



## **Ferramentas de BI**

**Bootcamp: Desenvolvedor Business Intelligence**

Davidson Ignácio Oliveira

2020

## **Ferramentas de BI**

Bootcamp: Desenvolvedor Business Intelligence

Davidson Ignácio Oliveira

© Copyright do Instituto de Gestão e Tecnologia da Informação.

Todos os direitos reservados.

## Sumário

---

Capítulo 1. Business Intelligence .....	4
Principais Conceitos.....	4
A importância do BI.....	9
As empresas podem sobreviver sem BI?.....	11
Origens e aplicações de Business Intelligence.....	12
Capítulo 2. Arquitetura e soluções de Business Intelligence .....	15
Benefícios e vantagens das soluções de BI .....	15
Soluções de BI .....	16
Partes da arquitetura de BI.....	19
Como definir a melhor arquitetura .....	22
Capítulo 3. Aplicações ETL.....	26
Definição de ETL.....	26
Como funciona o processo de ETL .....	27
Vantagens de uma aplicação de ETL .....	30
Plataformas codificadas x Low Code .....	31
Ferramentas ETL e principais fornecedores .....	34
Capítulo 4. As Aplicações de BI Front-End .....	47
O que são aplicações front-end.....	47
Plataformas e soluções de BI .....	49
Ferramentas front-end: Qlik Sense .....	53
Ferramentas front-end: Tableau.....	57
Ferramentas front-end: Power BI .....	58
Referências.....	62

## Capítulo 1. Business Intelligence

O que você pensa quando ouve o termo BI? Talvez você pense que é de comer, passar no cabelo ou apenas um dashboard bonito com botões e gráficos. Neste capítulo vamos trabalhar os principais conceitos, benefícios, aplicações e origens de BI para melhor entendimento do tema.

### Principais Conceitos

#### ▪ Business Intelligence

O Grupo Gartner é a principal empresa no mundo de pesquisa, consultoria e prospecção de Mercado de TI. Ele define BI como um conceito que inclui aplicações, infraestrutura, ferramentas e boas práticas que nos possibilitam otimizar decisões e performances.

**Figura 1 – Exemplo de Dashboard.**

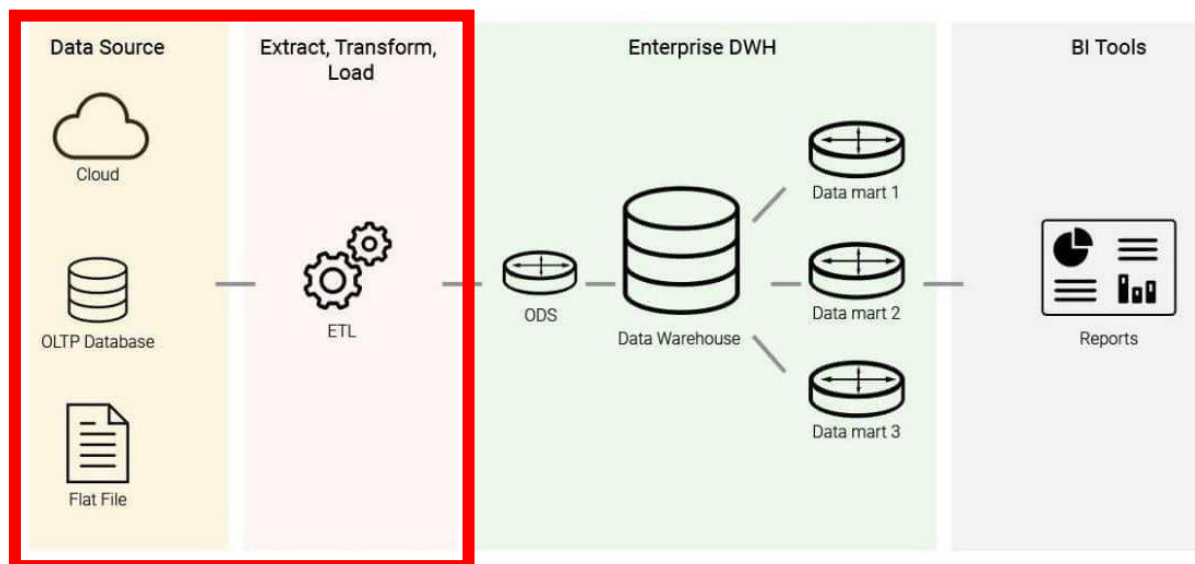


Fonte: <https://community.powerbi.com/t5/Data-Stories-Gallery/Executive-dashboard/td-p/439858>.

## ▪ ETL

Extract, Transform and Load (Extração, Transformação e Carga).

**Figura 2 – ETL.**

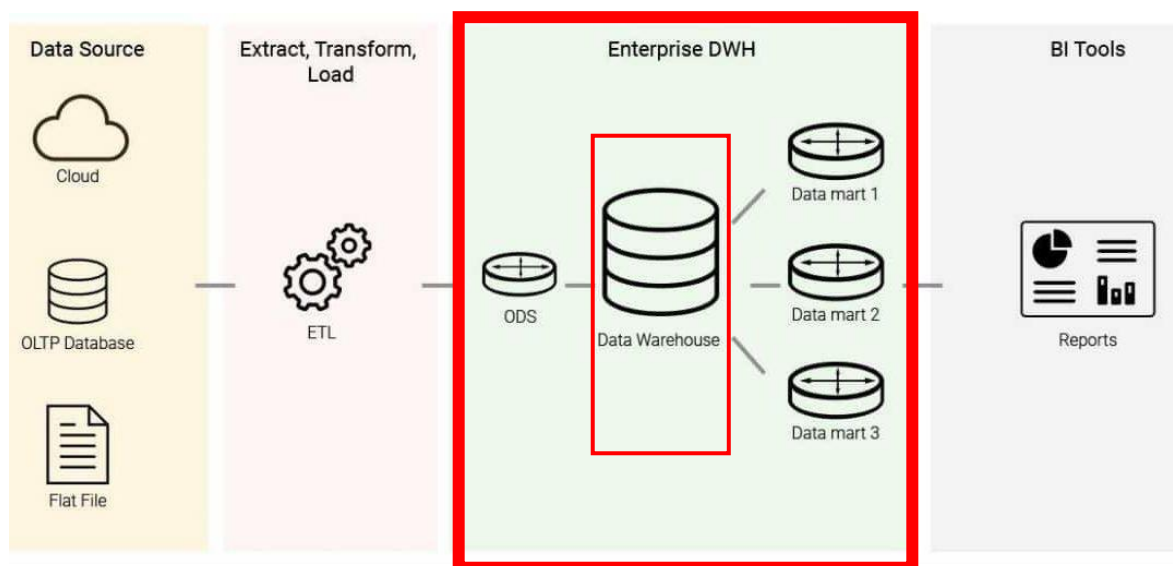


Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

## ▪ Data Warehouse

É literalmente um armazém de dados integrados, padronizados, detalhados, inteligíveis, correntes e históricos, para prover a **visão única da verdade** e suportar a tomada de decisão nos níveis estratégicos, táticos e operacionais.

**Figura 3 – Data Warehouse.**

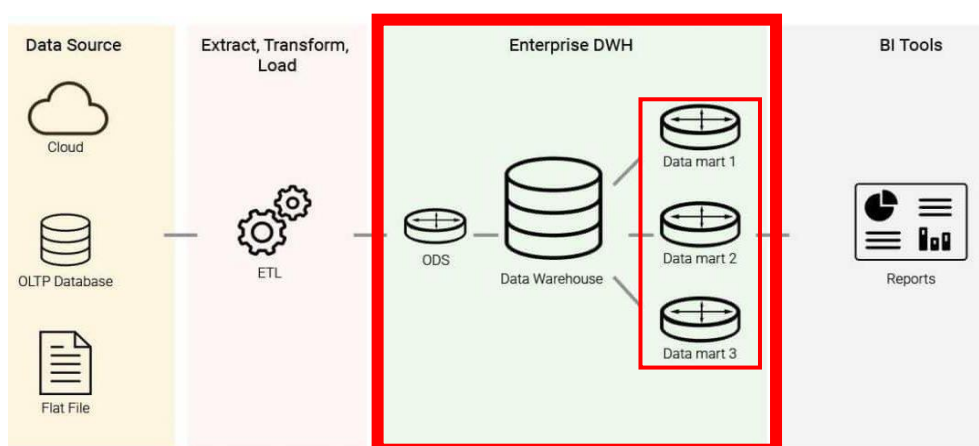


Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

#### ▪ Data Mart

É um pequeno Data Warehouse, que abrange uma determinada área de assunto e oferece informações mais detalhadas sobre o mercado (ou departamento) em questão.

**Figura 4 – Data Mart.**

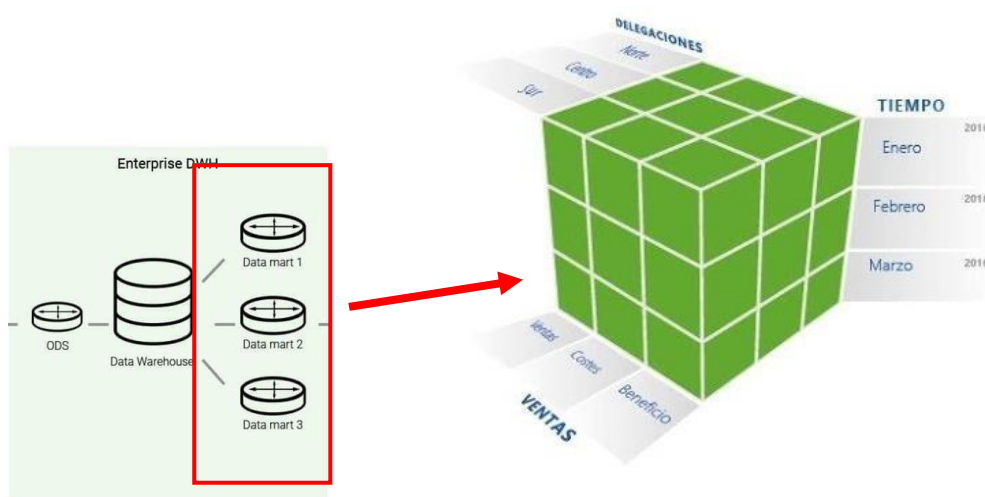


Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

## ▪ Modelagem de Dados

Prática de organização para visualizar dados abstratos de forma simples e relacionar informações de diferentes assuntos de maneira eficaz.

**Figura 5 – Modelagem de Dados.**

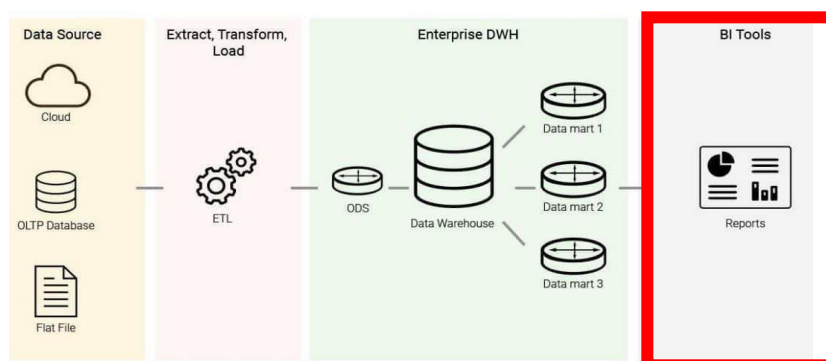


Fonte: Adaptado de <http://www.factorysoft.com.ve/>.

## ▪ OLAP

Online Analytical Processing se trata da capacidade de analisar grandes volumes de informações nas mais diversas perspectivas dentro de DW. Faz referência às ferramentas para visualização das informações gerenciais.

**Figura 6 – OLAP.**



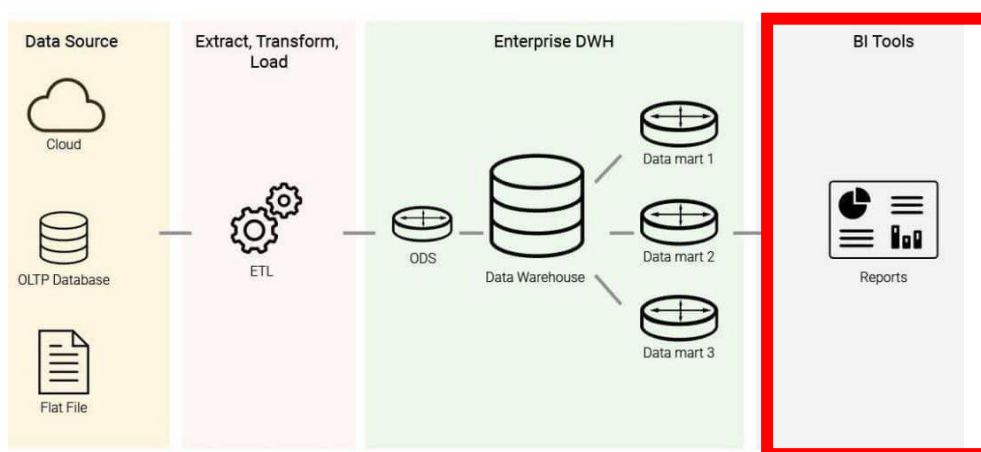
Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.



## ▪ Ferramentas Front-End

Soluções de análise de negócios que permitem conectar às fontes de dados, visualizá-los de forma amigável e dar vida a eles com painéis e relatórios ativos que apoiem tomadas de decisão.

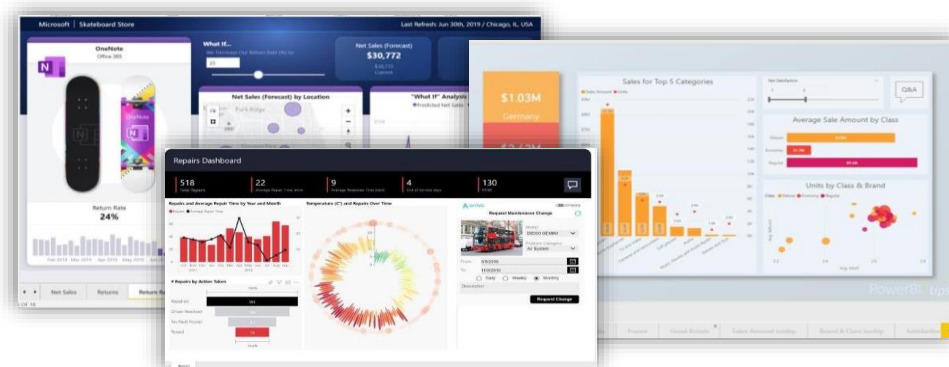
**Figura 7 – Ferramentas Front-End.**



Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

As ferramentas front-end possuem um objetivo muito claro de proporcionar análises intuitivas, manipulação fácil, acessibilidade, multiusuários desenvolvendo e consumindo, compartilhamento, relatórios dinâmicos e interativos, transparência.

**Figura 8 – Trabalhos com Ferramentas Front-End.**



Fonte: <https://community.powerbi.com/t5/Data-Stories-Gallery/>.



## A importância do BI

---

### ▪ Benefícios de Business Intelligence

Existem alguns problemas no processo de análise de dados nas empresas que são abraçados e resolvidos com boas práticas de BI:

- Demora na consolidação e entrega de informações.
- Informações conflitantes e sem qualidade, obtidas de fontes diferentes.
- Grande demanda para a área de TI, que constrói relatórios pré-definidos.
- Relatórios estáticos e engessados que não facilitam a geração de insights gerenciais.

Estamos na era dos dados, portanto o excesso e variedade de dados disponibilizados em diversos lugares não pode ser um problema para as empresas, mas sim uma fonte de estratégias e vantagens competitivas.

BI tem o importante papel de usar de ferramentas, aplicações e infraestruturas para transformar o grande volume de dados disponíveis em inteligência. Com isso, é possível conhecer melhor a concorrência, os parceiros, os fornecedores ou até mesmo ingressar em novos negócios. Um dos grandes benefícios também é a possibilidade de conhecer os sucessos e falhas internos e externos para aprender pelo exemplo, se antecipar às mudanças do mercado e contornar situações com mais assertividade.

Os processos de coleta, organização, análise e disseminação da Business Intelligence têm o objetivo de ajudar os empresários a utilizar dados de maneira estratégica para ajudar na tomada de decisões e no planejamento dos negócios. **Monitoramento e análise de dados é determinante para o sucesso de qualquer empresa.**

## ▪ BI como ferramenta de negócio

Enquanto as fontes primárias de informação são recolhidas dentro da própria empresa, as secundárias referem-se:

- À necessidade do público que consome os produtos/serviços da empresa;
- Às decisões do cliente;
- Às condições industriais mais importantes;
- Aos desafios competitivos;
- Às questões econômicas, tecnológicas e às tendências culturais.

Trata-se de uma **comparação entre sistemas internos** ou mesmo de uma empresa concorrente — nesse caso, o BI torna-se ainda mais estratégico, já que permite a coleta e a interpretação de dados sobre a concorrência.

