A logo with pink flowers

Description automatically generated

Análise de dados para *Praceando*

Descobrindo, Conectando, Praceando – Blomera

# **Relatório do Bi da área web restrita**

Link para o BI <https://app.powerbi.com/links/IbnoyG38s0?ctid=b148f14c-2397-402c-ab6a-1b4711177ac0&pbi_source=linkShare>

## **1. Introdução**

### **1.1 Contexto e Objetivo do Projeto**

O *Praceando* é um aplicativo inovador desenvolvido para facilitar a organização, divulgação e descoberta de eventos em espaços públicos, como praças e parques. Seu principal objetivo é conectar pessoas interessadas em atividades ao ar livre, promovendo a interação social e o uso consciente dos espaços urbanos. O aplicativo permite que organizadores anunciem seus eventos de forma eficiente e que os usuários encontrem atividades de acordo com seus interesses e localização.

Este relatório apresenta uma análise detalhada dos dados coletados por meio de um formulário online, visando compreender profundamente o perfil dos usuários, suas preferências, comportamentos e necessidades. A análise dos dados é fundamental para aprimorar as funcionalidades do aplicativo, melhorar a experiência do usuário e direcionar estratégias de marketing mais eficazes.

### **1.2 Objetivo da Criação da Base de Dados**

A base de dados foi criada com o propósito de alimentar análises de Business Intelligence (BI) e servir como fundamento para o desenvolvimento futuro de uma Inteligência Artificial (IA). Essa IA terá a capacidade de predizer se o usuário seria um futuro anunciante

### **1.3 Metodologia de Coleta e Análise de Dados**

Os dados foram coletados através de um formulário estruturado, distribuído aos potenciais usuários do aplicativo. O formulário continha perguntas que abrangiam informações demográficas, preferências por tipos de eventos, comportamentos de organização e participação em eventos, além de dificuldades enfrentadas na divulgação e busca por eventos.

As respostas foram importadas para o Power BI, onde passaram por um processo de transformação e limpeza. A modelagem dos dados envolveu a normalização e a criação de tabelas fato e dimensão, seguindo as melhores práticas de Data Warehousing. As visualizações foram desenvolvidas utilizando gráficos interativos que facilitam a compreensão dos dados e a extração de insights relevantes.

## **2. Estrutura da Base de Dados**

### **2.1 Tabelas Fato e Dimensão**

A base de dados foi estruturada seguindo o modelo dimensional, que é amplamente utilizado em sistemas de BI por facilitar a consulta e a análise dos dados. Nesse modelo, os dados são organizados em tabelas fato e tabelas dimensão.

#### **Tabela Fato:**

* FPraceando: Esta tabela registra cada resposta individual do formulário. Cada linha representa uma submissão única e contém chavesestrangeiras que referenciam as tabelas dimensão. A tabela fato captura eventos ou transações, neste caso, as respostas dos usuários.

#### **Tabelas Dimensão:**

* **DFaixaEtaria**: Contém informações sobre as faixas etárias dos usuários. Permite analisar os dados segmentados por idade.
* **DCidade**: Armazena as cidades de residência dos usuários. Fundamental para análises geográficas.
* **DZonaSP:** Especifica a zona de residência dentro da cidade de São Paulo (Norte, Sul, Leste, Oeste, Centro). Detalha a distribuição geográfica intraurbana.
* **DTipoUsuario:** Classifica os usuários quanto ao seu envolvimento com eventos (organizador, participante, ambos, nenhum).
* **DDificuldades:** Registra as dificuldades que os usuários enfrentam na divulgação ou busca de eventos.
* **DTipoEvento:** Lista os tipos de eventos que os usuários organizam ou participariam (alimentação, música, lazer etc.).
* **DFatoresEscolhaApp**: Contém os fatores que influenciam a escolha do aplicativo pelos usuários (facilidade de uso, custo-benefício, alcance de público etc.).
* **DDisposicaoPagamento:** Informa a disposição dos usuários em pagar um valor simbólico para anunciar eventos.

