

**UNIVERSIDADE DE BRASÍLIA**

**Faculdade do Gama**

**Sistemas de Banco de Dados 2**

**Trabalho Final (TF)**

**Data Visualization**

**Representação Gráfica Significativa de Grande Volume de  
Dados para Tomada de Decisões**

**Guilherme Brito Vilas – Bôas - 190108011**

Brasília, DF

2023

## Definição

A visualização de dados é uma prática que utiliza representações visuais para que as pessoas compreendam as informações de forma intuitiva e significativa. Essa combinação de habilidades científicas e artísticas vai além da simples apresentação de dados, oferecendo insights valiosos e revelando padrões inovadores. As boas representações visuais têm o poder de enfatizar a mensagem da visualização (C WILKE, 2019). Nesse contexto, surge o conceito de Alfabetização Visual, que envolve a capacidade de interpretar, criar e comunicar-se por meio de imagens visuais (OLIVEIRA, [s.d.]).

No entanto, os profissionais dessa área enfrentam o desafio de equilibrar a arte e os dados, representando as informações de forma esteticamente satisfatória, clara e precisa (C WILKE, 2019). Um estudo realizado em 2019 pelas empresas 3M Corporation e Zabisco revelou que 90% das informações transmitidas ao cérebro são visuais (SIBLEY, [s.d.]). Considerando isso, o uso de figuras visuais esteticamente agradáveis, claras e matematicamente corretas tem um impacto significativo na tomada de decisões corporativas. Para facilitar o acompanhamento das informações, uma equipe coletaria os dados relevantes, transformando-os em elementos visuais e os apresentando aos gestores, sendo os responsáveis pela tomada de uma decisão.

Existem diversas formas de apresentar essas informações, cada uma com sua aplicação específica. Alguns tipos de visualizações já são amplamente difundidos na comunidade de análise de dados, o que beneficia essa tecnologia. Cada situação requer uma solução distinta, garantindo clareza ao receptor da mensagem.

## **Tipos de Visualização <sup>1</sup>**

Os tipos de visualização mencionados desempenham um papel fundamental na análise de dados e na tomada de decisões. Cada categoria oferece uma maneira específica de representar e interpretar informações.

### **Números**

Essa opção permite a representação de dados numéricos, sendo especialmente útil para destacar números relevantes no contexto da decisão.

### **Tendência**

As tendências são ideais para mostrar a variação de um número ao longo do tempo, essa opção é essencial na análise da evolução de um valor em um período específico.

### **Detalhes**

Os detalhes oferecem uma visualização simplificada de um único resultado, facilitando a leitura e compreensão dos dados.

### **Barra de Progresso**

Barras de Progresso são usadas para comparar um número com uma meta estabelecida, proporcionando uma compreensão abrangente do desempenho em relação ao objetivo.

### **Gauges**

Gráficos circulares ou semicirculares que mostram o desempenho, progresso ou status de uma métrica específica.

### **Tabelas**

As tabelas são uma opção comum para representar dados de forma organizada e tabular, facilitando a comparação, identificação de padrões e análise de relações.

### **Tabelas Pivotadas**

As tabelas pivotadas permitem resumir e reorganizar informações tabulares de maneira mais compacta e significativa, facilitando a análise exploratória e a identificação de padrões.

---

1 Nesse tópico, considerei as opções fornecidas pelo Metabase. Uma ferramenta de BI, open source, responsável pela gestão e análise de dados. Os termos utilizados se referem também a plataforma e mais informações sobre eles podem ser encontrados neste link: <https://www.metabase.com/docs/latest/questions/sharing/visualizing-results>.

## **Gráficos de Linha**

Gráficos de Linhas são importantes para visualizar tendências, variações e padrões ao longo do tempo ou de uma sequência ordenada.

## **Gráfico de Barras**

Gráficos de Barras permitem a comparação visual de quantidades ou valores agrupados por algum atributo.

## **Gráfico de Área**

Gráficos de área mostram a distribuição e evolução de valores ao longo do tempo ou de uma sequência ordenada, facilitando a identificação de padrões e análise de crescimento ou declínio.

## **Gráficos Combinados**

Gráficos Combinados combinam barras, linhas (ou áreas) em um único gráfico, permitindo a análise de múltiplas métricas no mesmo eixo.

## **Histogramas**

Histogramas são gráficos que representam a distribuição de frequência dos dados em intervalos. Eles revelam padrões, agrupamentos e a forma da distribuição dos dados, permitindo uma compreensão rápida e visualmente impactante.

