Estudo de caso

Empresa de compartilhamento de bicicletas

**Set/2023 a Ago/2024**

## Estudo elaborado para conclusão do Certificado Profissional de Google Data Analytics

# Cenário

A **Cyclistic** é uma empresa (fictícia) de compartilhamento de bicicletas em Chicago e que conta com uma frota de 5.824 bicicletas com rastreamento geográfico e bloqueio dentro de uma rede de 692 estações na cidade. O diretor de marketing acredita que o sucesso futuro da empresa depende da maximização do número de planos anuais contratados, por isso encomendou a análise para entender como os ciclistas casuais e os membros anuais usam as bicicletas da Cyclistic de forma diferente, a fim de que, a partir dos *insights* gerados, possa ser desenvolvida uma nova estratégia de marketing para converter passageiros casuais em membros anuais, uma vez que estes são mais lucrativos para a empresa.

# Tarefa de negócios

Analisar o conjunto de dados disponibilizado com a finalidade de identificar as diferenças entre usuários casuais e membros do sistema de compartilhamento, no que se refere à utilização das bicicletas, e buscar *insights* para nortear decisões de negócios por parte da direção da empresa com vistas ao incremento do número de usuários membros, ou seja, aqueles que realizam a assinatura anual do serviço.

# Partes interessadas

São partes diretamente interessadas a Diretora de Marketing Lily Moreno e sua gerente, a equipe de análise de marketing e a equipe executiva da Cyclistic.

# Fonte dos dados

Os dados foram disponibilizados pela Motivate International Inc. sob esta [licença](https://www.divvybikes.com/data-license-agreement), estando os arquivos disponíveis no endereço: <https://divvy-tripdata.s3.amazonaws.com/index.html>.

Encontram-se disponíveis arquivos mensais a partir de abril de 2020, porém, para fins deste estudo de caso, foram baixados e utilizados os relatórios referentes aos 12 últimos meses (set/2023 a ago/2024).

# Processamento dos dados

Para viabilizar a análise do período os arquivos mensais foram consolidados em um único data frame, utilizando para isso recursos da linguagem de programação R, com auxílio de ferramentas de IA generativa (ChatGPT e Copilot).

Numa avaliação preliminar, observou-se que os dados eram abrangentes, somando um total de 5.699.639 registros de viagens de bicicleta, não sendo identificadas inicialmente inconsistências ou omissões de dados nos campos chave.

Aprofundando a análise, verificou-se que uma parte pouco significativa dos registros, no total de 1263 viagens, representado 0,02% do total, tinha a data/hora de retirada da bicicleta na estação maior ou igual à data/hora de devolução. Por se tratar de dado reconhecidamente inconsistente, decidiu-se pela sua exclusão, restando um total de 5.698.376 registros.

Verificou-se ainda que uma parte dos registros não possuía a identificação (ID) e/ou o nome da estação de retirada ou de devolução. Considerando que essa omissão não significa necessariamente que as viagens deixaram de ocorrer, decidiu-se pela manutenção dos registros e foi dado tratamento para preenchimento do nome da estação quando o ID estava presente.

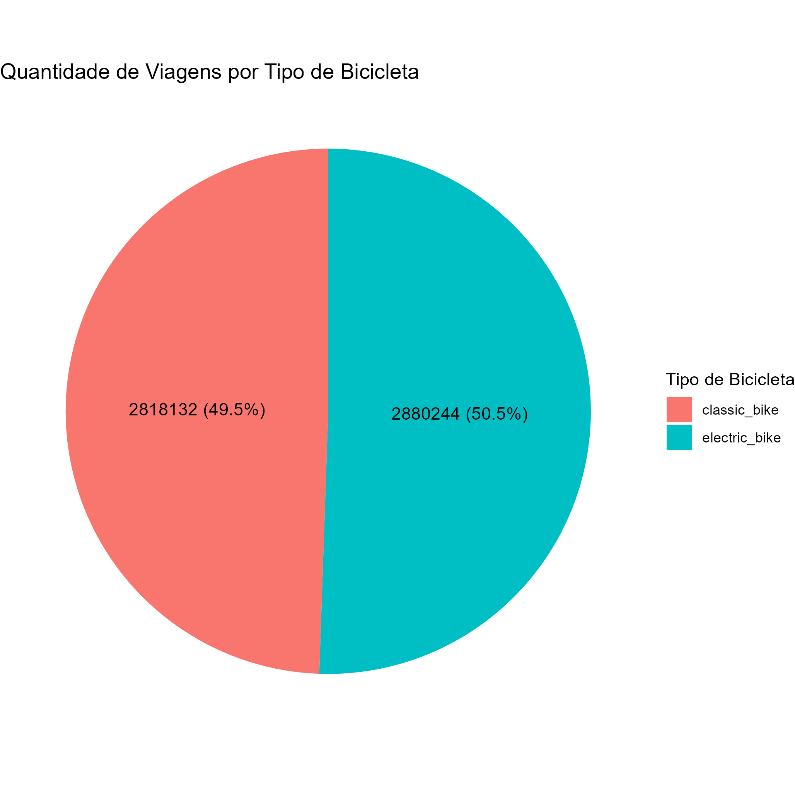
Para carga do nome da estação, foram utilizados os dados de outros registros do mesmo conjunto de dados. Quando o nome estava associado a mais de um ID, optou-se por aquele nome com maior número de registros.

Foram descartadas nessa etapa as colunas de latitude e longitude, que não seriam utilizadas, a fim de manter o conjunto de dados mais enxuto.

