

# **GOOGLE Data Analytics Professional Capstone Project**

## **CYCLIST Bike Share**

Uma análise de como os clientes da empresa utilizaram as bicicletas no ano de 2022 em Chicago.

Análise de dados realizada por Lucas Bellini

GitHub: <https://github.com/Lucas-Bellini>

LinkedIn: <https://www.linkedin.com/in/lucasbellini/>

# **QUEM É LUCAS BELLINI?**

## **Analista de Dados**

Prazer, meu nome é **Lucas Bellini**.

Neste momento da minha vida, sou Policial Militar, formado em Gestão Financeira, graduando em Ciência de Dados e realizando meu **MBA em CIÊNCIA DE DADOS PARA NEGÓCIOS: AI, Data Science e Big Data** pela PUC-RS.

Me interessei pela área de dados ainda na faculdade de Gestão Financeira e hoje sou um aspirante à Ciência de Dados.

Neste projeto de conclusão de curso da Google irei realizar uma análise de dados de uma empresa chamada "Cyclistic" e aplicarei uma habilidade que acredito ser veemente em analistas de dados: **Storytelling**.

# CENÁRIO

Equipe de Análise de Marketing

Neste projeto realizado pela Google, eu sou um **analista de dados júnior** da empresa Cyclistic, uma empresa de compartilhamento de bicicletas em Chicago.

O diretor de marketing acredita que o futuro da empresa está em **maximizar a quantidade de assinantes do plano de assinatura anual**.

A equipe de marketing precisa saber como os **usuários casuais divergem dos assinantes anuais** e a partir das observações que eu realizar com os dados eles poderão montar estratégias de marketing.

Para isso, os executivos da empresa precisarão **aprovar minhas recomendações** que devem estar baseadas nos dados originais e em boas visualizações.

## SOBRE A EMPRESA

Cyclistic

Em 2016, a Cyclistic lançou uma oferta bem-sucedida de compartilhamento de bicicletas. Desde então, o programa cresceu a uma frota de 5.824 bicicletas geolocalizadas e interligadas a uma rede de 692 estações em Chicago. As bicicletas podem ser desbloqueadas em uma estação e devolvidas em qualquer outra estação no sistema a qualquer momento. Até agora, a estratégia de marketing da Cyclistic baseava-se na construção de uma consciência geral e no apelo aos consumidores.

Uma abordagem que ajudou a tornar essas coisas possíveis foi a flexibilidade de seus planos de preços: passes de viagem única, passes de dia inteiro e assinaturas anuais.

Os clientes que compram passes de viagem única ou de dia inteiro são chamados de passageiros **CASUAIS**.

Os clientes que adquirem assinaturas anuais são membros **ASSINANTES**.

Os analistas financeiros da Cyclistic concluíram que os membros anuais são muito mais lucrativos do que pilotos casuais. Embora a flexibilidade de preços ajude a Cyclistic a atrair mais clientes, a direção acredita que maximizar o número de membros anuais será fundamental para o crescimento futuro.

Em vez de criar uma campanha de marketing voltada para novos clientes, a direção acredita há uma oportunidade sólida de converter passageiros casuais em membros.

# PREFÁCIO

Sobre as visualizações

A partir da próxima página estarão disponíveis a visualização dos dados:

**Área Principal**, dashboard geral dos dados

**Estações**, em geral e para cada cliente

**Semana**, para cada cliente

**Horário**, para cada cliente

**Modelo e bicicleta**, para cada cliente

**Período destaque**

**Mapa de calor**

# GOOGLE Data Analytics Professional Capstone Project

## CYCLIST Bike Share

Uma análise de como os clientes da empresa utilizaram as bicicletas no ano de 2022 em Chicago.

Cliente

Seleção...

