

Apostila do Professor Bernardo Prates

Palavras iniciais

Olá, queridos alunos! Estou super animado para compartilhar com vocês minha paixão e conhecimento em Visualização de Dados e Introdução a SQL. Primeiro, deixa eu contar um pouco sobre mim e as apostilas que estou preparando durante nossas aulas.

Sou gaúcho, lá do Rio Grande do Sul, e hoje moro em Passo Fundo, bem no norte do estado. Minha aventura no mundo da tecnologia começou lá por volta de 2008. Vocês acreditam que tudo começou com umas revistas sobre HTML e CSS que comprei em uma banca de jornal? (Sim, revistas! Parece coisa de outro mundo hoje em dia, né? 😊). Lembro como se fosse ontem: criando meus primeiros sites com cores vibrantes e fontes nada convencionais.

Logo depois, mergulhei de cabeça em um curso técnico, onde me apaixonei pela programação ao aprender C#. Essa paixão me levou a cursar Sistemas de Informação na UFSM (Universidade Federal de Santa Maria), onde também fiz meu primeiro estágio. Formei-me em 2014 e, desde então, venho construindo uma carreira que equilibra academia e mercado de trabalho. Em 2018, concluí meu mestrado em Ciência da Computação, também pela UFSM. Fui professor substituto na universidade, lecionei programação no Senac-RS e preparei alunos para concursos.

Em 2019, senti falta de programar e mergulhei de volta ao mercado de trabalho, passando por empresas como Dell, B2W e, atualmente, atuando como Desenvolvedor Backend na Sicredi. Hoje, meu dia a dia é recheado de Java, Kotlin, SQL, Oracle, Mongo, Kafka, GitLab, WebFlux, programação reativa, testes unitários e várias outras tecnologias.

Minha jornada até aqui não foi fácil. Como muitos de vocês podem descobrir, o caminho na TI envolve desafios, responsabilidades e, às vezes, até burnout. Mas acreditem, é uma jornada incrível de aprendizado e crescimento. Estou aqui para ajudar vocês nessa caminhada, compartilhando não apenas conteúdo técnico, mas também experiências e lições que aprendi em mais de 10 anos de atuação entre academia e indústria.

No decorrer do semestre eu vou ir desenvolvendo algumas apostilas que irão guiar as nossas aulas. A apostila nada mais é do que um guia do que veremos em sala de aula. Prestem atenção nos momentos de reflexão, por quê? São momentos que eu peço para vocês pensarem no que estudamos, infelizmente, hoje em dia, pensar está ficando cada vez mais difícil, e algo que eu quero insistir muito nesse semestre é fazer vocês pensarem, para não sofrerem tanto quando chegarem no mercado de trabalho.

Estou sempre disponível no Teams para qualquer ajuda que precisarem. Não sou o mais ativo nas redes sociais, mas quem quiser pode me acompanhar e conectar-se comigo no LinkedIn e Instagram, a qual são as plataformas que mais uso atualmente.

LinkedIn: [Bernardo Prates](#)

Instagram: [@be_prates](#)

Estou ansioso para embarcarmos juntos nessa aventura de aprendizado! Vamos nessa! 🚀

Dado, Informação e Conhecimento: Entendendo as Diferenças.....	2
1. Dado: O Ponto de Partida.....	2
2. Informação: Dados com Contexto.....	2
3. Conhecimento: Informação em Ação.....	2
A Ponte Entre Eles: Tecnologia e Experiência.....	2
Quantidade de dados por minutos na web.....	3
Análise e visualização de dados.....	3
O que é Google Looker Studio?.....	4
Características Principais.....	5
Importância na Visualização de Dados.....	5
Dashboards - O que é isso?.....	5
O Papel dos Dashboards na Interpretação de Dados.....	6
Conceitos Básicos do Google Looker.....	7
Fontes de Dados.....	7
Conectores.....	7
Visualizações.....	7
Arquivos CSV.....	8
Estrutura de um Arquivo CSV.....	8
Exemplo.....	8
Uso em Análise de Dados.....	9
Demonstração Prática.....	9
Interface.....	10
Controle de acesso.....	13
Editando o schema dos dados.....	15
Entendo os dados e ferramenta explorar.....	18
Operações com o nosso dataset.....	22
Relatórios.....	23

Combinado inicial:

- Se liguem nas datas de entregas do TPs e AT - As datas estão no calendário no moodle.
- A entrega de PDFs vazios, zips corrompidos ou outras tentativas de burlar as regras são enquadradas no código disciplinar, podendo causar repetência na disciplina, bloco ou desligamento.
- O Microsoft Teams é uma ferramenta de uso essencial para sua graduação.
- Nossas comunicações serão na grande maioria via teams.
- Usem o monitor da disciplina que é um cara qualificado para responder às perguntas, eu posso estar com outras pendências e demorar para responder.
- O plágio é enquadrado no código disciplinar, podendo causar repetência na disciplina ou bloco, suspensão ou desligamento.
- O regime disciplinar é aplicado tanto ao aluno que copiou quanto ao aluno que deixou seu trabalho ser copiado.
 - É sua responsabilidade cuidar do sigilo de seu trabalho: se você deixá-lo disponível publicamente ou permitir acesso para um colega, está assumindo o risco de tê-lo plagiado.

Dado, Informação e Conhecimento: Entendendo as Diferenças

1. Dado: O Ponto de Partida

O que é: Dados são fatos brutos, como números, caracteres ou imagens, que por si só não têm significado. São como as peças de um quebra-cabeça espalhadas.

Exemplo: Imagine que você tem uma série de números: 20, 35, 45. Isso são dados. Eles ainda não nos dizem nada específico.

2. Informação: Dados com Contexto

O que é: Quando você organiza os dados e lhes dá contexto, eles se transformam em informação. A informação é útil e tem significado.

Exemplo: Se eu te disser que esses números (20, 35, 45) são as idades dos membros de uma família, agora eles se tornam informação. Você entende que são idades e tem um contexto.

3. Conhecimento: Informação em Ação

O que é: Conhecimento é o que você faz com a informação. Quando você aplica a informação, usando-a para tomar decisões ou resolver problemas, isso é conhecimento.

Exemplo: Sabendo as idades dos membros da família, você pode deduzir relações ou necessidades (como a possibilidade de um dos membros ser pai/mãe dos outros). Isso é usar a informação para gerar conhecimento.

A Ponte Entre Eles: Tecnologia e Experiência

Tecnologia: Em programação, por exemplo, usamos bancos de dados para armazenar dados. Com SQL, transformamos esses dados em informação através de consultas. E, finalmente, aplicamos essa informação em nossos programas para criar conhecimento útil.

Experiência Pessoal: Uma vez eu trabalhei em um grande e-commerce do mercado brasileiro e tínhamos muitos dados dos clientes. Um desses dados disponível era a quantidade de cliques que aconteciam quando a gente colocava uma notificação no aplicativo. Resumidamente, os cliques perdidos não representavam nada para nós, quando começamos a agrupar os cliques pelo o horário em que aconteciam os cliques, descobrimos algo muito interessante para o negócio. Descobrimos que mandar uma notificação as 12:42 (horário fictício, mas muito próximo do real) da promoção de 5 kit-kat por 10 reais (hehehe) era mais efetivo do que mandar as 12:30, por exemplo.

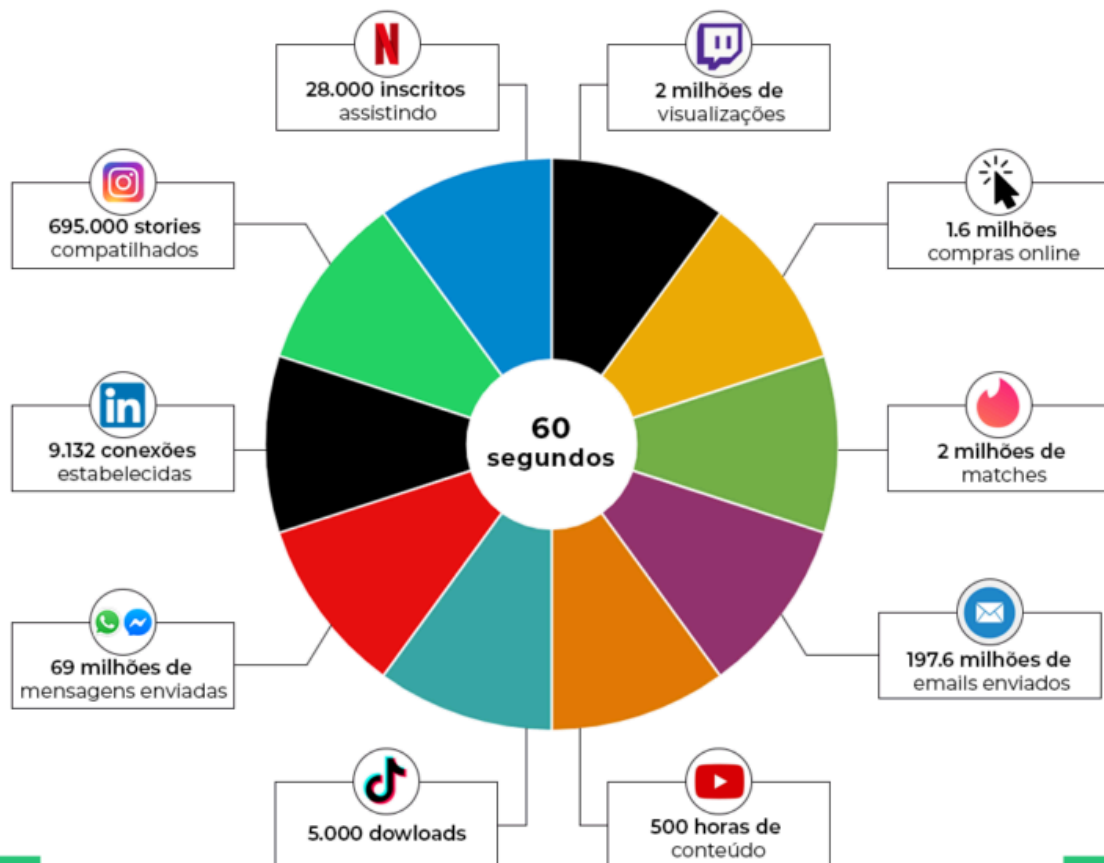
Quantidade de dados por minutos na web



O que acontece a cada minuto na Internet em 2021?

Fonte: Statista

aaa.inovação.com.br
@aaa.inovação



Para refletir? Quantos livros poderia ser gerados em 60 segundos com a quantidade de informações compartilhadas por e-mail?

Análise e visualização de dados

Imagine que você é o gerente de uma grande rede de supermercados. Todos os dias, milhares de transações acontecem em suas lojas. Essas transações geram uma enorme quantidade de dados brutos: o que foi comprado, quando, por quanto, e assim por diante. Isso é o que chamamos de "dados brutos". Sozinhos, eles são como um amontoado de peças de um quebra-cabeça. Agora, vamos ver como a análise e a visualização de dados podem ajudar a montar esse quebra-cabeça.

Análise de Dados

- **Limpeza de Dados:** Primeiro, os dados brutos são limpos e organizados. Isso significa remover erros ou dados irrelevantes e padronizar os formatos.
- **Processamento e Exploração:** Em seguida, usamos ferramentas de processamento de dados (como SQL, Python, R) para explorar esses dados. Isso pode envolver a execução de algoritmos estatísticos para encontrar padrões ou tendências.
- **Insights e Padrões:** Por exemplo, a análise pode revelar que as vendas de sorvete aumentam significativamente em dias quentes. Ou que um determinado produto vende mais quando colocado próximo a outro (a famosa cerveja com fraldas).