## ▪ Área de TI x Área de negócio

Relacionamento que precisa de uma ponte que transmita com clareza as necessidades das duas áreas. Por isso o Self-Service BI hoje tem tanta força.

**Figura 9 – Área de TI x Área de Negócio.**

ÁREA DE TI	ÁREA DE NEGÓCIO
Redução do tempo gasto na criação de relatórios.	Liberdade para manipular dados em seu próprio computador.
Liberação dos recursos para análises mais avançadas.	Experiência com diversos relatórios que podem ser construídos a partir dele.

**Fonte: Elaborado pelo autor.**

## As empresas podem sobreviver sem BI?

### ▪ BI é indispensável?

Existem três principais motivos que tornam BI indispensável para as empresas, e eles estão ligados à esfera de dados, informação e conhecimento.

**Figura 10 – Dados, informação e conhecimento.**



**Fonte: Elaborado pelo autor.**

Na perspectiva dos dados, há uma necessidade de eles estarem sempre atualizados, disponíveis, confiáveis e com qualidade. Essas são as quatro características que tornam BI indispensável, pois garante essas propriedades e uma melhor transformação estratégica.

A perspectiva de informação consiste na capacidade de obter dados de diversas origens para possibilita trabalhar com informações consolidadas. Isso quer dizer que é preciso saber realizar integrações para aliar os dados a uma estratégia de consolidação que permita acessar isso tudo de uma forma otimizada.

Por fim, a perspectiva do conhecimento diz respeito à descoberta de conhecimento em meio a dados crus. Isso pressupõe que os dados já estejam trabalhados e as informações consolidadas. Tudo isso não faz sentido se não tiver o objetivo claro e direto de **criar soluções que gerem resultado**. Resultado quer dizer mais controle, clareza, insights, visão de histórico, análises, possibilidades de previsão, maior gestão e suporte à tomada de decisão.

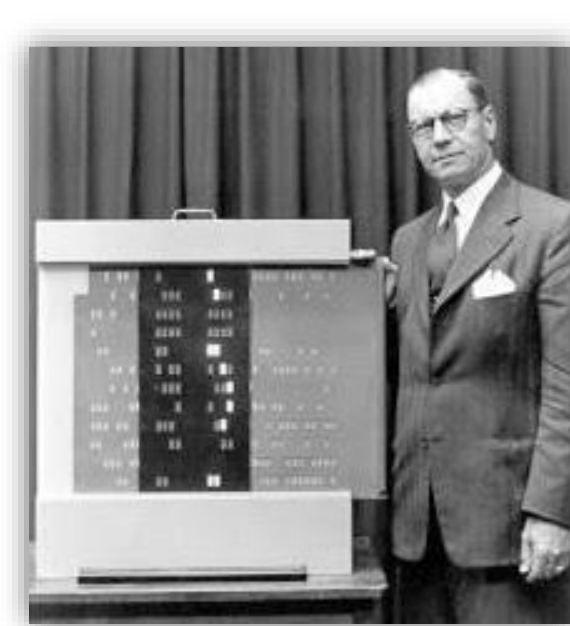
## Origens e aplicações de Business Intelligence

---

### ▪ Como surgiu o BI

O termo “Business Intelligence”, criado em 1865 por Richard Millar Devens em “Cyclopaedia of Commercial and Business Anecdotes”, foi usado para descrever como um bancário lucrava baseado em informações recuperadas do ambiente antes da concorrência, que é o funcionamento real do BI.

**Figura 11 – Richard Millar Devens.**

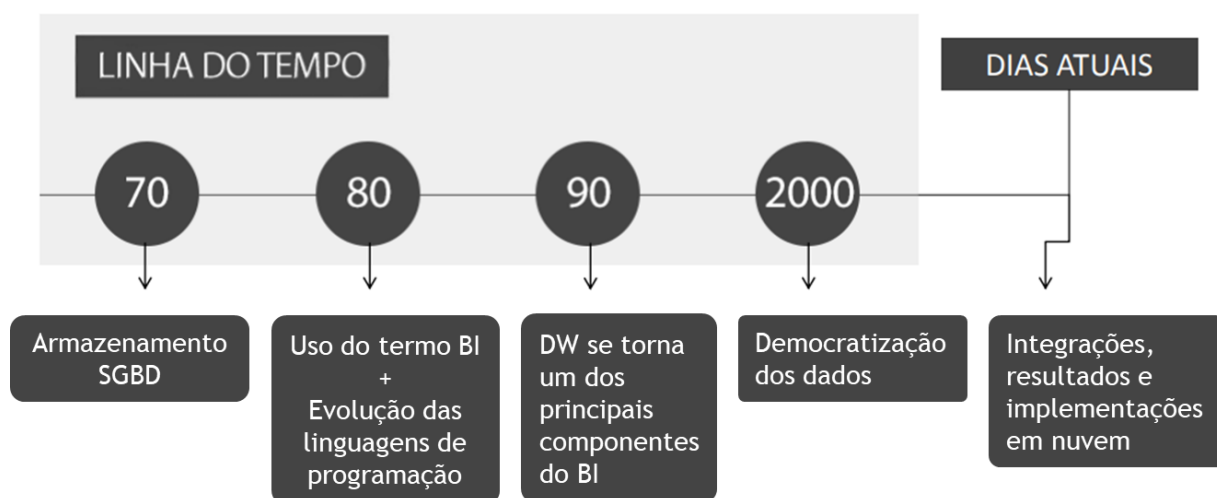


**Fonte: Google Imagens.**

Apesar das grandes mudanças, **o conceito continua o mesmo**, e com a Revolução Industrial, a alta do capitalismo e a tecnologia, foi necessário que o mundo se organizasse e analisasse a situação para que os recursos **se adaptassem a maiores produções com mais agilidade e qualidade**.

## ▪ Breve histórico sobre BI

**Figura 12 – Business Intelligence Time Line.**



**Fonte: Elaborado pelo autor.**

## ▪ 3 Cases reais

O primeiro case foi no Liverpool Futebol Clube, da Inglaterra. Um clube que num período de 3 ou 4 anos conseguiu ganhar todos os títulos mais importantes na Europa.

Dentro e fora de campo, o trabalho do time de Analytics em conjunto com o time de Pesquisa influenciou diretamente na contratação do técnico Jurguen Klopp, do meia Keyta e do atacante Salah, por exemplo, que foram frutos de cruzamentos de bases e sugestão do time de dados.

O segundo case foi a criação da série Stranger Things, que foi um produto de um cruzamento de dados que mapearam a receita da audiência. Todas as referências que abrangem todas as gerações, personagens, cenários, enredos, tudo foi fruto de análise de dados. O resultado foi a quebra de dezenas de recordes de audiência e hoje é considerada uma das series de maior sucesso da Netflix.

Por último, o terceiro case é um dos mais famosos: o do Walmart. Análises no Data Warehouse apontaram certa relação entre fraldas e cervejas, mas as

variáveis eram aparentemente desconexas. No entanto, adotaram a estratégia e tomaram a decisão de aproximar esses dois produtos fisicamente, por causa dessas análises realizadas. O resultado foi a alavancagem de vendas e atualmente esse é um case modelo de aplicação de análise em DW para tomada de decisão orientada a dados. Detalhe: Isso aconteceu há quase 30 anos.

## Capítulo 2. Arquitetura e soluções de Business Intelligence

---

Neste capítulo vamos trabalhar os benefícios e vantagens das soluções de BI, o que na prática são essas soluções e quais as principais arquiteturas aplicadas para elas.

### Benefícios e vantagens das soluções de BI

---

- **Geração de Valor**

A interpretação e análise de dados para identificar oportunidades e riscos é uma forma de integrar o negócio, o gerenciamento e a TI. Estruturar um planejamento estratégico constrói vantagens competitivas sustentáveis que agregam valor ao negócio e às soluções oferecidas aos clientes. Essa é uma premissa de qualquer solução de Business Intelligence.

**Figura 13 – BI como intercessão de TI e negócio.**



**Fonte: Elaborado pelo autor.**

- **10 benefícios e vantagens de soluções de BI**

- 1) Fim do achismo, sem muitas dificuldades é possível mapear os riscos e adiantar decisões baseadas em acontecimentos concretos.



- 2) Usuários não chegarão numa reunião tendo que explicar divergências de números. Isso traz segurança nas análises e um foco naquilo que importa: Resultado.
- 3) A velocidade e a disponibilidade dos processos que envolvem a consolidação das informações reflete na tomada de decisão estratégica.
- 4) Independente do setor, categoria, região, cargo, todos verão o mesmo número em última instância.
- 5) Conhecendo o próprio negócio, há mais transparência e uma clareza maior ao analisar o mercado, principalmente ao usar recursos de extração de dados externos à empresa.
- 6) O retorno sobre o investimento chega principalmente na economia de recursos proveniente da automatização e gestão de indicadores.
- 7) Quanto maior a qualidade dos dados, menor a chance de erros analíticos. Dados incorretos refletem decisões fatais, dependendo.
- 8) A visão histórica e presente interna possibilita o autoconhecimento do negócio. O cruzamento com dados públicos externos e cenário trazem o conhecimento do mercado e rivais
- 9) Cada vez mais na palma da mão, com agilidade, facilidade e responsividade.
- 10) Um processo que 5 funcionários demoram 2 semanas para fazer pode facilmente virar uma rotina de dois dias feita por 2 funcionários. Isso é economia de dinheiro.

## Soluções de BI

---

- **O que as soluções precisam ter**

Soluções de BI aplicam métodos de agregação, análise e disponibilização de dados coesos para auxiliar em decisões estratégicas ou operacionais. Portanto, ao cruzar dados, abre-se uma visão panorâmica e global do negócio e é proporcionada uma “inteligência” que não seria possível se apenas olhássemos para um único conjunto de dados isolado.

Cruzamento e agregação de dados de fontes externas, internas e de vários tipos (financeiros, produção, vendas, demográficos, econômicos, etc.) para ganhar visão sobre novos mercados, processos, serviços e produtos. Com esta visão estratégica, entre várias potencialidades, o BI permite às organizações, por exemplo, saber como se enquadram no mercado, ganhar visão sobre novos mercados, controlar processos, analisar a procura de serviços e produtos para novos segmentos de mercado.

Soluções precisam possuir bases para previsão de tendências e impacto de ações comerciais. Ao estabelecer um histórico sólido, soluções de Analytics podem ser aplicadas com mais facilidade, de acordo com o aumento do amadurecimento da organização com relação aos seus dados.

- **Principais soluções de mercado**

As principais soluções de mercado no que diz respeito a plataformas de Business Intelligence e Analytics são: Power BI, QlikView, QlikSense, Tableau, TIBCO Spotfire, MicroStrategy, Salesforce, Google DataStudio, IBM Cognos, SAP BO, Sisense, SAS, Thoughtspot e Amazon QuickSight.

**Figura 14 – Quadrante mágico do Gartner – BI e Analytics.**



**Fonte: Gartner (Fevereiro 2020).**

#### ▪ Iniciação gratuita

As soluções que considero mais recomendadas para iniciar de maneira gratuita e que possibilitam um aprendizado sólido, guiado e intuitivo, são as ferramentas Google Data Studio, Tableau, Qlik Sense e principalmente Power BI.

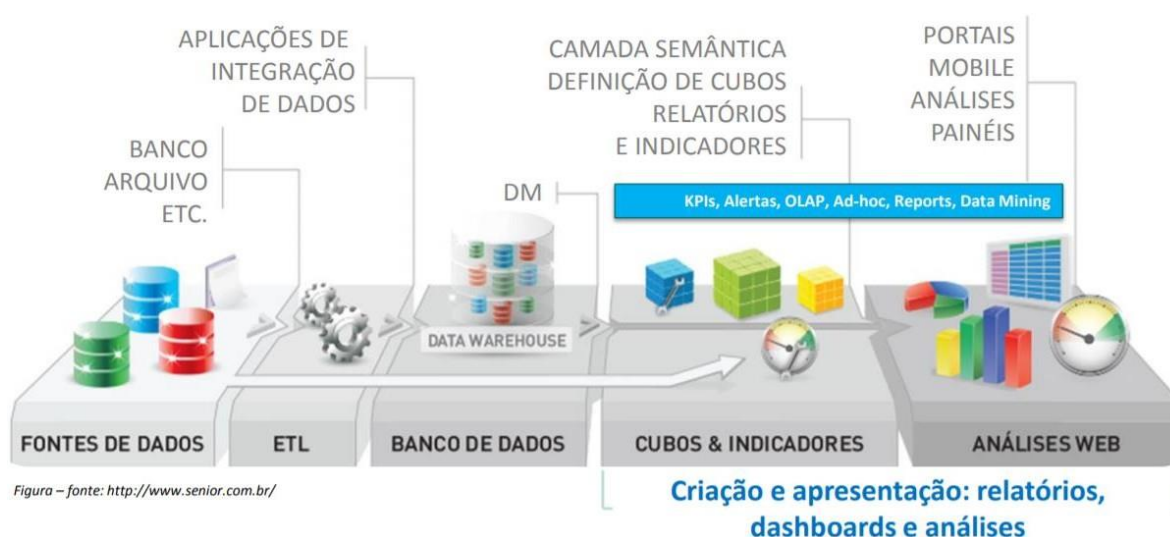
Essas são ferramentas que possuem comunidades fortes ao redor do mundo, documentações detalhadas e muitos materiais de consulta, além de fóruns e templates para soluções locais. Possuem ambientes em nuvem para práticas e também muitos mecanismos que agregam valor a aqueles que são menos autodidatas.

## Partes da arquitetura de BI

### ▪ Sobre a arquitetura de BI

A arquitetura de BI é composta por vários processos, infraestruturas e boas práticas para concretizar tudo o que vimos a respeito da transformação de dados em conhecimento.

**Figura 15 – Arquitetura de BI.**



Fonte: <https://www.senior.com.br/>.

A parte de ETL conta com várias ferramentas como Pentaho, SSIS, Alteryx (mesmo não sendo nativamente de ETL), Talend, entre outras. O armazenamento é muito comum em bancos relacionais como SQL Server, Oracle, MySQL, Firebird, Postgree e também não relacionais como NoSQL. Outro componente dessa arquitetura é a modelagem semântica, exemplificada por cubos e indicadores. O SQL Server Analysis Services é um ambiente muito recomendado para essa parte. A última camada é a mais visual, retratada por reports, dashboards e ferramentas já citadas anteriormente e que iremos detalhar em capítulos posteriores.

A seguir trabalharemos resumidamente as principais características e motivações para utilização de Staging, Data Lake, Data Warehouses, Data Marts e Data Sandboxes.

## ▪ Staging Area

**Figura 16 – Resumo Staging Area.**

CARACTERÍSTICAS	POR QUE USAR?
Área onde os dados são colocados após a extração a partir dos sistemas de origem.	<b>Acesso ao sistema fonte pode ser feito uma única vez</b> , as leituras necessárias para o armazenamento dos dados do DW são realizadas a partir da Staging Area.
Mantém uma “foto” parcial, temporária ou permanente dos dados dos sistemas fontes.	Os dados da Staging Area são úteis para validar a integridade da informação do DW quando o sistema fonte sobrescreve o histórico.
Dedicada para a integração de dados e não disponível para os usuários finais do DW.	Após a gravação dos dados na Staging Area <b>não é necessário acessar o sistema fonte novamente</b> no caso de falha.

Fonte: Elaborado pelo autor.

## ▪ Data Lake

**Figura 17 – Resumo Data Lake.**

CARACTERÍSTICAS	POR QUE USAR?
Repositórios de Big Data que armazenam dados brutos e fornecem funcionalidade para integração sob demanda com a ajuda de descrições de metadados.	<b>Suportar dados não estruturados e estruturados e</b> prover uma estrutura de fácil adaptação quanto a mudanças.
Armazena dados na forma mais pura e pode atender a vários tipos de usuários.	Há demandas Adhoc em que não é possível definir quais dados ou quais requisitos serão necessários para implementação de relatórios avançados.
	Suportar demandas de diferentes tipos de usuários.

Fonte: Elaborado pelo autor.

## ▪ Data Warehouse e Data Marts

**Figura 18 – Resumo Data Warehouse e Data Marts.**

CARACTERÍSTICAS	POR QUE USAR?
Armazém de dados centralizado que irá prover informações táticas e estratégicas detalhadas de toda empresa.	Visão única da verdade. Flexibilidade. Alta disponibilidade e escalabilidade.
É um sub-conjunto de dados de um Data Warehouse, referentes a um assunto em especial ou diferentes níveis de sumarização que foca em uma ou mais áreas específicas.	Foco no desempenho, por atuar num assunto específico.

**Fonte: Elaborado pelo autor.**

“O Data Warehouse não é nada mais do que a união dos Data Marts que o constituem. ” - Ralph Kimball.

