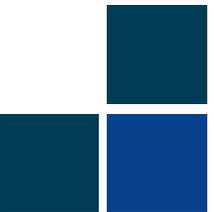
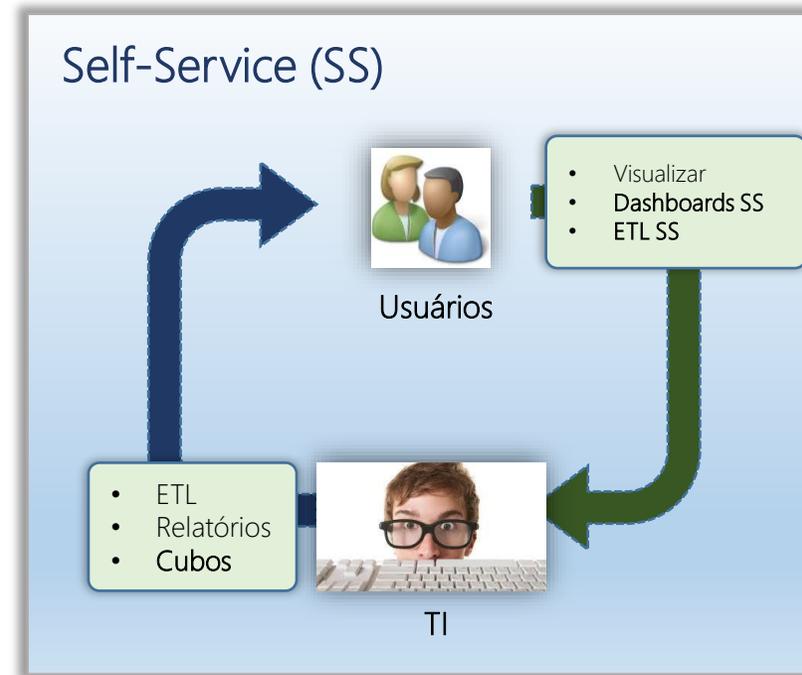
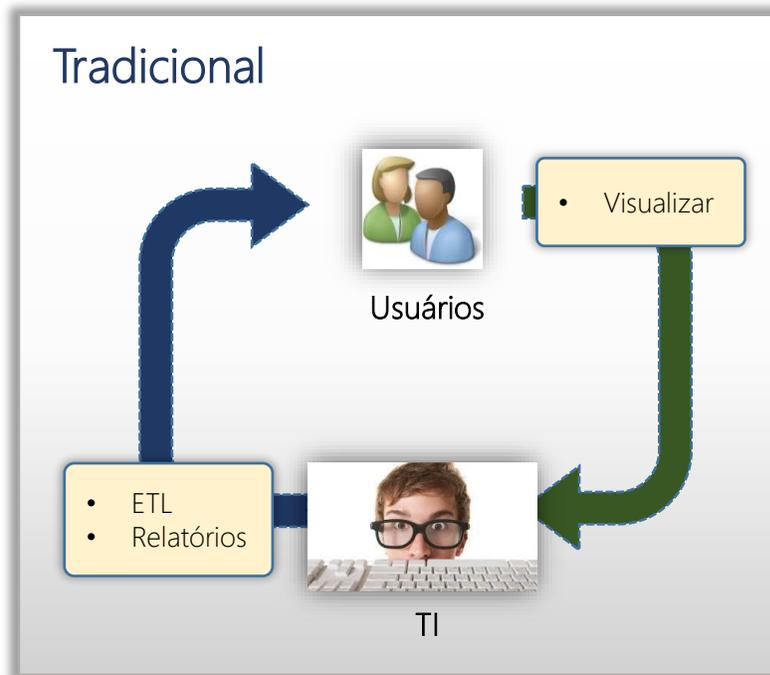
- Facilitar tomada de decisão
- “Fonte única da verdade”
- Eliminação de retrabalhos
- Flexibilidade de painéis gerenciais
- Segurança simplificada