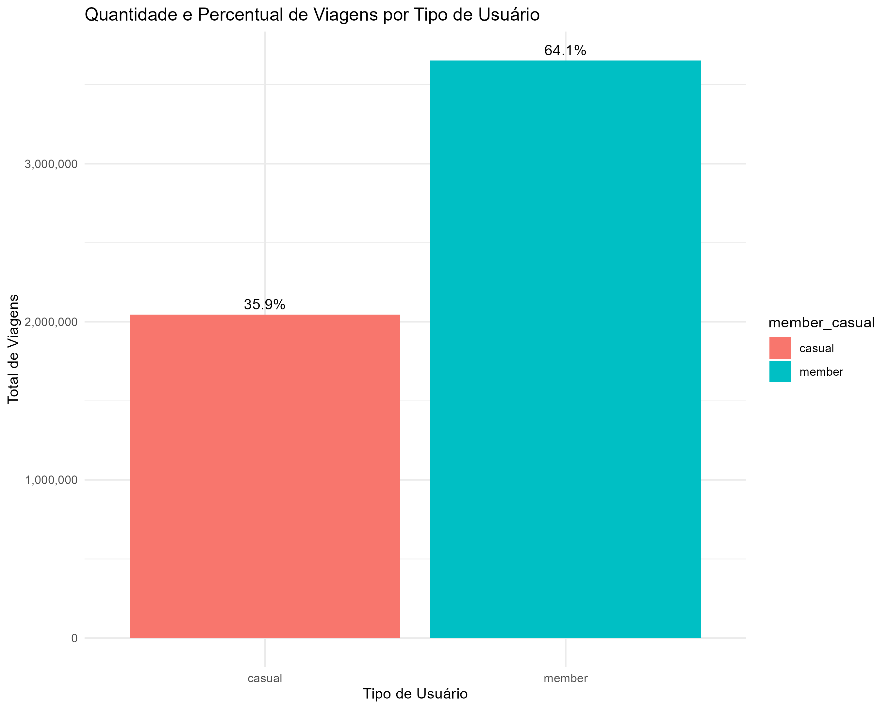
# Análise

Inicialmente, procurou-se obter uma visão geral dos dados, mediante a criação de alguns gráficos por tipo de bicicleta (clássica ou elétrica) e tipo de usuário (casual ou membro).

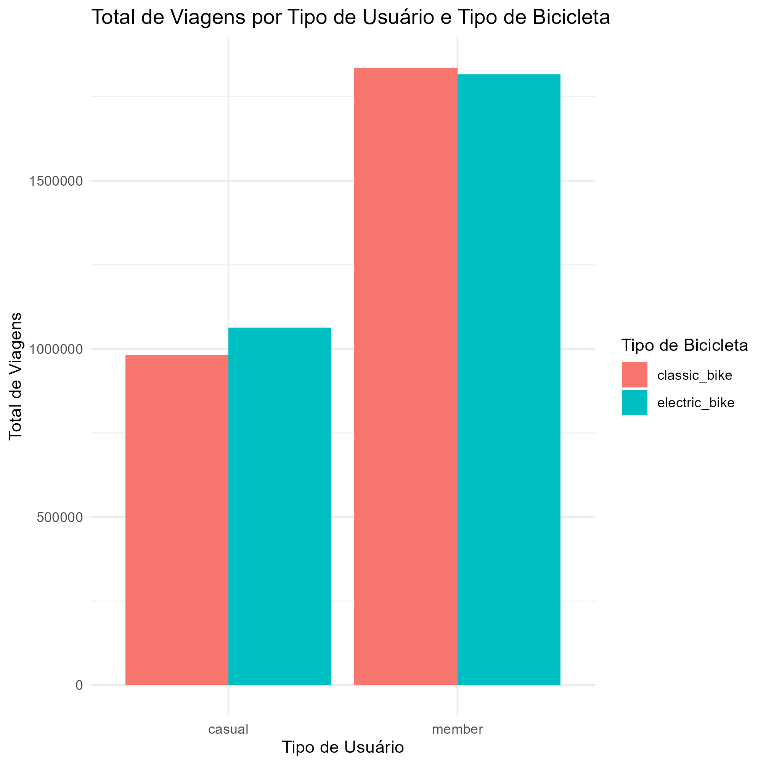
Verificou-se de início que a quantidade de viagens de cada tipo foi idêntica, conforme demonstrado no gráfico a seguir, diferindo da informação contida no material disponibilizado, que mencionava que “*a maioria dos ciclistas opta por bicicletas tradicionais; cerca de 8% dos motociclistas usam as opções assistivas*”.



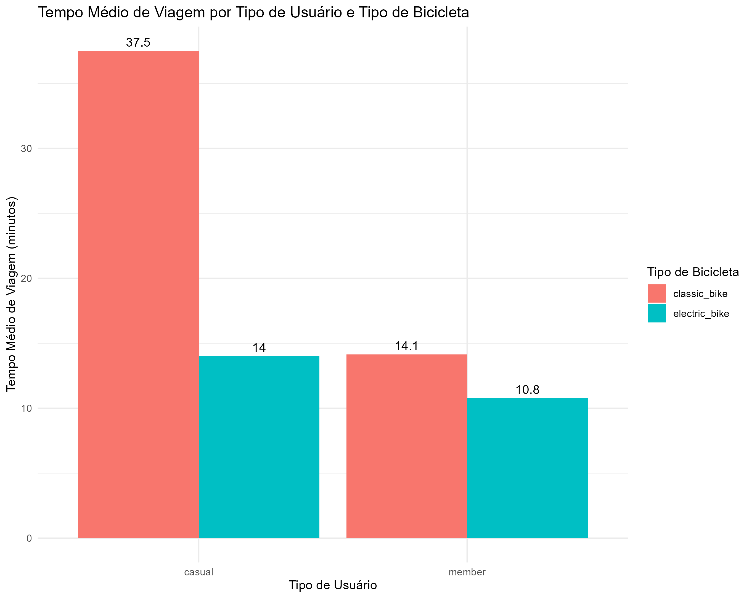
Em relação à quantidade de viagens por tipo de usuário (membro e casual), a proporção é de quase 2 para 1. No entanto, os usuários casuais ainda são responsáveis por mais de 2 milhões de viagens, sendo pertinente, portanto, o objetivo de negócio definido de buscar a conversão de usuários casuais em membros.



