

Estudo de caso: Como um compartilhamento de bicicletas possibilita o sucesso rápido?

Visão geral do projeto

Este projeto consiste em um estudo de caso de análise de dados baseado no programa de compartilhamento de bicicletas da Cyclistic, uma empresa fictícia localizada em Chicago. O estudo foi desenvolvido como parte da formação Google Data Analytics, com o objetivo de aplicar, na prática, todas as etapas do processo de análise de dados: Perguntar, Preparar, Processar, Analisar, Compartilhar e Agir.

O contexto do projeto está relacionado a um desafio estratégico da área de marketing da empresa: aumentar o número de assinaturas anuais. Para isso, foi realizada uma análise comparativa entre o comportamento de usuários casuais e membros anuais, utilizando dados históricos de trajetos de bicicleta.

Ao longo do projeto, foram aplicadas técnicas de organização, limpeza, transformação e análise de dados em diferentes ferramentas, além da criação de visualizações interativas para apoiar a tomada de decisão. O foco principal foi gerar insights açãoáveis, capazes de auxiliar a equipe de marketing na definição de estratégias voltadas à conversão de usuários casuais em membros anuais.

Problema de negócio

O principal desafio da empresa é identificar formas de aumentar o número de assinaturas anuais, utilizando de maneira mais eficiente os dados já disponíveis. Para isso, é necessário compreender como os diferentes tipos de usuários utilizam o serviço, permitindo que a equipe de marketing desenvolva estratégias mais direcionadas e eficazes.

Perguntas de análise

- Como os membros anuais e os usuários casuais utilizam as bicicletas da Cyclistic de forma diferente?
- Quais padrões de uso podem indicar oportunidades de conversão para planos anuais?
- Como os dados podem apoiar estratégias de marketing mais direcionadas?

Stakeholders

- Diretoria de Marketing da Cyclistic.
- Equipe de Análise de Marketing.
- Equipe Executiva da Cyclistic.

Tarefa de negócio

Analisar os dados históricos de trajetos da Cyclistic para identificar as diferenças de comportamento entre membros anuais e usuários casuais, com o objetivo de gerar insights que apoiem decisões estratégicas de marketing voltadas à conversão de usuários casuais em membros anuais.

Preparação dos dados

Fonte dos dados

Os dados utilizados neste estudo de caso são dados públicos de trajetos de bicicletas da Cyclistic, disponibilizados pela Motivate International Inc. Os conjuntos de dados representam os registros de uso do serviço de compartilhamento de bicicletas ao longo de um período de 12 meses.

Os dados foram fornecidos exclusivamente para fins educacionais e não contêm informações pessoais identificáveis, estando em conformidade com as diretrizes de privacidade e segurança de dados.

Estrutura e organização dos dados

Os dados estão organizados em arquivos mensais no formato CSV, em que cada arquivo representa um mês de utilização do serviço. Todos os arquivos possuem a mesma estrutura de colunas, incluindo informações sobre:

- início e término das viagens;
- tipo de usuário (membro ou casual);
- estações de retirada e devolução;
- coordenadas geográficas.

Inicialmente, os arquivos foram explorados no Excel para compreensão da estrutura, validação das colunas e entendimento do conteúdo. No entanto, devido ao alto volume de registros (mais de 5 milhões de linhas no total) e às limitações do Excel, optou-se por realizar a consolidação e o processamento final dos dados em ambiente mais adequado.

Confiabilidade e credibilidade dos dados

Os dados utilizados são considerados confiáveis por apresentarem as seguintes características:

- Originais: provenientes diretamente da fonte responsável pelo serviço;
- Abrangentes: incluem registros completos de uso ao longo de um ano;
- Atuais: refletem um período recente de operação;
- Confiáveis: amplamente utilizados em estudos e análises educacionais.

