

Fabio Dias Rezende Carvalho

email:fabior.carvalho@hotmail.com

Github: <https://github.com/FabioDiasRC>

CYCLISTIC - Análise dos dados de uma empresa de bicicletas

A idéia é simular o trabalho de uma analista de dados júnior que atua na equipe de marketing da Cyclistic, uma empresa de compartilhamento de bicicletas de Chicago. O foco aqui é entender como os usuários anuais divergem dos usuários casuais e tentar otimizar o número de clientes anuais, uma vez que a empresa afirma que a maior parte de seu rendimento vem de usuários anuais, chamados de membro Cyclistic. Para isso eu devo fazer análises e ter insights convincentes com boas visualizações de dados.

A empresa possui 3 tipos de bicicletas (bicicletas reclináveis, triciclos manuais e bicicletas de carga), as bicicletas assistivas, que atendem pessoas com deficiência atende 8% do total. A empresa afirma que na média 30% dos usuários anualmente utilizam para trabalhar diariamente, mas no geral os usuários utilizam o serviço por lazer (farei uma análise específica para verificar essa informação da empresa).

A equipe de marketing pede a respostas de três perguntas específicas:

- Como os membros anuais e os ciclistas casuais usam as bicicletas da Cyclistic de forma diferente?
- Por que os passageiros casuais iriam querer adquirir planos anuais da Cyclistic?
- Como a Cyclistic pode usar a mídia digital para influenciar os passageiros casuais a se tornarem membros?

Análise exploratória dos dados:

Aqui podemos observar alguns pontos relevantes dentro do nosso conjunto de dados:

- Alguns dados da estação de entrega, nas colunas de nomes de estação e id de estação (`end_station_name`, `end_station_id`, `start_station_name`, `start_station_id`), não possuem valores definidos, como isso ocorre em muitas linhas dessas colunas eu decidi por remover essas colunas completamente, tendo em vista que os valores de longitude e latitude estão preenchidos, acredito que elas possam ser trabalhadas para calcular a distância percorrida pelo usuário caso necessário e tomarei essa decisão mais a frente;
- Os dados de retirada e entrega estão como 'object', eles serão passados para datas de forma correta e ajudarão na análise de comportamento de consumo futuramente;
- Também serão criadas duas colunas, uma sendo o tempo de duração das viagens (momento da entrega da bicicleta menos o momento da retirada da bicicleta) e outra sendo definido o dia da semana que essa viagem ocorreu.

As perguntas que norteiam o projeto de análise

- **Como os membros anuais e os ciclistas casuais usam as bicicletas da Cyclistic de forma diferente?**

Com nossa análise podemos notar que a média de tempo de utilização dos usuários casuais são 28 minutos, contra a média de 12 minutos dos usuários membros da empresa. Também temos uma mudança nítida pela utilização durante a semana, uma vez que a maior parcela de usuários casuais se encontra nos dias do fim de semana, sexta, sábado e domingo, logo, podemos definir que a utilização dos usuários casuais se relacionam muito mais com passeios e lazer do que os usuários membros, que tendem a utilizar as bicicletas mais durante os dias da semana. Além disso, temos esses valores muito diferentes durante os meses do ano, uma vez que os meses de verão, e seus dois meses adjacentes, encontram os maiores valores de usuários casuais durante o ano.

| | |
|--|----------------|
| | Dias da semana |
| | Meses |

- **Por que os passageiros casuais iriam querer adquirir planos anuais da Cyclistic?**

Os usuários casuais poderiam utilizar os planos anuais se forem influenciados a começarem a ir trabalhar de bicicleta, mas é algo mais difícil de se influenciar com propagandas diretas da empresa, pois é um hábito mais complexo de ser modificado. A utilização desses usuários ocorre em maior peso durante as estações do ano que possuem melhor temperatura para utilizar a bicicleta como um lazer, onde a motivação para esses usuários se tornarem membros pode ocorrer pelo próprio incentivo da empresa, com o impulsionamento de marketing digital correto nas mídias digitais, as datas indicadas estão descritas abaixo, assim como incentivo financeiro para épocas do ano, por exemplo, uma promoção para os 5 meses do ano que são mais utilizados pelo usuários casuais, apesar de não ser um plano para o ano todo pode-se chegar a um maior lucro caso os usuários casuais assinem esse plano sazonal da Cyclistic, onde essa promoção deve ser impulsionada nos meses abril, maio e junho, onde os usuários se sentiriam mais tentados a comprar esse pacote sazonal por terem mais tempo para utilizar esse pacote.

- **Como a Cyclistic pode usar a mídia digital para influenciar os passageiros casuais a se tornarem membros?**

Os esforços por parte da Cyclistic em mídia digital devem levar em conta os comportamentos dos usuários durante o ano, então existe a possibilidade de impulsionar os anúncios em mídia digital da empresa durante os meses com melhor clima para os usuários (maio, junho, julho, agosto e setembro), que são os meses que mais usuários casuais utilizam o serviço da empresa. Também é interessante colocar esses anúncios focados nos dias da semana com mais usuários casuais, sendo sexta, sábado e domingo a recomendação para esses dias.