Visualização de Dados

- **Gráficos e Diagramas:** Agora, transformamos esses insights em gráficos e diagramas. Isso pode incluir gráficos de barras, gráficos de linha, mapas de calor, etc.
- **Interpretação Intuitiva:** A visualização torna a interpretação dos dados mais intuitiva. Um gráfico de linha mostrando as vendas de sorvete ao longo do ano, com picos no verão, é fácil de entender.
- **Tomada de Decisão:** Essas visualizações ajudam os gestores a tomar decisões informadas. Por exemplo, com base nos dados, o gerente pode decidir estocar mais sorvete no verão ou fazer promoções cruzadas de produtos que são frequentemente comprados juntos (caso do Kit Kat descrito acima).

Aplicação na Ciência e Negócios

- **Negócios:** No nosso exemplo do supermercado, o gerente pode usar esses insights para otimizar o estoque, planejar promoções, ou reorganizar a disposição dos produtos.
- **Ciência:** Na ciência, pesquisadores podem usar análise e visualização de dados para validar hipóteses, descobrir novos fenômenos, ou apresentar resultados de pesquisa de forma clara.

Essencialmente, a análise e visualização de dados transformam montanhas de dados brutos em informações compreensíveis e acionáveis. É como resolver um quebra-cabeça complexo e, ao final, ter uma imagem clara que informa decisões críticas.

Para refletir: Como você acha que isso se aplica ao seu campo de interesse ou estudo?

O que é Google Looker Studio?

Google Looker Studio é uma ferramenta de visualização de dados e relatórios que permite transformar seus dados em relatórios interativos e dashboards personalizáveis. É uma plataforma incrivelmente útil para qualquer pessoa que deseje **extrair insights significativos** de seus dados.

Características Principais

- **Integração de Dados:** Looker Studio pode se conectar a uma variedade de fontes de dados, como Google Analytics, Google Sheets, bancos de dados SQL e muitas outras, permitindo uma integração de dados fluida.
- **Personalização e Flexibilidade:** Oferece uma ampla gama de opções de personalização para seus relatórios e dashboards. Você pode arrastar e soltar diferentes elementos, escolher entre vários tipos de gráficos e tabelas, e personalizar o design para se adequar ao seu estilo ou à identidade visual da sua empresa.
- **Relatórios Interativos:** Os relatórios são não apenas visuais, mas também interativos. Você pode configurar filtros e controles para que os usuários finais possam explorar os dados por conta própria.
- **Colaboração em Tempo Real:** Assim como outras ferramentas do Google, permite a colaboração em tempo real, facilitando o trabalho em equipe na construção e análise de relatórios.
- **Acessibilidade:** É uma ferramenta gratuita, tornando-a acessível para pequenas empresas, freelancers, e educadores.

Importância na Visualização de Dados

- **Transforma Dados em Histórias:** Com o Looker Studio, você não está apenas apresentando números e gráficos; você está contando uma história com seus dados. Isso é essencial para comunicar insights de forma eficaz para diferentes públicos, seja em reuniões de negócios, conferências ou salas de aula.
- **Tomada de Decisão Baseada em Dados:** Ao visualizar dados de forma clara e interativa, o Looker Studio apoia a tomada de decisões informadas. Isso é crucial em ambientes empresariais e de pesquisa, onde decisões baseadas em dados podem levar a melhores resultados.
- **Educação e Aprendizado:** Para estudantes e profissionais que estão aprendendo sobre análise de dados, é uma ferramenta valiosa para entender como os dados podem ser explorados e apresentados de forma significativa.
- **Democratização dos Dados:** O Looker Studio torna a análise de dados mais acessível para pessoas sem formação técnica avançada, democratizando o acesso a insights poderosos que antes eram restritos a analistas de dados ou cientistas de dados.

Para refletir: Como você acha que poderia aplicar o Looker Studio em seus projetos ou aprendizado?

Dashboards - O que é isso?

Conceito chave para o Looker. Ah, dashboards! Eles são uma peça fundamental no quebra-cabeça da interpretação de dados. Pense em um dashboard como o painel de controle de um carro. Assim como você olha para o painel para obter informações instantâneas sobre a velocidade, o nível de combustível ou a temperatura do motor, os dashboards em análise de dados fornecem uma visão rápida e clara do que está acontecendo com seus dados. Vamos detalhar isso:

O Papel dos Dashboards na Interpretação de Dados

- **Visão Geral Imediata:** Dashboards são projetados para oferecer uma visão compreensível à primeira vista. Eles usam gráficos, indicadores e resumos para mostrar os principais indicadores de desempenho (KPIs) e outras métricas importantes.
- **Tomada de Decisão Informada:** Ao fornecer dados em tempo real ou atualizados regularmente, os dashboards permitem que gestores e tomadores de decisão reajam rapidamente às mudanças e tendências.
- **Interatividade e Exploração de Dados:** Muitos dashboards modernos são interativos, permitindo que os usuários filtrem, perfurem (drill down) e explorem os dados para obter insights mais detalhados.
- **Facilita a Compreensão:** A visualização de dados em dashboards transforma números complexos e conjuntos de dados em formatos visuais mais fáceis de entender, como gráficos de barras, linhas, mapas de calor, etc.
- **Monitoramento Contínuo:** Dashboards são excelentes para monitoramento contínuo. Por exemplo, um dashboard de vendas pode mostrar o desempenho atual em comparação com as metas, enquanto um dashboard de mídia social pode rastrear o engajamento e o crescimento ao longo do tempo.
- **Personalização para Necessidades Específicas:** Eles podem ser personalizados para atender às necessidades específicas de diferentes departamentos ou funções dentro de uma organização. Um CFO, por exemplo, pode ter um dashboard financeiro, enquanto um gerente de marketing pode ter um focado em campanhas e análises de clientes.
- **Democratização dos Dados:** Os dashboards ajudam a democratizar o acesso aos dados dentro de uma organização, tornando-os acessíveis e compreensíveis para não especialistas.

Para refletir: Pense o dash mais importante da vida de vocês, eu já canto a pedra, entrem no Instagram de vocês, cliquem na foto no canto da direita abaixo, e vão nas configurações clicando no ícone de menu e vão em “Tempo gasto”. Reflitam sobre o tempo médio que vocês gastam no Instagram por dia, vou mostrar o meu tempo para vocês (hehehe):



Agora estou refletindo se não estou gastando tempo de mais, vamos fazer um pequeno cálculo em um ano, com média de 52 minutos, $365 \times 52 = 18980$, dividindo por 60 = em torno

de 316 horas, o que dariam 13 dias do ano só no Instagram, mais de um dia por mês. E vocês quanto gastam?

Conceitos Básicos do Google Looker

Adentrar no mundo do Google Looker Studio (antigamente conhecido como Google Data Studio) é como entrar em uma cozinha cheia de ingredientes e ferramentas para criar pratos deliciosos de dados! Vamos começar com os conceitos básicos: Fontes de Dados, Conectores e Visualizações.

Fontes de Dados

Imagine que você está preparando um bolo. Os ingredientes são as fontes de dados. No Looker Studio, as fontes de dados são onde suas informações estão armazenadas. Isso pode incluir:

- Google Sheets: como uma planilha com dados de vendas.
- Bancos de Dados: como SQL, BigQuery, onde você armazena dados transacionais.
- Google Analytics: para dados de tráfego do site.
- Outras fontes: como arquivos CSV, Google Ads, e muito mais.

Esses dados são os blocos de construção fundamentais para suas análises e relatórios.

Conectores

Voltando à analogia do bolo, se as fontes de dados são os ingredientes, os conectores são como as batedeiras ou colheres que misturam esses ingredientes. No Looker Studio, conectores são as ferramentas que você usa para acessar e trazer dados das fontes de dados para dentro da sua "cozinha" (o Looker Studio).

Existem conectores nativos do Google, como para o Google Sheets, Analytics, e BigQuery, e conectores de terceiros para serviços como Facebook, Twitter, ou até bancos de dados SQL. O ponto chave é que eles facilitam a conexão entre a fonte dos seus dados e o Looker Studio.

Visualizações

Agora, você tem os ingredientes (dados) e as ferramentas para misturá-los (conectores), então o que segue? A apresentação do prato (seus dados) - isto é, as visualizações. As visualizações são como você apresenta seus dados de uma forma que seja fácil de entender e atraente visualmente.

No Looker Studio, você pode criar:

- Gráficos de Barras e Colunas: ótimos para comparações.
- Linhas do Tempo e Gráficos de Linhas: perfeitos para mostrar tendências ao longo do tempo.
- Mapas de Calor e Gráficos Geográficos: para dados relacionados a localizações.
- Tabelas e Listas: para exibir dados em formato bruto, mas organizado.
- Gráficos de Pizza e Donut: para mostrar proporções e percentuais.

Cada um desses tipos de visualizações ajuda a contar uma história diferente com seus dados, e a beleza do Looker Studio é que você pode personalizar essas visualizações para atender às suas necessidades específicas.



Fonte: https://learning.oreilly.com/library/view/data-storytelling-with/9781800568761/B16568_04.xhtml

Para refletir: Como você se sente sobre começar a usar o Looker Studio? Há algum projeto ou conjunto de dados específico que você tem em mente para explorar com essas ferramentas?

Arquivos CSV

Os arquivos CSV são como os cadernos de anotações do mundo dos dados, simples, mas extremamente úteis. CSV significa "Comma-Separated Values" (Valores Separados por Vírgula).

Estrutura de um Arquivo CSV

Imagine um arquivo CSV como uma tabela simples. Aqui está como ele é estruturado:

- **Linhas e Colunas:** Cada linha do arquivo representa uma linha da tabela, e cada coluna é separada por uma vírgula (ou outro delimitador, como ponto e vírgula em algumas regiões).
- **Cabeçalho:** A primeira linha geralmente contém cabeçalhos, que são os nomes das colunas. Isso é como os rótulos no topo de cada coluna em uma tabela.
- **Dados:** A partir da segunda linha, você encontrará os dados reais. Cada linha representa um registro ou uma observação.
- **Formato de Texto Simples:** Os arquivos CSV são salvos em formato de texto simples, o que os torna fáceis de serem lidos por humanos e máquinas.

Exemplo

Vamos dizer que você tem um arquivo CSV com dados de vendas de uma loja. Pode parecer algo assim:

```
Data,Produto,Quantidade,Preço
2024-01-01,Camiseta,10,19.99
2024-01-01,Calça,5,49.99
2024-01-02,Camiseta,7,19.99
```

Neste exemplo, cada linha depois do cabeçalho representa uma venda, com a data, o produto vendido, a quantidade e o preço.

Uso em Análise de Dados

- **Flexibilidade:** Arquivos CSV podem ser usados em praticamente qualquer ferramenta de análise de dados, desde planilhas simples como o Excel até linguagens de programação mais avançadas como Python e R.
- **Importação e Exportação:** A maioria dos sistemas de banco de dados e ferramentas de análise de dados suporta importação e exportação de dados no formato CSV, tornando-o um formato padrão para transferência de dados.
- **Análise e Manipulação:** Você pode abrir um arquivo CSV em uma ferramenta de análise de dados e realizar várias operações, como filtragem, ordenação, agregação (como somas e médias), e visualização (criação de gráficos e tabelas).
- **Integração de Dados:** CSVs são ótimos para integrar dados de diferentes fontes. Por exemplo, você pode combinar dados de vendas de diferentes lojas, cada uma exportando seus dados como um arquivo CSV.
- **Simplicidade e Eficiência:** Por serem simples e leves, os CSVs são eficientes para armazenar e manipular grandes quantidades de dados tabulares.

Cuidados

Formato de Dados: Tenha cuidado com os formatos, especialmente datas e números, pois podem variar dependendo da região.