## **Gráficos de Barras Horizontais**

Gráficos de barras horizontais exibem a distribuição de dados de forma visual e concisa em forma de barras horizontais. Eles facilitam a comparação entre categorias ou grupos, fornecendo insights claros e rápidos.

## **Gráficos de Cascata**

Gráficos de Cascata são uma forma de visualização que representa mudanças incrementais em um valor ao longo de categorias ou períodos. Eles mostram as contribuições positivas e negativas para o valor total, permitindo uma análise clara das alterações.

## **Gráficos de Dispersão**

Gráficos de Dispersão são uma forma de visualização que exibe a relação entre duas variáveis em um plano cartesiano. São úteis para identificar padrões, tendências ou relações entre as variáveis, permitindo uma análise visual da correlação entre os dados.

## Gráficos de Pizza

Os gráficos de pizza são uma forma de visualização que representa a distribuição de uma variável em relação a um todo. Esses gráficos são úteis para destacar a proporção de cada categoria em relação ao conjunto total e permitem uma comparação visual rápida entre as partes.

## Gráficos de Funil

Os gráficos de funil são usados para representar visualmente a progressão sequencial dos dados em um processo. Eles ajudam a identificar pontos de melhoria, gargalos e acompanhar o desempenho do processo de maneira visualmente impactante.

## Mapas

Os mapas permitem visualizar dados geográficos de forma interativa. É possível identificar padrões espaciais, distribuição geográfica e insights baseados em localização. Com marcadores, regiões coloridas ou camadas personalizadas, os mapas facilitam a análise de vendas, demografia e localização dos dados.

Todas essas métricas podem ser unidas em um painel de visualização chamado de Dashboard.

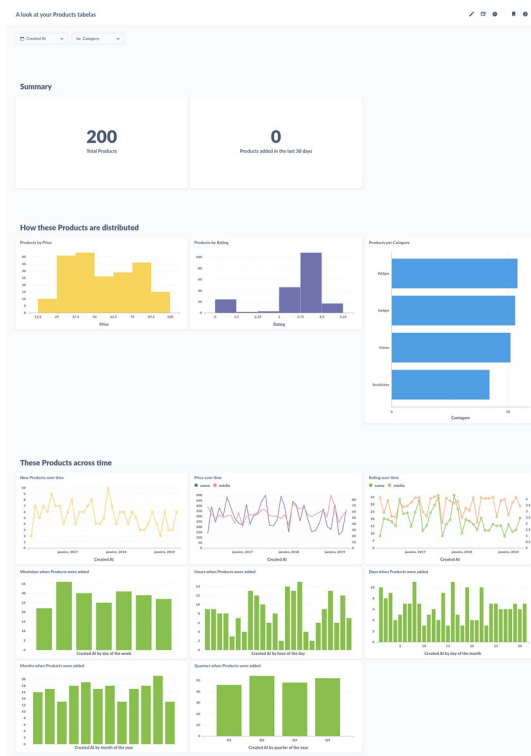


Figura 1: Dashboard criado a partir de uma tabela de Produtos unindo métricas

Concluindo, os diferentes tipos de visualização desempenham um papel crucial na análise de dados e na tomada de decisões. Cada categoria oferece uma forma única de representar informações, permitindo uma compreensão mais clara e intuitiva dos dados. Desde a visualização de números importantes até a análise de tendências, detalhes e comparações, essas opções de visualização fornecem insights valiosos para embasar decisões informadas. Seja por meio de gráficos de linha, barras, áreas ou combinações, ou por meio de tabelas, mapas ou gráficos específicos, cada tipo de visualização possui suas características distintas e benefícios. Ao explorar e utilizar essas opções de visualização de forma adequada, é possível extrair informações valiosas dos dados, identificar padrões, tendências e oportunidades de melhoria, e, assim, tomar decisões mais embasadas e eficazes.

## **Objetivos**

A utilização de Data Visualization tem como objetivo principal facilitar a compreensão e interpretação das informações complexas presentes nos conjuntos de dados. Por meio de representações gráficas e visuais, busca-se transformar os dados em elementos visuais claros e concisos, permitindo uma visualização mais intuitiva e acessível. Dessa forma, a tomada de decisões torna-se mais informada e embasada, uma vez que os gestores e analistas podem identificar padrões, tendências e insights relevantes de forma mais ágil.

A Data Visualization permite uma análise mais aprofundada dos dados, facilitando a detecção de relações, comparações entre métricas e avaliação do desempenho. Além disso, ao comunicar visualmente os dados, a visualização de dados facilita a colaboração entre equipes e promove uma compreensão comum, possibilitando uma tomada de decisões mais eficiente e alinhada.