Tipo de Bicicleta

Todos

Mês

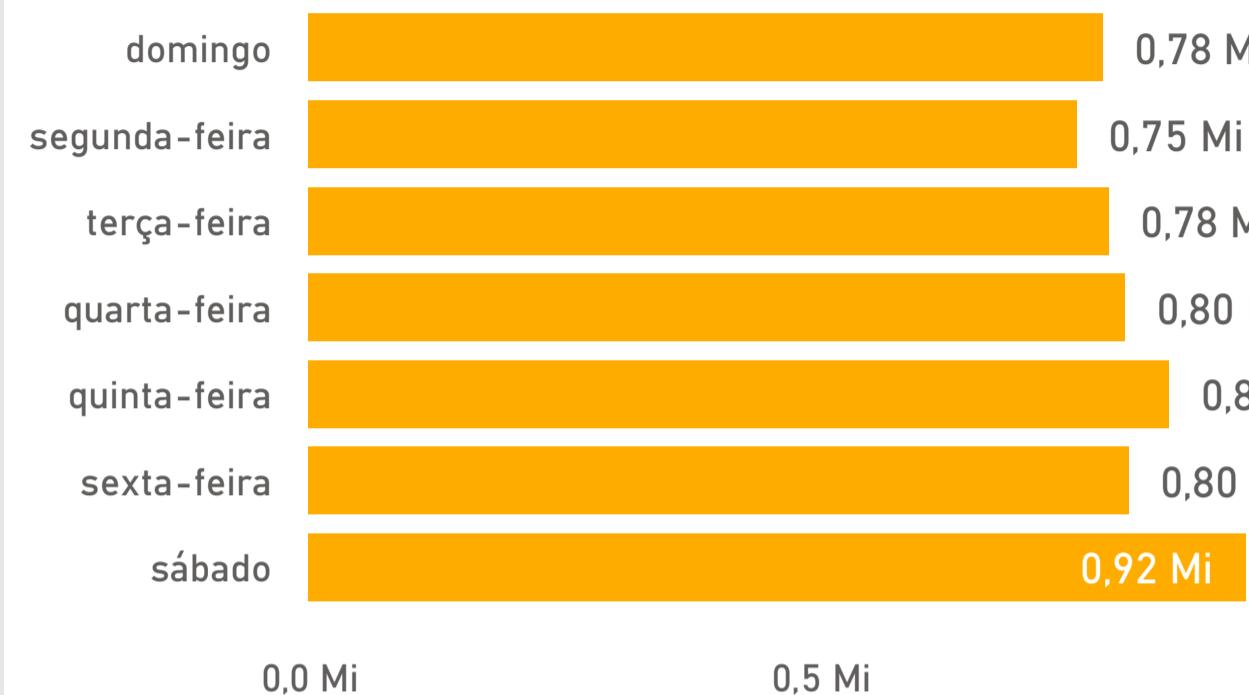
Todos

Dia da Semana

Todos

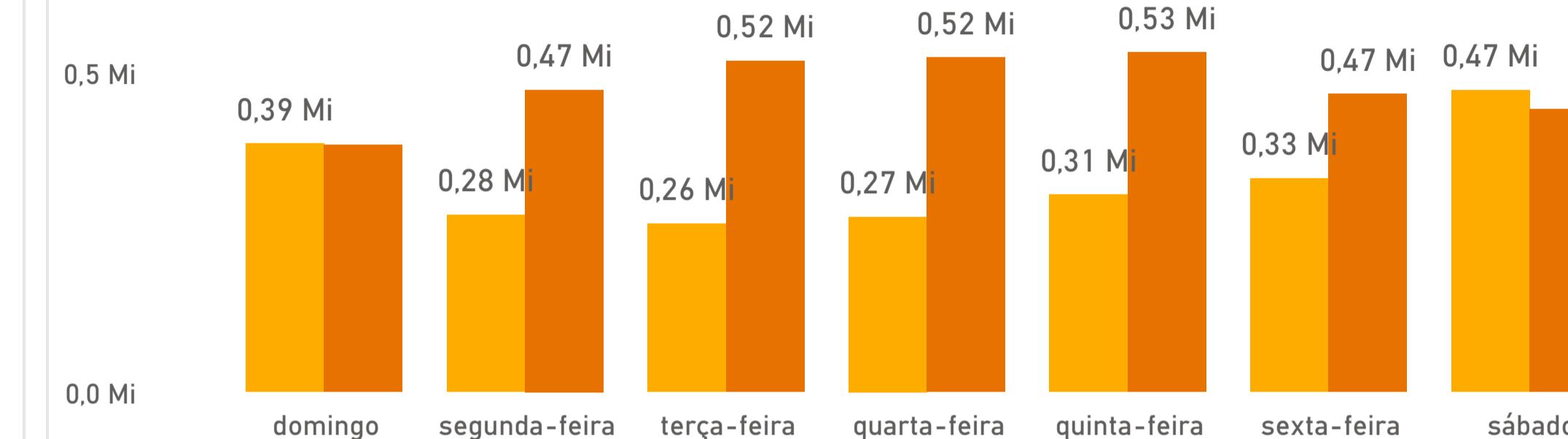
Projeto elaborado por Lucas Bellini, 2024.