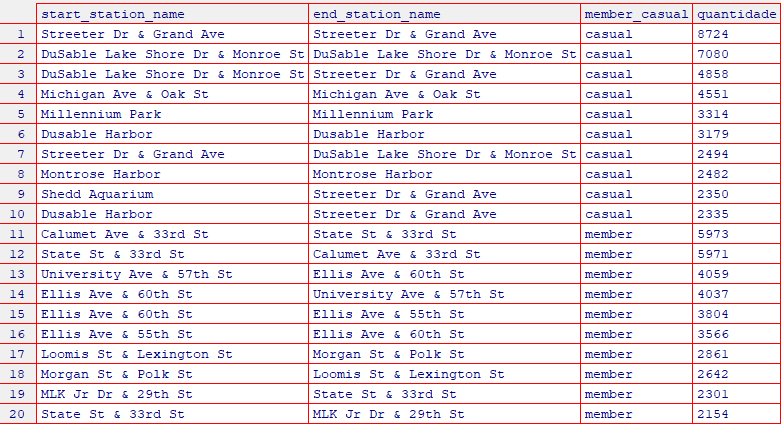
Na sequência das análises, procurou-se verificar se os dois tipos de usuário tinham alguma predileção por algum dos tipos de bicicleta, no entanto constatou-se que há um certo equilíbrio na utilização de bicicletas convencionais e elétricas, havendo apenas uma pequena propensão de usuários casuais por bicicletas elétricas, conforme demonstrado no gráfico a seguir.



Na sequência, buscou-se conhecer a duração média das viagens dos usuários casuais e dos membros, com vistas a identificar eventuais diferenças na utilização das bicicletas, sendo observado que o tempo médio de utilização dos usuários casuais é cerca de 30% maior quando utilizam bicicletas elétricas (14 minutos contra 10,8 minutos) e 166% maior no caso de bicicletas convencionais (37.5 minutos contra 14,1 minutos), o que pressupõe que os dois grupos fazem uso das bicicletas para finalidades distintas.



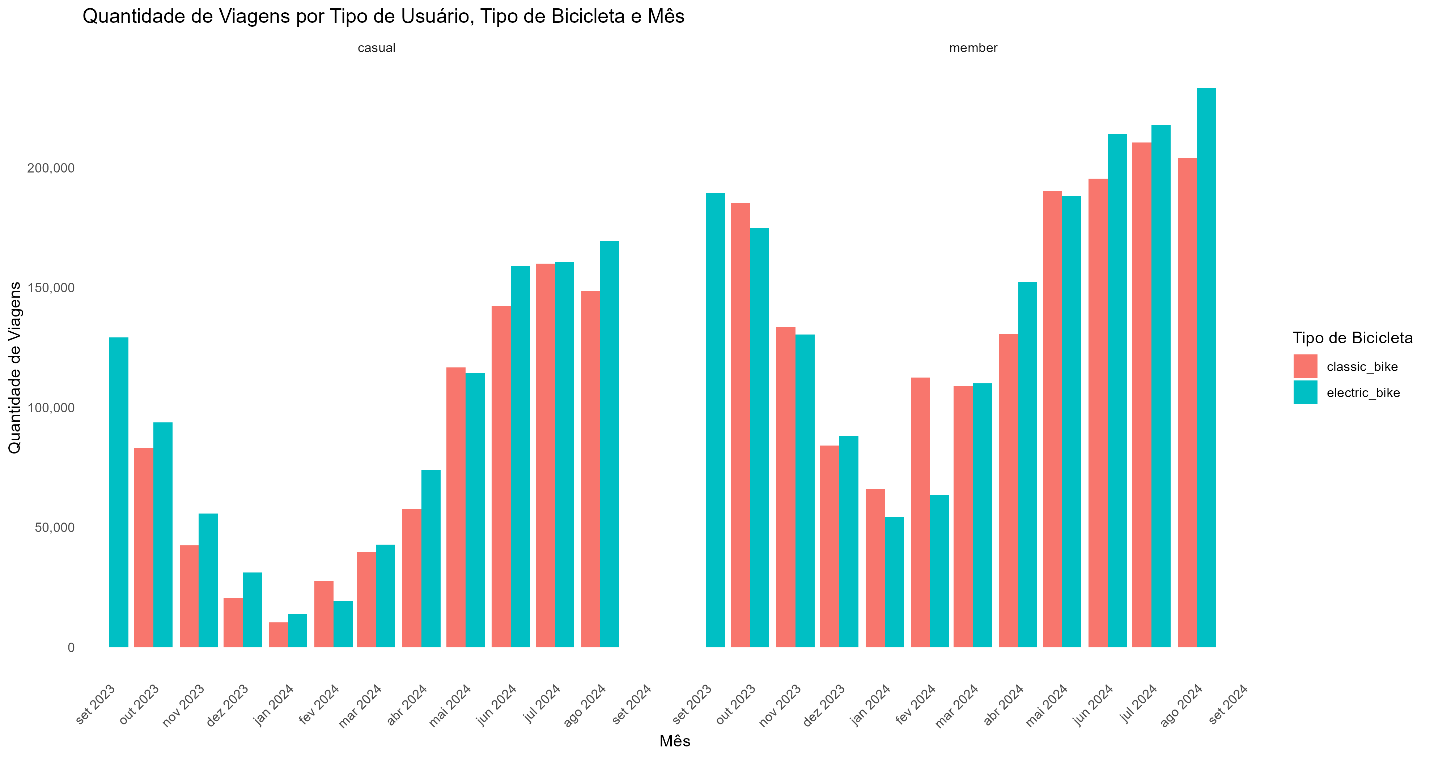
Seguindo essa importante informação revelada pelos dados, buscou-se verificar os trajetos mais realizados pelos dois grupos, mediante o levantamento dos pares estação de origem / estação de destino com maior frequência. O resultado apontou que há uma frequência de retirada e devolução da bicicleta na mesma estação entre os usuários casuais, fato que não é observado para os membros. Essa constatação veio corroborar a premissa de que usuários membros utilizam as bicicletas mais para deslocamentos cotidianos (trabalho, escola, etc.), enquanto que usuários casuais fazem uso das bicicletas para passeio.



