

ANÁLISE DE DADOS DO GOOGLE ANALYTICS

• Desvendando o Comportamento da Audiência

Nicole Pessoa

[saiba mais](#)



Tópicos

01 Introdução

- Objetivo da análise

02 Metodologia

- Ferramentas utilizadas
- Fonte dos dados

03 Análise Exploratória

- Desafios com dados ausentes
- Tratamento de dados

04 Resultados

- Distribuição de visitas por canal de marketing
- Comportamento por dispositivo
- Tendências ao longo do tempo

05 Conclusões e Recomendações

- Insights principais
- Sugestões de ações

06 Próximos Passos

- Como essa análise pode evoluir

Nicole Pessoa

Introdução

Esta análise foi conduzida com o objetivo de explorar e entender o comportamento da audiência por meio de dados do Google Analytics. Buscamos desvendar padrões de tráfego, identificar tendências de engajamento e detectar oportunidades de melhoria para otimizar a experiência do usuário e o desempenho do site.

Nicole Pessoa

ALÉM DISSO, O PROJETO TEVE COMO FOCO:

Compreender a distribuição de visitas por canais de marketing:

como busca orgânica, redes sociais e tráfego direto

Analisar o comportamento da audiência por tipo de dispositivo:

desktop, mobile e tablet

Nicole Pessoa

ALÉM DISSO, O PROJETO TEVE COMO FOCO:

Identificar mudanças ao longo do tempo:

comparando os anos de 2016 e 2017

Propor ações estratégicas:

para aumentar o engajamento e melhorar a eficácia das campanhas de marketing.

Ao transformar dados brutos em insights acionáveis, esta análise visa apoiar decisões informadas e impulsionar resultados tangíveis para o negócio.

Metodologia

Python

```
/arquitetura_projeto/  
| — pandas/  
| — processamento.py # Manipulação e limpeza de dados  
| — analise.py # Análises estatísticas básicas  
|  
| | — numpy/  
| | — operacoes.py # Operações matemáticas e  
transformações de arrays  
| | — manipulacao.py # Modificação e ajustes de dados  
numéricos  
|  
| | — matplotlib/  
| | — graficos_basicos.py # Plotagens simples  
| | — customizacoes.py # Personalização de gráficos
```

Metodologia

Python

```
| | ——— seaborn/
| | ——— visualizacoes.py # Gráficos estatísticos avançados
| | ——— estilos.py # Configuração de estilos para visualização
|
| | ——— re/
| | ——— validacoes.py # Validações com expressões regulares
| | ——— limpeza.py # Remoção de caracteres especiais
|
| | ——— scikit-learn/
| | ——— imputer.py # Tratamento de valores ausentes
(SimpleImputer)
| | ——— regressao.py # Modelos de regressão (LinearRegression)
```


Metodologia



O Pandas é a principal ferramenta para manipulação de dados em tabelas. Com ele, é possível ler arquivos como CSV e Excel, organizar informações e realizar cálculos rapidamente. Se você precisa limpar dados, remover valores duplicados ou preencher informações ausentes, o Pandas facilita muito esse processo.



Essas bibliotecas são essenciais para visualizar dados. O Matplotlib permite criar gráficos altamente personalizáveis, enquanto o Seaborn, que é construído sobre ele, facilita a criação de gráficos estatísticos mais bonitos e informativos.

Metodologia



O NumPy é a base para cálculos matemáticos eficientes em Python. Ele permite trabalhar com grandes volumes de dados de forma otimizada, sendo muito utilizado em estatística e machine learning.

reg[ular]
expr[essio]n

Quando se trabalha com textos, às vezes é necessário encontrar padrões específicos, como e-mails ou números de telefone. As expressões regulares ajudam a localizar e padronizar essas informações de forma eficiente.

Nicole Pessoa

Metodologia



Essa biblioteca é uma das mais usadas para machine learning. Dois recursos úteis são o SimpleImputer, que preenche valores ausentes, e o LinearRegression, que ajuda a prever valores com base em tendências.