### Usabilidade ao longo da semana



### Usabilidade ao longo da semana por cliente

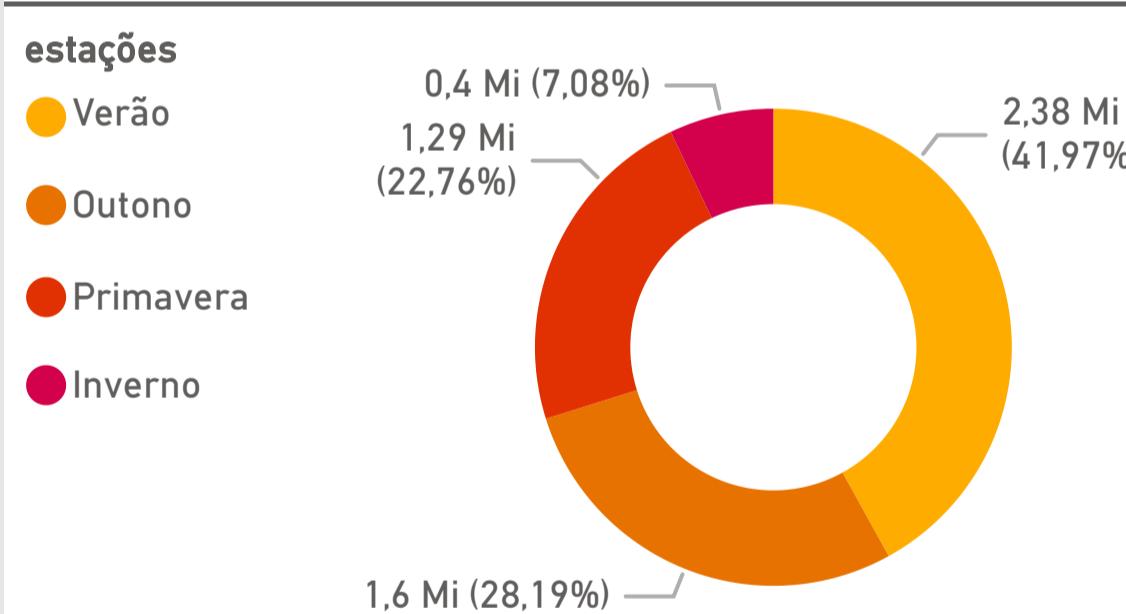
tipo de usuário ● casual ● member