No quadro acima pode-se observar que dos top 10 pares estação de origem / estação de destino referentes às viagens dos usuários casuais, em 6 deles a estação de origem e destino é a mesma.

Na sequência, buscou-se o aprofundamento da análise com a finalidade de identificar distinções relevantes entre os dois tipos de usuários, considerando meses do ano, dias da semana e horários do dia.

Na comparação mês a mês, observa-se que há uma redução significativa no período do inverno, tanto no uso de bicicletas convencionais quanto elétricas, com maior intensidade no grupo dos usuários casuais. Em janeiro de 2024, mês em que houve maior redução no número de viagens, os passageiros casuais apresentaram redução de 87% na quantidade de viagens em relação à média mensal (10451 viagens contra 81814 de média) nas bicicletas convencionais e de 84% de redução em relação à média mensal (13995 viagens contra 88623 de média) nas bicicletas elétricas.



Os usuários membros, por sua vez, também tiveram redução significativa na utilização das bicicletas nos meses do inverno, porém em grau um pouco inferior aos usuários casuais, conforme quadro abaixo, o que pode ser justificado pelo fato de os usuários membros já terem uma rotina de uso de bicicletas e o desembolso com a assinatura já ter sido feito para o ano todo.

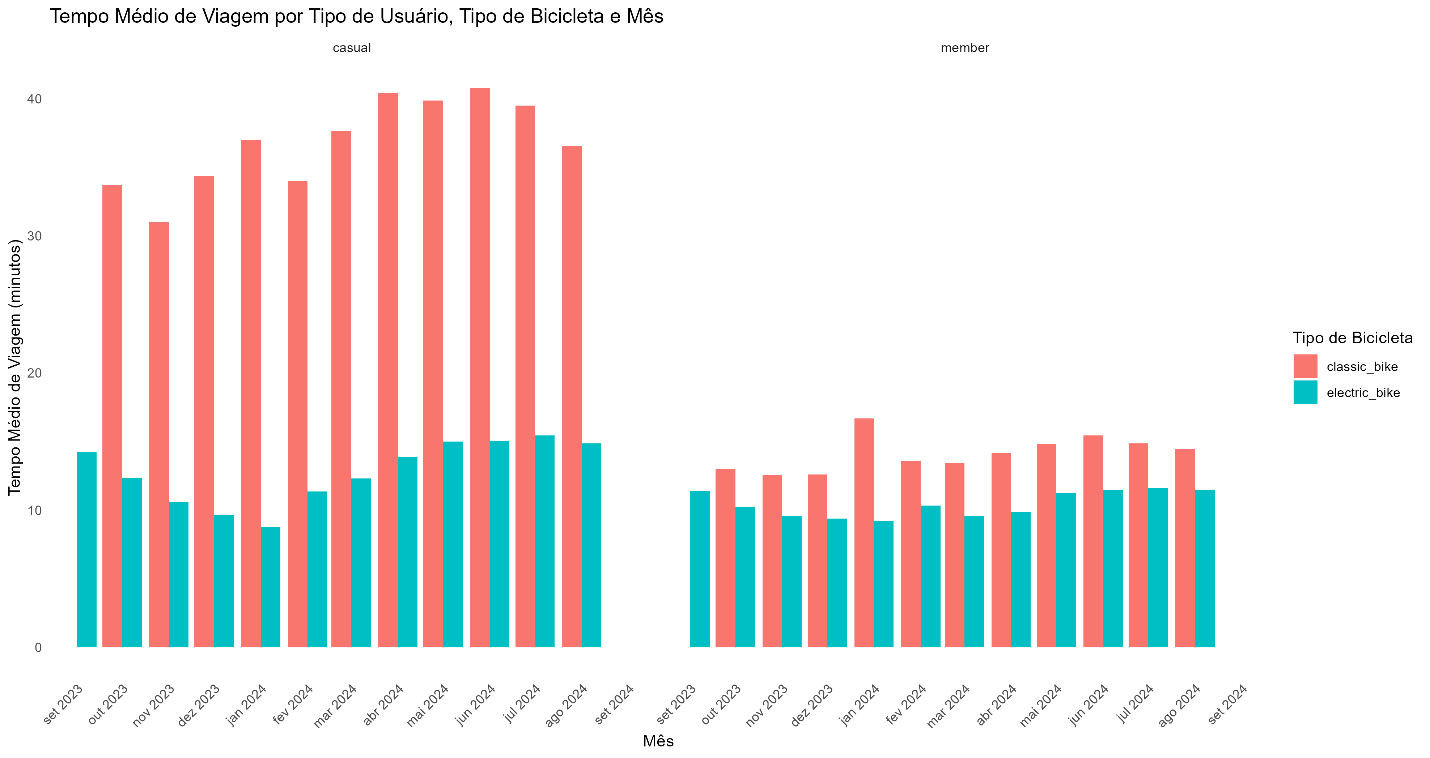
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tipo de Usuário** | **Tipo de Bicicleta** | **Média Mensal de Viagens** | **Qtd Viagens jan 2024** | **% Redução jan 2024** |
| Membro | classic\_bike | 153030 | 66065 | -57 |
| Membro | eletric\_bike | 151397 | 54265 | -64 |
| Casual | classic\_bike | 81814 | 10451 | -87 |
| Casual | eletric\_bike | 88623 | 13995 | -84 |

Isso indica que os usuários de modo geral se sentem menos encorajados a utilizar bicicletas durante os meses mais frios do ano, o que é justificável, uma vez nesse período do ano as pessoas em geral curtem menos atividades ao ar livre.