Em resumo, o objetivo da utilização de Data Visualization é fornecer uma representação visual dos dados que seja clara, intuitiva e informativa, visando apoiar a tomada de decisões fundamentadas.

## **Vantagens**

Como visto, as ferramentas de Data fornecem funcionalidades para a visualização de dados de forma gráfica. Essa representação oferece diversas vantagens para a tomada de decisões, como por exemplo:

- Maior agilidade de resposta
  - Os gestores enfrentam constantemente o desafio de tomar decisões importantes de forma rápida e precisa. Nesse contexto, a visualização de dados desempenha um papel fundamental. Através da representação visual clara e concisa das informações, os gestores

podem compreender e interpretar melhor os dados, facilitando a identificação de padrões, tendências e insights relevantes. A visualização de dados permite uma compreensão mais intuitiva e rápida do panorama geral, proporcionando uma base sólida para a tomada de decisões informadas. Ao transformar dados complexos em gráficos, tabelas e outros elementos visuais, os gestores podem explorar relações, comparar métricas e avaliar o desempenho de forma mais eficiente. Além disso, a visualização de dados facilita a comunicação e o compartilhamento de informações entre equipes, promovendo uma compreensão comum e alinhamento estratégico.

- Progresso contínuo da empresa
  - A análise de dados e a representação visual são essenciais no mundo dos negócios para compreender a complexidade do mercado. Ao utilizar técnicas de visualização de dados, é possível identificar padrões e tendências, facilitando a tomada de decisões ágeis e precisas. Isso permite às empresas antecipar mudanças, identificar oportunidades e ajustar suas estratégias de forma mais assertiva. A visualização de dados é uma ferramenta poderosa para compreender a dinâmica do mercado e reagir de forma eficaz às suas demandas.
- Vantagem competitiva
  - De acordo com a BI-Survey, uma pesquisa recente, constatou-se que 58% das empresas baseiam pelo menos metade de suas decisões em sentimentos ou experiência, em vez de serem orientadas por dados relevantes (JANOSCHEK, 2016). No entanto, os gestores que adotam uma abordagem baseada em dados têm uma vantagem significativa. Ao utilizar dados relevantes, esses gestores conseguem entender as nuances do mercado e, conseqüentemente, estabelecer estratégias de negócio mais eficazes. Os dados atuam como parâmetros para compreender o que está ocorrendo externamente à organização, permitindo que a empresa aprenda e execute ações competitivas. Em um ambiente empresarial cada vez mais complexo e competitivo, a utilização de dados confiáveis se torna essencial para impulsionar o sucesso e a tomada de decisões informadas.

Em resumo, a visualização de dados desempenha um papel crucial na tomada de decisões empresariais, fornecendo uma representação gráfica das informações complexas que facilita a compreensão e interpretação dos dados. Por meio dessa abordagem, os gestores são capazes de identificar padrões,

tendências e insights relevantes, permitindo uma tomada de decisão informada e ágil.

A visualização de dados também promove a comunicação e o compartilhamento de informações entre equipes, fomentando a compreensão mútua e o alinhamento estratégico. Além disso, a análise e representação visual de dados são instrumentos essenciais para compreender a complexidade do mercado, antecipar mudanças, identificar oportunidades e ajustar estratégias de negócios de maneira assertiva. A adoção de uma abordagem baseada em dados confiáveis confere uma vantagem competitiva às empresas, permitindo-lhes compreender as nuances do mercado e tomar decisões embasadas em dados relevantes.

Em um contexto empresarial cada vez mais complexo e competitivo, a visualização de dados se torna fundamental para impulsionar o sucesso organizacional e a tomada de decisões informadas.

## **Desvantagens**

Embora a visualização de dados ofereça uma série de benefícios para a tomada de decisões, também apresenta algumas desvantagens que devem ser consideradas. Algumas das principais desvantagens do uso de Data Visualization são:

- **Simplificação excessiva**
  - Ao representar dados complexos de forma visual, há o risco de simplificar demais as informações, perdendo detalhes e nuances importantes. Isso pode levar a uma compreensão superficial dos dados e a interpretações errôneas.
- **Viés na interpretação**
  - A interpretação dos gráficos e visualizações de dados pode ser influenciada por vieses subjetivos. Os usuários podem interpretar os dados de acordo com suas próprias percepções e preconceitos, resultando em conclusões tendenciosas ou imprecisas.
- **Manipulação enganosa**
  - A manipulação incorreta ou enganosa dos elementos visuais pode distorcer a representação dos dados, levando a interpretações falsas ou equivocadas. É importante garantir a integridade e a precisão das visualizações de dados para evitar conclusões errôneas.