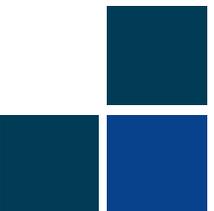
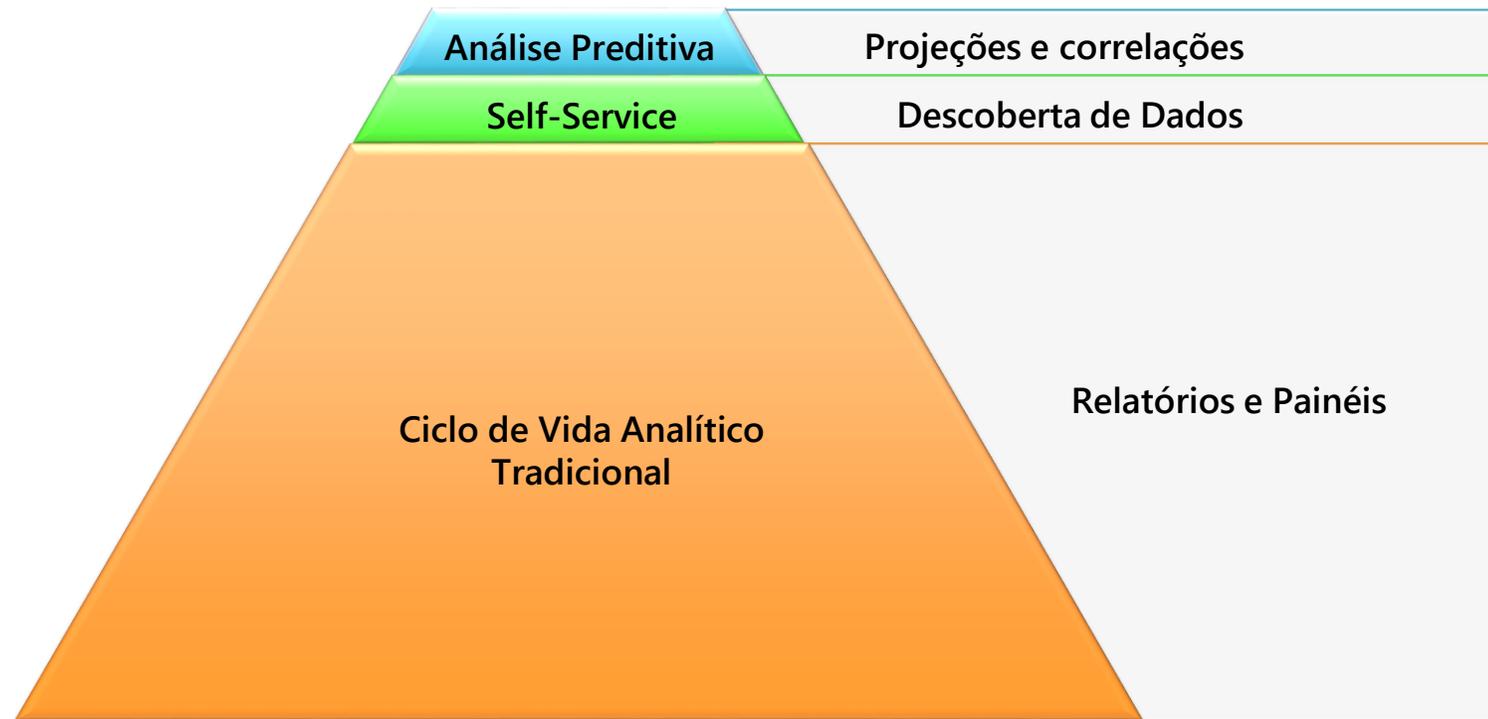
G.I.G.O. Garbage In, Garbage Out