Metodologia

Fonte de Dados

[Google Analytics](#)

Faça o download para utilizá-lo

Nicole Pessoa

Metodologia

Pandas

- Utilizado para manipulação e limpeza dos dados, incluindo a leitura do dataset, tratamento de valores ausentes (NaNs) e transformação de colunas.
- **Exemplo:** Extração do ano da coluna date e cálculo de métricas como visitas, páginas visualizadas e tempo no site.

Matplotlib e Seaborn:

- Empregados para a criação de visualizações claras e informativas, como gráficos de distribuição de visitas por canal de marketing e comportamento por dispositivo.
- **Exemplo:** Gráficos de barras para comparar métricas entre 2016 e 2017.

Metodologia

Expressões Regulares (re)

- Utilizado para processar e limpar textos em colunas específicas, garantindo a consistência dos dados.
- **Exemplo:** Padronização de nomes de canais de marketing ou categorias de dispositivos.

Scikit-learn (SimpleImputer e LinearRegression):

- **SimpleImputer:** Aplicado para tratar valores ausentes (NaNs) em colunas numéricas, utilizando a média como estratégia de preenchimento.
- **LinearRegression:** Utilizado para explorar relações entre variáveis e prever tendências (se aplicável).
- **Exemplo:** Imputação de valores ausentes na coluna `total_time_on_site` e análise de correlação entre visitas e tempo no site.

Metodologia

NumPy

- Utilizado para operações matemáticas e manipulação de arrays, especialmente em cálculos estatísticos e transformações de dados.
- **Exemplo:** Cálculo de porcentagens e normalização de métricas.

Análise Exploratória

Renomeie as colunas para facilitar a compreensão e a manipulação dos dados. Para que pudessem entender claramente o que cada campo representava. As colunas foram renomeadas da seguinte forma:

Date → **Data**

Representa a data em que os dados foram coletados. Essa coluna é essencial para analisar tendências temporais, como sazonalidade ou crescimento ao longo do tempo.

channel_grouping → **canal_de_marketing**

Indica o canal de marketing que trouxe o visitante ao site (ex.: orgânico, direto, redes sociais, e-mail). Essa informação é vital para entender quais estratégias de marketing estão funcionando melhor.

source → **origem**

Especifica a fonte exata de tráfego (ex.: Google, Facebook, Bing). Complementa o canal de marketing, fornecendo detalhes mais granulares.

Nicole Pessoa

Análise Exploratória

device_category → **categoria_de_dispositivo**

Mostra o tipo de dispositivo usado pelo visitante (mobile, desktop, tablet). Isso ajuda a entender o comportamento do usuário em diferentes plataformas.

browser → **navegador**

Indica o navegador utilizado (ex.: Chrome, Safari, Firefox). Pode revelar insights sobre a compatibilidade do site com diferentes tecnologias.

total_visits → **total_visitas**

Contabiliza o número total de visitas ao site. É uma métrica fundamental para medir o engajamento.

Nicole Pessoa

Análise Exploratória

total_pageviews → **total_paginas_visualizadas**

Mostra quantas páginas foram visualizadas pelos visitantes. Essa métrica ajuda a entender a profundidade da interação dos usuários.

total_timeonsite → **tempo_total_no_site**

Mede o tempo gasto pelos usuários no site. Um tempo maior pode indicar um conteúdo mais envolvente.

Análise Exploratória

total_users



total_usuarios

Representa o número total de usuários únicos que visitaram o site. Essa métrica é importante para entender o alcance.

year



ano

Refere-se ao ano da coleta de dados. Útil para análises temporais de longo prazo.

Análise Exploratória

A análise exploratória é um processo essencial para compreender padrões ocultos nos dados. Neste estudo, mergulhei nos registros do Google Analytics para entender melhor o comportamento da audiência ao longo do tempo. Vamos explorar os desafios enfrentados, as descobertas feitas e os insights extraídos dessa investigação.