## ▪ Data Sand Box

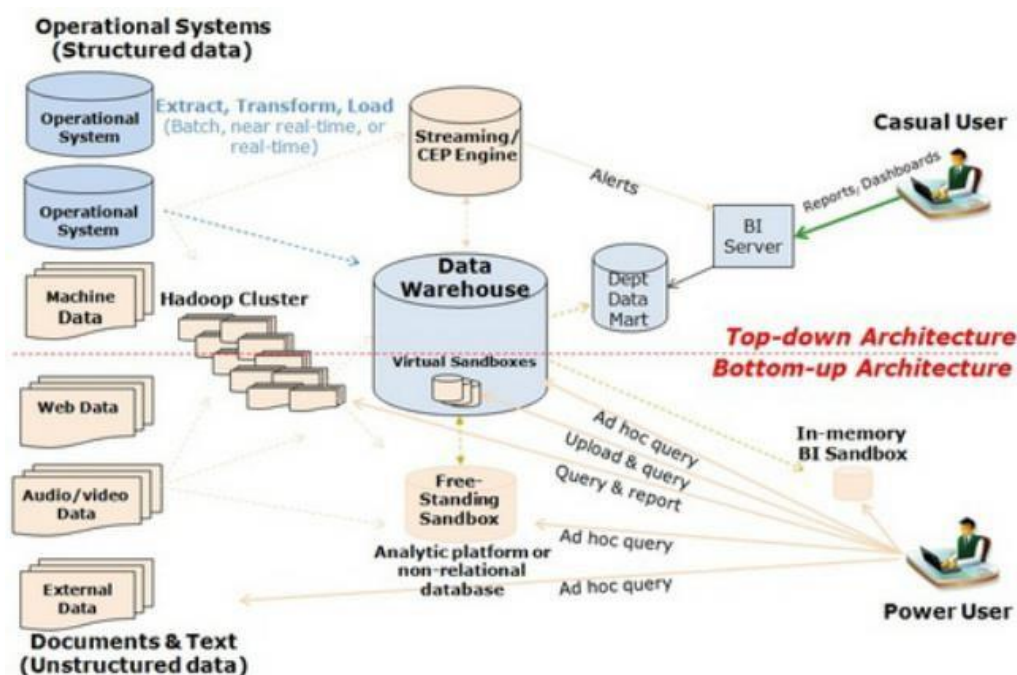
**Figura 19 – Resumo Data Sand Box.**

CARACTERÍSTICAS	POR QUE USAR?
Não há backup.	Usuários de BI avançados fazem exploração intensiva dos dados disponíveis.
Não há suporte dedicado.	As vezes a camada semântica, relatórios ou dashboards pré-definidos não atendem suas demandas.
Quando esgota a capacidade, 'algo' deve deixar a Sand Box.	Manter as demandas de <b>usuários comuns e usuários avançados</b> no mesmo componente arquitetural pode ser muito arriscado.
Não há obrigatoriedade do dado ser limpo ou mesmo modelado de forma específica: depende da função.	

**Fonte: Elaborado pelo autor.**

Uma peculiaridade da implantação de Data Sand box, pois essa arquitetura permite uma abordagem mais Bottom-Up de desenvolvimento de relatórios por parte do usuário, trazendo uma estrutura em memória e requisições personalizadas e livres, empoderando os usuários de negócio.

**Figura 20 – Arquitetura de Analytics Sand box.**



Fonte: <https://twitter.com/bigdatablender/status/726509010191585280>.

## Como definir a melhor arquitetura

### ▪ Clusterização corporativa

Para utilizar melhor a tecnologia no auxílio ao negócio, é preciso analisar o cenário e distinguir os tipos de empresas. Por exemplo, vamos pegar três tipos de empresas:

- Padaria do Tião (3 unidades);
- Ricardo Eletro;



- Microsoft.

O Tião tem 3 unidades da padaria dele. A arquitetura não vai ser a mesma do Ricardo Eletro que não vai ser a mesma da Microsoft. No entanto, todas precisam utilizar soluções de BI, mas essa segmentação deve ter relação mais direta com a quantidade que as empresas **processam e trabalham os dados**.

#### ▪ Data Literacy

O Gartner define alfabetização de dados como a capacidade de ler, escrever e comunicar dados em contexto e a capacidade de descrever o caso de uso, a aplicação e o valor resultante.

A maturidade da empresa na área de dados é medida pela capacidade e o desejo de um funcionário de usar a tecnologia existente e emergente para gerar melhores resultados nos negócios.

Ou as empresas tem legados ou estão iniciando um novo ciclo de desenvolvimento.

Algumas perguntas podem ajudar a determinar uma clusterização mais otimizada e definir uma arquitetura mais adequada para cada tipo:

- Qual o tamanho da necessidade da empresa em conhecer seus dados e trabalhar com eles?
- Isso deve ser feito mensalmente, semanalmente, diariamente ou em tempo real?

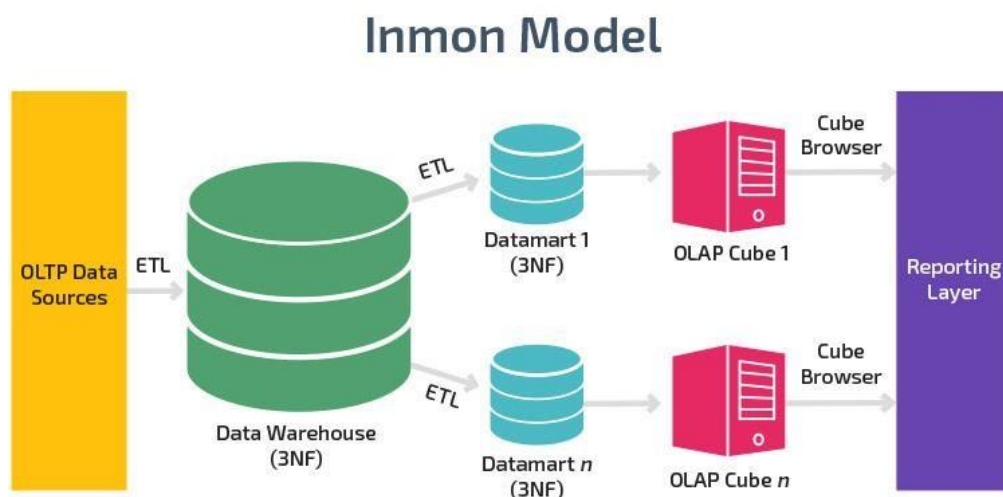
Então, teremos três tipos de empresas: as que precisam de BI para se inserirem na cultura data driven, as que precisam com menos latência e as que precisam em tempo real.

#### ▪ Definições de arquitetura

Existem muitas delas disponíveis no mercado, e as mais conhecidas são as do Inmon e do Kimball. Enquanto o Inmon olha para uma solução mais corporativa com

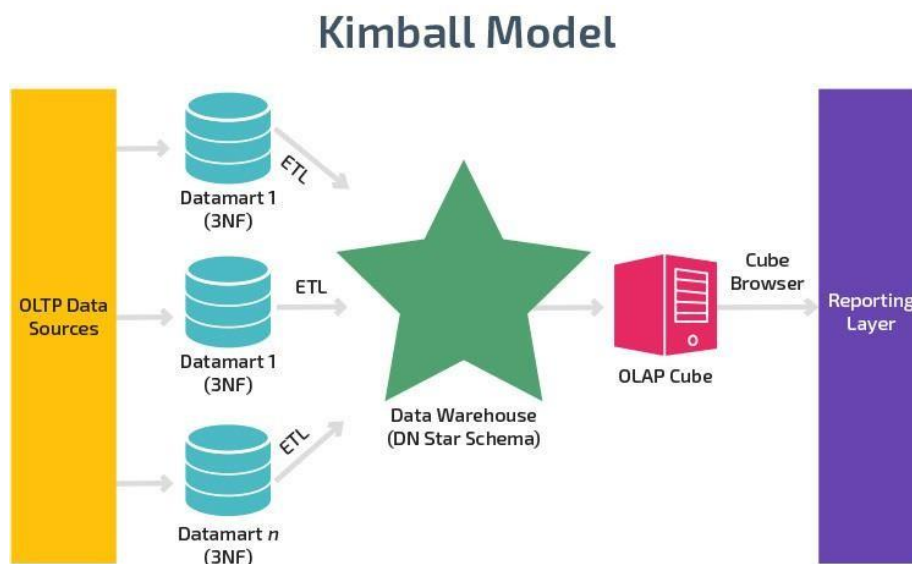
diversos componentes e mais complexa, o Kimball simplifica a criação, mas minimiza nas possibilidades.

**Figura 21 – Modelo Inmon.**



Fonte: <https://www.quora.com/>.

**Figura 22 – Modelo Kimball.**



Fonte: <https://www.quora.com/>.

Temos, ainda, vários componentes tecnológicos que são adicionados quando tratamos de Data Lake, NoSQL, Big Data, Hadoop, IoT e ML, mas como solucionar as questões de Inteligência de Negócios para todos esses cenários?

Por isso existem diversas consultorias no mundo. Depende muito do volume de dados, do investimento, da distribuição dos dados, da governança, da qualidade, latência, etc.

A padaria do Tião, por exemplo, pode precisar apenas de um acesso aos sistemas transacionais, de uma arquitetura simplificada e de acesso direto aos dados com ferramentas OLAP. Porém o problema está as empresas com sistemas críticos e excesso de dados, pois necessitam de técnicas de replicação que minimizem o impacto dos sistemas transacionais, além de determinar como e onde armazenar os dados para as diversas análises.

Não existe uma receita de bolo para definir a melhor arquitetura, a questão é entender as necessidades e planejar as ações.

## Capítulo 3. Aplicações ETL

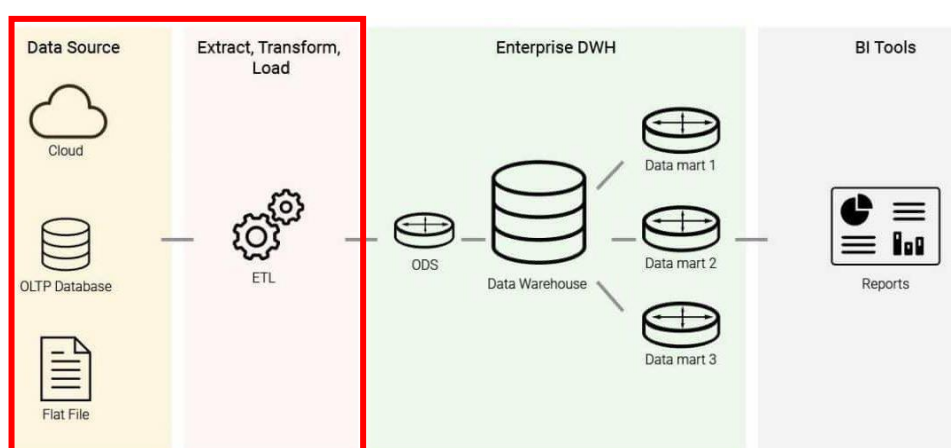
Neste capítulo vamos trabalhar o que é o processo de ETL, a sua importância nos projetos de BI e entender as principais aplicações, benefícios, vantagens, metodologias, boas práticas e implementações reais de mercado que tangem essa área que representa 70% de um projeto de BI.

### Definição de ETL

#### ▪ O que é o processo de ETL

- **EXTRAÇÃO:** Fase em que os dados são extraídos dos OLTPs e conduzidos para a staging area (área de transição ou área temporária), onde são convertidos para um único formato.
- **TRANSFORMAÇÃO:** É nesta etapa que realizamos os devidos ajustes, podendo assim melhorar a qualidade dos dados e consolidar dados de duas ou mais fontes.
- **CARGA:** Consiste em fisicamente estruturar e carregar os dados para dentro da camada de apresentação, seguindo o modelo dimensional.

**Figura 23 – O processo de ETL.**

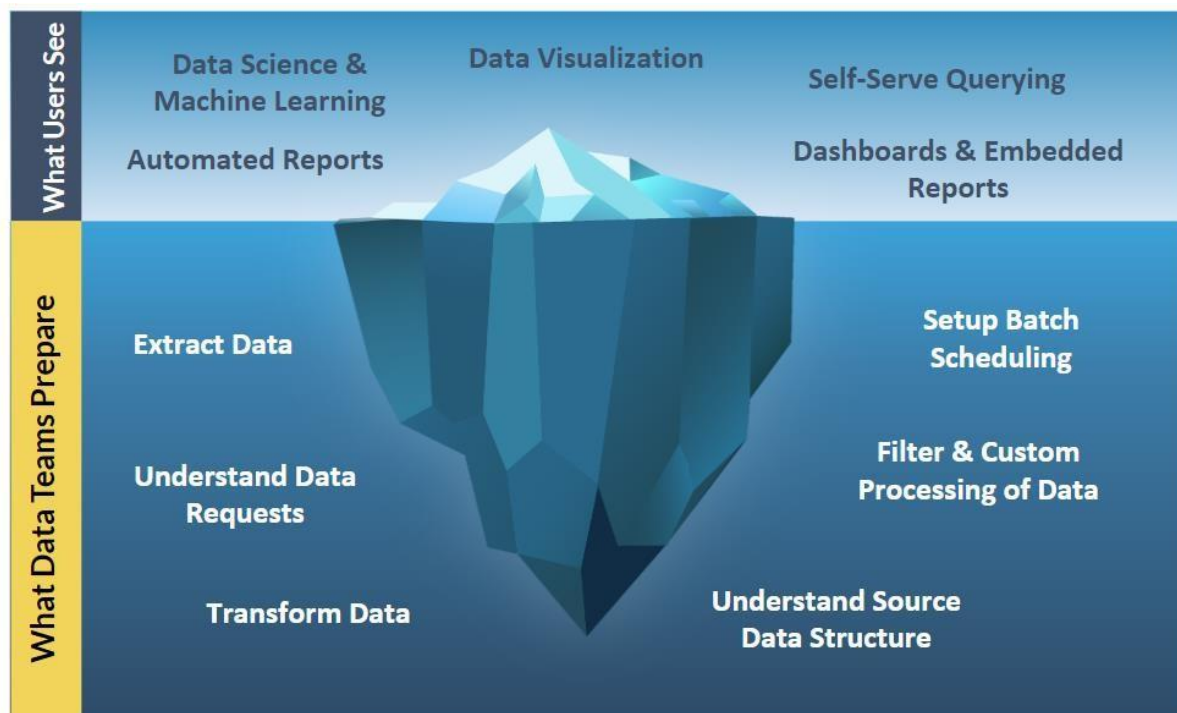


Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

## ▪ A importância do ETL nos projetos de BI

ETL é um dos processos mais importantes em uma solução de BI, chegando às vezes a corresponder a 70% de um projeto.

**Figura 24 – ETL iceberg.**



Fonte: <https://www.holistics.io>.

Para ajudar no desenvolvimento ETL temos hoje infinitas possibilidades de ferramentas no mercado, todas com o objetivo de padronizar e facilitar esse desenvolvimento.

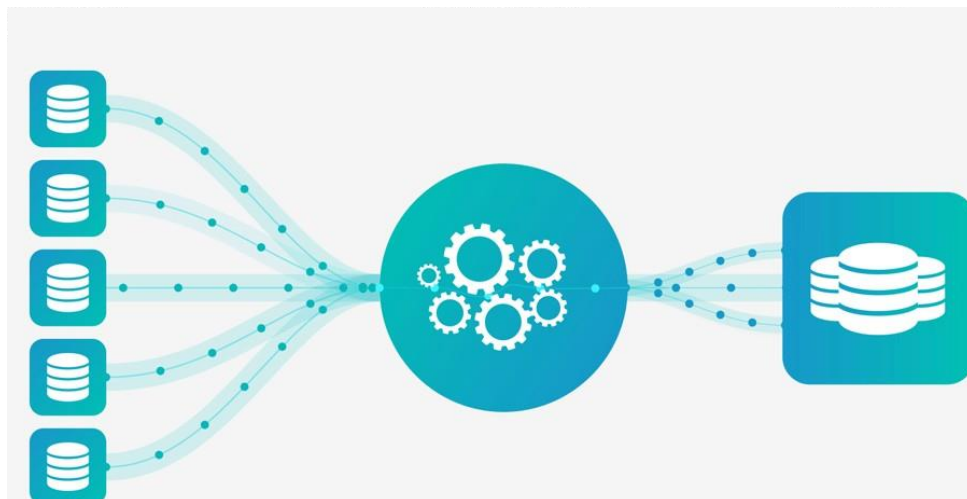
Devemos escolher bem as ferramentas que darão suporte ao processo, pois são essenciais para a correta execução das atividades do ETL.

## Como funciona o processo de ETL

### ▪ Extração

1º passo: Entendimento das origens e fontes de extração. Dados podem vir de diversas fontes diferentes, sejam servidores, arquivos avulsos, terminais, web, rotinas, nuvem ou outras origens.