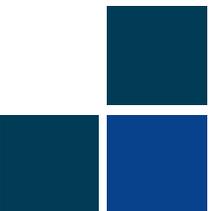
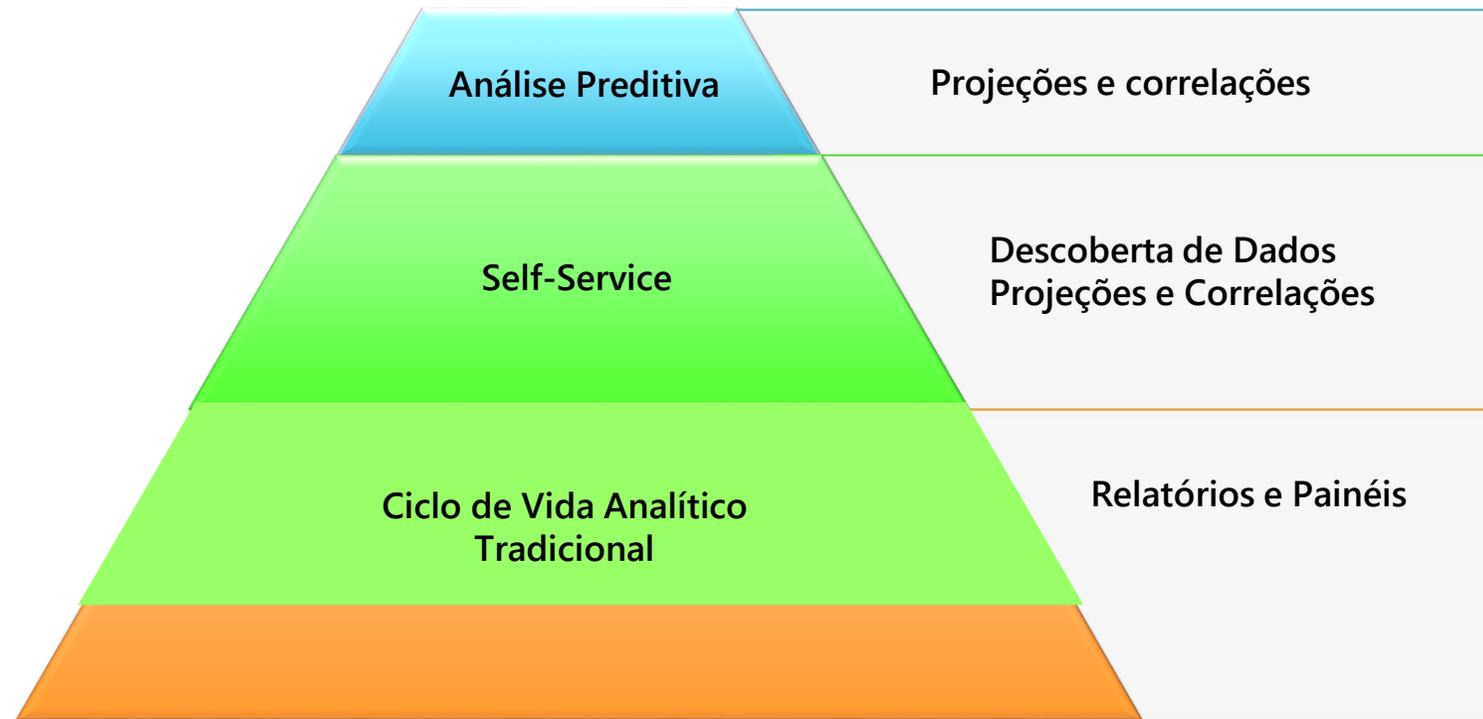
Ciclo de Vida Analítico



Usuários de BI



Usuários de BI



Estágios da Produção do Conhecimento

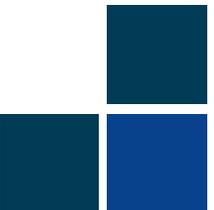
1 – Funcionários



2 – Pessoas



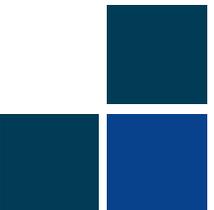
3 – Máquinas



Volume e Velocidade

- Até 2002: 2 exabytes (1 EB = 1.000.000.000 GB)
- Em 2003: 2 exabytes em 1 ano
- Em 2011: a cada 2 dias
- Junho/2013: a cada 10 minutos
- gigabyte | terabyte | petabyte
exabyte | zettabyte | yottabyte