Total Páginas Visualizadas

1

Tempo Total no Site

12645

Esses eram únicos valores ausentes

O Desafio Inicial: Dados Incompletos

Ao iniciar a análise, encontramos um desafio clássico da ciência de dados: **valores ausentes (NaNs)**. Para lidar com esse problema, foi realizado um censo detalhado dos dados, mapeando a proporção de informações ausentes.

A solução?

Imputação de dados, substituindo os valores faltantes por **estatísticas robustas**, como a **média**.

Nicole Pessoa

Resultados: Desvendando o Comportamento da Audiência

Agora que os dados estão tratados e organizados, é hora de explorar os resultados. Vamos responder a algumas perguntas-chave que nos ajudarão a entender melhor a audiência e identificar oportunidades de melhoria.

Distribuição de visitas por canal de marketing

2016

PESQUISA ORGÂNICA

36,2%

Direct

25,2%

Referral

8,5%

Social

7,7%

2017

PESQUISA ORGÂNICA

38,8%

Direct

24,4%

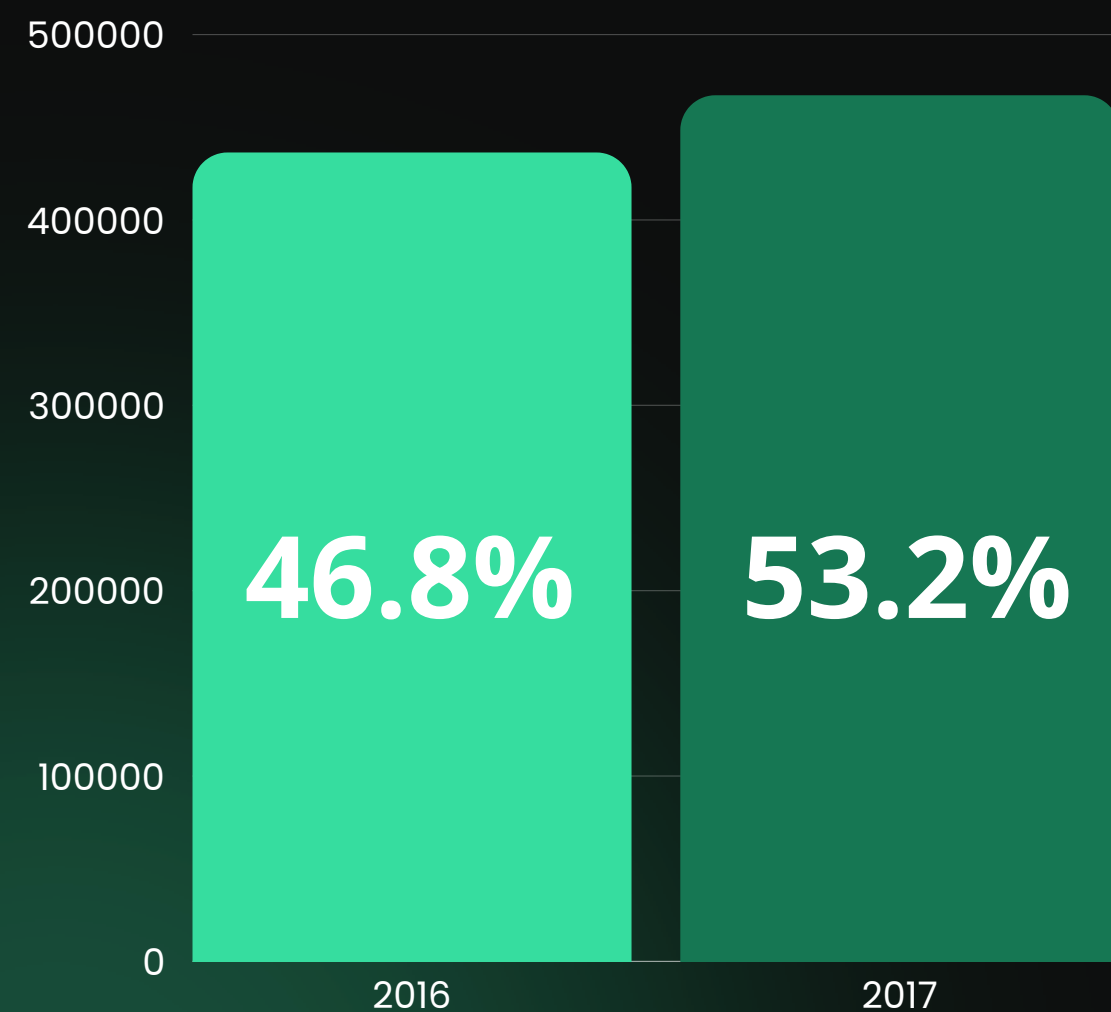
Referral

9,9%

Social

6,6%

Resultados



A busca orgânica dominou o tráfego, consolidando sua posição como principal fonte de audiência. Entretanto, a queda das redes sociais pode indicar uma necessidade de ajuste na estratégia de marketing.

O Caminho da Audiência: De Onde Vêm os Visitantes?

Um dos primeiros insights revelados foi a distribuição das visitas por canal de marketing. Os dados mostraram que a Pesquisa Orgânica sempre teve grande relevância, mas a importância das redes sociais diminuiu ao longo do tempo, dando espaço para um aumento no acesso direto ao site.