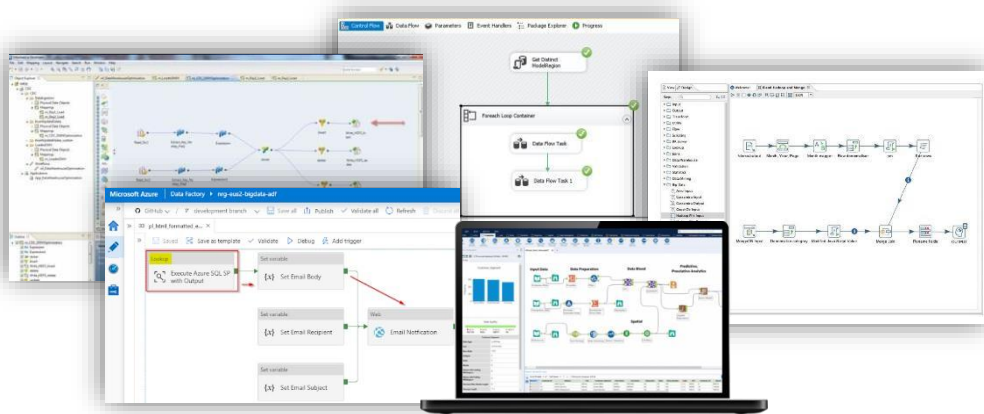
**Figura 25 – 1º passo ETL.**



Fonte: <https://blog.bismart.com/en/what-do-we-do-etl>.

2º passo: Definir a ferramenta que vai fazer a extração.

**Figura 26 – Definição de ferramentas.**



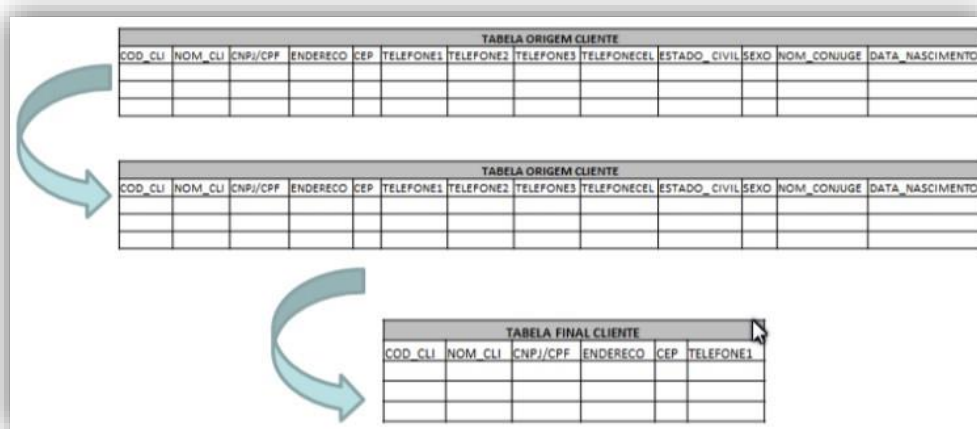
Fonte: Elaborado pelo autor.

É importante avaliar a necessidade da empresa, o custo da ferramenta, o custo da mão de obra, os serviços, as plataformas, os recursos, as integrações e as compatibilidades.

### ▪ Transformação

É nesta etapa que realizamos os devidos ajustes, podendo assim melhorar a qualidade dos dados e consolidar dados de duas ou mais fontes.

**Figura 27 – Transformação e regras de negócio.**



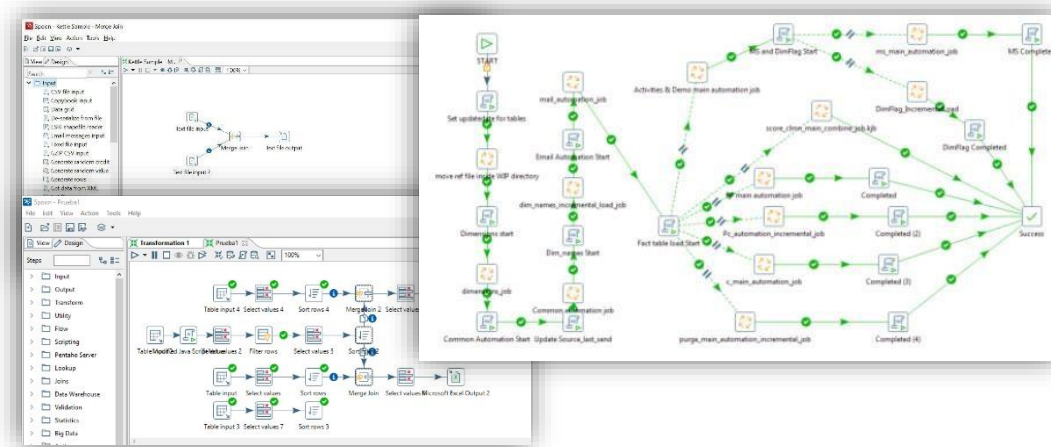
Filtros de campos, períodos, regras de negócio, conversões de tipos e valores, cruzamento de origens, classificações, agrupamentos e definições de carga são ajustados, por exemplo configuração full ou incremental de carga, entre outras propriedades.

### ▪ Carga

Fase responsável por estruturar e carregar dados na Staging Area, ODS, Data Warehouse ou Data Mart. Essa é uma etapa que contém amplas variações e realiza toda a parte de estruturação lógica do processo, desde os inputs até as transformações, cálculos e outputs.



**Figura 28 – Processo de Carga.**



Fonte: Adaptado de <https://www.hitachivantara.com>.

## Vantagens de uma aplicação de ETL

- **Características de padronização**

Quanto mais a área de BI cresce, observamos diferentes aplicações de arquiteturas. Características comuns de arquiteturas e quatro pilares, que dão sentido e tornam ETL um processo vantajoso para qualquer empresa, são:

- **Agilidade** tanto na parte de automatização de rotinas quanto de alinhamento estratégico (metodologia ágil).
- Foco 100% na transformação de dado para conhecimento, insumos para **tomada de decisão**
- **Volume** de dados (Big Data ou não).
- **Variedade de origens** de dados, sejam não estruturados ou estruturados, marketing, logística, independente do negócio.

## ▪ Benefícios de aplicações

Desenvolver em aplicações ETL traz benefícios bem interessantes por causa da intuitividade de interface, tornando fácil o gerenciamento, a documentação e a padronização dos fluxos. Além disso, há a questão dos controles especiais de carga, realizando a estruturação, os agendamentos, a documentação, a manutenção, os diagnósticos e os reports de carga.

O valor agregado também é refletido na redução de tempo e custo, uma vez que quanto menos tempo se gasta com o desenvolvimento, mais economia gera para os projetos. Tempo é dinheiro de fato. A performance também é um outro fator que as aplicações oferecem, pois os processos rodam de maneira ágil, integrando centenas de conectores de origens variadas operando em perfeita sintonia e compatibilidade.

## Plataformas codificadas x Low Code

---

### ▪ Vantagens e desvantagens de “codar”

São seis vantagens que podem ser identificar ao se trabalhar com plataformas que exigem conhecimentos de programação:

**Controle:** Desenvolvendo tudo do zero, há um controle completo sobre o processo que evita qualquer tipo de desconhecimento de como uma determinada parte da carga funciona. Todo o processo é conhecido e controlado de ponta a ponta, e pode ser modificado da forma desejada quando necessário.

**Customização:** As possibilidades de customização são infinitas, pois, todo o código do processo de carga estará disponível para eventuais mudanças. Isso inclui novas funcionalidades na carga, adaptação a algum framework interno da empresa e até integrações individuais.

**Convergência de Plataforma:** Não será necessário adquirir hardware, software ou um sistema operacional à parte para adaptar a ferramenta de ETL ao

direcionamento tecnológico da organização. A criação de um processo de ETL "hard code", permite a escolha da linguagem de programação que a organização julgar mais adequada.

**Convivência com o legado:** O desenvolvimento interno sempre se adaptará aos sistemas legados e nunca será impositiva a mudança do legado para adaptar-se às cargas e (ou) construção de pequenos módulos para realizar essa integração.

**Suporte:** A construção própria dispensa contratos de suporte e manutenção com o fabricante. Você também não será surpreendido pelo fato de um fabricante simplesmente descontinuar o produto e não comercializar mais suporte.

**Debug:** As atividades de debug com alta disponibilidade, facilidade de manutenção e testes, pois não haverá, no processo, partes cujo código fonte não esteja acessível.

A parte ruim, classifico como a manutenção, que é complexa e mais demorada; a disseminação do conhecimento, que é limitado, pois depende de skills técnicas muito específicas; e a evolução difícil e também limitada.

#### ▪ **Vantagens e desvantagens de *Low Code***

É possível elencar quatorze vantagens muito claras e presentes em aplicações Low Code.

**Desenvolvimento das cargas:** Desenvolver uma rotina de carga em uma ferramenta de ETL é muito mais fácil e rápido que codificá-la. Dependendo da facilidade da ferramenta, é possível inclusive que usuários não técnicos a utilizem para cargas mais simples.

**Manutenção das cargas:** As tarefas de manutenção de uma rotina de carga são mais fáceis de realizar em relação à manutenção de código.

**Desempenho:** As ferramentas de ETL utilizam métodos mais performáticos para trabalhar com grandes volumes, e normalmente conseguem extrair, transformar

e carregar dados com mais velocidade e menos utilização de recursos. Isso inclui operações não logadas, gravações em bloco, etc.

**Execução em paralelo:** Ferramentas de ETL possuem recursos de paralelização nativos e facilmente implementáveis.

**Escalabilidade:** Ferramentas de ETL podem ser transferidas de servidor mais facilmente e até, eventualmente, distribuir sua carga entre vários servidores.

**Diversidade de conectores:** A conexão de uma ferramenta de ETL com múltiplas fontes de dados é transparente. Caso apareça alguma fonte não trivial como o SAP, Mainframe, Hadoop, Twitter, etc., **é possível adquirir o conector sem a necessidade de codificar um.**

**Funcionalidade x manipulação:** Uma ferramenta de ETL já possui suas funcionalidades disponíveis (Lookup, Merge, Split, Expressões calculadas, etc.). Só é necessário concentrar-se em como fluir os dados dentro da carga e não codificar cada tarefa da carga.

**Reusabilidade:** Uma carga normalmente pode ser reaproveitada dentro de outras cargas, ou sobre a forma de um template.

**Reinicialização:** Ferramentas de ETL possuem a capacidade de reiniciar a carga de onde pararam sem a necessidade de codificar essa inteligência.

**Manutenção de metadados:** Os metadados são gerados e mantidos automaticamente com a ferramenta, evitando que problemas de conversão gerem dados não íntegros ao final do processo. A manutenção de metadados também evita ou alerta para alterações de esquema que invalidem a carga.

**Documentação:** As ferramentas de ETL possuem mecanismos de documentação (quando não são autoexplicativas). Isso pode ser um diferencial significativo, principalmente para equipes de alta rotatividade.

**Qualidade dos dados:** Ferramentas de ETL podem disponibilizar meios para **trabalhar a qualidade dos dados através de algoritmos complexos** (lógica fuzzy, IA, etc.).

**Auditoria e Tracking:** É possível implementar recursos de auditoria e tracking para **conhecer de onde veio o registro, que transformações sofreu e como foi carregado**.

**Segurança:** É possível tornar a segurança mais modular, dividindo-se os papéis (criação de cargas, execução de cargas, agendamento, etc.).

Temos três desvantagens muito presentes no mercado quanto às soluções Low Code.

**Alto custo:** O preço de uma licença, dependendo da ferramenta, pode chegar a 30% do custo total de um projeto de ETL. Tendo em vista que os patrocinadores de um projeto normalmente são usuário de negócio, **justificar a compra de um ETL talvez fique inviável**.

**Profissionais capacitados:** Além de termos que encontrar profissionais com conhecimento em SQL, algumas ferramentas são bastante complexas e precisamos de profissionais capacitados nelas, o que pode encarecer o custo do recurso. Em alguns casos, é mais eficiente termos um ETL baseado em SQL e contratarmos um bom profissional na linguagem.

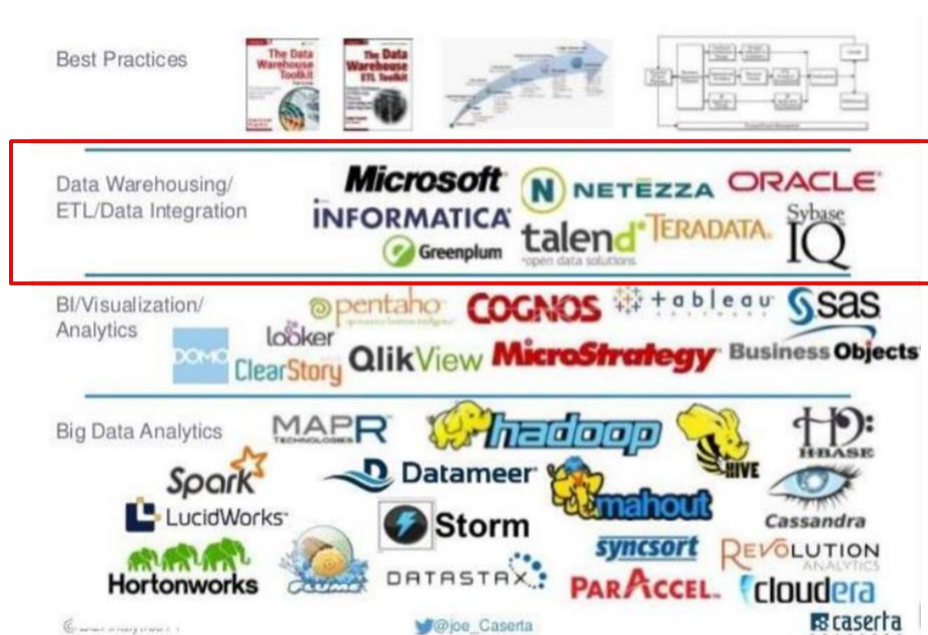
**Reduzir a flexibilidade:** Uma solução fornecida obviamente fica dependente da capacidade de desenvolvimento de seu fabricante, assim como da dependência na resolução de problemas encontrados e atualizações.

## Ferramentas ETL e principais fornecedores

---

- **Ferramentas ETL e fornecedores**

Figura 29 – Ferramentas ETL.



Fonte: Adaptado pelo autor.

Se considerarmos os principais fornecedores ainda temos centenas de ferramentas, como mostra a figura abaixo:

Figura 30 – Ferramentas ETL 2 – Zoom In.



Fonte: Adaptado pelo autor.



O Quadrante mágico do Gartner nos mostra que ferramentas como Microsoft, Qlik Sense, Knime Analytics, Rapid Miner e até mesmo o Alteryx estão em alta no mercado, considerando as avaliações de negócio.

**Figura 31 – Quadrante mágico Gartner ferramentas de ETL.**



**Fonte: Gartner, 2019 (agosto).**

#### ▪ Microsoft – SSIS e Azure

O Microsoft SQL Server Integration Service é uma plataforma para criar integração de dados em nível corporativo e soluções de transformações de dados. Possui algumas características importantes, como as elencadas a seguir:

- Solucionar problemas empresariais complexos.
- Copiar ou baixar arquivos.
- Enviar mensagens de e-mail em resposta a eventos.



- Atualizar o Data Warehouse.
- Realizar limpeza e mineração de dados.
- Gerenciar objetos e dados do SQL Server.
- Extrair e transformar dados de uma ampla variedade de fontes, como arquivos de dados XML, arquivos simples e fontes de dados relacionais.
- Transferir dados para um ou mais destinos.

**Figura 32 – Microsoft no Quadrante mágico Gartner.**

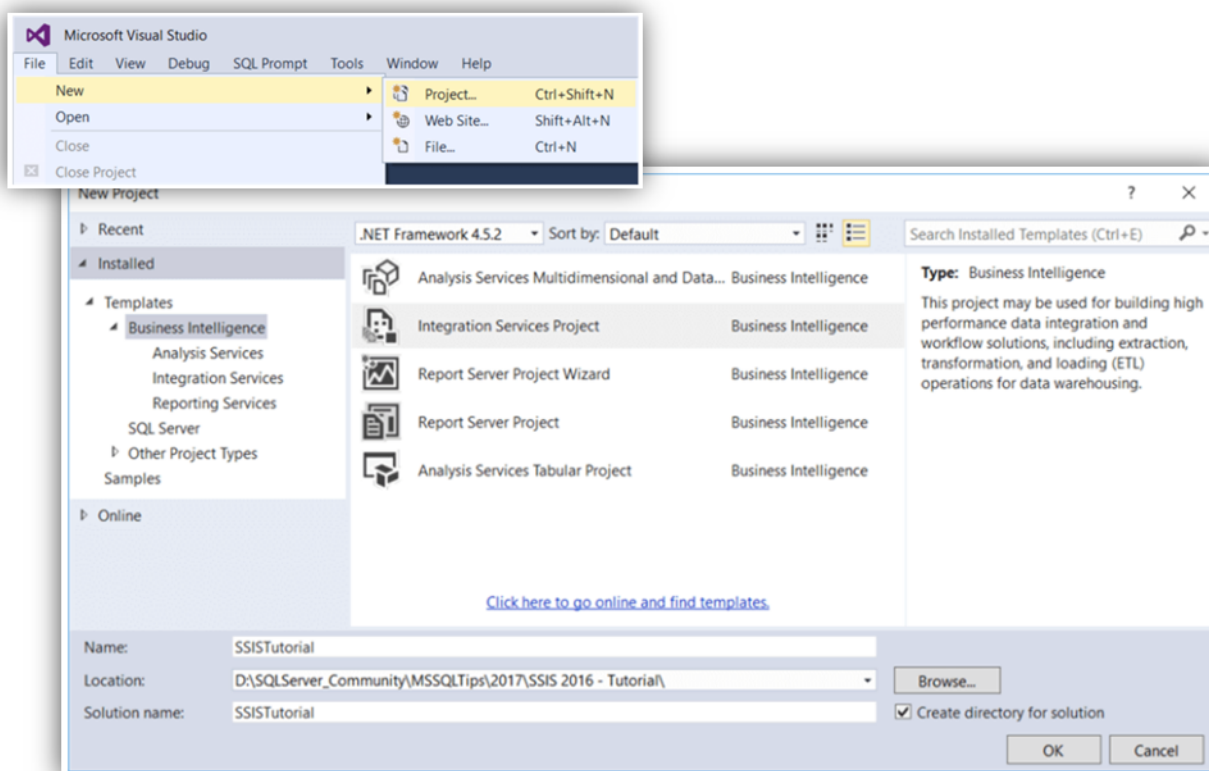
Figure 1. Magic Quadrant for Data Integration Tools



**Fonte: Gartner, 2019 (agosto).**

A interface do SSIS está contida no Microsoft Visual Studio ao criar um projeto de serviço de integração.