Fonte: Kirk Borne @ TEDx, 06/04/2013



Variedade

- NoSQL
 - Conteúdo de redes sociais
 - Tweets
 - Gravações telefônicas
 - Músicas e Vídeos
 - E-mails
 - Cliques e padrões de navegação



Novas Profissões

The image shows a screenshot of a web browser displaying an article on the Glassdoor platform. The article is from Harvard Business Review, dated October 2012, and is titled "Data Scientist: The Sexiest Job of the 21st Century" by Thomas H. Davenport and D.J. Patil. The article preview includes a key icon and text indicating that full access requires sign-in or registration. Below the title, there are social media sharing icons and a "Comments (89)" link. To the right of the article, there is a "RELATED" section with an "Executive Summary" and an "ALSO AVAILABLE" section with a "Buy PDF" option. The price for the PDF is listed as \$116,840. Below the price, there are two rows of data: "1,736" and "4.1", and "116" and "4.7".

Percentage of Matching Job Postings (%)

Year	Percentage of Matching Job Postings (%)
2012	0.00
2013	0.05
2014	0.10
2015	0.15
2016	0.20
2017	0.25
2018	0.30
2019	0.35
2020	0.40
2021	0.45

glassdoor Search Jobs or Companies... Employers | Try Free Job Postings

Harvard Business Review

THE MAGAZINE BLOGS VIDEO BOOKS CASES WEBINARS COURSES

Guest Subscribe today and get access to all current articles and HBR online archive.

THE MAGAZINE
October 2012

ARTICLE PREVIEW To read the full article, [sign-in](#) or [register](#). HBR subscribers, click [here to register for FREE access](#) »

Data Scientist: The Sexiest Job of the 21st Century
by Thomas H. Davenport and D.J. Patil

Comments (89)