Limitações dos dados

Algumas limitações foram identificadas:

- Ausência de informações pessoais dos usuários, devido a restrições de privacidade;
- Impossibilidade de acompanhar o comportamento individual de um mesmo usuário ao longo do tempo;
- Limitações técnicas do Excel para consolidação de grandes volumes de dados.

Essas limitações foram consideradas durante a análise e não comprometem os objetivos do estudo.

Justificativa de adequação ao problema de negócio

Mesmo com as limitações identificadas, os dados são suficientes para analisar padrões de uso, diferenças comportamentais e tendências temporais entre membros anuais e usuários casuais, atendendo ao objetivo do estudo de caso.

Processamento e limpeza

Após a etapa de preparação, os dados passaram por um processo de estruturação, limpeza e transformação, com o objetivo de garantir consistência, qualidade e confiabilidade antes da fase de análise.

Inicialmente, os arquivos mensais foram explorados no Excel para validação da estrutura, identificação de campos relevantes e entendimento das variáveis disponíveis. Contudo, devido ao elevado volume de dados, o processamento principal foi realizado no RStudio, ferramenta mais adequada para lidar com grandes conjuntos de dados.

No RStudio, todos os arquivos CSV mensais foram importados e combinados em um único conjunto de dados. Durante essa etapa, foi identificada uma inconsistência no formato de leitura de um dos arquivos, que resultou na criação indevida de uma coluna adicional. Esse problema foi corrigido ao forçar explicitamente o delimitador correto durante a importação, garantindo a padronização das colunas.

As colunas de data e hora de início e término das viagens foram convertidas do formato texto para o formato apropriado de data e hora. A partir dessas conversões, foram criadas variáveis derivadas para apoiar a análise, incluindo:

- duração dos passeios, calculada em minutos e horas;
- dia da semana de início da viagem;
- mês de referência, para análises temporais.

Além disso, registros com duração inválida (igual ou inferior a zero) foram removidos, assegurando maior confiabilidade aos dados analisados. Todo o processo de limpeza e transformação foi aplicado de forma padronizada, resultando em uma base de dados consistente e pronta para a etapa de análise.

Análise

A análise foi conduzida com foco em identificar diferenças de comportamento entre usuários casuais e membros anuais, considerando frequência de uso, duração das corridas e variações temporais.

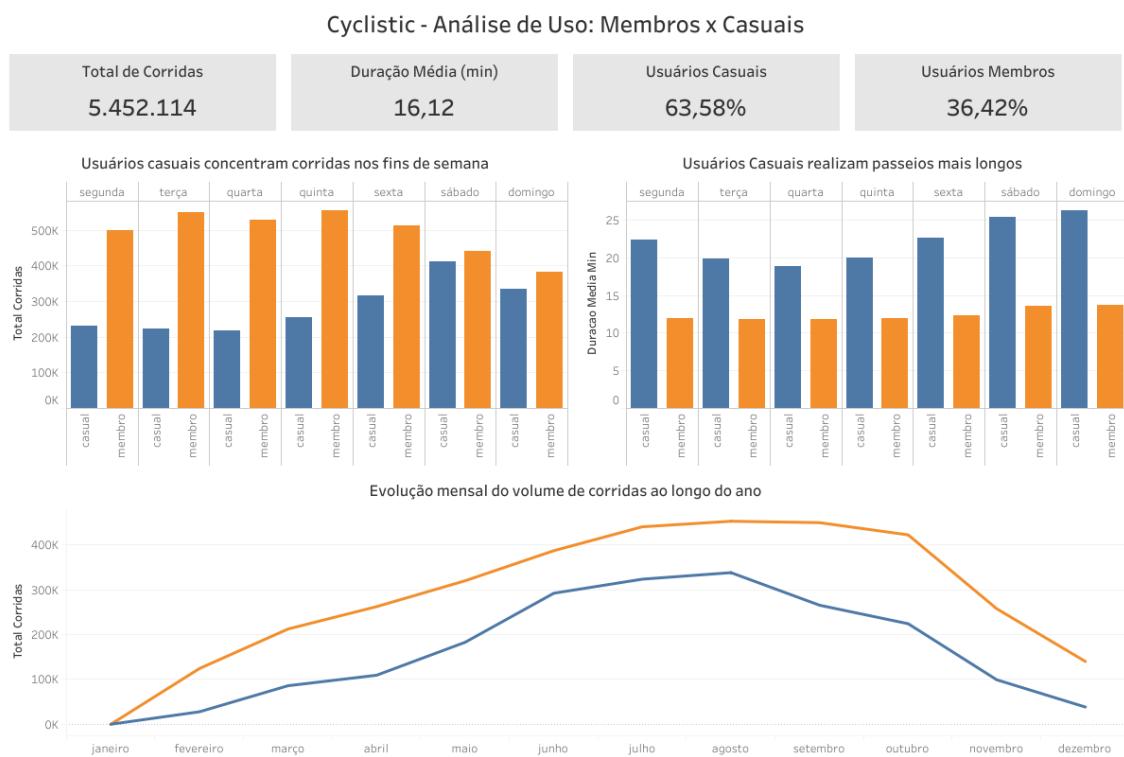
Foram realizados cálculos descritivos e agregações para:

- total de corridas;
- duração média dos passeios;
- distribuição de corridas por dia da semana;
- evolução do volume de corridas ao longo dos meses.

Os dados analisados indicaram padrões distintos entre os dois tipos de usuários, permitindo a identificação de oportunidades estratégicas para ações de marketing direcionadas.

Visualizações

As visualizações foram desenvolvidas no Tableau, com foco em clareza, comparabilidade e suporte à tomada de decisão.



Recomendações

Com base nos padrões identificados na análise dos dados, são propostas as seguintes ações estratégicas para apoiar a conversão de usuários casuais em membros anuais:

Campanhas de conversão focadas em fins de semana

Os dados indicam que usuários casuais concentram maior volume e maior duração de uso aos sábados e domingos, sugerindo um perfil de uso recreativo.

Recomendação: implementar campanhas promocionais específicas para fins de semana, oferecendo benefícios temporários, como descontos na assinatura anual ou períodos de teste gratuito, direcionados a usuários casuais ativos nesses dias.

Planos de associação flexíveis para usuários recorrentes

Usuários casuais apresentam recorrência de uso ao longo do tempo, ainda que não utilizem o serviço diariamente.

Recomendação: criar planos de associação intermediários ou flexíveis, com menor custo inicial e benefícios progressivos, visando reduzir a barreira de entrada para a assinatura anual.

Ações sazonais nos períodos de maior volume

A análise mensal evidencia crescimento significativo no volume de corridas durante os meses mais quentes do ano.

Recomendação: intensificar ações de marketing e ofertas promocionais no início desses períodos sazonais, aproveitando o aumento natural da demanda para incentivar a conversão de usuários casuais.

Conclusão e aprendizados

Este estudo de caso demonstrou como a análise de dados pode apoiar decisões estratégicas de negócio. A partir da comparação entre membros anuais e usuários casuais, foi possível identificar padrões claros de uso, evidenciando diferentes perfis de comportamento.

O projeto reforçou a importância de:

- escolher ferramentas adequadas para grandes volumes de dados;
- documentar corretamente as etapas de preparação, processamento e limpeza;
- transformar dados em insights acionáveis por meio de visualizações eficazes.

Como aprendizado principal, destaca-se que dados bem tratados e analisados permitem transformar informações operacionais em estratégias concretas de negócio, apoiando o crescimento sustentável da empresa.

Ferramentas utilizadas

- Excel – exploração inicial dos dados

- RStudio – processamento, limpeza e análise de grandes volumes de dados
- Tableau – visualização e comunicação dos insights

Materiais

<https://github.com/JaderJFranca/cyclistic-case-study>