Delimitadores e Citações: Alguns dados podem conter vírgulas, o que pode confundir a estrutura do CSV. Nesses casos, é comum o uso de aspas para encapsular os dados.

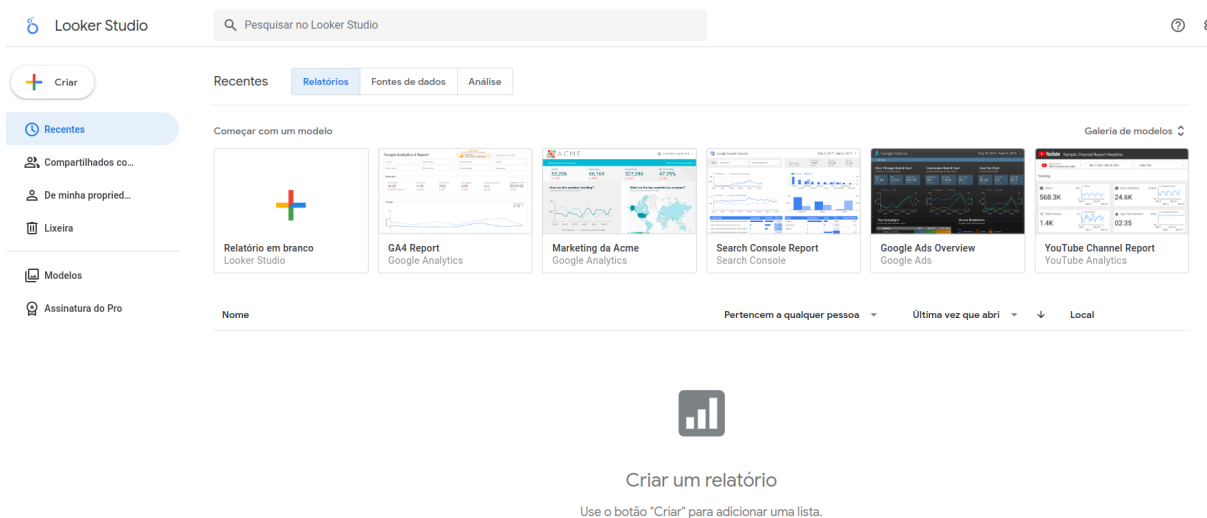
Demonstração Prática

Para a parte prática iremos utilizar um datasets público de vendas de jogos de videogame disponível em:

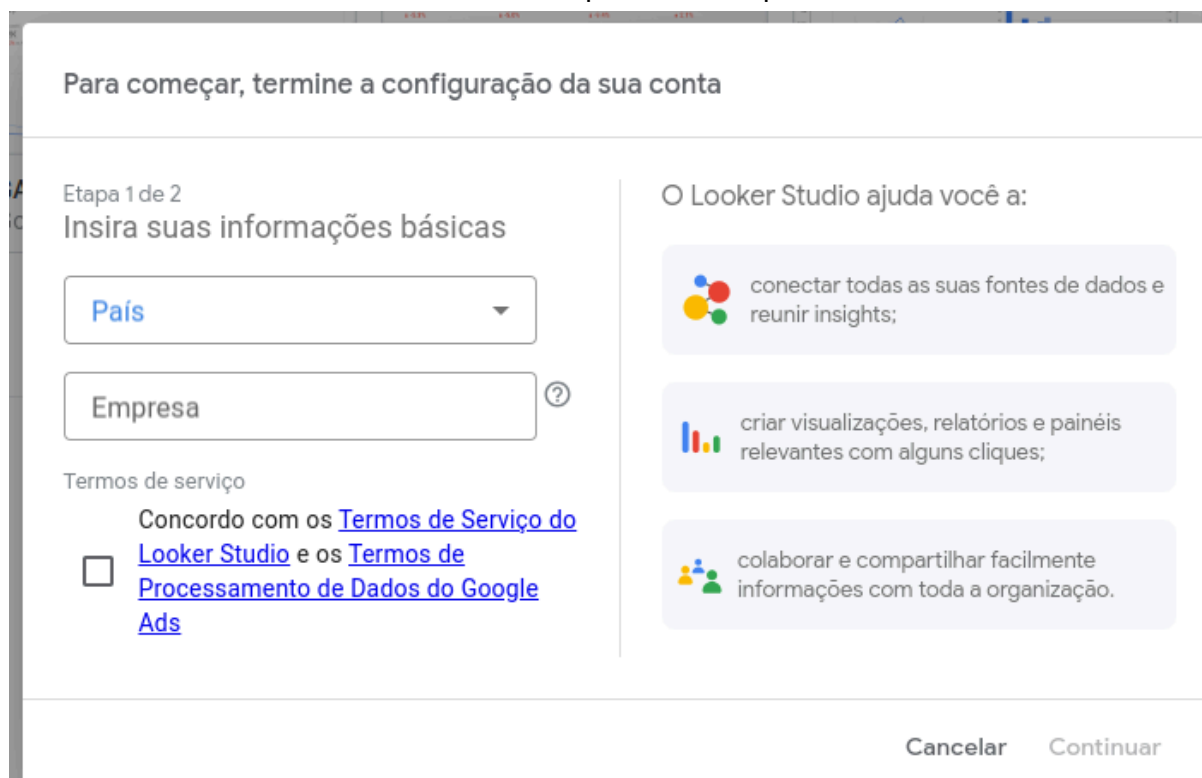
<https://www.kaggle.com/datasets/gregorut/videogamesales?resource=download>

E também será disponibilizado no moodle o dataset

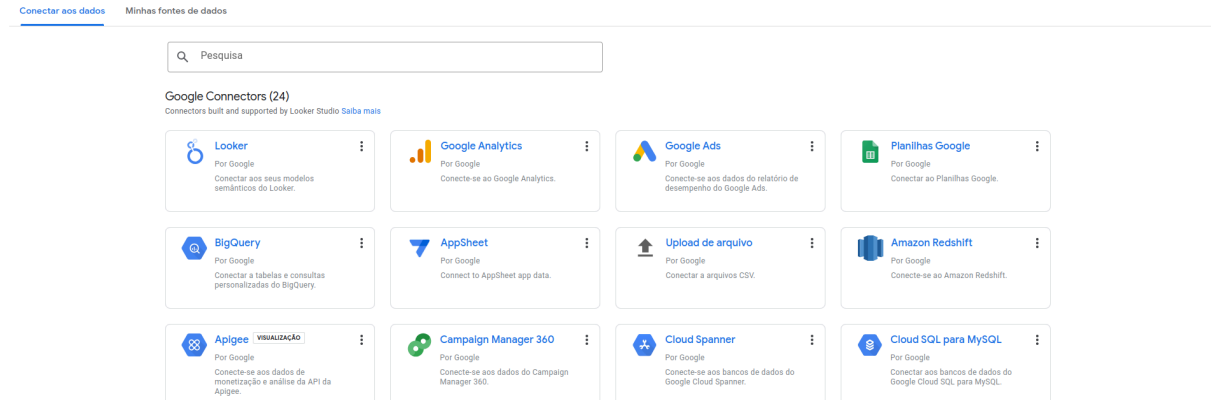
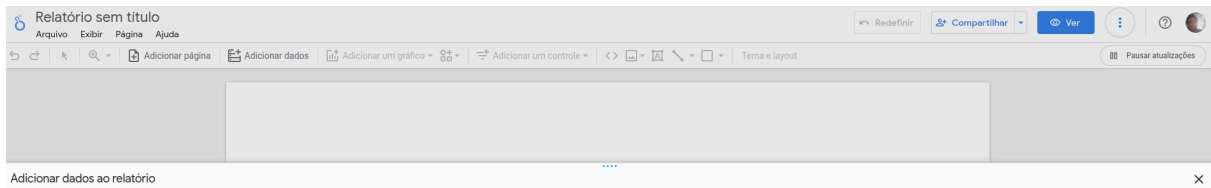
Interface



Quando somos novos no Looker Studio irá aparecer essa primeira tela:

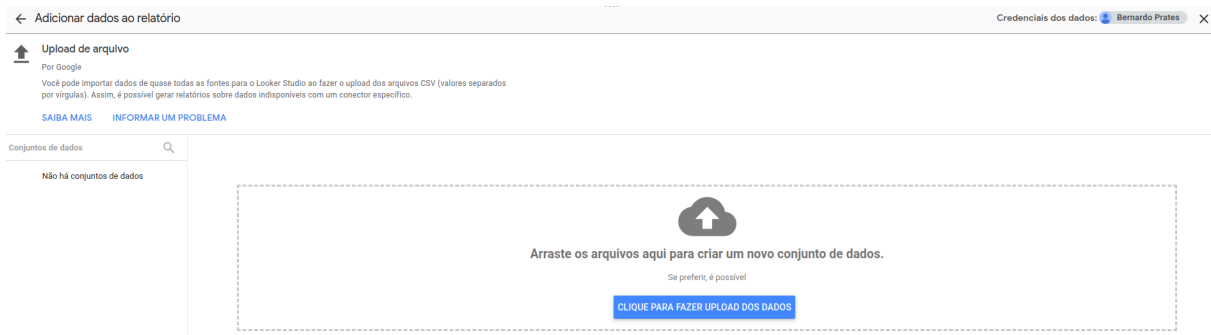


Após essa tela, quando vamos criar um novo relatório teremos algo mais ou menos assim:

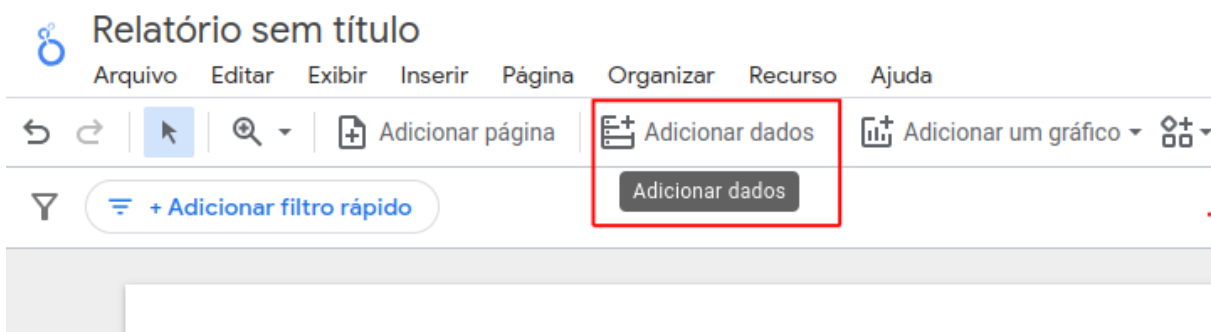


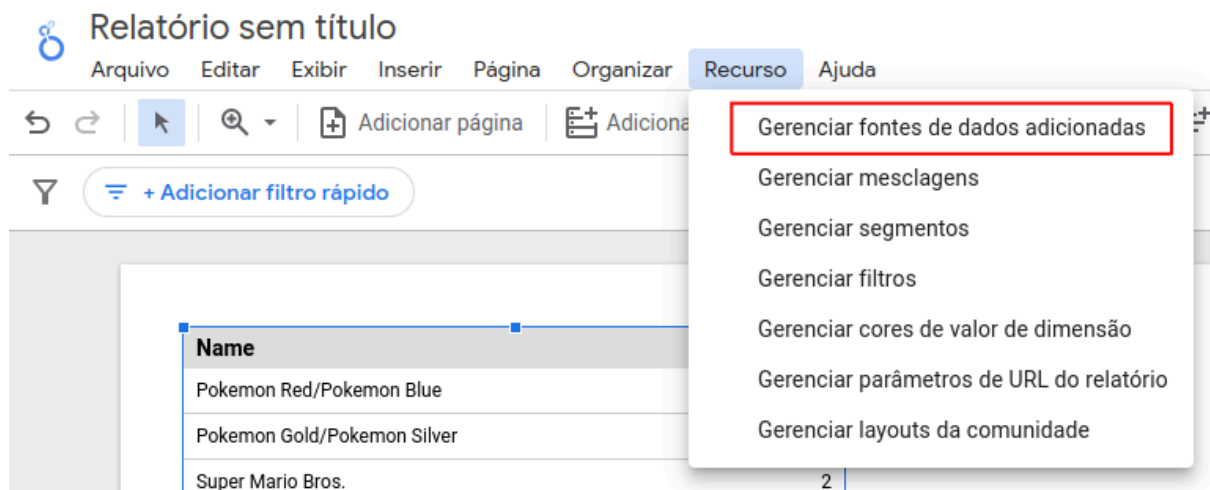
Temos que escolher o conector e também no nosso dataset!