Portanto, para maximizar os benefícios da visualização de dados e mitigar essas desvantagens, é essencial ter cuidado na seleção e design das



visualizações, garantir a precisão dos dados, promover a interpretação objetiva e considerar as limitações de cada tipo de visualização. O uso responsável e crítico da visualização de dados é fundamental para uma análise precisa e informada.

## **Caso de Sucesso – Ambev<sup>2</sup>**

A implementação da Control Tower pela BRQ na Ambev proporcionou uma solução eficaz para os desafios enfrentados pela empresa em relação à centralização de informações e ao uso do Big Data em suas operações logísticas. Anteriormente, a falta de visibilidade e integração entre as diferentes áreas dificultava a compreensão do que estava acontecendo em cada setor. Isso levava a tomadas de decisões imprecisas, uma vez que os dados logísticos estavam dispersos e cada setor tomava suas próprias decisões, sem necessariamente envolver gestores e diretores. Além disso, a troca de dados era prejudicada nesse cenário fragmentado. A Control Tower desenvolvida pela BRQ atendeu a essas necessidades, permitindo que diretores e vice-presidentes visualizassem os dados de forma centralizada e segmentada, respeitando as restrições e regras de controle de dados. Essa abordagem proporcionou uma visão integrada e abrangente do processo logístico, facilitando a tomada de decisões informadas e promovendo maior agilidade e eficiência na gestão logística da Ambev (BRQ E AMBEV: CONFIRA OS RESULTADOS DESSA PARCERIA!, 2020).

---

2 <https://blog.brq.com/brq-e-ambev/>

## Base de Dados Escolhida

A base de dados disponível no link: <https://www.kaggle.com/datasets/lukeidip/nba-boxscore-dataset> é chamada de "NBA Boxscore Dataset" e está hospedada na plataforma Kaggle. Essa base de dados contém informações sobre jogos da National Basketball Association (NBA), uma das principais ligas de basquete profissional dos Estados Unidos.

A base de dados contém várias tabelas relacionadas a boxscores de jogos da NBA, que são registros estatísticos detalhados de cada partida. Esses registros incluem informações sobre equipes, jogadores, estatísticas individuais e coletivas, como pontuação, rebotes, assistências, bloqueios, roubos de bola e muito mais. Os boxscores são amplamente utilizados para analisar o desempenho dos jogadores e equipes, avaliar estatísticas e realizar análises estatísticas avançadas.

O projeto original consistia em 3 tabelas, mas para melhor mapeamento dos relacionamentos e organização dos dados, o projeto final contemplou com a presença de 7 tabelas, sendo derivadas dessas 3.

As tabelas são:

- TEAM: Descreve os times que participaram dos jogos.
- PLAYER: Descreve os jogadores que participaram dos jogos.
- GAME: Descreve os jogos que ocorreram nesses 10 anos.
- GAME\_INFO: Descreve informações adicionais sobre cada jogo.
- plays\_for: Tabela intermediária originária do relacionamento N:M entre PLAYER e TEAM, pois um jogador pode ter jogado em diversos times durante esses 10 anos, desde que tenha sido em jogos diferentes.
- PLAYER\_STATS: Descreve diversas métricas de um jogador em um jogo.
- TEAM\_STATS: Descreve diversas métricas e indicadores de um time em um jogo.

## Modelagem<sup>3</sup>

Após a análise da base de dados, o seguinte projeto final foi modelado.