#### **Detalhamento das Colunas Principais:**

* **ID\_Resposta:** Identificador único de cada resposta no formulário.
* **Data\_Hora\_Inicio e Data\_Hora\_Termino**: Marcam o início e o término do preenchimento do formulário, permitindo analisar o tempo médio gasto pelos usuários.
* **Email e Nome:** Dados de contato dos usuários, úteis para comunicação e estratégias de engajamento.
* **Chaves Estrangeiras:** Referências às tabelas dimensão, permitindo a ligação entre as informações descritivas e os eventos registrados na tabela fato.

### **2.2 Processo de Desnormalização**

A desnormalização foi aplicada para otimizar a performance das consultas e simplificar a estrutura dos dados para análises no Power BI. Isso envolveu a junção de tabelas relacionadas e a redução de complexidade nas relações, permitindo um acesso mais rápido aos dados relevantes durante a criação das visualizações.

## **3. Análise Detalhada dos Dados e Visualizações**

### **3.1 Perfil Demográfico dos Usuários**

**Visualização:** Gráfico de pizza de usuários por faixa etária

A diagram with numbers and a circle

Description automatically generated

**Descrição:**

* O gráfico apresenta as diferentes faixas etárias no e o número de usuários em cada faixa no eixo horizontal.
* **As faixas etárias estão categorizadas em:**
  + Menor que 18 anos
  + Entre 18 e 30 anos
  + Entre 31 e 45 anos
  + Entre 46 e 60 anos
  + Mais de 60 anos

**Insights:**

* Predominância de Jovens Adultos: A maior concentração de usuários está na faixa entre 18 e 30 anos, indicando que o público-alvo principal do aplicativo é composto por jovens adultos.
* Representatividade Significativa de Adultos: A faixa entre 31 e 45 anos também possui um número considerável de usuários.
* Baixa Participação de Menores de 18 anos e Maiores de 60 anos: Essas faixas etárias apresentam menor engajamento, o que pode ser devido a preferências diferentes ou menor familiaridade com tecnologias móveis.

### **3.2 Análise Geográfica**

**Visualização 1:** Gráfico de colunas destacando a distribuição dos tipos usuários por cidade.

A graph of different colored squares

Description automatically generated

**Descrição:**

* O mapa mostra as cidades onde os usuários residem, com colorações variando de acordo com a densidade de usuários em cada local.
* As principais cidades identificadas são São Paulo, Osasco e Guarulhos.

**Insights:**

* Concentração na Cidade de São Paulo: A maioria dos usuários reside na capital, sugerindo que as funcionalidades do aplicativo devem atender às necessidades desse público urbano.
* Potencial de Expansão Regional: A presença de usuários em cidades próximas indica oportunidades para expansão dos serviços do aplicativo para essas áreas metropolitanas.

**Visualização 2:** Gráfico de setores mostrando a distribuição dos usuários pelas zonas da cidade de São Paulo.

**Descrição:**

* O gráfico divide os usuários de São Paulo em Zona Norte, Zona Sul, Zona Leste, Zona Oeste e Centro.
* Cada setor representa a porcentagem de usuários em cada zona.

**Insights:**

* **Predominância na Zona Oeste:** A Zona Oeste concentra a maior parte dos usuários, o que pode estar relacionado a fatores socioeconômicos ou disponibilidade de espaços públicos atrativos.
* **Distribuição Equilibrada nas Outras Zonas:** As demais zonas apresentam uma distribuição mais uniforme de usuários.
* **Estratégias Localizadas:** Conhecer a distribuição geográfica permite o planejamento de eventos e ações promocionais direcionadas.

### **3.3 Tipos de Usuários e Participação em Eventos**

**Visualização:** Gráfico de barras empilhadas mostrando o envolvimento dos usuários com eventos (organizam, participam, ambos ou nenhum).

A green and white graph

Description automatically generated

**Descrição:**

* As barras representam o total de usuários, segmentadas pelas categorias de envolvimento.
* **As categorias incluem:**
  + Apenas participam de eventos
  + Apenas organizam eventos
  + Tanto organizam quanto participam
  + Não organizam nem participam

**Insights:**

* **Maioria dos Usuários Participa de Eventos:** A grande maioria dos usuários declara que costuma participar de eventos em praças e parques.
* **Menor Proporção de Organizadores:** Há um número menor de usuários que organizam eventos, indicando um potencial para incentivar a organização de eventos pela comunidade.
* **Usuários Ativos em Duas Frentes:** Uma parcela dos usuários tanto organiza quanto participa, demonstrando alto nível de engajamento com atividades em espaços públicos.