Nesse contexto, é interessante que a Cyclistic desenvolva campanhas de marketing em períodos do ano em que a utilização de bicicletas em geral é maior, estimulando os usuários ao uso de bicicletas como uma prática saudável e não poluente.

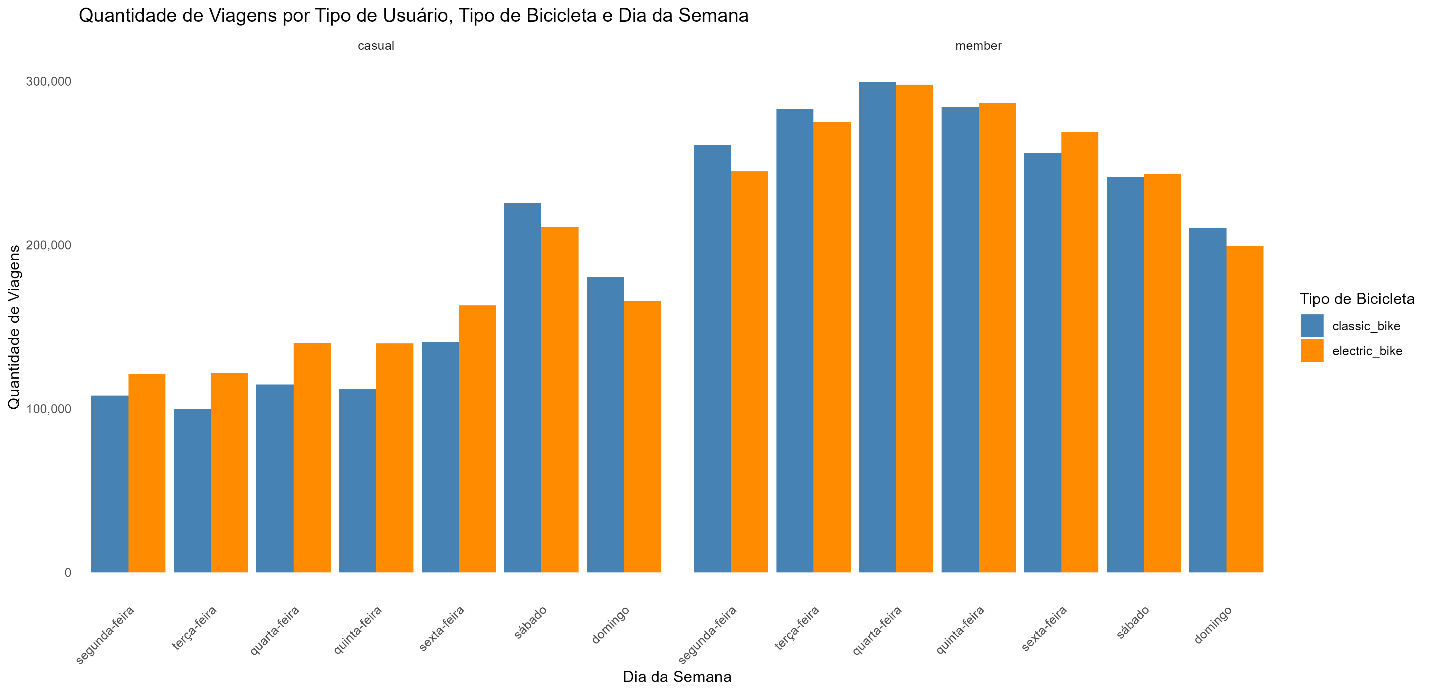
Com relação ao tempo médio de utilização das bicicletas elétricas, não há variação significativa considerando os períodos do ano, embora se tenha constatado uma pequena redução no tempo de uso destas no inverno pelos usuários casuais e também por parte dos usuários membros, embora estes últimos em grau menor.

Já em relação às bicicletas convencionais, surpreendentemente se verificou um aumento no tempo médio de uso no mês de janeiro de 2024 (mês com menor quantidade de viagens) em relação aos meses anteriores, tanto por parte dos usuários casuais quanto pelos usuários membros, estes últimos com variação mais acentuada.