### DER

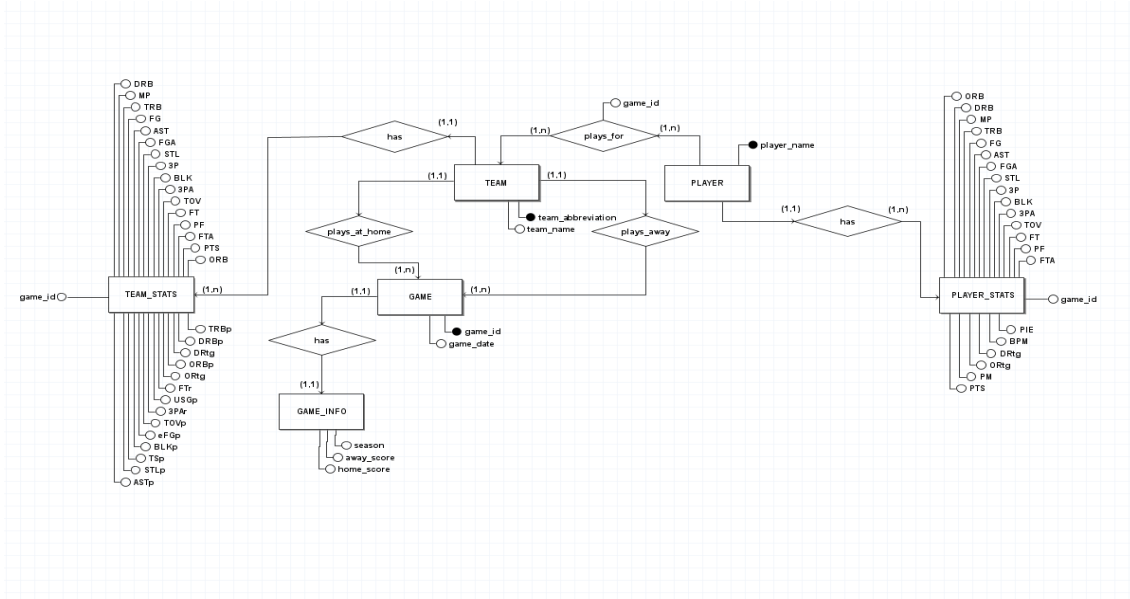


Figura 2: DE-R referente a base de dados escolhida para o Projeto Prático

### DLD

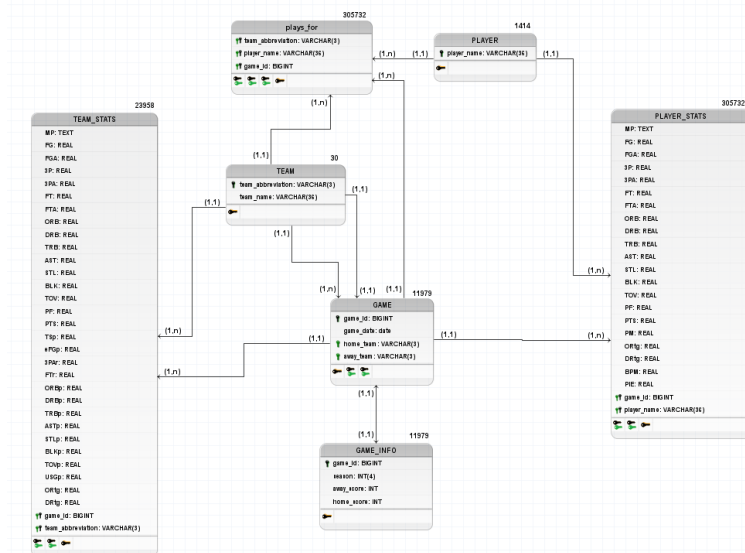


Figura 3: DLD referente a Base de Dados

3 Para melhor visualização, os arquivos podem ser baixados no link:  
<https://drive.google.com/drive/folders/113VlgNKJFPiXPTQn15df1Ns8fwlkF48?usp=sharing>

## Referências Bibliográficas

BRQ e Ambev: confira os resultados dessa parceria! 2020. Disponível em: <https://blog.brq.com/brq-e-ambev/>. Acesso em: 11 jun. 2023.

C WILKE. Fundamentals of data visualization : a primer on making informative and compelling figures. Sebastopol, Ca: O'reilly Media, 2019.

Dados para tomada de decisão: saiba usá-los da maneira correta! 2020. Disponível em: <https://blog.brq.com/dados-para-tomada-de-decisao/>. Acesso em: 11 jun. 2023.

JANOSCHEK, Nikolai. Data-Driven Decision-Making: 14 Recommendations on how to benefit. 2016. Disponível em: <https://bi-survey.com/data-driven-decision-making-business>.

OLIVEIRA, Gabriela Pintar De. A importância da alfabetização visual. [s.d.]. Disponível em: <https://inovadocente.com.br/blog/a-importancia-da-alfabetizacao-visual>. Acesso em: 11 jun. 2023.

SIBLEY, Amanda. 19 Reasons You Should Include Visual Content in Your Marketing [Data]. [s.d.]. Disponível em: <https://blog.hubspot.com/blog/tabid/6307/bid/33423/19-Reasons-You-Should-Include-Visual-Content-in-Your-Marketing-Data.aspx>.