Nicole Pessoa

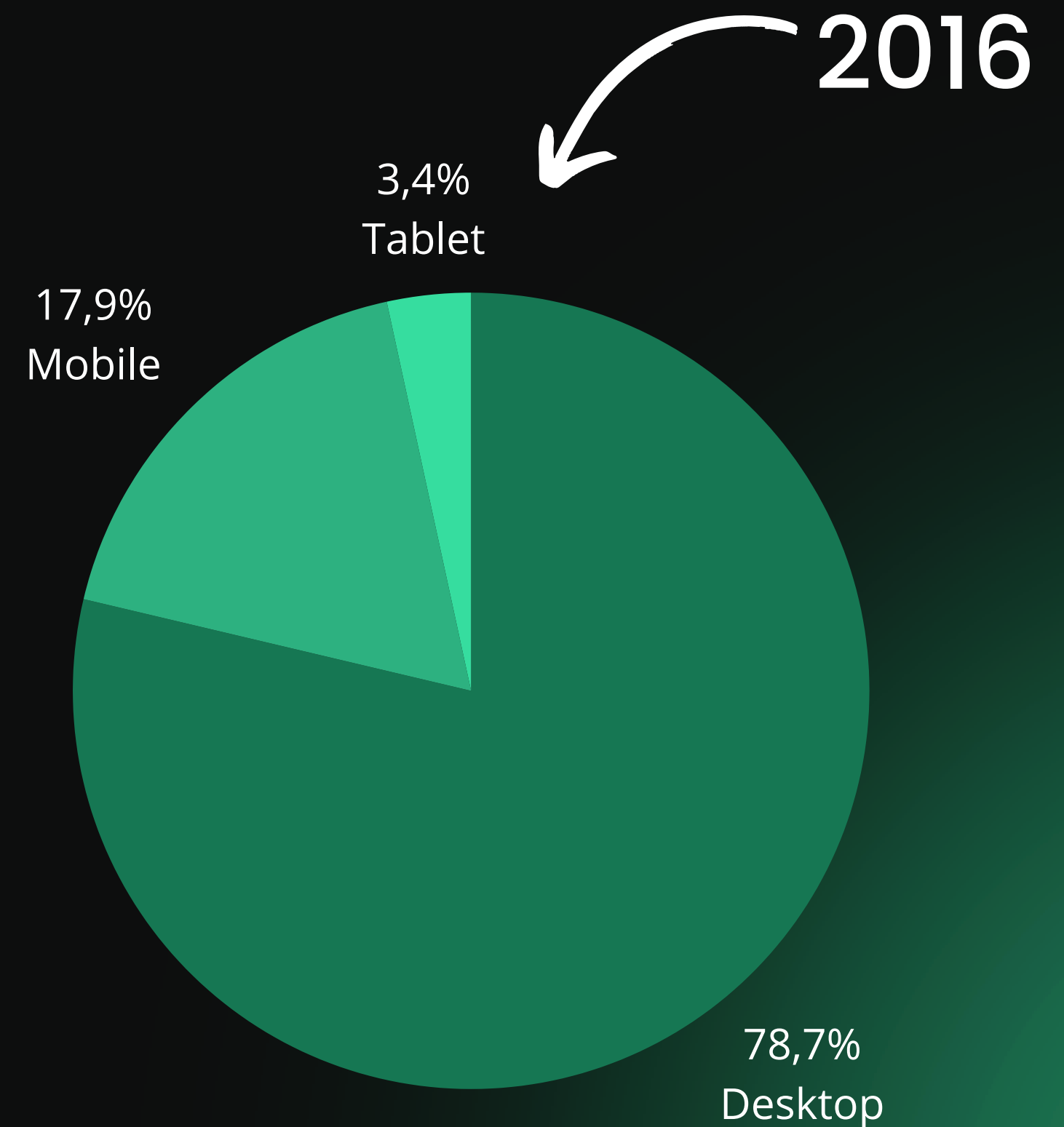
Concentração da Audiência

A **Pesquisa Orgânica** é o canal dominante, seguido pelo **Direct**. As **redes sociais**, embora importantes, têm uma participação menor e em declínio.

Audiência está concentrada em dois principais aspectos: canais de marketing e dispositivos