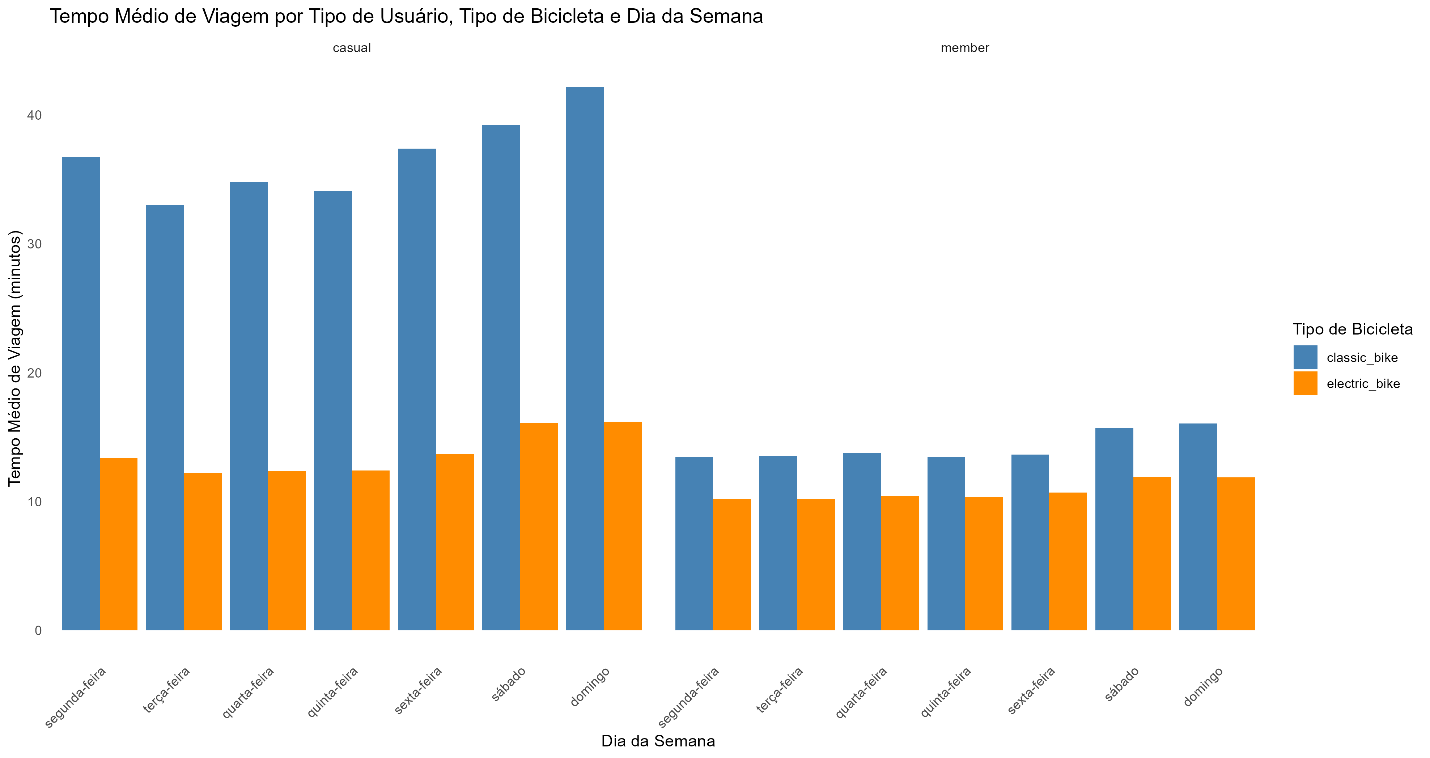
Na comparação do uso das bicicletas nos diferentes dias da semana, nota-se que há maior utilização das bicicletas por parte dos usuários casuais em finais de semana, principalmente no sábado, ocorrendo o oposto com usuários membros, que fazem uso de bicicletas principalmente durante a semana, com redução do uso em final de semana, especialmente no domingo.

Os usuários casuais utilizam mais bicicletas elétricas durante os dias de semana e mais bicicletas convencionais no final de semana.



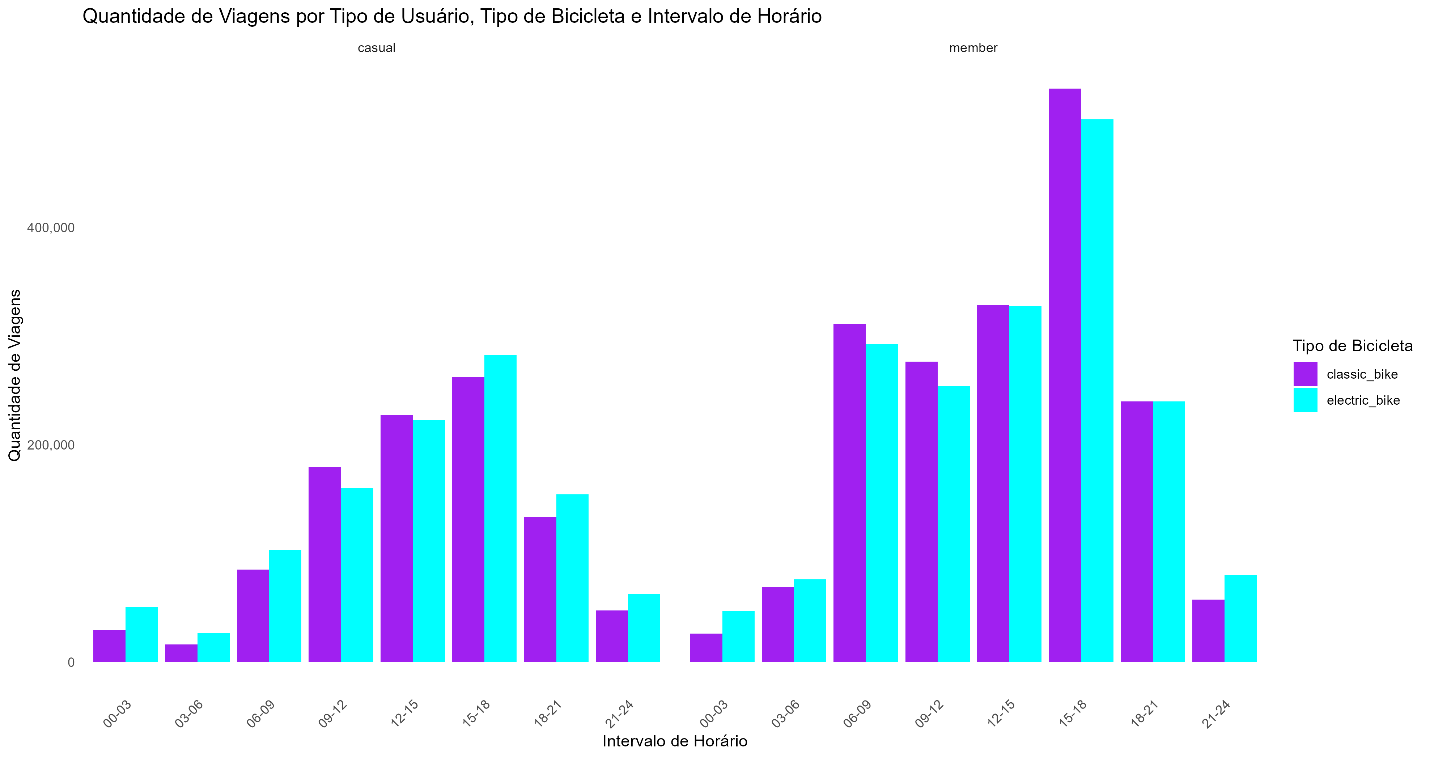
Com relação à duração das viagens, observa-se uma duração bem superior das viagens dos usuários casuais quando utilizam bicicletas convencionais, em qualquer dia da semana, mas de forma um pouco mais destacada nos finais de semana. Isso reforça que esse grupo utiliza bicicletas principalmente para passeio e atividades não rotineiras, o que os leva a aproveitar melhor o tempo de que dispõem.