RELATED

Executive Summary

ALSO AVAILABLE

- Buy PDF

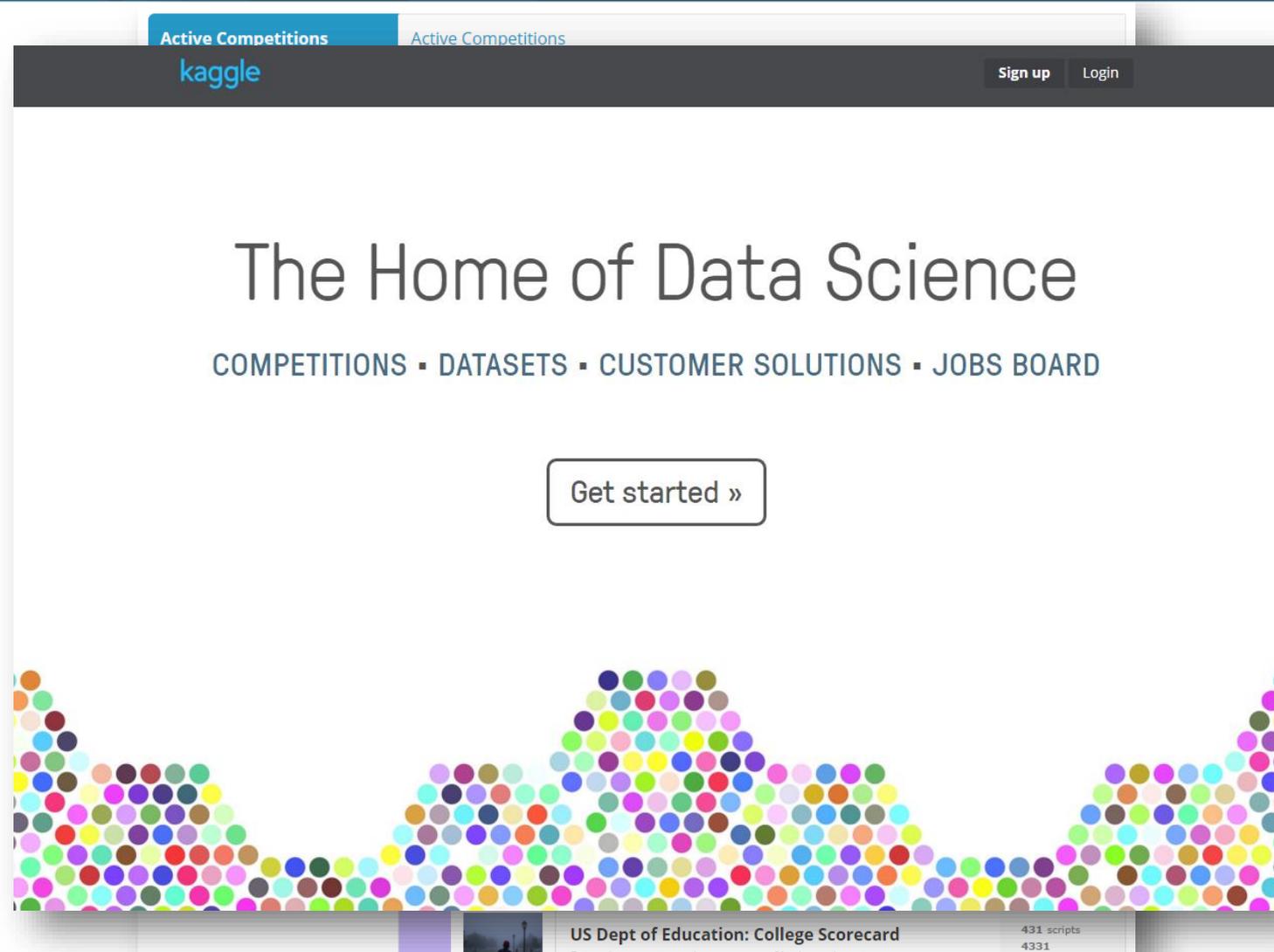
1,736
\$116,840
4.1
4.7

Job Score

Additional Resources
[Award FAQ](#)

Fonte: indeed.com

Kaggle



The image shows a screenshot of the Kaggle website's homepage. At the top, there is a dark grey navigation bar with the Kaggle logo on the left and "Sign up" and "Login" buttons on the right. Below the navigation bar, the main content area features the heading "The Home of Data Science" in a large, grey font. Underneath this heading is a horizontal list of menu items: "COMPETITIONS", "DATASETS", "CUSTOMER SOLUTIONS", and "JOBS BOARD", separated by small square icons. A prominent "Get started »" button is centered below the menu. The bottom of the page is decorated with a colorful, pixelated graphic consisting of many small circles in various colors, arranged in a wave-like pattern. At the very bottom, a partial view of a dataset card is visible, showing the title "US Dept of Education: College Scorecard" and the number of scripts "4331".

Active Competitions Active Competitions

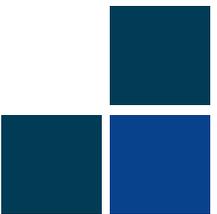
kaggle Sign up Login

The Home of Data Science

COMPETITIONS ▪ DATASETS ▪ CUSTOMER SOLUTIONS ▪ JOBS BOARD

Get started »