Vamos escolher a opção **upload de arquivo** para importar o nosso dataset:

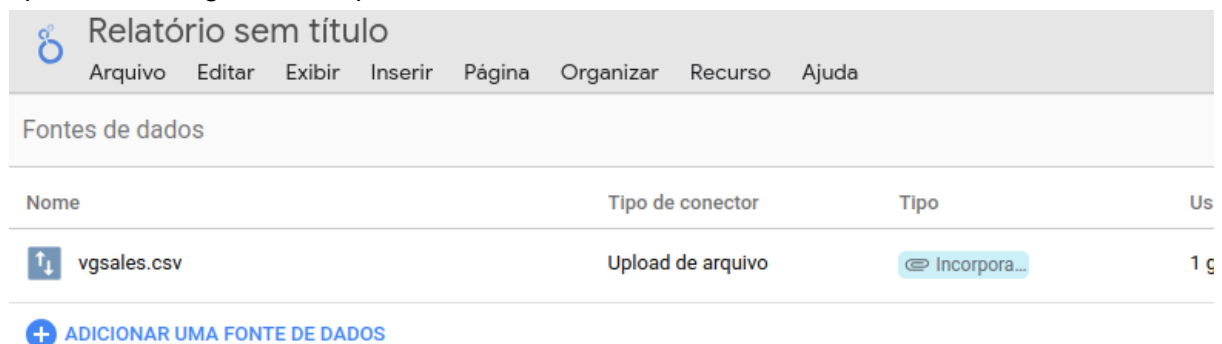


Caso a gente queira adicionar outras fontes de dados é possível também:

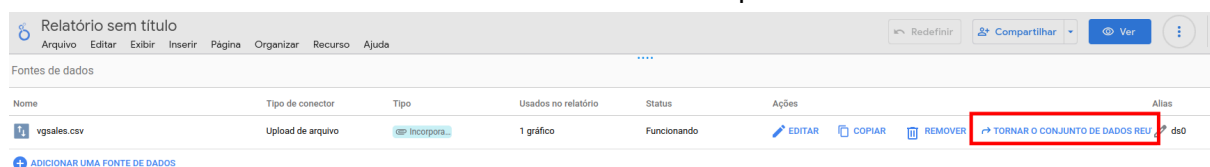




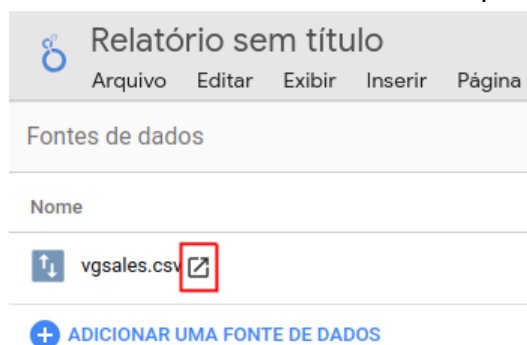
Aparecerá a seguinte tela, podemos ir em “Adicionar uma fonte de dados”



Nessa tela temos a opção “Tornar o Conjunto de Dados Reutilizável” que é uma funcionalidade que permite aos usuários criar e configurar conjuntos de dados que podem ser usados várias vezes em diferentes relatórios ou componentes dentro do Looker Studio.



Tornando esse conjunto reutilizável podemos acessar ele para fazer algumas configurações, uma das formas de acessar esse arquivo é pelo atalho que aparece no nome do arquivo:



Então vai aparecer a seguinte tela onde iremos trabalhar com os dados do nosso arquivo que foi importado:

Campo	Tipo	Agregação padrão	Descrição
DIMENSÕES (11)			
EU_Sales	Número	Soma	
Genre	Texto	Nenhum	
Global_Sales	Número	Soma	
JP_Sales	Número	Soma	
NA_Sales	Número	Soma	
Name	Texto	Nenhum	
Other_Sales	Número	Soma	
Platform	Texto	Nenhum	
Publisher	Texto	Nenhum	
Rank	Número	Soma	
Year	Ano (AAAA)	Nenhum	
MÉTRICAS (1)			
Record Count	Número	Automático	

Importante ressaltar que os resultados das consultas são geralmente armazenados em um cache temporário no Looker Studio para melhorar o desempenho durante a exploração de dados, construção de visuais ou interação com relatórios. O cache é atualizado automaticamente em intervalos regulares, importante escolher adequadamente o tempo de atualização dos dados uma vez que há um impacto no desempenho do relatório enquanto a atualização da fonte de dados está acontecendo. Dependendo do tamanho do relatório e do número de consultas geradas, a atualização pode demorar um pouco.

Podemos modificar esse tempo de atualização apertando no campo “Atualização dos dados” da tela acima onde aparecerá o seguinte modal:

Atualização de dados

Upload de arquivo: vgsales.csv

Esta fonte de dados verifica se há dados novos a cada 12 horas. [Saiba mais](#)

CANCELAR

Obs: No Google Looker Studio, quando você conecta um arquivo CSV como fonte de dados, as opções para a frequência de atualização dos dados são mais limitadas em comparação com outras fontes de dados, como o Google Sheets ou o BigQuery. Isso se deve à natureza da conexão com um arquivo CSV, que é essencialmente uma importação de dados estática, ao invés de uma conexão ao vivo.

Controle de acesso

Quando você compartilha uma fonte de dados ou um relatório com outras pessoas que usam esta fonte de dados, elas poderão visualizar os dados, mesmo que não tenham acesso ao conjunto de dados subjacente. Você pode configurar a fonte de dados para usar

as credenciais do Visualizador, o que exige que qualquer pessoa que use esta fonte de dados tenha suas próprias credenciais para acessar os dados.

The screenshot shows a data source configuration interface for 'vgsales.csv'. At the top, there's a header with 'Escopo: Reutilizável', 'Credenciais dos dados: Bernardo Prates', 'Atualização dos dados: 12 horas', and 'Acesso aos layouts da comum'. Below this is a table with columns 'Campo', 'Tipo', 'Agregação padrão', and 'Descrição'. The table lists dimensions (DIMENSÕES) and metrics (MÉTRICAS). A modal dialog titled 'Atualizar credenciais da fonte de dados' is open, showing two options: 'Credenciais do proprietário' (selected) and 'Credenciais do leitor'. The 'Credenciais do proprietário' option includes a description and a dropdown for 'Proprietário das credenciais' showing 'Bernardo Prates'. The 'Credenciais do leitor' option also has a description. At the bottom of the dialog are 'Cancelar' and 'Atualizar' buttons.

Campo	Tipo	Agregação padrão	Descrição
DIMENSÕES (11)			
EU_Sales	123 Número	Soma	
Genre	RBC Texto	Nenhum	
Global_Sales	123 Número		
JP_Sales	123 Número		
NA_Sales	123 Número		
Name	RBC Texto		
Other_Sales	123 Número		
Platform	RBC Texto		
Publisher	RBC Texto		
Rank	123 Número		
Year	Ano (AAAA)		
MÉTRICAS (1)			
Record Count	123 Número		

Atualizar credenciais da fonte de dados



As credenciais determinam como essa fonte de dados acessa as informações ?



Credenciais do proprietário

Permite que essa fonte de dados se conecte ao conjunto de dados subjacentes usando as credenciais que pertencem à pessoa abaixo

Proprietário das credenciais: Bernardo Prates



Credenciais do leitor

Requer que todos os usuários dessa fonte de dados tenham as próprias credenciais para acessar o conjunto de dados subjacente.



Se você alterar as credenciais, as pessoas poderão perder o acesso aos dados.

[Saiba mais](#)

Cancelar

Atualizar

Podemos implementar também o recurso de filtragem de e-mail para a fonte de dados, que permite escolher um campo no conjunto de dados que contém o e-mail do usuário que pode ver cada linha de dados.

Escopo: → Reutilizável

← EDITAR CONEXÃO | FILTRAR POR E-MAIL

Campo ↓		Tipo ↓		Agregação
DIMENSÕES (11)				
EU_Sales	⋮	123	Número	▼ Soma
Genre	⋮	ABC	Texto	▼ Nenhuma
Global_Sales	⋮	123	Número	▼ Soma
JP_Sales	⋮	123	Número	▼ Soma
NA_Sales	⋮	123	Número	▼ Soma
Name	⋮	ABC	Texto	▼ Nenhuma

vgsales.csv
Escopo: → Reutilizável | Credenciais dos dados: [Leitor](#) | Atualização dos dados: 12 horas | Acesso aos dados: [Ver](#)

← TODOS OS CAMPOS

Filtrar por e-mail

Se seus dados contêm endereços de e-mail, você pode filtrar esta fonte para exibir somente informações correspondentes às permissões do endereço de e-mail do usuário conectado no momento. Os visualizadores de relatórios precisam autorizar que esta fonte de dados acesse o endereço de e-mail deles. [Saiba mais](#)

☒ Filtrar dados por e-mail do visualizador

ABC Genre
🔍 Digite para pesquisar

Grupo padrão

ABC Genre

ABC Name

ABC Platform

ABC Publisher

📅 Year

Este campo contém os endereços a serem filtrados.

Editando o schema dos dados

Dimensão e Métrica

- **Dimensão:** O campo de dimensão contém um valor para cada linha do conjunto de dados e é tipicamente usado para agrupar os dados em gráficos.
- **Métrica:** Uma métrica é um campo de resumo calculado em várias linhas de dados em tempo real quando usado em gráficos. Campos de dimensão podem ser adicionados como métricas na configuração do gráfico, aplicando a agregação

"Contagem Distinta" para campos não numéricos e o método de agregação padrão definido para campos de dimensão numéricos.

DIMENSÕES (11)						
EU_Sales	:	123	Número	▼	Soma	▼
Genre	:	ABC	Texto	▼	Nenhum	
Global_Sales	:	123	Número	▼	Soma	▼
JP_Sales	:	123	Número	▼	Soma	▼
NA_Sales	:	123	Número	▼	Soma	▼
Name	:	ABC	Texto	▼	Nenhum	
Other_Sales	:	123	Número	▼	Soma	▼
Platform	:	ABC	Texto	▼	Nenhum	
Publisher	:	ABC	Texto	▼	Nenhum	
Rank	:	123	Número	▼	Nenhum	▼
Year	:	📅	Ano (AAAA)	▼	Nenhum	
MÉTRICAS (1)						
Record Count	:	123	Número	▼	Automático	

O Looker Studio detecta automaticamente o tipo de dado de cada campo com base no conjunto de dados subjacente. Para cada campo, você pode definir também um método de agregação padrão na fonte de dados. A agregação padrão é o método de agregação (soma, média, mediana, etc.) que é calculado visualmente por padrão quando um campo é adicionado a ele. Por padrão, o Looker Studio define a agregação padrão para Nenhum para todos os campos de dimensão não numéricos, Soma para todos os campos de dimensão numéricos e Automático para todas as métricas.