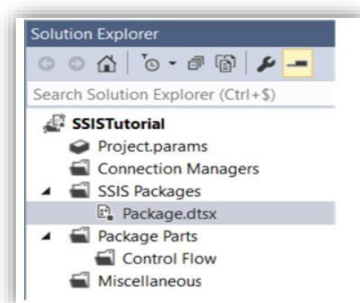
**Figura 33 – Criando um projeto no SSIS.**



**Fonte: Microsoft, 2020.**

Dentro do Visual Studio, o projeto é criado como serviço de integração, e é possível adicionar vários projetos a uma mesma solução, onde um pacote vazio será adicionado ao projeto. A estrutura do projeto na janela Solution Explorer fica da seguinte maneira:

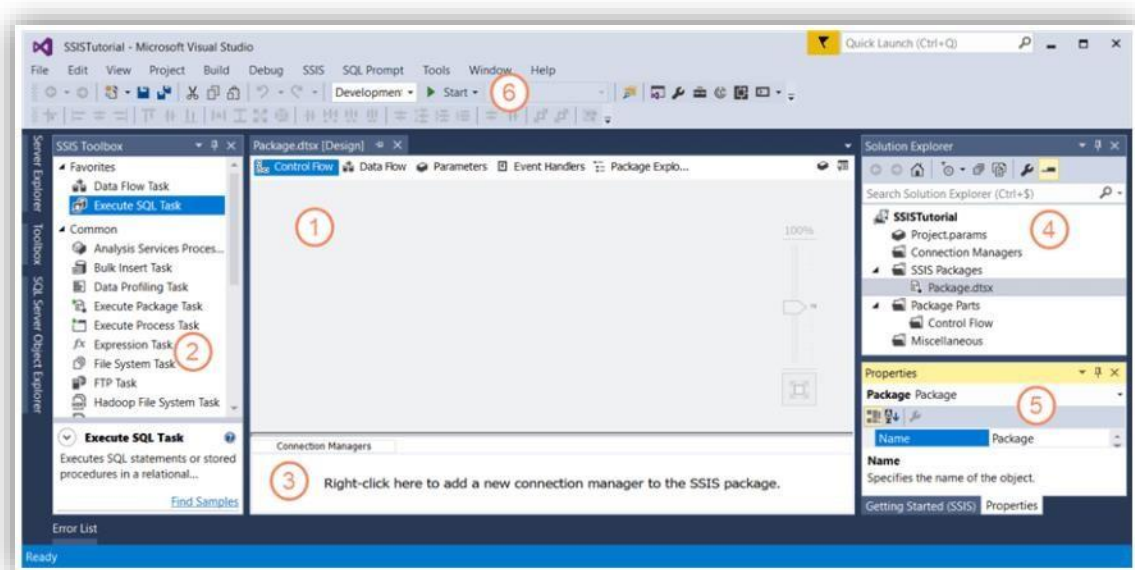
**Figura 34 – Solution Explorer SSIS.**



**Fonte: Microsoft, 2020.**

Uma solução que é bastante usada em conjunto na construção de um fluxo ETL e que também está integrada ao ambiente do Visual Studio é o SQL Server Data Tools (SSDT). Com ele, é possível criar fluxos de controles e fluxos de dados em um pacote, além de adicionar manipuladores de eventos ao pacote e aos objetos do pacote e exibir o conteúdo e o processo de execução do pacote.

**Figura 35 – Interface SSIS.**



**Fonte: Microsoft, 2020.**

Esta é a interface da ferramenta, onde é possível arrastar itens da caixa de ferramentas conectá-los um ao outro para criar um fluxo de trabalho. A tela do pacote possui várias guias:

- 1) **O fluxo de controle.** Nele há várias tarefas com as quais pode se conectar. O fluxo de controle é importante, pois define o que o pacote realmente faz.

**O fluxo de dados.** Essa é uma tarefa especial do fluxo de controle. Os dados são movidos entre as fontes e os destinos e pode transformar os dados enquanto estão na memória.

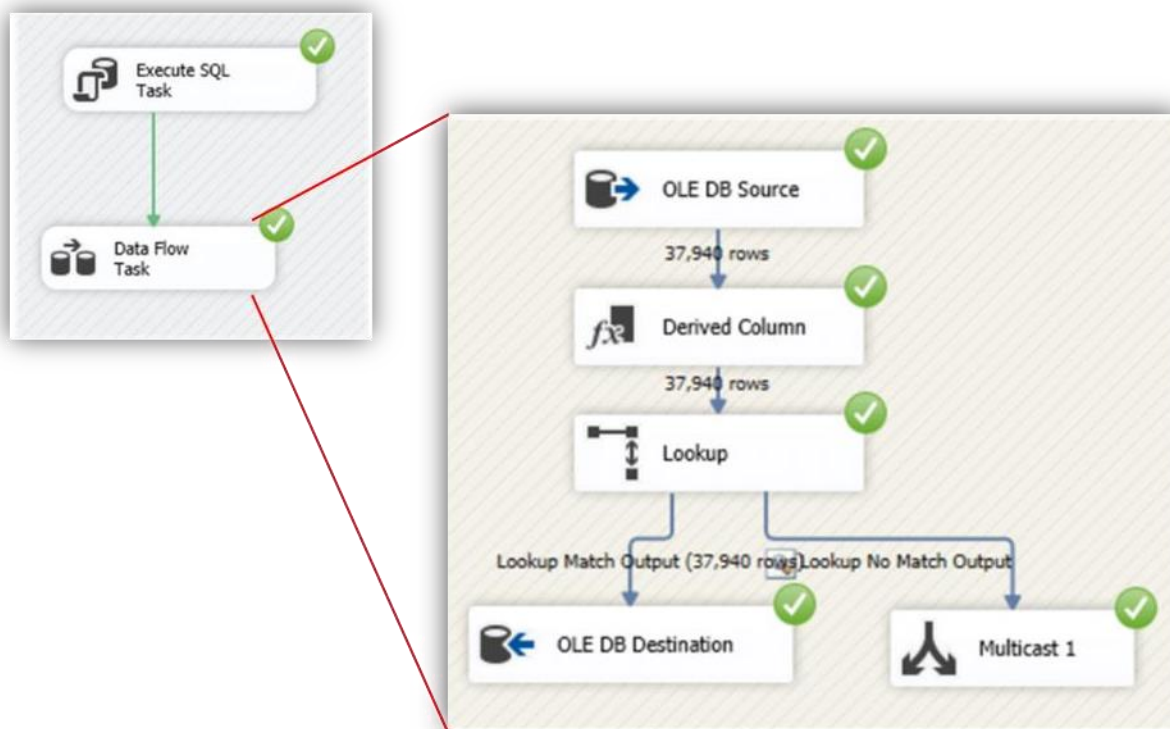
**Parâmetros.** Definição de parâmetros podem tornar seu pacote mais flexível.

**Manipuladores de eventos.** São telas especiais do tipo “fluxo de controle”, nas quais é possível definir tarefas que somente serão executadas se ocorrer um evento específico. Os manipuladores de eventos estão fora do escopo deste tutorial.

**Explorador de Pacotes.** Uma visualização em árvore de todos os objetos dentro do seu pacote.

- 2) A caixa de ferramentas do SSIS. Nela é possível encontrar todas as tarefas e transformações para o controle e o fluxo de dados, além de arrastá-los da caixa de ferramentas para a tela.
- 3) Os gerenciadores de conexão. Um gerenciador de conexões define uma conexão com um objeto específico. Pode ser um arquivo simples, um banco de dados, uma pasta e assim por diante. Tarefas e transformações usam um gerenciador de conexões para criar uma conexão com o objeto
- 4) A solução Explorer. Uma exibição em árvore de todos os objetos no projeto ou solução.
- 5) A janela de propriedades. Nela é possível visualizar e alterar as propriedades de quase todos os objetos em um pacote SSIS.
- 6) As barras de ferramentas. O item mais importante é a seta verde, que é usado para iniciar o depurador. O depurador executará o pacote SSIS no Visual Studio.

**Figura 36 – Fluxo de Controle e Fluxo de Dados.**

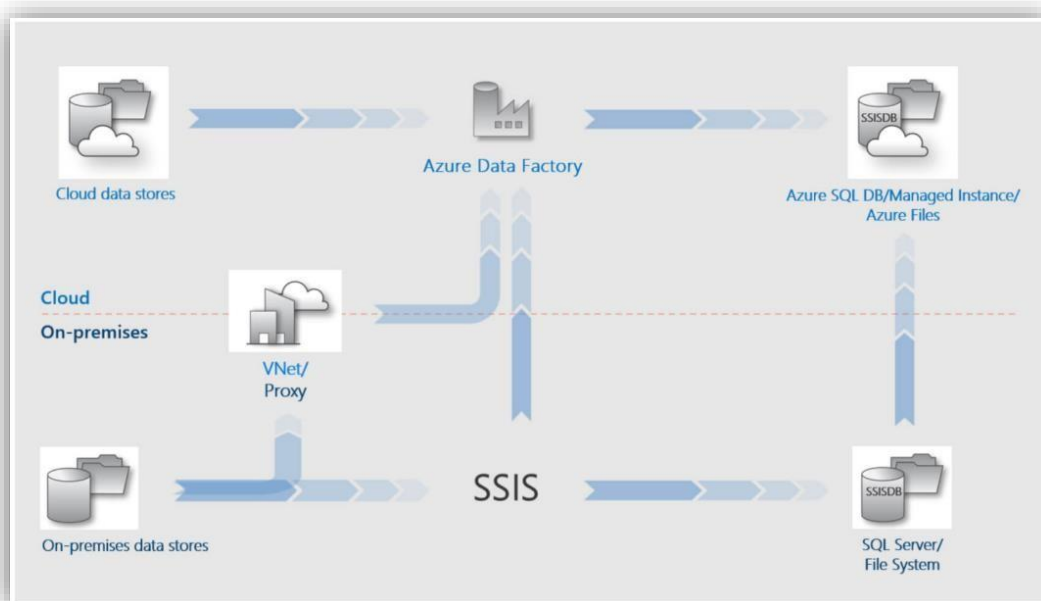


**Fonte: Microsoft, 2020.**

O Azure Data Factory é o serviço ETL da Microsoft na nuvem para integração e transformação de dados sem servidor em expansão. Ele oferece uma interface do usuário livre de código para criação intuitiva e gerenciamento e monitoramento em painel único. Também é possível migrar pacotes SSIS existentes por lift-and-shift para o Azure e executá-los com total compatibilidade.

O Data Factory contém uma série de sistemas interconectados que fornecem uma plataforma de ponta a ponta para os engenheiros de dados.

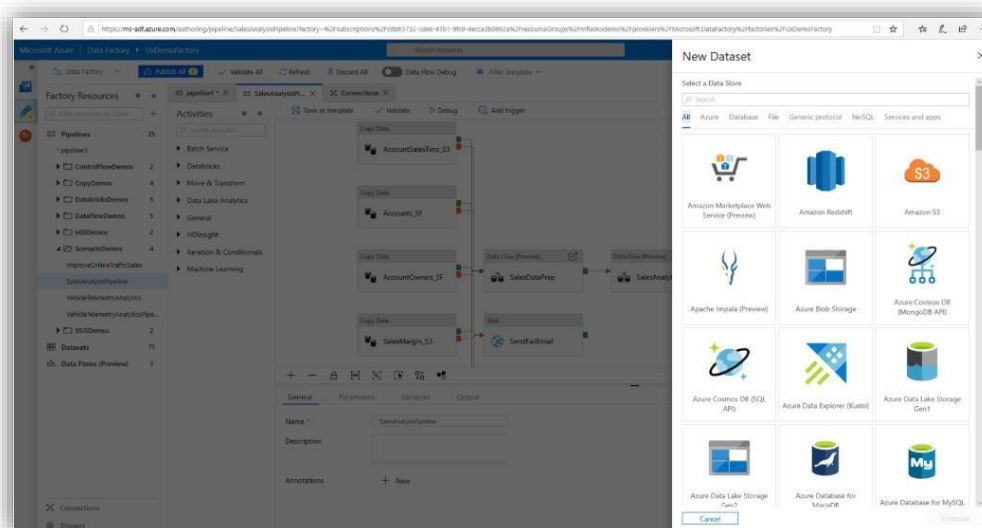
**Figura 37 – Arquitetura Azure Data Factory.**



**Fonte: Microsoft, 2020.**

A interface do ambiente do Azure Data Factory possui semelhanças com o Data Tools no Visual Studio, porém é necessário ter uma conta, mesmo gratuita, na Azure e criar os armazenamentos.

**Figura 38 – Interface Azure Data Factory.**



**Fonte: Microsoft, 2020.**

## ▪ Pentaho Data Integration (PDI)

Ferramenta que realiza tanto a integração de dados quanto os processos de ETL que alimentam Data Warehouses.

- Mais de trinta formatos de SGDB, como Oracle, PostgreSQL e SQLServer.
- Importa arquivos texto (csv ou fixo) e planilhas Excel.
- Base de dados ODBC.
- Ambiente gráfico no qual conexões com fontes de dados são estabelecidas e sequência de passos executam a extração de dados.
- Integra dados entre empresas e sistemas.

**Figura 39 – Microsoft no Quadrante mágico Gartner.**

Figure 1. Magic Quadrant for Data Integration Tools

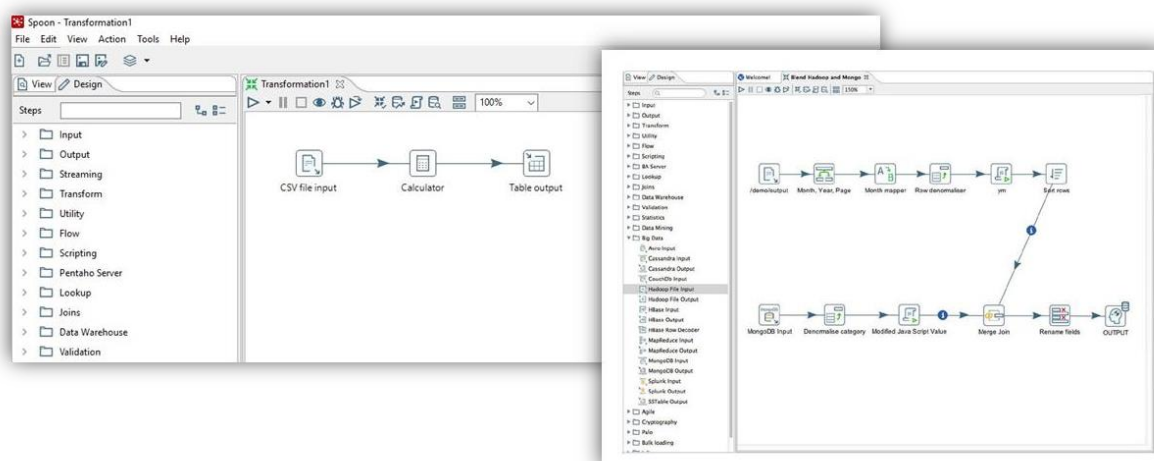


Fonte: Gartner, 2019 (agosto).



Por ser uma ferramenta Open-Source, a avaliação feita pelo mercado não é tão justa, por isso o PDI está nos players de nicho, apesar de ser muito utilizado.