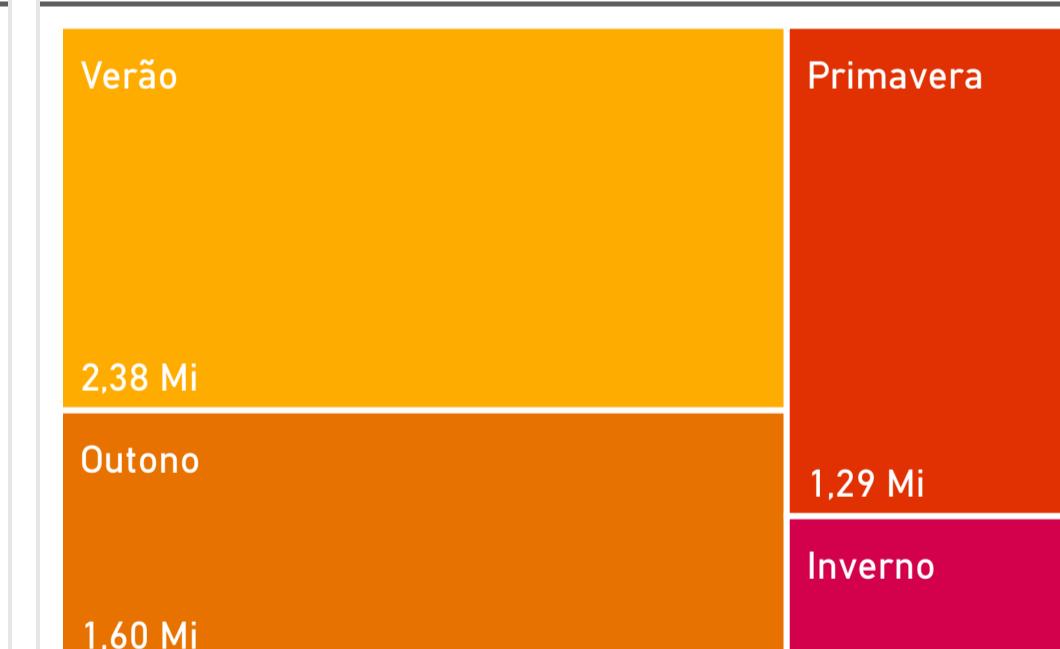
**5,67 Mi**

Contagem de tipo de usuário

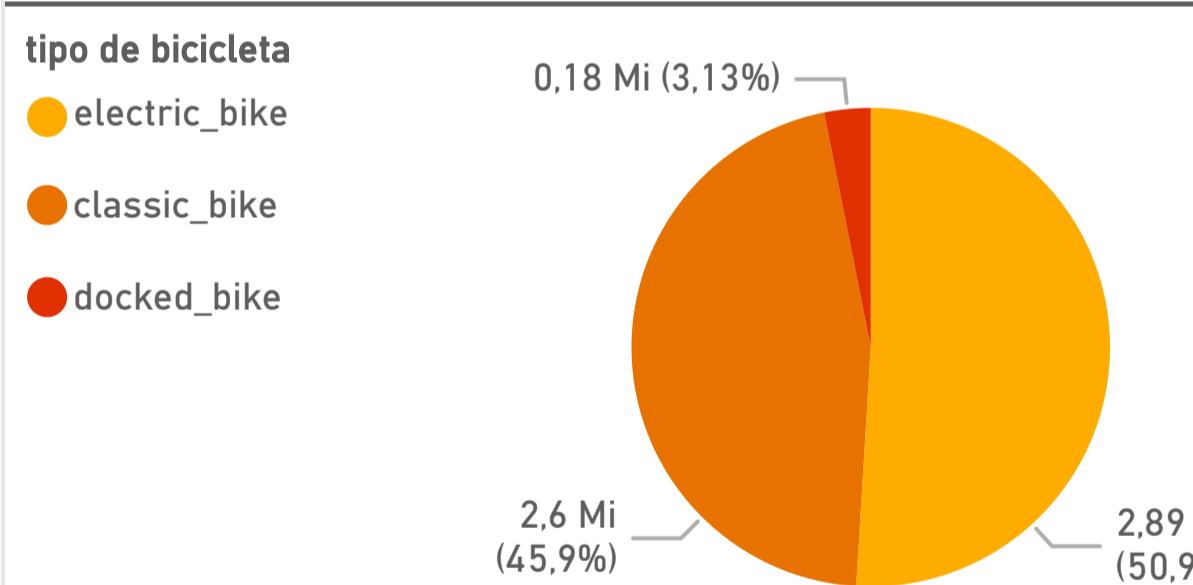
### Contagem de viagens por estação (gráfico 1)