Vamos ajustar nossa base para termos nomes em português e deixar uma descrição em cada campo, vamos alterar os nomes para o seguinte padrão, porém você pode se sentir a vontade de colocar os nomes que desejar o padrão abaixo é [nome anterior] - [nome alterado] - [descrição]:

- EU_Sales - Vendas_Europa - Vendas do jogo na Europa, expressas em milhões
- Genre - Genero - O gênero do jogo. Este campo descreve a categoria ou estilo do jogo, como ação, aventura, esportes, etc.
- Global_Sales - Vendas_Globais - Vendas totais do jogo em todo o mundo
- JP_Sales - Vendas_Japao - Vendas do jogo no Japão, em milhões
- NA_Sales - Vendas_AN - Vendas do jogo na América do Norte, expressas em milhões
- Name - Nome - O nome do jogo
- Other_Sales - Vendas_Outros - Vendas do jogo em outras partes do mundo
- Platform - Plataforma - A plataforma na qual o jogo foi lançado (por exemplo, PC, PS4, Xbox, Nintendo Switch, etc.)
- Publisher - Empresa - A empresa que publicou o jogo

- Rank - Rank - O ranking do jogo em termos de vendas totais
- Year - Ano - O ano de lançamento do jogo

Em Métricas, vamos chamar o record count de Quantidade de registros, pois essa métrica é uma função padrão que conta o número de linhas (ou registros) do nosso conjunto de dados.

DIMENSÕES (11)					
Vendas_Europa	123	Número	▼	Soma	Vendas do jogo na Europa, expressas em milhões
Genero	ABC	Texto	▼	Nenhum	O gênero do jogo. Este campo descreve a categoria ou estilo do jogo
Vendas_Globais	123	Número	▼	Soma	Vendas totais do jogo em todo o mundo
Vendas_Japao	123	Número	▼	Soma	Vendas do jogo no Japão, em milhões
Vendas_AN	123	Número	▼	Soma	Vendas do jogo na América do Norte, expressas em milhões
Nome	ABC	Texto	▼	Nenhum	O nome do jogo
Vendas_Outros	123	Número	▼	Soma	Vendas do jogo em outras partes do mundo
Plataforma	ABC	Texto	▼	Nenhum	A plataforma na qual o jogo foi lançado (por exemplo, PC, PS4, Xbox)
Empresa	ABC	Texto	▼	Nenhum	A empresa que publicou o jogo
Rank	123	Número	▼	Nenhum	O ranking do jogo em termos de vendas totais
Ano	📅	Ano (AAAA)	▼	Nenhum	O ano de lançamento do jogo
MÉTRICAS (1)					
Quantidade de registros	123	Número	▼	Automático	

Podemos ocultar um campo:

DIMENSÕES (11)			
Vendas_Globais	123	Número	⋮
Ano	📅	Ano (AAAA)	⋮
Empresa	ABC	Texto	⋮
Genero	ABC	Texto	⋮
Nome			⋮
Plataforma			⋮
Rank			⋮
Vendas_AN			⋮
Vendas_Europa			⋮
Vendas_Japao			⋮
Vendas_Outros	123	Numero	⋮

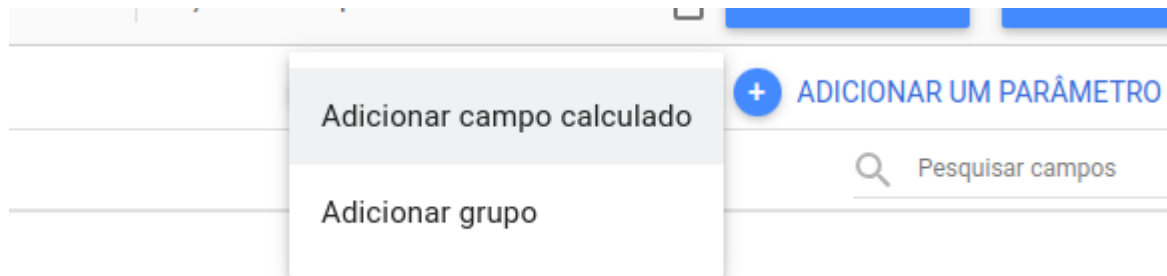
Contar
 Contar os diferentes
 Duplicar
 Remover
 Ocultar

Entendo os dados e ferramenta explorar

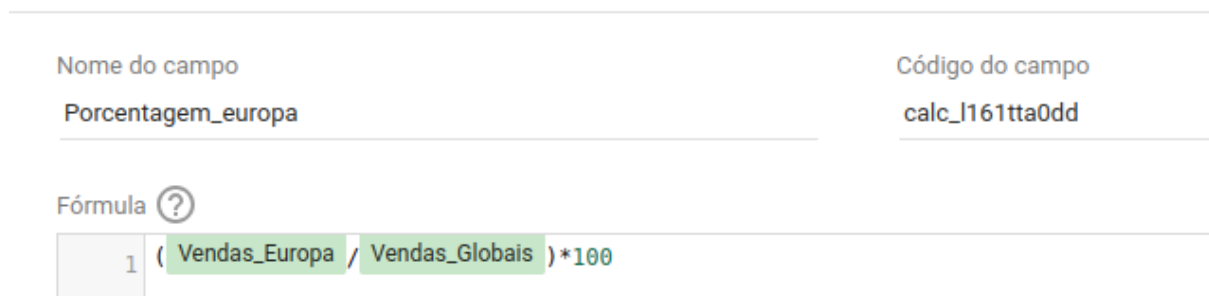
Vamos adicionar um campo calculado para a gente começar a se ligar em certas particularidades e certos pensamentos que vamos ter que ter para entender os dados, vamos fazer de uma forma, que pode não fazer muito sentido, gerando um provável “erro” e corrigir isso depois, só para a gente pegar a forma de pensar, pode ser? Não me xinguem hehehe.

Vamos lá o desafio é o seguinte, criar um campo que mostre a porcentagem de vendas na Europa em relação às vendas globais, antes de ver a resposta, tente fazer.

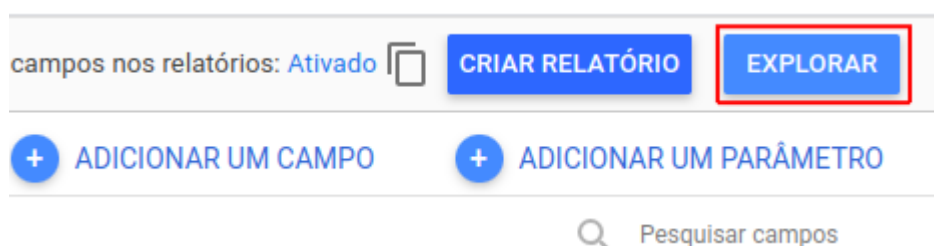
Primeiro precisamos adicionar um campo calculado:



Sabemos que para calcular uma porcentagem temos que pegar o valor das vendas na Europa, dividir pelas vendas no mundo e multiplicar por 100 para achar a porcentagem. Porém, replicando essa fórmula literalmente teríamos isso:



Que, pode dar errado ou pode dar certo, aí que mora o perigo. Existe uma opção nos dados que se chama “Explorar” ela fica bem em cima a direita:



Explorar é uma ferramenta de exploração que permite que você examine dados rapidamente e obtenha insights de forma mais ágil. Ele fornece uma área temporária onde você pode facilmente explorar os dados e realizar análises ad hoc. No Explorer, você cria explorações em uma ou mais abas. Qualquer trabalho que você faça no Explorer **não é salvo automaticamente**, ao contrário dos relatórios e fontes de dados. No entanto, você pode salvar sua exploração, se desejar, para uso posterior. A tela inicial em relação ao nosso dataset vai aparecer assim:

Adicionar um gráfico -

Filtro:

	Nome	Quantidade de registros
1.	Need for Speed: Most Wanted	12
2.	Madden NFL 07	9
3.	Ratatouille	9
4.	LEGO Marvel Super Heroes	9
5.	FIFA 14	9
6.	Madden NFL 08	8
7.	Lego Batman 3: Beyond Gotham	8
8.	Monopoly	8
9.	LEGO Jurassic World	8
10.	Terraria	8
11.	Cars	8
12.	Angry Birds Star Wars	8
13.	The LEGO Movie Videogame	8
14.	LEGO Harry Potter: Years 5-7	8
15.	FIFA 15	8
16.	LEGO Star Wars II: The Original Trilogy	8
17.	LEGO The Hobbit	8
18.	FIFA Soccer 13	8
19.	Tomb Raider: Legend	7
20.	Major League Baseball 2K11	7

1 - 100 / 11493

Guia 1

Gráfico > Tabela

Fontes de dados: **COMBINAR DADOS** **Adicionar dimensão**

Dimensão de período: **Adicionar dimensão**

Dimensão: **Nome** **Adicionar dimensão**

Detalhar: ☐

Métrica: **Quantidade de r...** **Adicionar métrica**

Métricas opcionais: ☐

Campos disponíveis:

- Ano
- Empresa
- Genero
- Nome
- Plataforma
- Rank
- Vendas_AN
- Vendas_Europa
- Vendas_Globais
- Vendas_Japao
- Vendas_Outros
- Percentagem_europa
- Quantidade de registr...

Ele agrupa os nomes e coloca a quantidade de registros conforme o nome. Então galera, temos que conhecer muito bem os dados aos quais estamos trabalhando, por exemplo, Need for Speed, temos jogos que tem o mesmo nome, porém muda a plataforma, se deixarmos configurado o padrão e colocar a dimensão apenas por nome, olha os resultados que temos:

	Nome	Quantidade de registros	Porcentagem_europa	Vendas_Europa	Vendas_Globais
1.	Need for Speed: Most Wanted	12	490,87	5,85	14,08
2.	Need for Speed: ProStreet	7	275,15	2,56	8,58
3.	Need For Speed: Undercover	7	286,63	2,55	6,92
4.	Need for Speed Carbon	7	259,61	2,14	6,42
5.	Need for Speed Underground 2	6	230,26	4,41	10,97
6.	Need for Speed Underground	5	218,72	3,98	10,46
7.	Need for Speed: The Run	5	242,94	2,07	4,27
8.	Need for Speed Rivals	5	240,52	2,39	5,11
9.	Need for Speed: Hot Pursuit	4	191,26	2,67	5,89
10.	Need for Speed: Shift	4	236,47	2,81	5,42
11.	Need for Speed (2015)	3	210,87	1,85	3,1
12.	Need for Speed: Hot Pursuit 2	3	39,36	0,51	3,73
13.	Need for Speed: Shift 2 Unleashed	3	160,07	0,69	1,32
14.	Need for Speed Carbon: Own the City	2	49,69	1	2,65
15.	Need for Speed: Nitro	2	69,8	0,44	1,24
16.	Need for Speed: Porsche Unleashed	2	64,75	0,42	1,25
17.	Need for Speed III: Hot Pursuit	1	27,56	0,86	3,12
18.	Need For Speed: High Stakes	1	27,71	0,64	2,31
19.	Need for Speed: Most Wanted 5-1-0	1	6,19	0,13	2,1
20.	Need for Speed Underground Rivals	1	28,79	0,38	1,32

1 - 22 / 22

CONFIGURAÇÃO ESTIL

Adicionar dimensão

Dimensão: **Nome** **Adicionar dimensão**

Detalhar: ☐

Métrica: **Quantidade de r...** **Adicionar métrica**

Métricas opcionais: ☐

SUM Percentagem_eu...
SUM Vendas_Europa
SUM Vendas_Globais

A porcentagem de “Need for Speed: Most Wanted” está visivelmente incorreta, note que o valor é 490,87 como já dividimos por 100 essa seria a porcentagem final. Agora eu vou adicionar na dimensão além do nome a plataforma. Note:

	Nome	Plataforma	Quantidade de registros	Porcentagem_europa	Vendas_Europa	Vendas_Globais
1.	Need for Speed: Most Wanted	X360	2	63,81	0,97	2,87
2.	Need for Speed: Most Wanted	PC	2	151,67	0,29	0,38
3.	Need for Speed Underground	PS2	1	39,31	2,83	7,2
4.	Need for Speed (2015)	PC	1	100	0,03	0,03
5.	Need for Speed: Most Wanted	PS3	1	52,25	1,51	2,89
6.	Need for Speed: Hot Pursuit	PS3	1	44,48	1,25	2,81
7.	Need for Speed: Shift	PS3	1	52,61	1,41	2,68
8.	Need for Speed Carbon: Own the City	PSP	1	40,16	0,98	2,44
9.	Need for Speed: ProStreet	PS2	1	0	0	2,41
10.	Need for Speed (2015)	PS4	1	62,34	1,49	2,39
11.	Need For Speed: High Stakes	PS	1	27,71	0,64	2,31
12.	Need for Speed Underground 2	XB	1	35,09	0,8	2,28
13.	Need for Speed: Hot Pursuit	X360	1	44,14	0,98	2,22
14.	Need for Speed: Hot Pursuit 2	PS2	1	14,69	0,31	2,11
15.	Need for Speed: Most Wanted S-1-0	PSP	1	6,19	0,13	2,1
16.	Need For Speed: Undercover	PS3	1	49,04	1,02	2,08
17.	Need for Speed: ProStreet	X360	1	38,83	0,8	2,06
18.	Need for Speed: ProStreet	PS3	1	44,88	0,92	2,05
19.	Need for Speed: The Run	PS3	1	51,96	1,06	2,04
20.	Need for Speed: Shift	X360	1	50,52	0,97	1,92

Já deu uma melhorada, mas note que ainda temos “erros” (na porcentagem), tem dois jogos que ainda possuem 2 registros semelhantes, vou adicionar o ano de lançamento dentro da dimensão, vamos observar:

	Nome	Ano	Plataforma	Quantidade de registros	Porcentagem_europa	Vendas_Europa	Vendas_Globais
1.	Need for Speed Underground	2003	PS2	1	39,31	2,83	7,2
2.	Need for Speed Underground 2	2004	PS2	1	43,77	3,02	6,9
3.	Need for Speed: Most Wanted	2005	PS2	1	40,96	1,79	4,37
4.	Need for Speed III: Hot Pursuit	1998	PS	1	27,56	0,86	3,12
5.	Need for Speed: Most Wanted	2012	PS3	1	52,25	1,51	2,89
6.	Need for Speed: Hot Pursuit	2010	PS3	1	44,48	1,25	2,81
7.	Need for Speed: Shift	2009	PS3	1	52,61	1,41	2,68
8.	Need for Speed Carbon: Own the ...	2006	PSP	1	40,16	0,98	2,44
9.	Need for Speed: ProStreet	2007	PS2	1	0	0	2,41
10.	Need for Speed (2015)	2015	PS4	1	62,34	1,49	2,39
11.	Need For Speed: High Stakes	1999	PS	1	27,71	0,64	2,31
12.	Need for Speed Underground 2	2004	XB	1	35,09	0,8	2,28
13.	Need for Speed: Hot Pursuit	2010	X360	1	44,14	0,98	2,22
14.	Need for Speed: Hot Pursuit 2	2002	PS2	1	14,69	0,31	2,11
15.	Need for Speed: Most Wanted S-1-0	2005	PSP	1	6,19	0,13	2,1
16.	Need For Speed: Undercover	2008	PS3	1	49,04	1,02	2,08
17.	Need for Speed: ProStreet	2007	X360	1	38,83	0,8	2,06
18.	Need for Speed: ProStreet	2007	PS3	1	44,88	0,92	2,05
19.	Need for Speed: The Run	2011	PS3	1	51,96	1,06	2,04
20.	Need for Speed: Shift	2009	X360	1	50,52	0,97	1,92

Agora, sim, então aprendemos que na nossa base existem jogos que tem mesmo nome, na mesma plataforma, porém lançados em anos diferentes:

29.	Need for Speed: Most Wanted	2012	X360
30.	Need for Speed: Most Wanted	2005	PS2
31.	Need for Speed: Most Wanted	2013	WiiU
32.	Need for Speed: Most Wanted	2005	X360

Para refletir: Fica nosso primeiro questionamento, será que é possível? Os dados estão corretos?

Porém, vamos imaginar que queremos ter a porcentagem tendo apenas a dimensão nome mesmo, queremos juntar todas as plataformas e todos os anos de lançamentos e dividir os nossos dados apenas pelo nome mesmo. Então poderíamos melhorar a nossa fórmula para obter a porcentagem de vendas na Europa. Uma ideia é pegar a soma das vendas na Europa e dividir pela soma das vendas globais, nossa fórmula ficaria mais ou menos assim:

Nome do campo

Porcentagem_europa

Fórmula ?

1 (SUM(Vendas_Europa)/SUM(Vendas_Globais))*100

Agora adicionando apenas o campo nome na dimensão, teríamos provavelmente o valor correto:

Nome	Quantidade de registros	Porcentagem_europa
Need for Speed: Most Wanted	12	41,55
Madden NFL 07	9	4,39
Ratatouille	9	30,72
LEGO Marvel Super Heroes	9	41,99
FIFA 14	9	64,99
Madden NFL 08	8	2,82
Lego Batman 3: Beyond Gotham	8	44,98
Monopoly	8	32,34
LEGO Jurassic World	8	43,37
Terraria	8	50,6
Cars	8	14,14
Angry Birds Star Wars	8	38,42
The LEGO Movie Videogame	8	42,78
LEGO Harry Potter: Years 5-7	8	41,59
FIFA 15	8	65,19
LEGO Star Wars II: The Original Trilogy	8	21,09
LEGO The Hobbit	8	56,06
FIFA Soccer 13	8	63,43
Tomb Raider: Legend	7	35,33
Major League Baseball 2K11	7	8,26

Vamos melhorar um pouco mais, vamos tirar a multiplicação por 100 e dizer que esse é um campo de porcentagem, primeira vamos deixar a fórmula assim:

Nome do campo

Porcentagem_europa

Fórmula ?

1 (SUM(Vendas_Europa)/SUM(Vendas_Globais))

E lá no campo vamos mudar de número para porcentagem:

MÉTRICAS (2)

Porcentagem_europa	fx	123	Número	Automático
Quantidade de registros		123	Número	Número
		123	Moeda	Porcentagem
				Duração

Nossos dados:

Nome	Quantidade de registros	Porcentagem_europa
Need for Speed: Most Wanted	12	41,55%
Madden NFL 07	9	4,39%
Ratatouille	9	30,72%
LEGO Marvel Super Heroes	9	41,99%
FIFA 14	9	64,99%
Madden NFL 08	8	2,82%
Lego Batman 3: Beyond Gotham	8	44,98%
Monopoly	8	32,34%
LEGO Jurassic World	8	43,37%
Terraria	8	50,6%
Cars	8	14,14%
Angry Birds Star Wars	8	38,42%
The LEGO Movie Videogame	8	42,78%
LEGO Harry Potter: Years 5-7	8	41,59%
FIFA 15	8	65,19%
LEGO Star Wars II: The Original Trilogy	8	21,09%
LEGO The Hobbit	8	56,06%
FIFA Soccer 13	8	63,43%
Tomb Raider: Legend	7	35,33%
Major League Baseball 2K11	7	8,26%

Agora ficou bem melhor não? Exploramos rapidamente alguns conceitos da ferramenta Explorar para a gente conseguir fazer rápidas consultas, etc. Você pode explorar uma fonte de dados ou um gráfico em um relatório. Ter um local separado para explorar os dados significa que você não precisa criar e gerenciar um relatório se tudo o que você precisa fazer é olhar para os dados e entendê-los. Da mesma forma, você pode explorar qualquer gráfico em um relatório e realizar análises ad hoc nele sem ter que editar o relatório.

Operações com o nosso dataset

Além de editar o esquema, gerenciar a atualização dos dados e o acesso, existem outras operações comuns que você pode realizar com uma fonte de dados no Google Looker Studio, incluindo duplicar ou copiar a fonte de dados, compartilhá-la e mais.

Voltando para a página inicial vamos na opção “Fonte de dados”

Looker Studio

Pesquisar no Looker Studio

Criar

Recentes

Relatórios Fontes de dados Análise

Nome	Pertencem a qualquer pessoa	Última vez que abriu	Local
vgsales.csv	Bernardo Prates	19:08	De minha prop...


Pertencem a qualquer pessoa Última vez que abriu Local


Bernardo Prates 19:08 De minha prop...

- Compartilhar
- Renomear
- Remover
- Gerenciar relatórios adicionados
- Revogar credenciais


Compartilhar: Você pode compartilhar uma fonte de dados com outros selecionando a opção Compartilhar e fornecendo os e-mails dos usuários e o nível de acesso. Ao

compartilhar, você pode gerenciar o acesso por meio de controles como desabilitar a cópia para visualizadores, impedir que os editores adicionem novas pessoas e assim por diante.


Compartilhar com pessoas e grupos 

Compartilhar como  Bernardo Prates

Pessoas com acesso

 Bernardo Prates Editor ▾

Configurações de link




 ▾ Restrito
Somente as pessoas adicionadas podem usar este link

[Copiar link](#) [Concluído](#)

Podemos também editar os detalhes da conexão para uma fonte de dados, atualizar os campos, renomear a fonte de dados, excluí-la e mais...

Relatórios

Vamos criar um novo relatório a partir da página inicial com a nossa fonte de dados de games. Na página inicial, vamos em fonte de dados e vamos clicar duas vezes no nosso arquivo csv aparecendo a seguinte tela que já demos uma modificada:

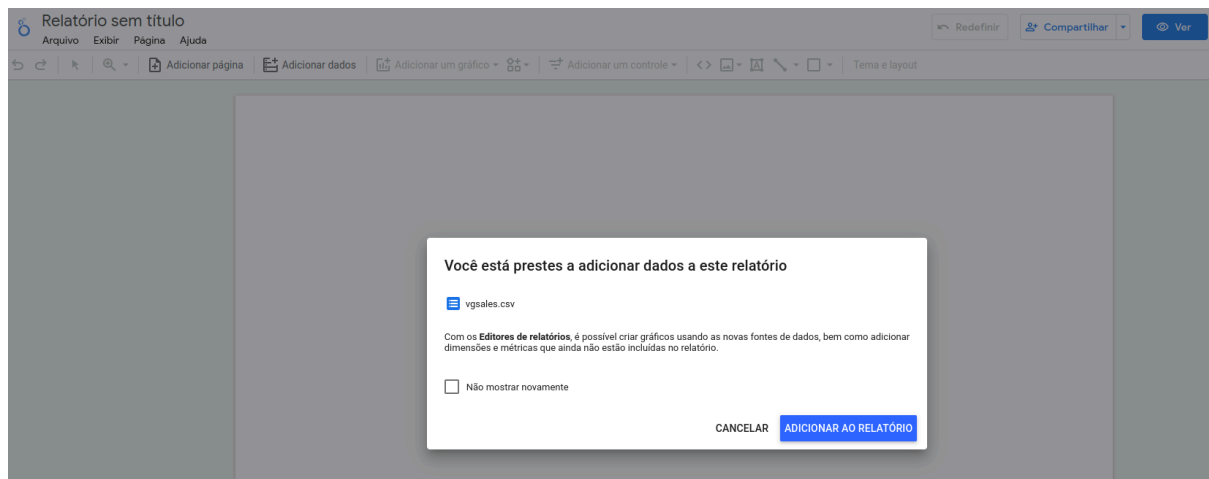
vgsales.csv  [Compartilhar](#)  