Já os usuários membros, estes realizam as viagens em bem menos tempo do que os usuários casuais, principalmente quando utilizam bicicletas convencionais. Na comparação da duração das viagens desse grupo, observa-se que é maior no caso de bicicletas convencionais do que nas elétricas. Isso pode ser explicado pelo fato desse grupo fazer mais uso de bicicletas para deslocamentos rotineiros, sendo normal levar mais tempo para se deslocarem com bicicletas clássicas do que com bicicletas elétricas.



Por último, procurou-se levantar os períodos do dia com maior número de viagens pelos dois grupos, sendo constatado que os usuários casuais tendem a utilizar as bicicletas em horário comercial, notando-se um aumento gradual na utilização no decorrer do dia, com pico no final do dia, no período entre 15 e 18h.

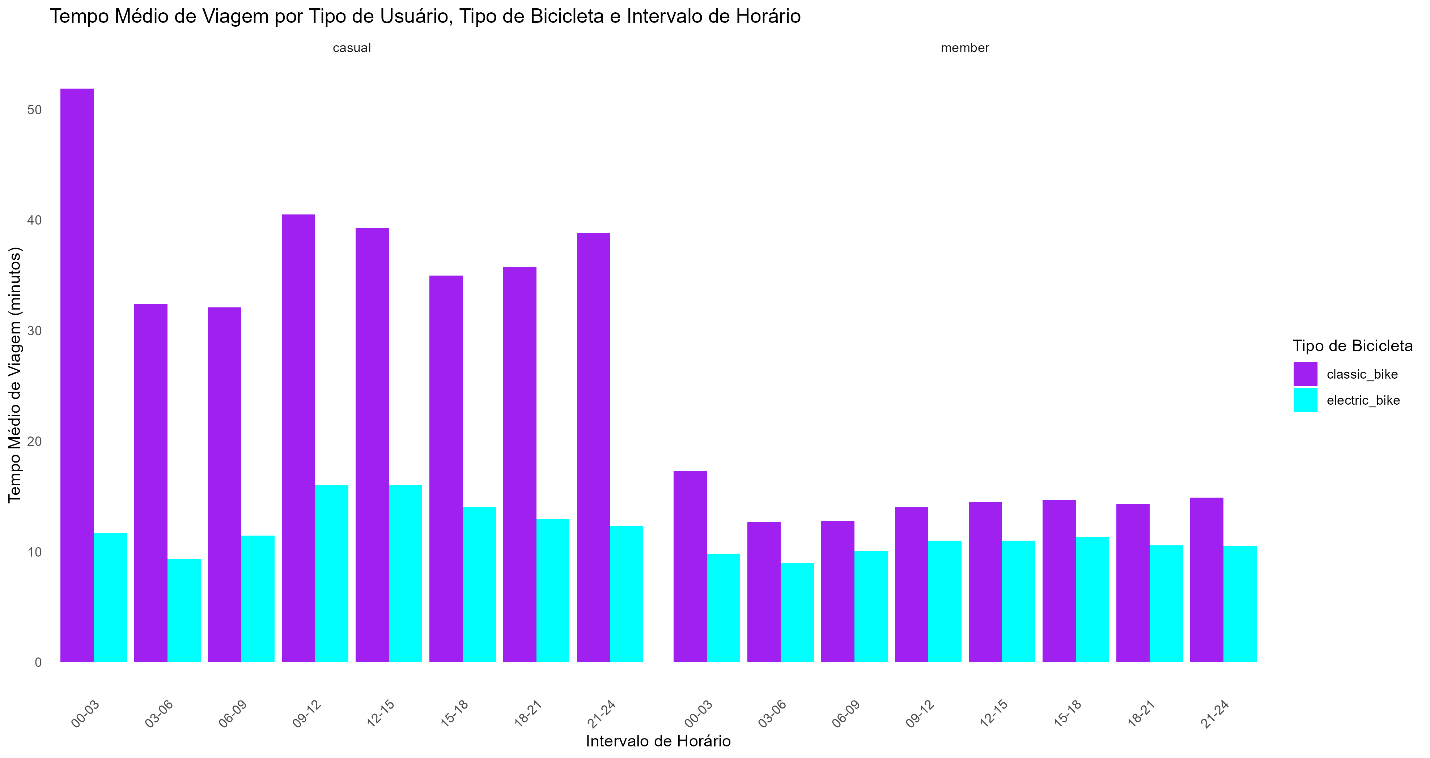
Os usuários membros, por sua vez, apresentam um elevado nível de utilização já no período compreendido entre 6 e 9h, havendo uma ligeira redução entre 9 e 12h, voltando a subir no período entre 12 e 15h, alcançando o pico entre 15 e 18h, mas ainda com bom nível utilização entre 18 e 21h, tendo redução significativa após 21h e até 6h do dia seguinte.