### **3.4 Preferências por Tipos de Eventos**

**Visualização**: Nuvem de palavras mostrando os tipos de eventos que os usuários organizam ou participariam.

A close up of words

Description automatically generated

**Descrição:**

* **A nuvem representa cada tipo de evento, como:**
  + Alimentação
  + Música
  + Lazer
  + Esportes
  + Artesanato
  + Outros
* O tamanho de cada palavra indica o número de usuários interessados em cada categoria.

**Insights:**

* **Alimentação como Principal Interesse:** Eventos relacionados à alimentação são os mais populares entre os usuários.
* **Música e Lazer também em Destaque**: Essas categorias também apresentam alto interesse, sugerindo que atividades culturais e de entretenimento são atrativas para o público.
* **Diversidade de Interesses:** Embora alguns tipos de eventos sejam mais populares, há interesse em uma variedade de atividades, indicando a importância de oferecer conteúdo diversificado no aplicativo.

### **3.5 Fatores de Decisão na Escolha do Aplicativo**

**Visualização:** Gráfico de linhas destacando os fatores mais importantes para os usuários na escolha de um aplicativo de divulgação de eventos.

A graph with a line

Description automatically generated

**Descrição:**

* **A linha representa os fatores listados, como:**
  + Facilidade de uso
  + Custo-benefício
  + Alcance de público
  + Funcionalidades avançadas
  + Suporte e atendimento
* A altura da linha indica a importância atribuída pelos usuários a cada fator.

**Insights:**

* **Facilidade de Uso é Primordial:** A maioria dos usuários considera a facilidade de uso como o fator mais importante.
* **Custo-Benefício em Segundo Plano:** O custo associado ao uso do aplicativo também é um fator significativo na decisão dos usuários.
* **Alcance de Público Relevante:** Usuários valorizam aplicativos que permitam alcançar um público amplo, aumentando a eficácia na divulgação de eventos.

### **3.6 Distribuição dos Usuários por Zonas de São Paulo**

A screenshot of a graph

Description automatically generated

**Descrição do Gráfico:**

* **Campos Categóricos:** Zonas da cidade de São Paulo (Oeste, Norte, Leste, Sul, Centro, e "Não mora em São Paulo").
* **Visual:** Barras verticais representando a contagem de usuários em cada zona.
* **Insights Potenciais:**
  + **Predominância na Zona Oeste:** A barra mais alta no gráfico indica que a Zona Oeste possui a maior concentração de usuários do aplicativo. Isso pode sugerir que esta zona tem uma alta atividade em eventos ou uma população jovem e ativa que utiliza o aplicativo com mais frequência.
  + **Menor Engajamento nas Outras Zonas:** As zonas Norte, Leste e Sul têm representações menores, indicando menor penetração do aplicativo nessas áreas. Isso pode ser devido à falta de promoções direcionadas, menor número de eventos listados nesses locais, ou características demográficas distintas que influenciam o uso do aplicativo.
  + **Oportunidades para Expansão e Marketing:** O conhecimento sobre a distribuição geográfica dos usuários permite estratégias direcionadas de marketing e promoções locais para aumentar o engajamento nessas áreas menos ativas. Além disso, pode-se investigar as razões por trás da baixa adesão nas zonas Norte, Leste e Sul para adaptar o aplicativo ou as atividades promocionais a essas especificidades locais.
  + **Engajamento fora de São Paulo:** A presença de usuários que não moram em São Paulo sugere um interesse potencial pelo aplicativo fora da região metropolitana, indicando uma oportunidade de expansão para outras cidades ou regiões.

### **3.7 Filtro por Tipo de Usuário**

**Visualização:** Filtro interativo que permite segmentar os usuários de acordo com seu nível de engajamento com eventos.