Escopo: [Realizador](#) | Credenciais dos dados: [Leitor](#) | Atualização dos dados: 12 horas | Acesso aos layouts da comunidade: Desativado | Edição dos campos nos relatórios: [Ativado](#) [CRIAR RELATÓRIO](#) [EXPLORAR](#)

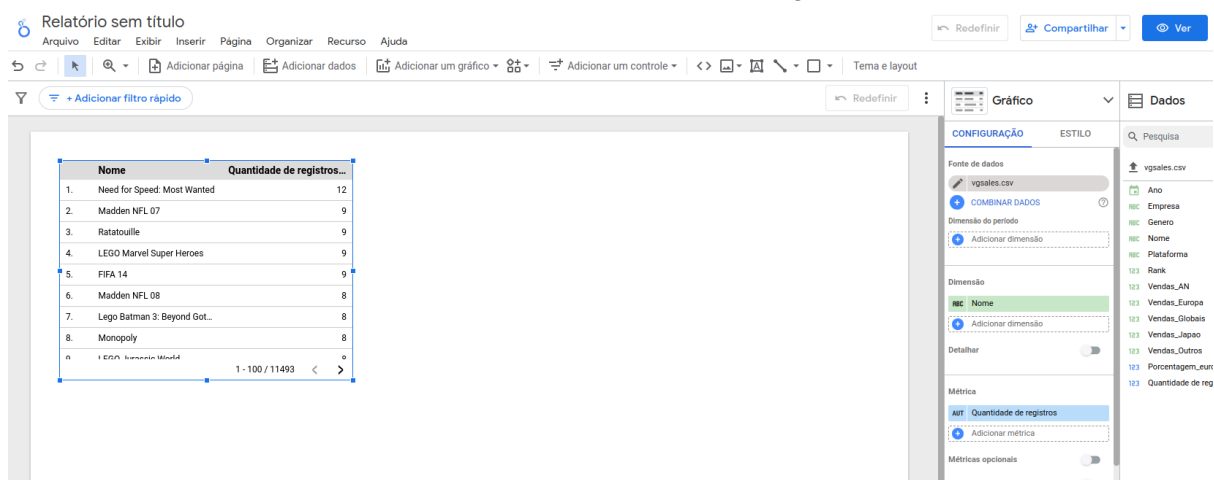
← EDITAR CONEXÃO | FILTRAR POR E-MAIL [+ ADICIONAR UM CAMPO](#) [+ ADICIONAR UM PARÂMETRO](#)

Campo	Tipo	Agregação padrão	Descrição
DIMENSÕES (11)			
Ano	ANO (AAAA)	Nenhum	O ano de lançamento do jogo
Empresa	RBC Texto	Nenhum	A empresa que publicou o jogo
Genero	RBC Texto	Nenhum	O gênero do jogo. Este campo descreve a categoria ou estilo do jogo, como ação, aventura, esportes, etc.
Nome	RBC Texto	Nenhum	O nome do jogo
Plataforma	RBC Texto	Nenhum	A plataforma na qual o jogo foi lançado (por exemplo, PC, PS4, Xbox, Nintendo Switch, etc.)
Rank	123 Número	Nenhum	O ranking do jogo em termos de vendas totais
Vendas_AN	123 Número	Soma	Vendas do jogo na América do Norte, expressas em milhões
Vendas_Europa	123 Número	Soma	Vendas do jogo na Europa, expressas em milhões
Vendas_Globais	123 Número	Soma	Vendas totais do jogo em todo o mundo
Vendas_Japao	123 Número	Soma	Vendas do jogo no Japão, em milhões
Vendas_Outros	123 Número	Soma	Vendas do jogo em outras partes do mundo
MÉTRICAS (2)			
Porcentagem_europa	123 Porcentagem	Automático	
Quantidade de registros	123 Número	Automático	

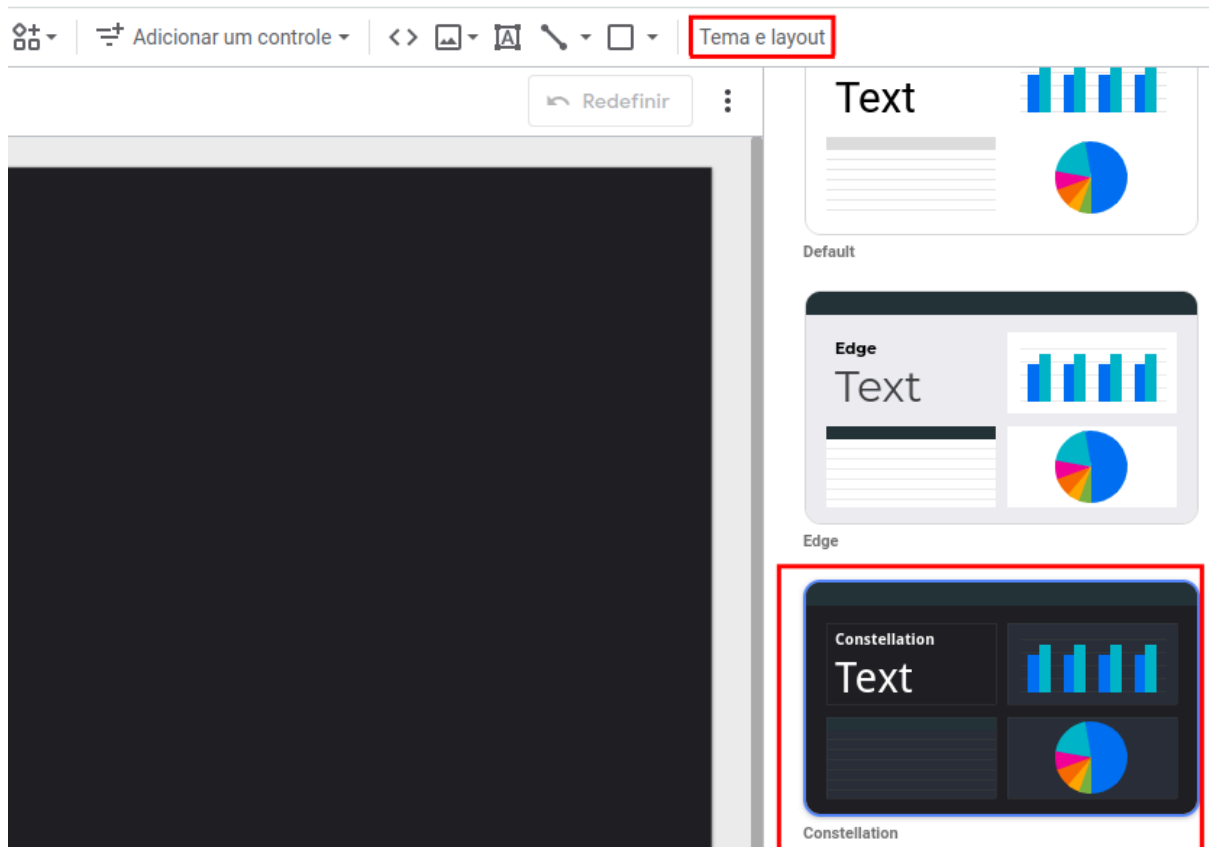
Vamos clicar na opção “Criar Relatório” no campo superior à direita. E teremos a seguinte tela:



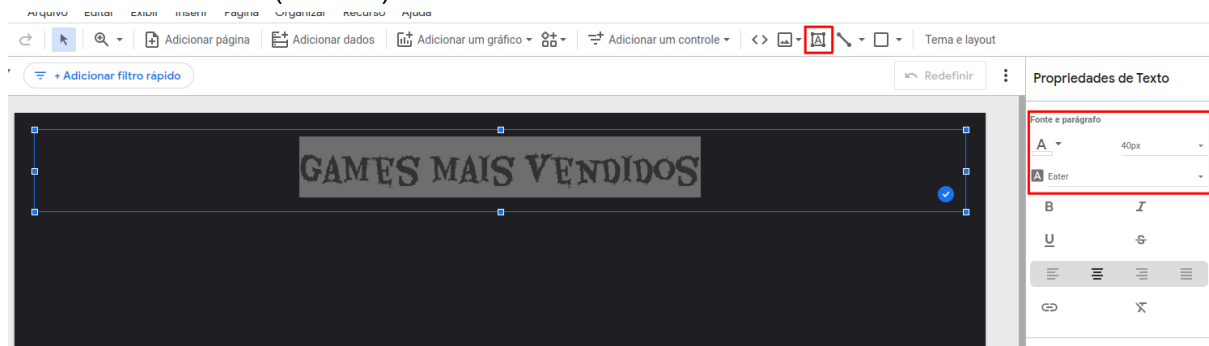
Após clicar em “ADICIONAR AO RELATÓRIO” teremos a seguinte tela:



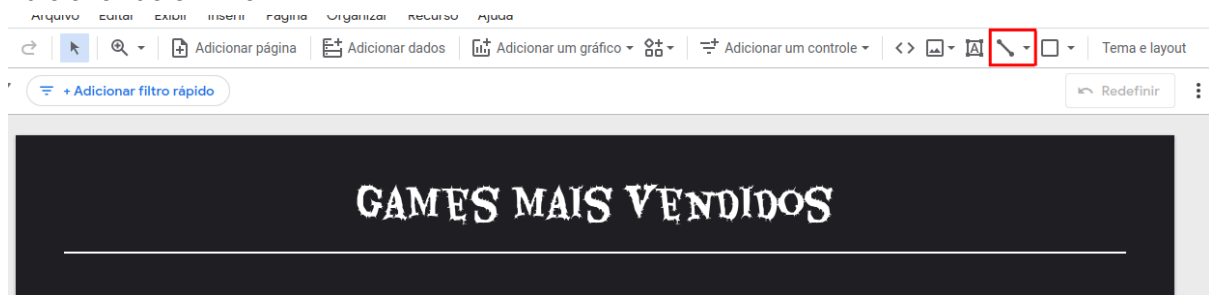
Primeiro vamos alterar o nome do relatório para “Games mais vendidos” e na opção Tema e layout vamos escolher um tema e um layout atrativo. Como é games, vamos escolher um modo mais dark (isso é um chute hehehe).