**Figura 40 – Interface Pentaho Data Integration (PDI).**



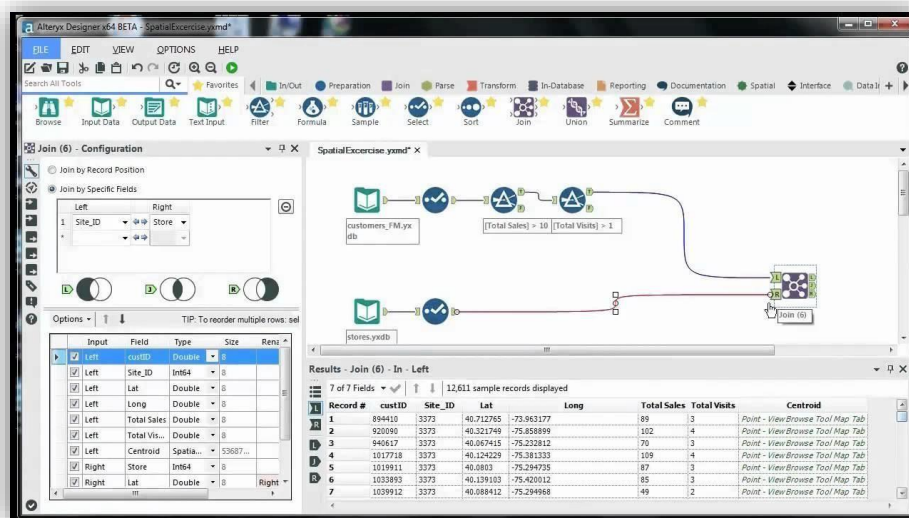
Fonte: <https://sourceforge.net/projects/pentaho/>.

Também aplica na prática o conceito de AgileBI, no qual se combinam em uma interface, a extração de dados, sua modelagem e relatórios. Essa integração permite que a equipe de BI gere resultados em dias ao invés de semanas.

Também são utilizadas no mercado as ferramentas Alteryx, Tableau Prep e Talend em muitos projetos.

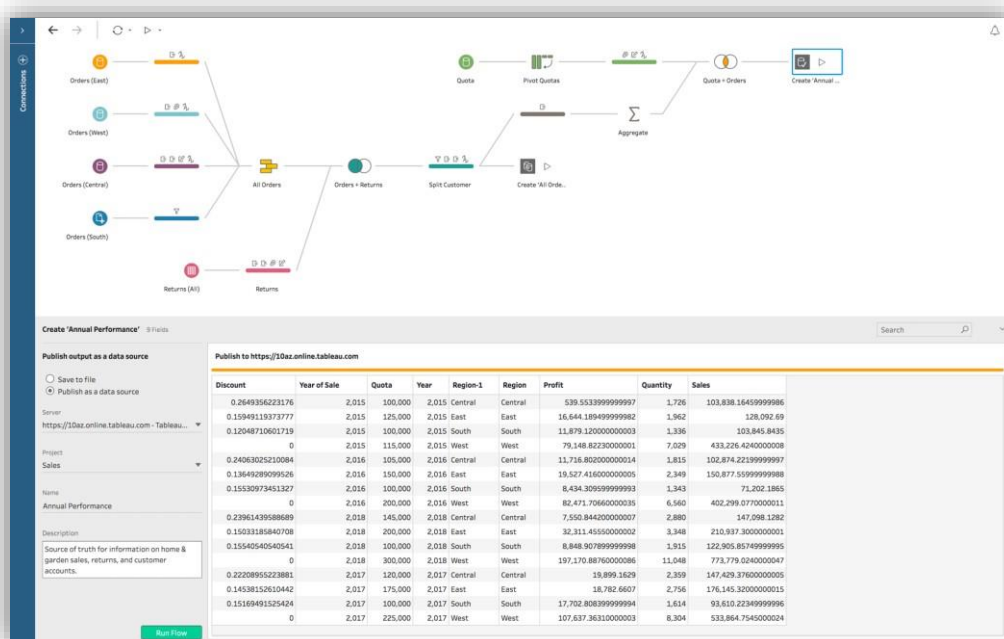


**Figura 41 – Interface Alteryx.**



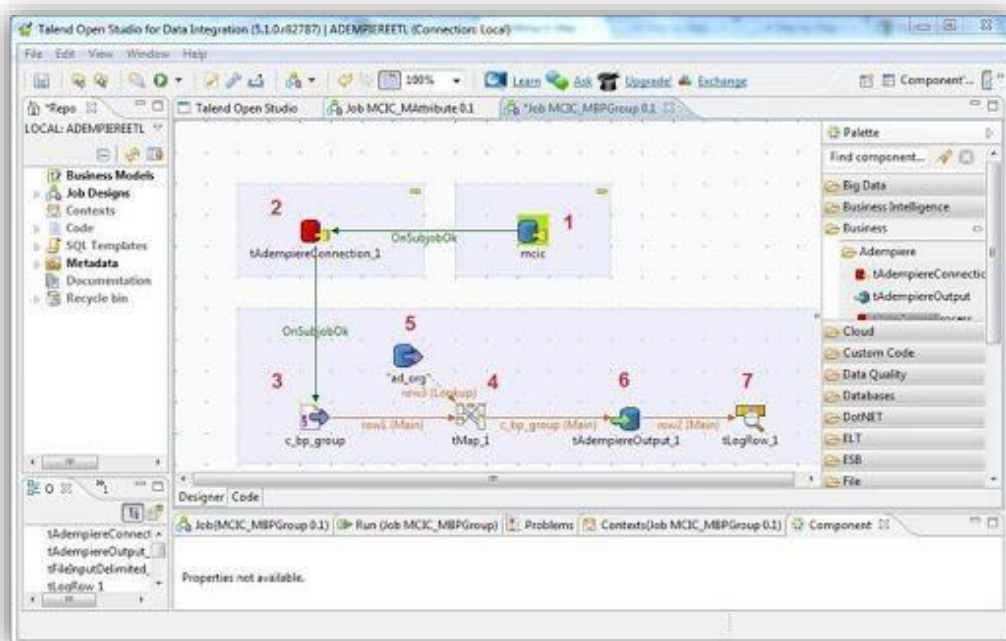
Fonte: <https://www.alteryx.com/pt-br/products/alteryx-platform/alteryx-designer>.

**Figura 42 – Interface Tableau Prep.**



Fonte: <https://www.tableau.com/learn/tutorials/on-demand/getting-started-tableau-prep>.

**Figura 43 – Interface Talend.**



Fonte: <https://www.talend.com/blog/2017/08/18/operationalize-machine-learning-talend/>.

## Capítulo 4. As Aplicações de BI Front-End

Neste capítulo vamos trabalhar o que são aplicações de BI front-end com mais detalhes, as suas características, governanças, arquiteturas, soluções, traçar comparativos e analisar as três principais forças de mercado: Tableau, Qlik Sense e Power BI.

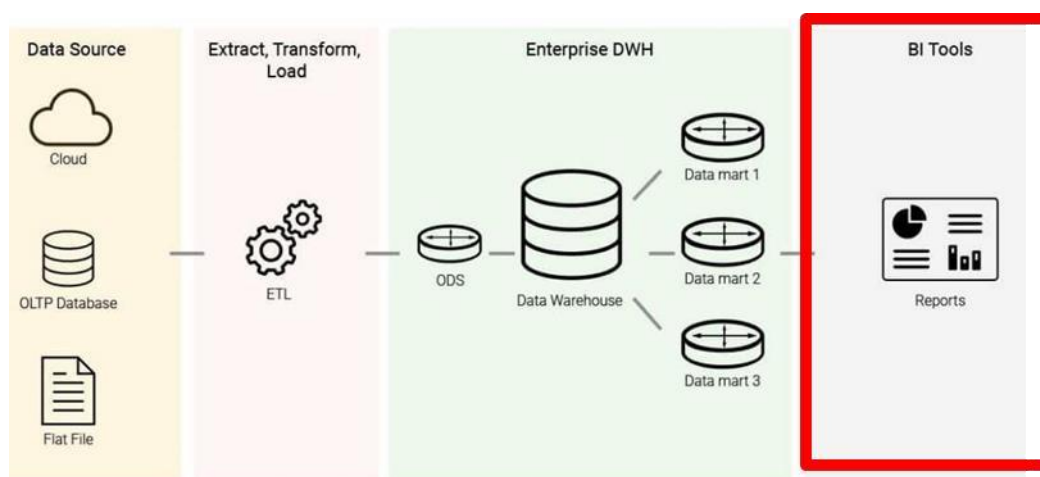
### O que são aplicações front-end

#### ▪ O que são:

Ferramentas front-end são aplicações que permitem ligar-se facilmente às origens de dados, visualizar e descobrir o que é importante, bem como partilhar os seus conteúdos com qualquer pessoa.

Disponibilizam, acessam e demonstram para as empresas e usuários as informações contidas em suas bases de dados e possibilitam um entendimento rápido das diferentes visões de análise.

**Figura 44 – Arquitetura de BI: Front-End.**



Fonte: <https://medium.com/@elisowski/data-warehouse-implementation-tips-400c3bc981e7>.

## Características

- Facilidade para análise e criação de conteúdo.
- Conexões diretas com diversas bases de dados.
- Customização de reports.
- Funcionalidade de arrastar e soltar, que proporciona facilidade e velocidade.
- Recursos de busca inteligente para descobrir a relação de dados.
- Colaboração e compartilhamento de informações com uma rica narração de dados.
- Painéis de relatórios interativos.
- Análises de qualquer lugar (WEB, Mobile).
- Self Service Analytics.
- Alto controle e segurança de acesso.

## Principais fornecedores

Figura 45 – Principais fornecedores.



Fonte: Elaborado pelo autor.

Não há outro objetivo na implementação de ferramentas front-end do que geração de valor e obtenção de resultado. Os resultados vêm através de análises intuitivas, manipulação fácil, acessibilidade, multiusuários desenvolvendo e consumindo, compartilhamento, relatórios dinâmicos, interativos e com transparência.

**Figura 46 – Resultados ferramentas front-end.**



Fonte: <https://community.powerbi.com/t5/Data-Stories-Gallery/bd-p/DataStoriesGallery>.

## Plataformas e soluções de BI

### ▪ Qlik

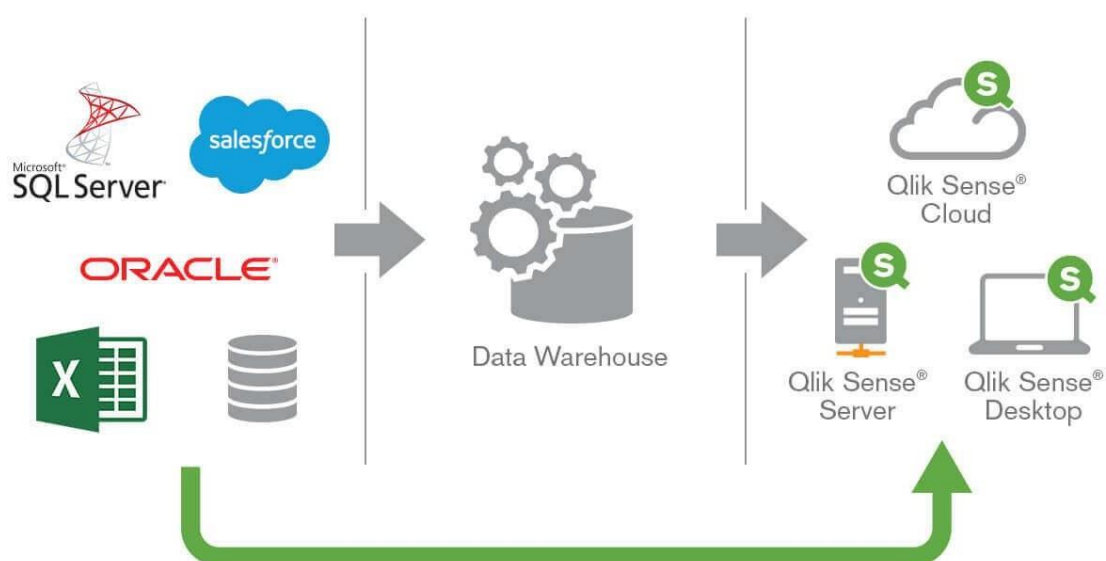
A arquitetura geral das soluções Qlik é composta de quase 15 produtos, segregados cada um com sua função específica dentro de um ecossistema organizado e coeso. A Qlik classifica suas soluções em três grupos com suas respectivas ferramentas:

- Data Integration:
  - Qlik Replicate;

- Qlik Compose;
  - Qlik Gold Client;
  - Qlik Enterprise Manager;
  - Qlik Catalog;
  - Qlik Connect.
- Data Analytics:
- Qlik Sense;
  - Qlik View;
  - Qlik N Printing (para relatórios);
  - Qlik Insignia;
  - Qlik Geo Analytics (para mapas, ícones, figuras, etc.).
- Developer Platforms:
- Core;
  - Analytics Platform;
  - Branch Playground;
  - Qlik Community;
  - Qlik Knowledge Hub.

O Qlik Sense possui sua orientação programática baseada numa linguagem associativa com otimização in-memory, que melhora o modelo de dados em quase 90%.

**Figura 47 – Arquitetura Qlik.**



Fonte: <https://www.traininginbangalore.com/qliksense-training-in-bangalore/>.

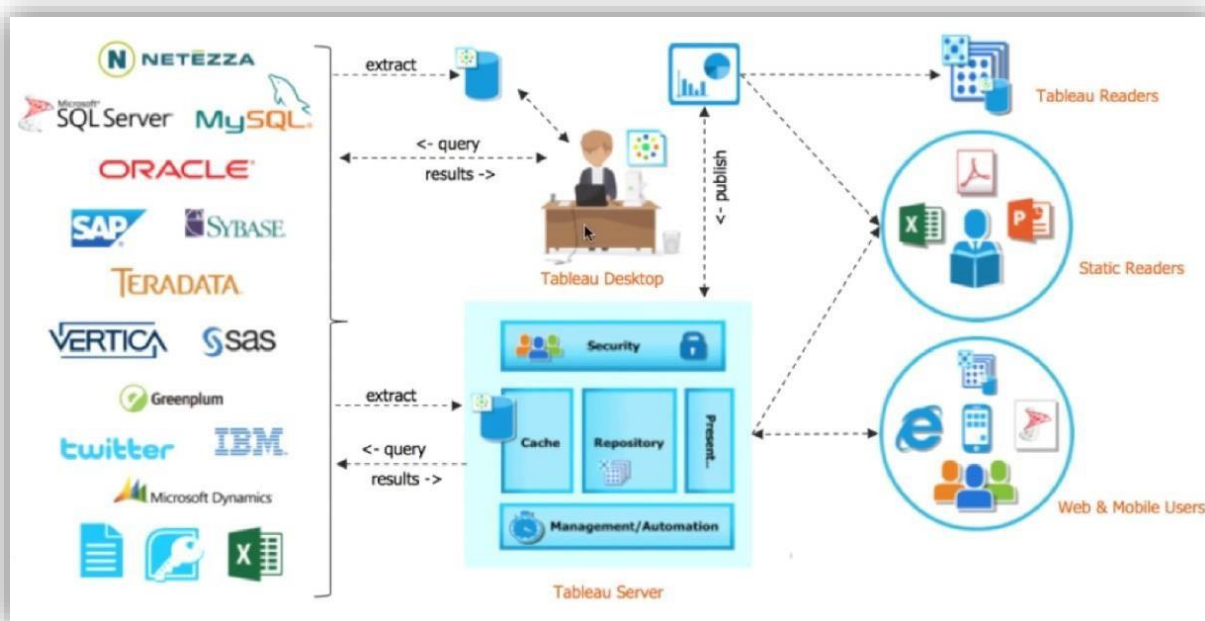
#### ▪ Tableau

A Tableau possui suas soluções concentradas em três frentes: a de desenvolvimento, por meio de sua ferramenta desktop (principal); a de Data Preparation, com o Tableau Prep, que faz o trabalho de integração e conectores; e de serviço.

A parte do serviço pode ser dividida em ambiente local e em nuvem. O ambiente local é tratado pelo Tableau Server, onde é você que administra o ambiente e toda sua arquitetura e governança de usuários. O ambiente em nuvem é o Tableau Online, que é uma licença do Tableau Server na nuvem, atuando como um software como serviço, conhecido como SaaS.



**Figura 48 – Plataforma Tableau.**



Fonte: <https://data-flair.training/blogs/tableau-architecture/>.