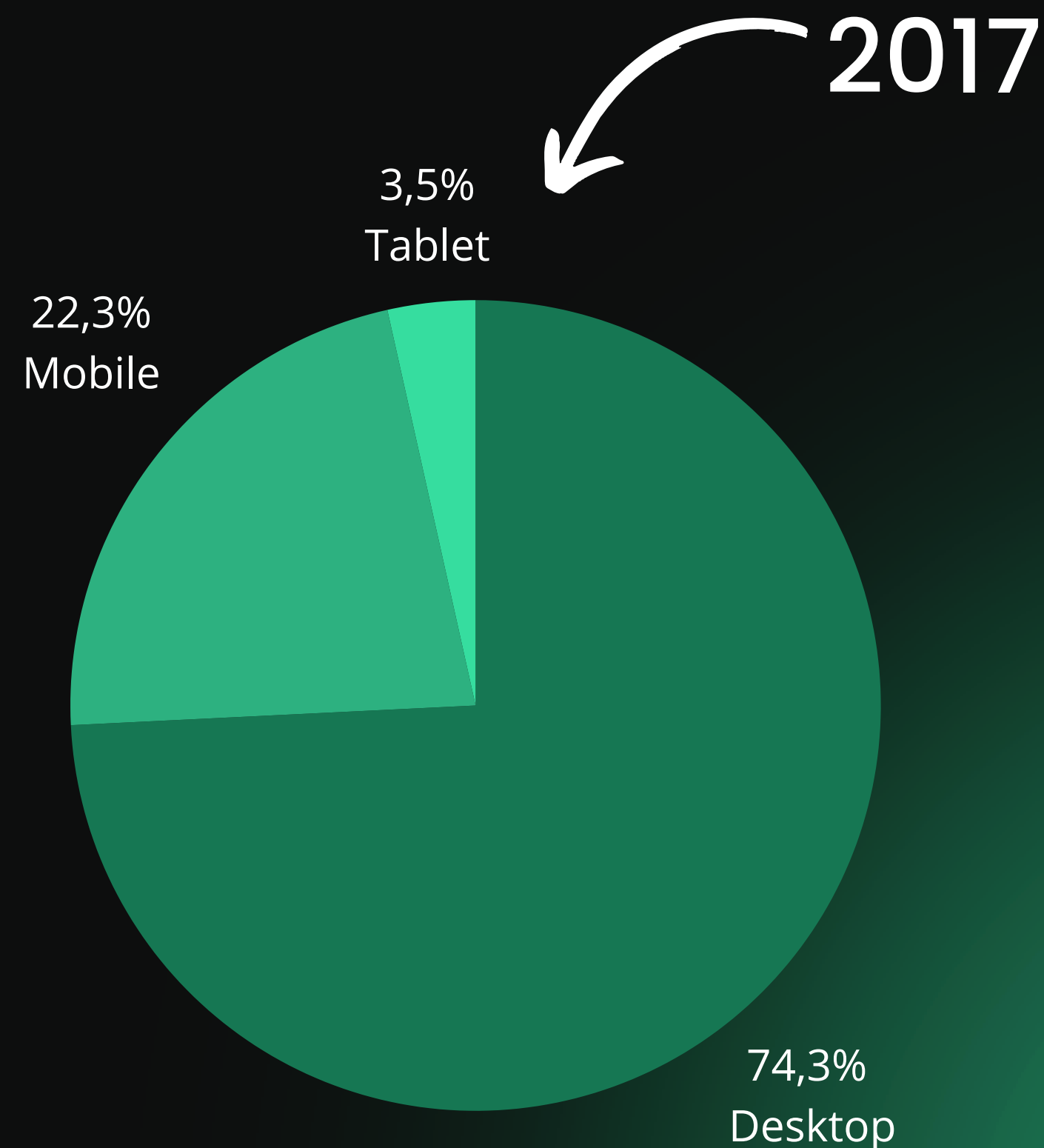
Distribuição de Visitas por Dispositivo

Agora, vamos ver como a audiência se comporta em relação aos dispositivos



Distribuição de Visitas por Dispositivo

Os **desktops** ainda dominam, mas os dispositivos **móveis** estão ganhando força, mostrando uma **mudança** no comportamento dos usuários.



Mudanças de Comportamento da Audiência

Mudanças nos Canais de Marketing

- **Redes Sociais:** A participação das redes sociais caiu de **10,3%** em **2016** para **5,7%** em **2017**. Isso pode indicar uma mudança na estratégia de marketing ou uma diminuição no alcance orgânico.
- **Pesquisa Orgânica:** Aumentou de **43,5%** para **47,4%**, consolidando sua posição como o principal canal de tráfego.

Mudanças nos Dispositivos

- **Mobile:** O uso de dispositivos móveis aumentou de **17,9%** para **22,3%**, enquanto os desktops diminuíram de **78,7%** para **74,3%**. Isso reflete a tendência global de migração para dispositivos móveis.

Nicole Pessoa

Conclusões e Recomendações

Insights Principais

1. **Concentração da Audiência:** A audiência está concentrada principalmente na Pesquisa Orgânica e no uso de desktops, mas os dispositivos móveis estão ganhando força.