A screenshot of a computer

Description automatically generated

**Descrição:**

**Campos Categóricos:**

**Ambos:** Usuários que tanto organizam quanto participam de eventos.

**Nenhum:** Usuários que não organizam nem participam de eventos.

**Organizo eventos:** Usuários que organizam eventos.

**Participo de eventos:** Usuários que apenas participam de eventos.

**Insights:**

Análise Segmentada do Comportamento: Esse filtro possibilita uma análise mais granular sobre o comportamento dos usuários, diferenciando entre os que apenas participam de eventos e os que também os organizam. Isso pode revelar padrões distintos de uso do aplicativo e ajudar a personalizar campanhas para incentivar mais usuários a se tornarem organizadores.

Aprimoramento das Funcionalidades: Ao identificar os grupos que mais organizam ou participam de eventos, o app pode ajustar suas funcionalidades para melhorar a experiência desses segmentos, potencialmente aumentando o engajamento e retenção.

### **3.8 Meta de Respostas Recebidas**

**Visualização:** KPI que mostra a meta de 200 respostas e o progresso atual de 87 respostas recebidas.

A white and yellow circle with black numbers

Description automatically generated

**Descrição:**

A visualização destaca o número de respostas recebidas até o momento e a meta total, indicando o percentual já alcançado em relação ao objetivo.

**Insights:**

Monitoramento de Progresso: O gráfico ajuda a acompanhar o desempenho da coleta de respostas. Com aproximadamente 43,5% da meta atingida, é possível inferir que ações adicionais podem ser necessárias para atingir a meta dentro do prazo, como aumentar a divulgação ou oferecer incentivos para mais participantes.

### **3.9 Usuários Interessados em ser notificado**

A number on a white background

Description automatically generated

**Visualização:** Big number mostrando o número de usuários que optaram por receber notificações sobre o aplicativo (45).

**Descrição:**

Exibe o total de usuários que demonstraram interesse em receber notificações, oferecendo uma visão clara sobre o segmento de usuários mais engajados.

**Insights:**

Engajamento Direto: Este número representa um grupo de usuários engajados que pode ser alvo de campanhas específicas, como feedback de novas funcionalidades ou atualizações. Eles estão mais propensos a interagir com o aplicativo, tornando-se valiosos em termos de retenção e crescimento orgânico.

Estratégia de Comunicação: A análise desse número permite à equipe de marketing planejar campanhas segmentadas, utilizando as notificações para manter esses usuários informados sobre eventos ou funcionalidades futuras.

### **3.10 Total de Usuários que Moram em São Paulo**

**Visualização:** Big number que exibe o total de 74 usuários residentes na cidade de São Paulo.

A number on a white background

Description automatically generated

**Descrição:**

Exibe o número de usuários que moram na capital paulista, indicando a base local de usuários que o aplicativo já conseguiu atingir.

**Insights:**

Foco Regional: Esse dado sugere que o aplicativo já tem uma base significativa de usuários em São Paulo, o que indica um foco inicial de marketing e atividades locais. Estratégias de expansão podem ser planejadas para aumentar a base de usuários em outras regiões metropolitanas ou no interior.

Oportunidades Locais: Este número também pode ser usado para planejar eventos específicos em São Paulo, aproveitando a alta concentração de usuários na cidade, garantindo uma maior participação e sucesso nos eventos.

## **4.Tratamento da Base**

A base de dados coletada pelo formulário foi submetida a um processo rigoroso de tratamento para garantir sua qualidade e integridade antes de sua utilização no BI. Esse tratamento resultou em duas versões da base: uma versão "suja", que reflete os dados originais, e uma versão "limpa", que passou por um processo de padronização e limpeza. A seguir, detalhamos as etapas de tratamento realizadas e suas justificativas.

### **4.1. Padronização de Colunas**

Todas as colunas foram padronizadas para manter uma nomenclatura uniforme e facilitar o manuseio dos dados.

* Os nomes das colunas foram convertidos para letras minúsculas e caracteres especiais foram substituídos por sublinhados ("\_").
* Exemplo: A coluna "Idade" foi renomeada para "idade" e "Nome Completo" para "nome\_completo".

*Justificativa:* A padronização de colunas evita erros ao realizar consultas e integrações, especialmente em ambientes de BI que requerem consistência nos nomes dos campos.

### **4.2. Tratamento de Valores Nulos**

Foram identificados valores nulos nas colunas numéricas e de texto. Para cada tipo de dado, aplicamos uma estratégia adequada:

* **Colunas Numéricas:** Preenchimento com a mediana dos valores existentes, evitando que outliers influenciem de forma inadequada a análise.
* **Colunas de Texto:** Preenchimento com o valor "Não informado" para indicar a ausência de resposta, garantindo que todas as linhas contenham informações completas.