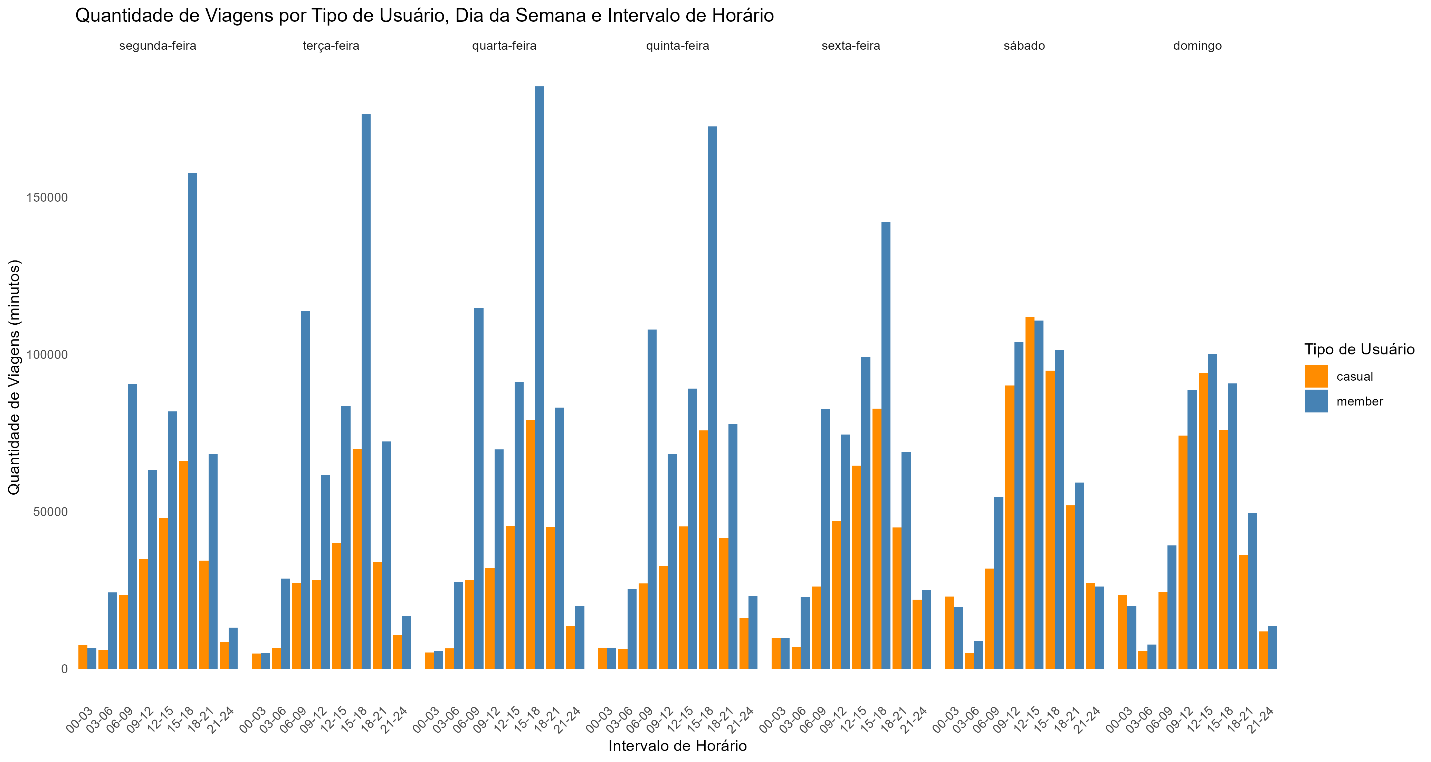
Com relação à duração das viagens por período do dia, verifica-se que, de modo surpreendente, o horário em que as viagens são mais longas, em bicicletas convencionais, é o compreendido entre 0 e 3h da manhã, tanto para usuários casuais quanto para usuários membros.

Embora a quantidade de viagens nessa hora da noite seja pequena, se comparada à quantidade de outros períodos do dia, presume-se usuários insones utilizam as bicicletas para percorrerem distâncias mais longas nessa hora da noite, aproveitando-se do trânsito mais tranquilo e da menor presença de pessoas nas ruas.

Já no caso das bicicletas elétricas, não há variações significativas na duração das viagens por qualquer dos grupos, levando a crer que a utilização destas seja para deslocamento de um ponto “A” a um ponto “B” no menor tempo possível.



Por fim, foi comparada quantidade de viagens por tipo de usuário, dias da semana e horário, ficando mais uma vez evidenciada a maior concentração de viagens pelos usuários membros em dias úteis e pelos usuários casuais em finais de semana, principalmente no horário compreendido entre 9 e 18h.



Desta forma, a abordagem dos usuários casuais com incentivo para migração para a condição de assinantes anuais deve ocorrer preferencialmente em finais de semana, entre 9 e 18h. Caso se busque reforçar a estratégia com abordagens em meio de semana, o horário mais indicado é entre 12 e 18h.

# Recomendações

Com base nas descobertas apresentadas acima, seguem propostas para incremento da quantidade de usuários membros, mediante o incentivo para que os usuários casuais migrem para o plano de assinatura anual:

1. Focalizar as campanhas no bem-estar, na vida saudável e prática regular de exercícios ao ar livre, considerando que os usuários casuais parecem prezar por esses elementos;
2. Realizar campanhas de adesão em períodos do ano de maior procura, evitando os meses de inverno, quando há uma redução significativa na quantidade de viagens;
3. Concentrar as abordagens em finais de semana, das 9 às 18h, ou, para expandir ainda mais, nos dias de semana no horário das 12 às 18, por serem os períodos em que os usuários casuais mais utilizam as bicicletas.