Vou excluir a tabela inicial, vou adicionar um título e colocar uma linha horizontal, bem clichê mesmo! Vou colocar fonte tamanho 40 e o tipo vai ser a “Eater”, não conheço, mas achei bonitinha (hehehe):



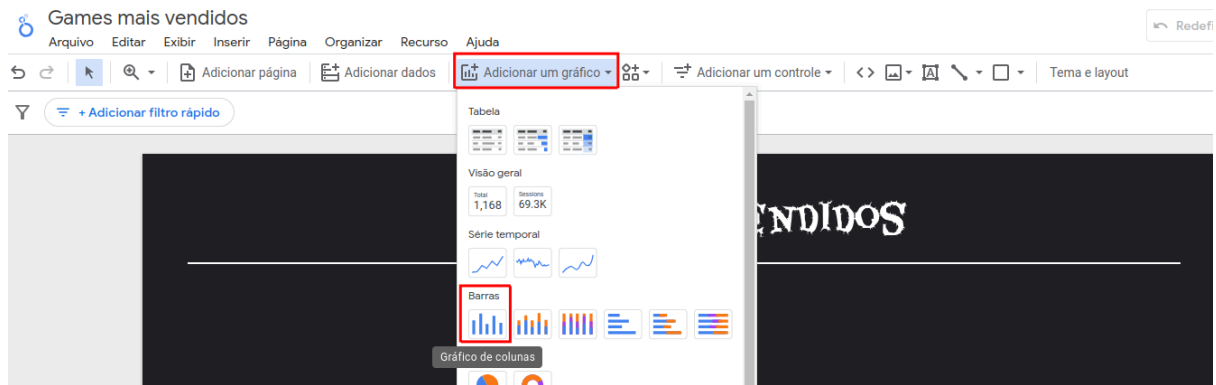
Adicionando a linha:



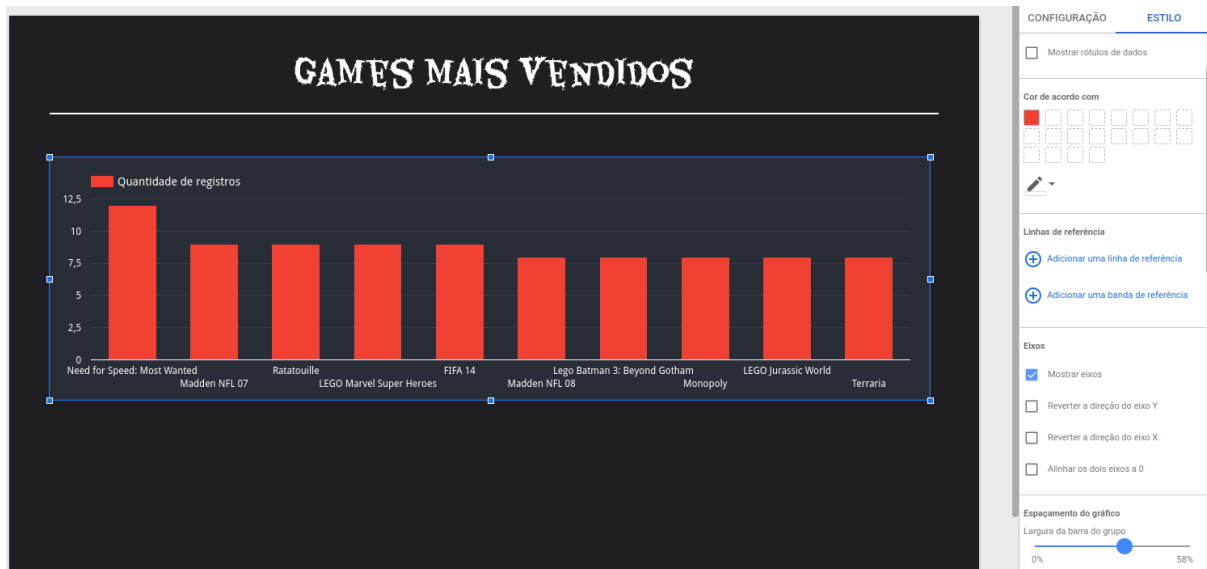
A ideia inicial é a seguinte, vamos criar dois graficos no nosso dashboard:

- **Gráficos de Barras:** Mostre as vendas globais dos jogos mais vendidos. Isso pode ajudar a visualizar quais jogos foram os mais bem-sucedidos em termos de vendas totais.

- **Gráfico de Linha do Tempo:** Exiba as vendas globais ao longo dos anos para ver tendências na popularidade de diferentes plataformas ou gêneros.



Ele adiciona um estilo consoante o tema, vamos mexer um pouco no estilo, podem colocar o estilo do gráfico conforme o que mais apetece vocês, o meu deixei assim:

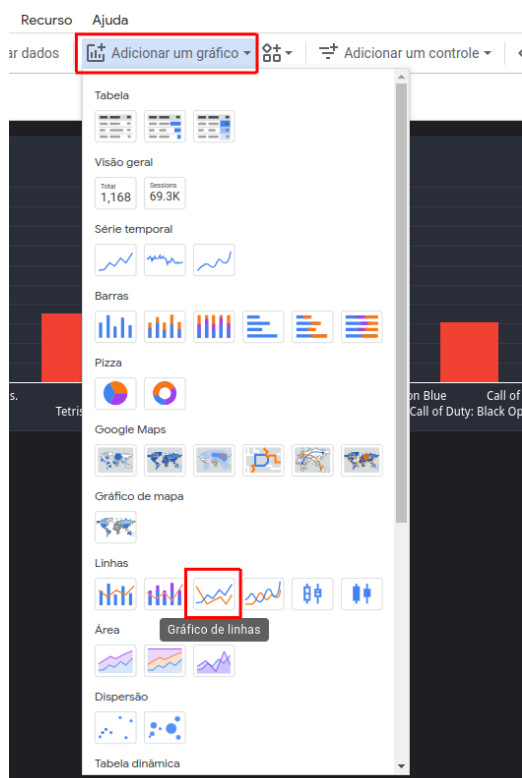


Agora vamos ajustar o gráfico, a dimensão dele, que no nosso caso representa o agrupamento, vamos deixar por nome mesmo, independente de plataforma e ano de lançamento. Porém, a métrica está conforme a quantidade de registros, queremos a quantidade de vendas globais:

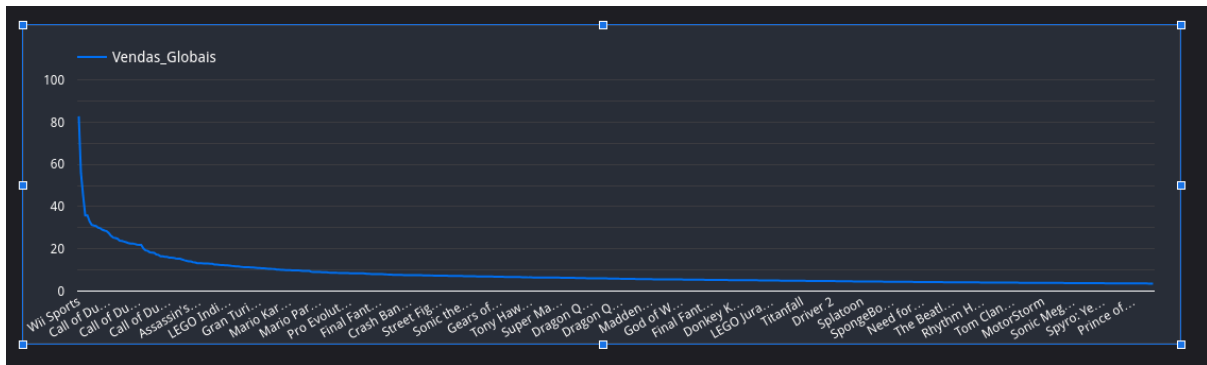


Para refletir: Agora vamos refletir sobre esses dados, você consegue ter uma noção do porquê o “Wii Sports” foi um dos mais vendidos no planeta inteiro? Eu lembro que, teve uma época que no Nintendo Wii foi um dos maiores fracassos da história dos videogames, porém olhem esse dado, conseguem me explicar? Teríamos que pesquisar, porém eu lembrei agora que muitos terapeutas ocupacionais estavam usando o Wii como ferramenta de trabalho, me lembro também de ir em alguns encontros onde a diversão era o Wii, jogar volei e boliche com os amigos, era divertido, não sei se vocês pegaram essa época.

Vamos então criar outro gráfico para finalizar essa primeira parte do nosso dashboard, vamos fazer um gráfico de linha do tempo para ver algumas tendências. Primeiro vamos adicionar esse gráfico logo abaixo do nosso gráfico de barras:



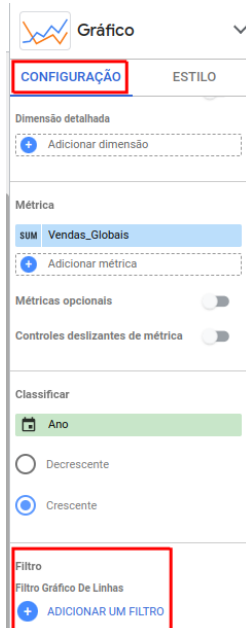
O gráfico inicial estará configurado automaticamente dessa forma:



Onde está configurada a dimensão nome e as métricas de vendas_globais, porém a gente quer ver quanto que foi vendido no decorrer do tempo, para isso, nossa dimensão primária (nesse caso o eixo x), será o Ano. Isso permitirá que a gente visualize as tendências ao longo do tempo. Já a métrica que nesse caso será nosso eixo y e vamos colocar as vendas globais, vamos alterar também o campo classificar, para ele classificar conforme o ano, nessas configurações nosso gráfico ficará da seguinte forma:



Vamos fazer um filtro direto nesse gráfico para retirar os valores N/A, o ideal é a gente tratar esses dados direto na base, porém essa base não é nossa, é só um exemplo, então vamos retirar os N/A apenas nesse gráfico mesmo criando um filtro, para isso vamos ir na aba “Configuração” que aparece quando clicamos no gráfico e ir na opção “Filtro”:



Vamos adicionar um Filtro com a seguintes condições:

Criar filtro

Nome
Excluir anos invalidos

↑ vgsales.csv

Excluir Ano Igual a (=) N/A

Clicar em Salvar. Vamos aproveitar já e criar um filtro de data, que daí podemos até remover esse filtro acima, vamos adicionar um filtro para pegarmos de 1985 até 2016 que é o que nos parece que tem as informações mais confiáveis (porém posso estar errado):

Criar filtro

Nome
Selecionar anos

↑ vgsales.csv

Incluir Ano Entre (>= && <=) 1985 2016

Então teremos um gráfico mais ou menos assim:



Para refletir: É estranho uma queda brusca das vendas em 2016, o que me leva a crer que talvez a nossa base não seja muito confiável para dados recentes. Porém, faz sentido o aumento de vendas, de 1985 até 2011 mais ou menos, porque ali pelos anos 2000 que começou a explosão de jogos de computadores etc. O que você acha sobre isso? Faz sentido? Ou falei abobrinhas?

Vamos melhorar só mais um pouco esse nosso gráfico para ver se conseguimos analisar algumas coisas melhores, vamos adicionar uma dimensão secundária (Faceta) para a gente comparar gêneros diferentes. Isso criará linhas separadas no gráfico para cada plataforma ou gênero.

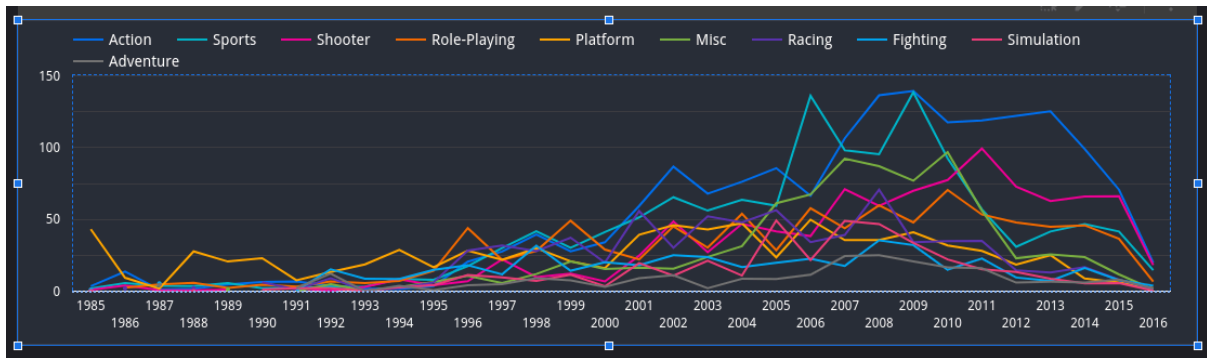
Dimensão

↑ Ano

Detalhar

Dimensão detalhada

ABC Genero



Podemos notar que os jogos de ação deram um salto a partir de 1996, 1997 e ficaram sempre no topo. Os de esporte tiveram um boom em 2006, o que vai bem ao encontro do lançamento do console wii e do wii sports que datam novembro de 2006.

Exploramos um pouco dos relatórios, gráficos, iremos aprofundar um pouco mais no decorrer das aulas.