*Justificativa:* A substituição de valores nulos por valores representativos (mediana ou "Não informado") permite análises sem distorções ou lacunas, fundamentais para visualizações corretas no Power BI.

### **4.3. Tratamento de Outliers**

Foi aplicada uma técnica de Interquartile Range (IQR) nas colunas numéricas para identificar e corrigir outliers:

* Valores abaixo do limite inferior ou acima do limite superior foram ajustados para esses limites.

*Justificativa:* A redução de outliers evita distorções em análises de média e mediana, garantindo que os valores sejam representativos do perfil real dos usuários.

### **4.4. Padronização dos Valores de Texto**

Para garantir consistência nos valores textuais, todas as entradas de texto foram padronizadas:

* Valores foram convertidos para letras minúsculas e espaços adicionais foram removidos.
* Exemplo: Gênero "Masculino" e "masculino" foram unificados como "m".

*Justificativa:* A padronização de valores de texto garante uma análise precisa, evitando a duplicidade de categorias que podem prejudicar a segmentação e visualizações no BI.

### **4.5. Normalização das Colunas Numéricas**

Aplicamos a técnica de normalização nas colunas numéricas, transformando seus valores para uma escala entre 0 e 1.

*Justificativa:* A normalização facilita a comparação de diferentes métricas no BI e evita que diferenças de escala influenciem visualizações ou análises comparativas.

## **5. Conclusão**

A análise detalhada dos dados coletados fornece uma compreensão profunda do público-alvo do *Praceando* e destaca diversas oportunidades para aprimorar o aplicativo e atender melhor às necessidades dos usuários.

**Perfil do Usuário:**

Os usuários do *Praceando* são majoritariamente jovens adultos, entre 18 e 30 anos, residentes na cidade de São Paulo, com concentração na Zona Oeste. Eles estão interessados principalmente em eventos relacionados à alimentação, música e lazer, indicando um público dinâmico e socialmente ativo.

**Preferências e Comportamentos:**

A facilidade de uso é o fator mais importante na escolha de um aplicativo, seguido pelo custo-benefício e alcance de público. Isso reforça a necessidade de um aplicativo intuitivo, com uma interface amigável e funcionalidades eficientes que permitam tanto a divulgação quanto a busca de eventos de forma simples.

**Desafios e Necessidades:**

As dificuldades encontradas pelos usuários na divulgação e na busca de eventos apontam para áreas críticas que necessitam de melhorias no aplicativo. Soluções que facilitem a promoção de eventos e aprimorem os mecanismos de busca podem aumentar significativamente a satisfação e o engajamento dos usuários.

**Potencial de Monetização:**

A disposição de uma parcela dos usuários em pagar por anúncios sugere que há potencial para monetização. No entanto, é importante considerar a sensibilidade ao preço e garantir que o valor cobrado seja percebido como justo e que agregue benefícios reais aos usuários.

**Estruturação dos Dados:**

A modelagem dos dados em tabelas fato e dimensão foi fundamental para realizar análises detalhadas e segmentadas. As tabelas dimensão permitiram categorizar e contextualizar os dados, enquanto a tabela fato registrou os eventos essenciais para a análise. Essa estrutura facilitou a extração de insights valiosos que podem orientar o desenvolvimento do aplicativo.

**Considerações Finais:**

O *Praceando* possui um potencial significativo para se tornar uma ferramenta indispensável para pessoas interessadas em eventos em espaços públicos. Com base nos insights obtidos, é possível direcionar esforços para aprimorar as funcionalidades que os usuários mais valorizam, resolver as dificuldades enfrentadas e explorar oportunidades de crescimento e monetização.

A compreensão aprofundada do público-alvo, obtida através desta análise, é um passo crucial para o sucesso contínuo do aplicativo. Ao alinhar as estratégias de desenvolvimento e marketing com as necessidades e preferências dos usuários, o *Praceando* poderá aumentar seu alcance, engajamento e satisfação do usuário, consolidando-se como uma referência no segmento de eventos em espaços públicos.