#### ▪ Microsoft

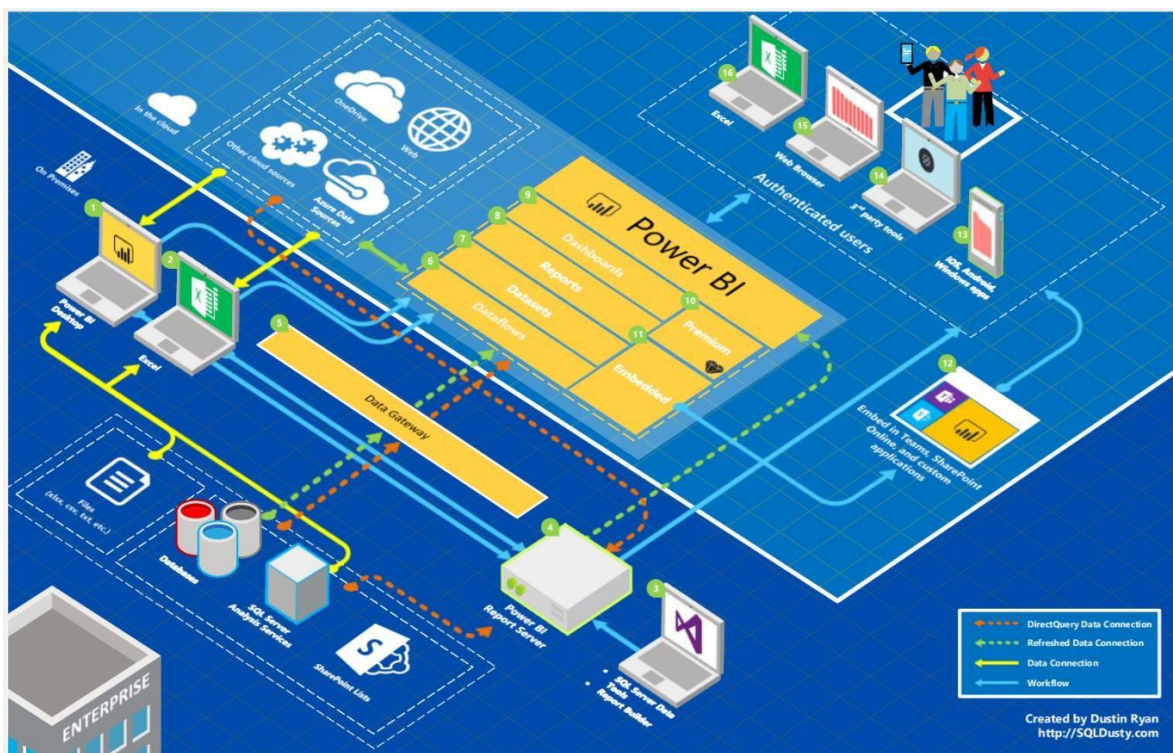
A Microsoft possui o Power BI como principal ferramenta front-end no contexto de BI e, principalmente, self-service BI. Com uma plataforma bem organizada, a Microsoft cobre todos os pontos de uma arquitetura de BI completa. O Power BI possui três tipos de utilização: a utilização desktop, a on-line e a licença de capacidade dedicada.

Gratuitamente é possível usufruir de todos os recursos da ferramenta, seus visuais, integração com gateways locais, bancos de dados, centenas de conectores e integrações. Tudo que toca na parte de compartilhamento seguro, passa a precisar de licença, e é nesse ponto que está o grande poder da plataforma.

Na figura a seguir percebe-se que o Power BI está na ponta de ETL, conectando com as origens, na nuvem para realizar processamentos e compartilhamentos e nos usuários finais.



**Figura 49 – Plataforma Microsoft.**



Fonte: <http://sqldusty.com>.

Assim como acontece com o SQL Server Integration Service, o serviço do Power BI também possui apoio do Azure Data Factory, mas somado ao Azure Synapsys e Databricks para atuação com Big Data.

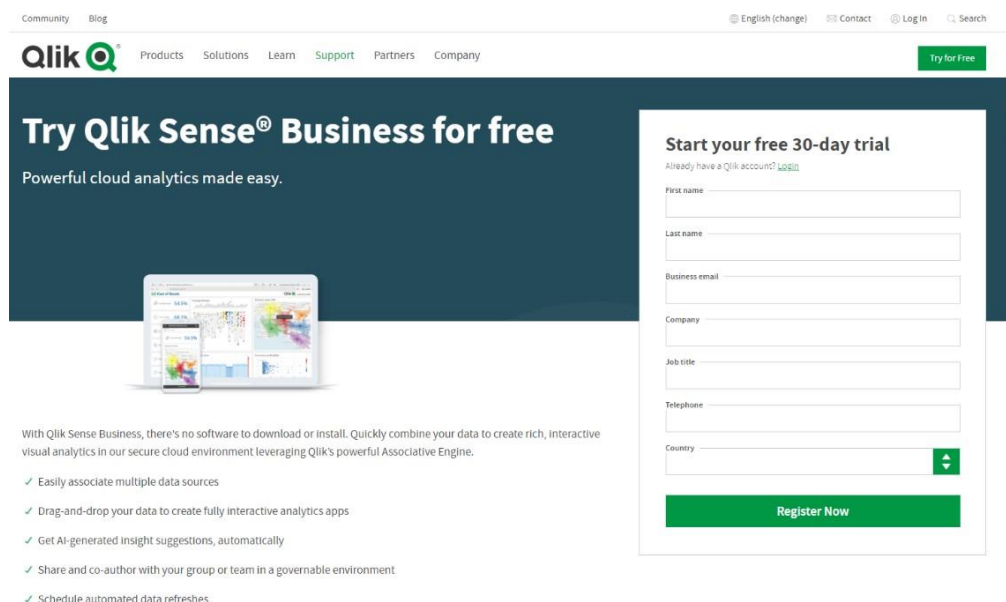
O processamento é muito otimizado no Power BI, uma vez que trabalha com a chamada Vertipaq Engine, que é um modelo tabular vindo do Analysis Services, com uma capacidade de compressão diferente de qualquer outra ferramenta da companhia.

## Ferramentas front-end: Qlik Sense

### ▪ Sobre o Qlik Sense

O Qlik Sense é um dos principais produtos Qlik para Data Analytics. Para iniciar de maneira gratuita, ele pede autenticações que possibilitam o acesso por 30 dias a todos os recursos. O primeiro passo é entrar no site oficial da Qlik.

**Figura 50 – Instale o Qlik Sense.**

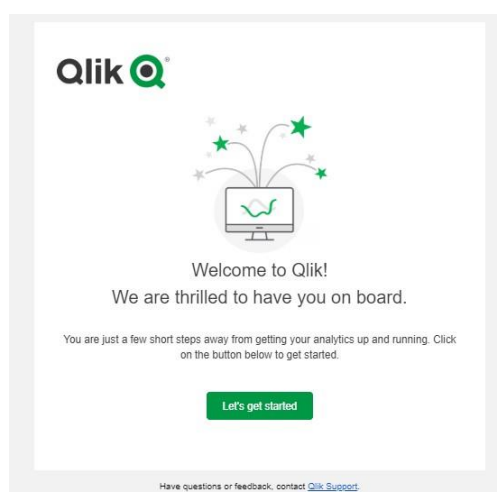


The screenshot shows the Qlik Sense website's trial registration page. The header includes links for Community, Blog, English (change), Contact, Log in, and Search. The main navigation bar lists Products, Solutions, Learn, Support, Partners, and Company. A green 'Try for Free' button is in the top right. The main content area features the headline 'Try Qlik Sense® Business for free' with the tagline 'Powerful cloud analytics made easy.' Below this is an image of a laptop displaying Qlik Sense dashboards. To the right is a registration form titled 'Start your free 30-day trial' with a link for existing users. The form fields include First name, Last name, Business email, Company, Job title, Telephone, and Country (with a dropdown arrow). A green 'Register Now' button is at the bottom of the form. Below the laptop image, a paragraph states: 'With Qlik Sense Business, there's no software to download or install. Quickly combine your data to create rich, interactive visual analytics in our secure cloud environment leveraging Qlik's powerful Associative Engine.' This is followed by a list of five benefits, each preceded by a green checkmark: 'Easily associate multiple data sources', 'Drag-and-drop your data to create fully interactive analytics apps', 'Get AI-generated insight suggestions, automatically', 'Share and co-author with your group or team in a governable environment', and 'Schedule automated data refreshes'.

**Fonte: Qlik (2020).**

Em seguida, basta clicar no link recebido no e-mail e criar o login.

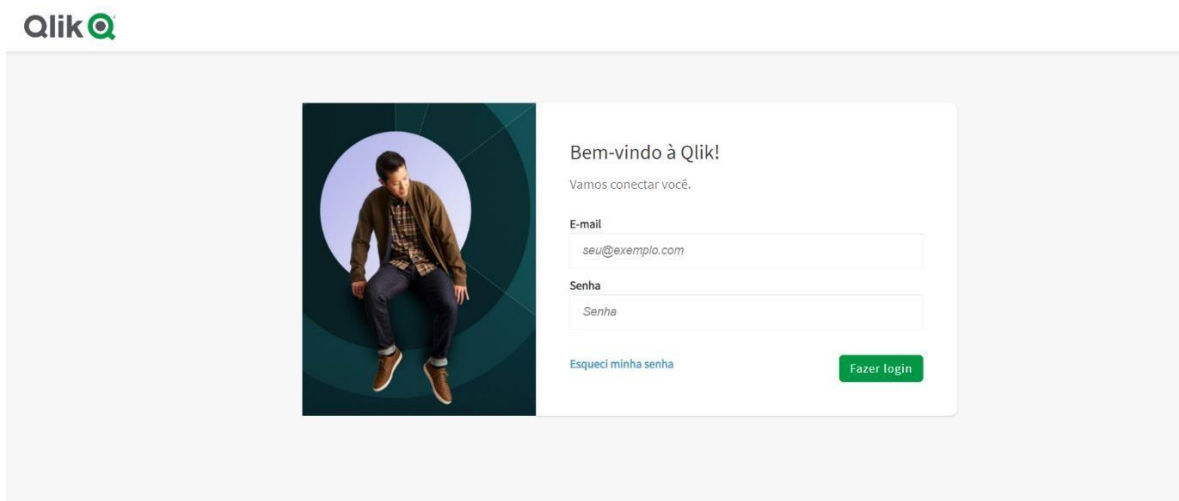
**Figura 51 – E-mail Qlik.**



**Fonte: Qlik (2020).**

O passo a passo chega até ao ponto de criar uma url personalizada no Qlik Cloud, com seu nome.us.qlikcloud.com. Toda a parte inicial passa pelo ambiente em nuvem da Qlik, desde o login até o download do Qlik Sense Desktop.

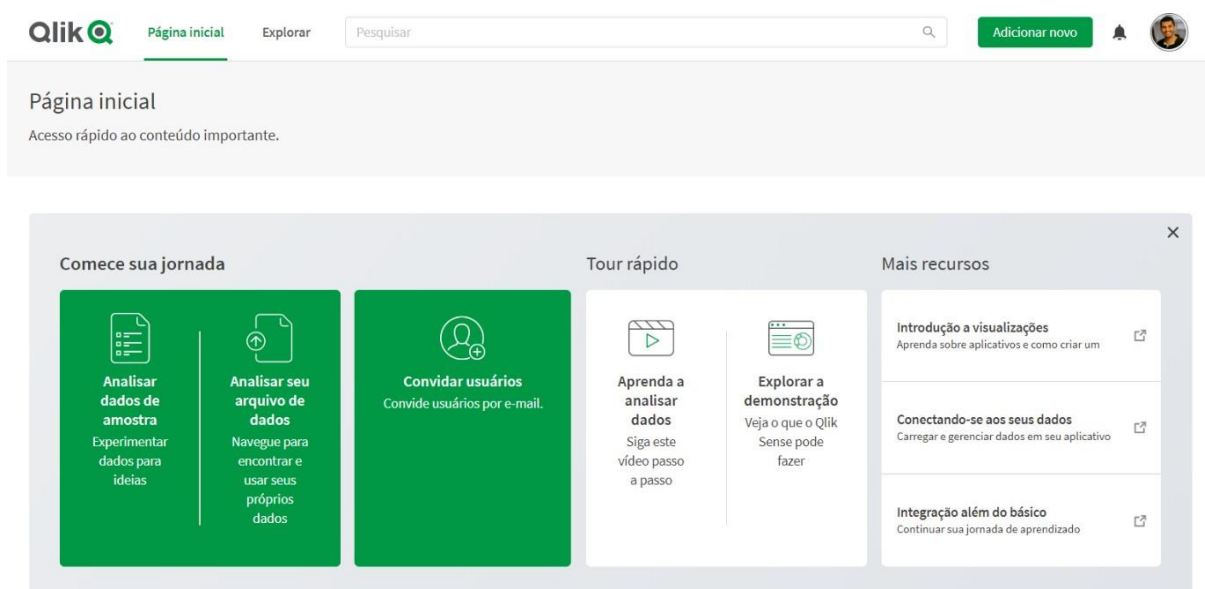
**Figura 52 – Tela de Login do Qlik.**



Fonte: Qlik (2020).

Com a autenticação realizada, o ambiente do Qlik Cloud fica disponível.

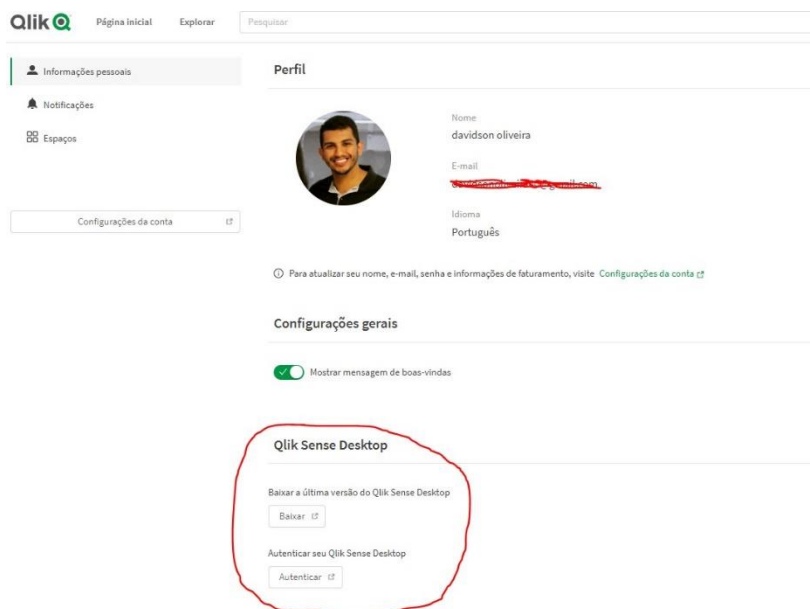
**Figura 53 – Página inicial Qlik Cloud.**



Fonte: Qlik (2020).

Na página de Configurações, é possível instalar e se autenticar no Qlik Sense Desktop.

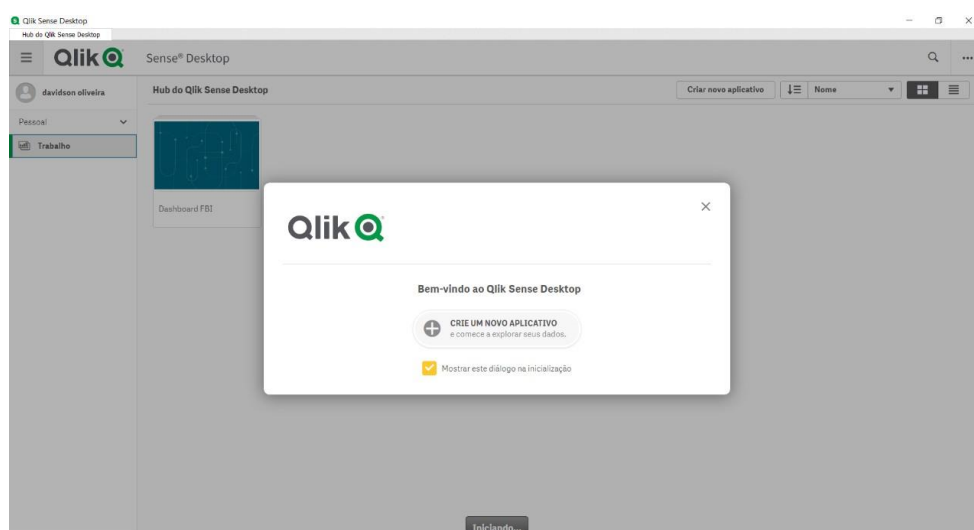
**Figura 54 – Página de Configuração.**



**Fonte: Qlik (2020).**

Com o Qlik Sense Desktop instalado e já logado no ambiente, essa é a interface inicial da ferramenta, que também tem sua opção on-line.

**Figura 55 – Página inicial Qlik Sense Desktop.**



**Fonte: Elaborado pelo autor.**

Dentro da ferramenta é possível criar aplicativos, importar dados de diversos conectores e compartilhar no Hub.

## Ferramentas front-end: Tableau

### ▪ Sobre o Tableau

O Tableau é menos burocrático, inicialmente, do que o Qlik Sense. Basta entrar no site oficial, na parte de downloads, e instalar a versão desktop.