US Dept of Education: College Scorecard 431 scripts 4331



Elementos de uma Plataforma de BI



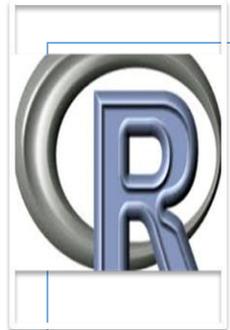
ETL

- Data Warehouse
- Corporativo / Self-service
- Tempo real



OLAP

- Cubos
- Linguagens DAX/MDX
- Alto desempenho e flexibilidade para consultas



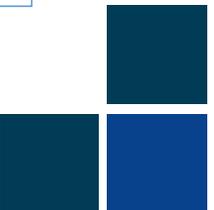
Análise Preditiva

- Data Mining
- Visualizações

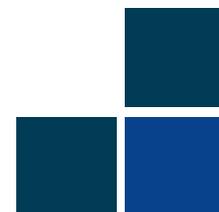


Visualização

- Servidor de Relatórios
- Data Discovery
- Self-Service



Qual a ferramenta de BI mais utilizada do mundo?



A person with a red backpack is captured mid-air, jumping over a large, reddish-brown rock formation in a desert canyon. The background shows a clear blue sky with a bright sun in the upper left corner, creating a lens flare effect. The foreground is filled with more red rock formations and sparse desert vegetation.

Demo

Sentiment Analysis no Excel

ÁVILA

A person with a red backpack is captured mid-air, jumping over a large, reddish-brown rock formation in a desert canyon. The background shows a clear blue sky with a bright sun in the upper left corner, creating a lens flare effect. The foreground is filled with more red rock formations and sparse desert vegetation.

Demo

Power BI Desktop

ÁVILA

Visualizações

- Em painéis, visualizações adequadas e interativas permitem percepções mais eficazes
- Análise de múltiplas métricas na mesma visualização
- Novo paradigma para visualizações: API aberta

Welcome to Power BI custom visuals

A gallery of visuals created by the Power BI community. Browse through the visuals or submit one of your own for others to use.

[Learn how to use custom visuals](#)

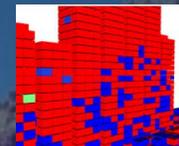
Visuals Sort by: Most Recent Search Visuals

The gallery displays a grid of various custom visualizations, including:

- SandDance, a Microsoft Gara...**: A 3D bar chart with red and blue blocks.
- LineDot Chart**: A line chart with green dots and a connecting line.
- Time Brush**: A dual-axis bar chart with green bars and a time range selector.
- Table Sorter**: A horizontal bar chart with yellow and red bars.
- Network Navigator**: A network diagram with nodes and connecting lines.
- Search**: A search interface showing results for #SQLServer, #powerpoint, #slide, #ppt, and #PowerBI.
- Stars**: A row of five orange stars.
- Texts**: A text box with a yellow circle and a grey triangle.

Demo

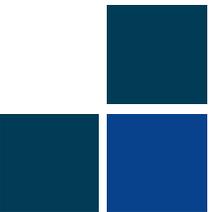
Sand Dance



ÁVILA

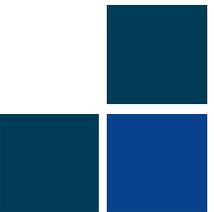
Projeto da Disciplina

- Equipes de 3 a 4 integrantes
- Construído de forma incremental
- Enunciado será passado em duas entregas: ETL e Cubo
- A entrega será por e-mail



Ferramentas para o Projeto

- SQL Server 2016 Developer (gratuito)
 - Transact-SQL
 - Integration Services (SSIS)
 - Analysis Services (SSAS)
- Power BI Desktop (gratuito)
 - Power Query (Linguagem M)
 - Linguagem DAX
 - Visualizações



Programação IDW 2016.2

Data	Conteúdo
16/ago	DW: Conceitos
23/ago	Mercado e Desempenho
30/ago	Modelagem DW
6/set	ETL: Transact-SQL e Integration Services
13/set	ETL: Integration Services ← Enunciado do projeto: primeira parte
20/set	Acompanhamento de Projetos
27/set	Analysis Services
4/out	MDX
11/out	Power BI ← Enunciado do projeto: segunda parte
18/out	Acompanhamento de Projetos
25/out	Acompanhamento de Projetos (opcional)
1/nov	Avaliação Prática