2. **Mudanças de Comportamento:** Houve uma migração significativa para dispositivos móveis e uma queda no tráfego das redes sociais.
3. **Oportunidades de Engajamento:** Melhorar a experiência mobile e reavaliar a estratégia de marketing nas redes sociais são as principais oportunidades.
4. **Queda de Alcance:** A queda no alcance das redes sociais é compensada pelo aumento no tráfego mobile e na busca orgânica.

Conclusões e Recomendações

Sugestões de Ação

1. **Otimização Mobile:** Investir em uma experiência mobile mais robusta e responsiva.
2. **Estratégia de Redes Sociais:** Reavaliar e intensificar a presença nas redes sociais com campanhas direcionadas e conteúdo interativo.
3. **SEO:** Continuar investindo em SEO para manter e aumentar o tráfego orgânico.
4. **Fidelização de Usuários:** Criar estratégias para fidelizar usuários e aumentar as visitas diretas.

Próximos Passos: Como Essa Análise Pode Evoluir?

1. **Análise de Cohort:** Para entender a retenção de usuários ao longo do tempo.
2. **Segmentação de Audiência:** Dividir a audiência em grupos com base em comportamento, demografia, etc.
3. **Testes A/B:** Para testar diferentes estratégias de marketing e otimização de site.
4. **Análise de Conversão:** Entender quais canais e dispositivos geram mais conversões (vendas, inscrições, etc.).