**Figura 56 – Interface Tableau Desktop.**



**Fonte: Elaborado pelo autor.**

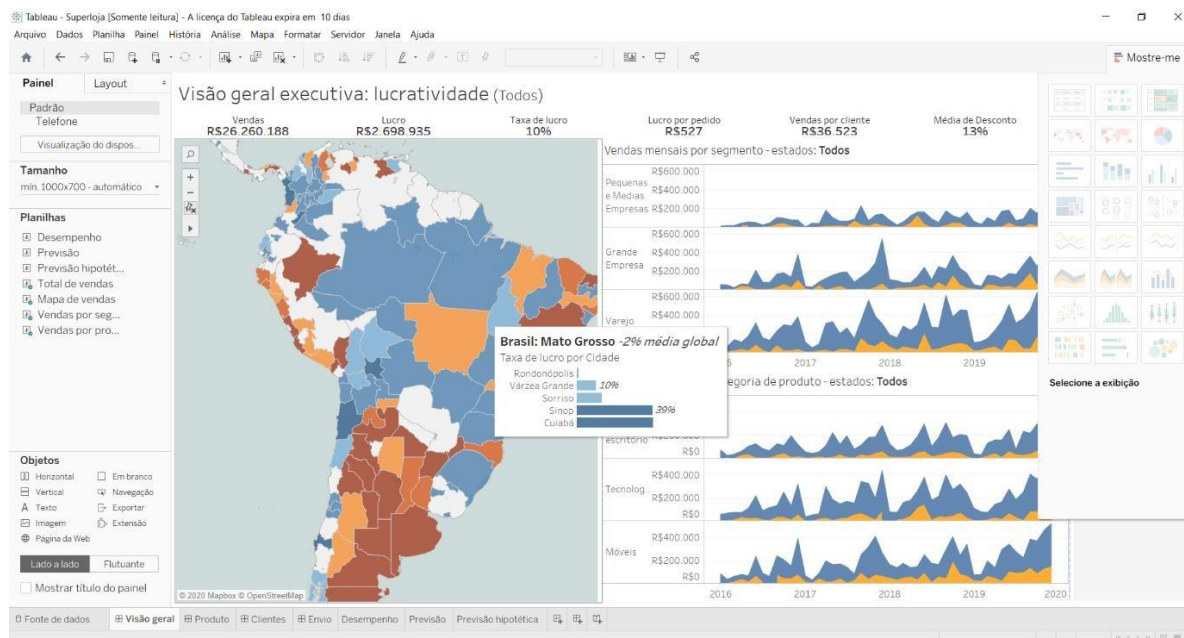
A integração para centenas de origens é dinâmica, rápida e fácil. O Tableau Desktop fornece 15 dias para testar a solução, então é necessário criar um login e autenticar no Tableau Online.

Para realizar importação de dados basta seguir o menu lateral e escolher um dos conectores, seja de arquivos, servidores ou fontes de dados. Outra opção para



ver o funcionamento é visualizando pastas de trabalho de exemplo, como é o caso do exemplo da Superloja.

**Figura 57 – Interface de desenvolvimento.**



**Fonte: Elaborado pelo autor.**

A interface de desenvolvimento oferece centenas de recursos, menus, planilhas, componentes, objetos, visualizações, opções de planilhas, painéis e exibições de storytelling. É realmente uma ferramenta muito completa, tendo um destaque muito grande na parte dos mapas nativos e interatividade com campos calculados.

## Ferramentas front-end: Power BI

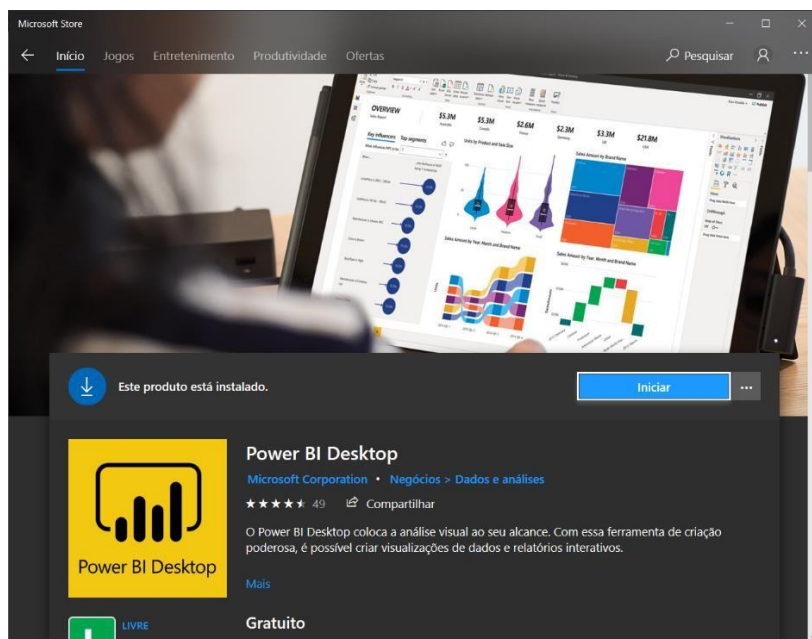
### ▪ Sobre o Power BI

O Power BI é a solução mais democrática dentre as principais. Não pede autenticação para baixar e instalar, tampouco limita a utilização para um período de testes de 15 ou 30 dias. Todos os recursos de desenvolvimento são liberados, ou

seja, é possível construir qualquer tipo de solução, por mais avançada que seja, de maneira totalmente gratuita e local, de qualquer computador que suporte.

Existem duas formas de instalar a ferramenta. A primeira é pela Microsoft Store, no próprio Windows 10, obtendo o produto com um clique.

**Figura 58 – Instalar Power BI pela Microsoft Store.**

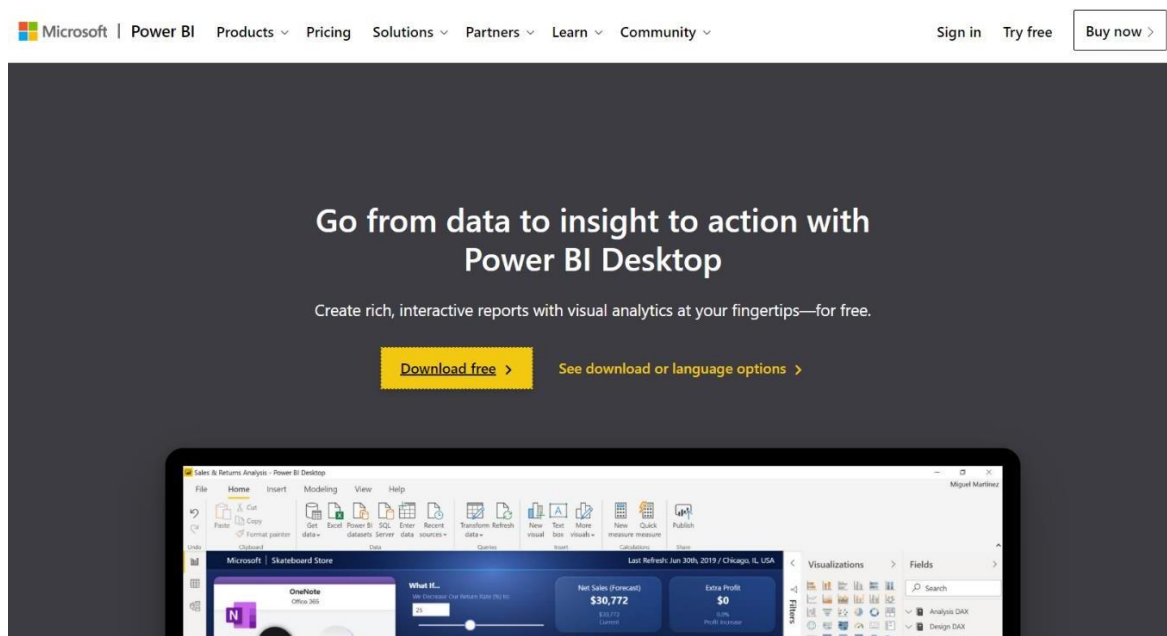


**Fonte: Microsoft Store (2020).**

A outra opção é instalar pelo próprio site da Microsoft na parte de downloads. É possível editar as preferências de download para definir o idioma de instalação e até qual instalador baixar.



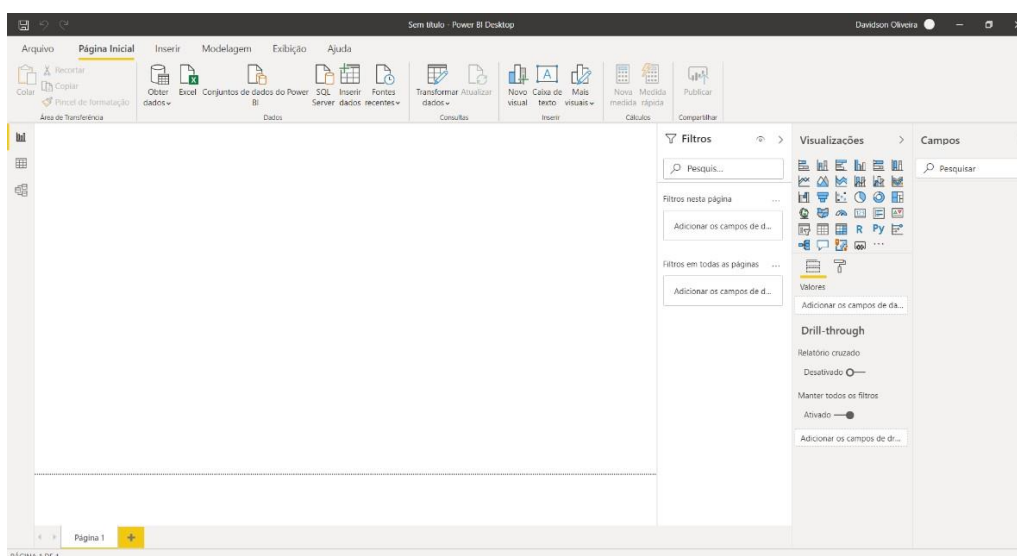
**Figura 59 – Instalar o Power BI pelo site.**



**Fonte: Power BI (2020).**

Com a ferramenta instalada e aberta, vemos a sua interface inicial dividida entre faixa de opções, as guias de Relatório, Dados e Modelo, a faixa lateral de Visualizações e a faixa de Campos.

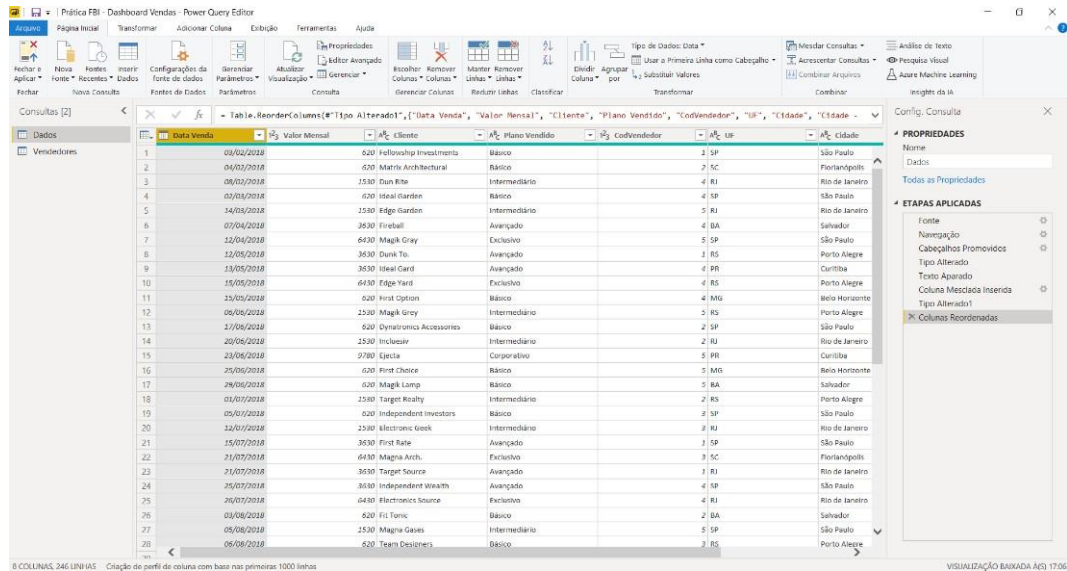
**Figura 60 – Interface Power BI Desktop.**



**Fonte: Elaborado pelo autor.**

Para importar dados, basta utilizar um dos conectores. O exemplo utilizado foi importando dados de uma planilha Excel com dados de vendas.

**Figura 61 – Interface Power Query.**



The screenshot displays the Power Query Editor window with a table of sales data. The table has 8 columns: Data Venda, Valor Mensal, Cliente, Plano Vendido, CodVendedor, UF, Estado, and Cidade. The data is organized into 27 rows, each representing a sales record. The interface includes a ribbon with various options like 'Transformar', 'Adicionar Coluna', and 'Exibição'. On the right, there is a 'Config. Consulta' pane with 'ETAPAS APLICADAS' and 'Colunas Reordenadas' sections.

	Data Venda	Valor Mensal	Cliente	Plano Vendido	CodVendedor	UF	Estado	Cidade
1	03/02/2018	620	Fellowship Investments	Básico	2	SP	São Paulo	
2	04/02/2018	620	Matrix Architectural	Básico	2	SC	Florianópolis	
3	09/02/2018	1530	Dun Rite	Intermediário	4	RJ	Rio de Janeiro	
4	03/03/2018	620	Ideal Garden	Básico	4	SP	São Paulo	
5	14/04/2018	1530	Edge Garden	Intermediário	5	RJ	Rio de Janeiro	
6	07/04/2018	3630	Fireball	Avançado	4	BA	Salvador	
7	12/05/2018	6430	Magik Gray	Exclusivo	5	SP	São Paulo	
8	12/05/2018	3630	Dunk To	Avançado	2	RS	Porto Alegre	
9	13/05/2018	3630	Ideal Card	Avançado	4	PR	Curitiba	
10	15/05/2018	6430	Edge Yard	Exclusivo	4	RS	Porto Alegre	
11	23/05/2018	620	First Option	Básico	4	MG	Belo Horizonte	
12	06/06/2018	2530	Magik Grey	Intermediário	5	RS	Porto Alegre	
13	11/06/2018	620	Dynamix Accessories	Básico	2	SP	São Paulo	
14	20/06/2018	2530	Incubator	Intermediário	2	RJ	Rio de Janeiro	
15	12/06/2018	9180	Glecia	Corporativo	5	PR	Curitiba	
16	25/06/2018	620	Font Choice	Básico	5	MG	Belo Horizonte	
17	24/06/2018	620	Magik Lamp	Básico	5	BA	Salvador	
18	01/07/2018	2530	Target Realty	Intermediário	2	RS	Porto Alegre	
19	05/07/2018	620	Independent Investors	Básico	3	SP	São Paulo	
20	12/07/2018	2530	Electronic Geek	Intermediário	3	RJ	Rio de Janeiro	
21	15/07/2018	3630	First Rate	Avançado	1	SP	São Paulo	
22	21/07/2018	6430	Magik Arch.	Exclusivo	3	SC	Florianópolis	
23	21/07/2018	3630	Target Source	Avançado	1	RJ	Rio de Janeiro	
24	25/07/2018	3630	Independent Wealth	Avançado	4	SP	São Paulo	
25	24/07/2018	6430	Electronic Source	Exclusivo	4	RJ	Rio de Janeiro	
26	03/08/2018	620	Fit Tonic	Básico	2	BA	Salvador	
27	05/08/2018	2530	Magik Games	Intermediário	5	SP	São Paulo	
28	06/08/2018	620	Team Designer	Básico	2	RS	Porto Alegre	

Fonte: Elaborado pelo autor.

## Referências

---

BARBIERI, Carlos. *BI: Business Intelligence: Modelagem e tecnologia*. Rio de Janeiro, Axcel Books, 2001.

DEVMEDIA. Disponível em: <<http://www.devmedia.com.br/>>. Acesso em: 15 nov. 2020.

GARTNER. Disponível em: <<https://www.gartner.com/en>> Acesso em: 15 nov. 2020.

GUALBERTO, Marcelo. *Você sabe o que é ETL?*. Disponível em: <<http://www.taticasolutions.com.br/?p=2529>>. Acesso em: 15 nov. 2020.

GUERRA, Bruno. *Nove critérios para escolher a melhor solução de business intelligence*. Disponível em: <<http://blog.in1.com.br/9-criterios-para-escolher-amelhor-solucao-de-business-intelligence>>. Acesso em: 29 jun.2020.

HEUSSER, Matthew. *Testing the Extract, Transform, and Load Process in Data Warehouses*. Disponível em: <<https://www.stickyminds.com/article/testingextract-transform-and-load-process-data-warehouses>>. Acesso em: 29 jun.2020.

INMON, William H. *Como construir o Data Warehouse?*. 2. ed., Editora Campus, Brasil, 1997.

KIMBALL GROUP. Disponível em: <<http://www.ralphkimball.com/>>. Acesso em: 15 nov. 2020.

MICROSOFT. *Conexões do SSIS (Integration Services)*. Disponível em: <<https://docs.microsoft.com/pt-br/sql/integration-services/connection-manager/integration-services-ssis-connections?view=sql-server-ver15>>. Acesso em: 15 nov. 2020.

MICROSOFT. *Introdução ao serviço do Power BI (app.powerbi.com)*. Disponível em: <<https://docs.microsoft.com/pt-br/power-bi/service-get-started>>. Acesso em: 15 nov. 2020.

QLIK. Disponível em: <<https://www.qlik.com/pt-br/>>. Acesso em: 15 nov. 2020.

TABLEAU. Disponível em: <<https://www.tableau.com/>>. Acesso em: 15 nov. 2020.

TURBAN, E. et al. *Business Intelligence: um enfoque gerencial para a inteligência do negócio*. São Paulo: Editora Bookman, 2008. p. 77.