### Contagem de viagens por estação (gráfico 2)



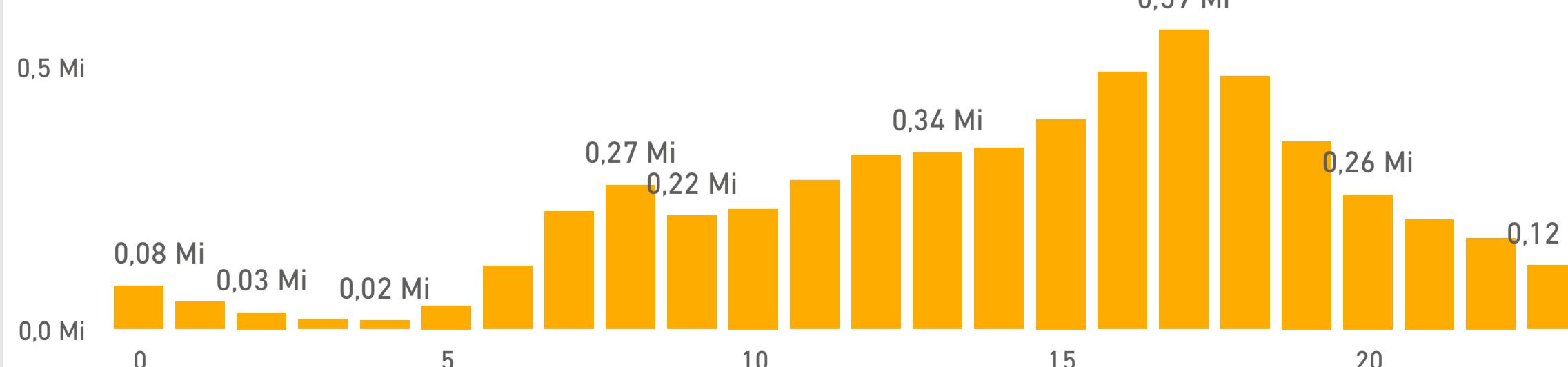
### Uso total por modelo de bicicleta



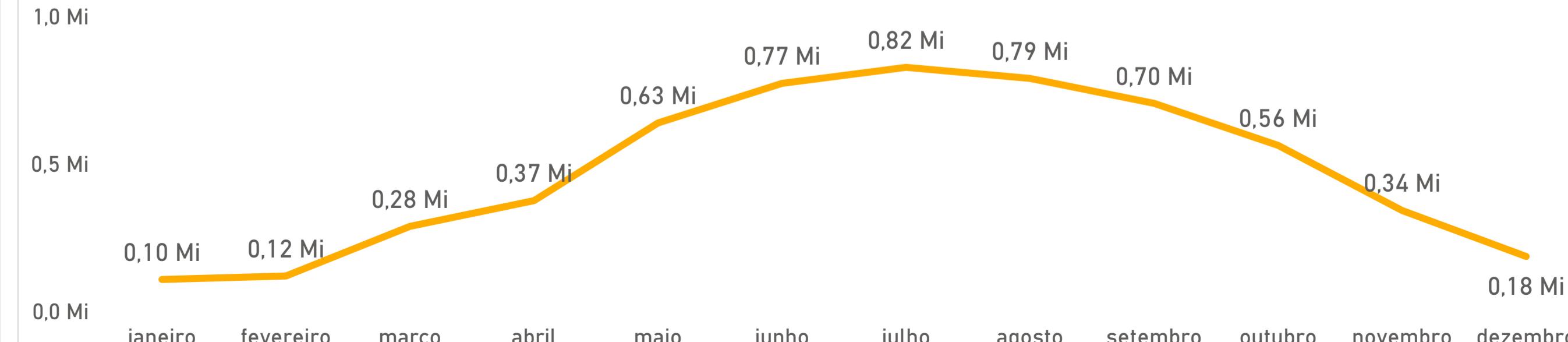
### Matrizes de Análise Rápida

horário	contagem	%	mês	contagem	%
0	82551	1,46%	janeiro	103770	1,83%
1	52340	0,92%	fevereiro	115609	2,04%
2	31580	0,56%	março	284042	5,01%
3	19167	0,34%	abril	371249	6,55%
4	16561	0,29%	maio	634858	11,20%
5	44719	0,79%	junho	769204	13,57%
6	120681	2,13%	julho	823488	14,53%
7	224031	3,95%	agosto	785932	13,87%
8	274293	4,84%	setembro	701339	12,37%
9	216137	3,81%	outubro	558685	9,86%
10	228800	4,04%	novembro	337735	5,96%
11	283803	5,01%	dezembro	181806	3,21%
12	331704	5,85%			
13	336421	5,94%			
14	344966	6,09%			
15	399776	7,05%			
16	489494	8,64%			
17	569593	10,05%			
18	482178	8,51%			
19	357740	6,31%			
20	256669	4,53%			
21	209362	3,69%			
22	173531	3,06%			
23	121620	2,15%			
			dia da semana	contagem	%
			domingo	776259	13,70%
			segunda-feira	751014	13,25%
			terça-feira	782372	13,80%
			quarta-feira	798223	14,08%
			quinta-feira	841591	14,85%
			sexta-feira	801787	14,15%
			sábado	916471	16,17%

### Usabilidade ao longo do dia

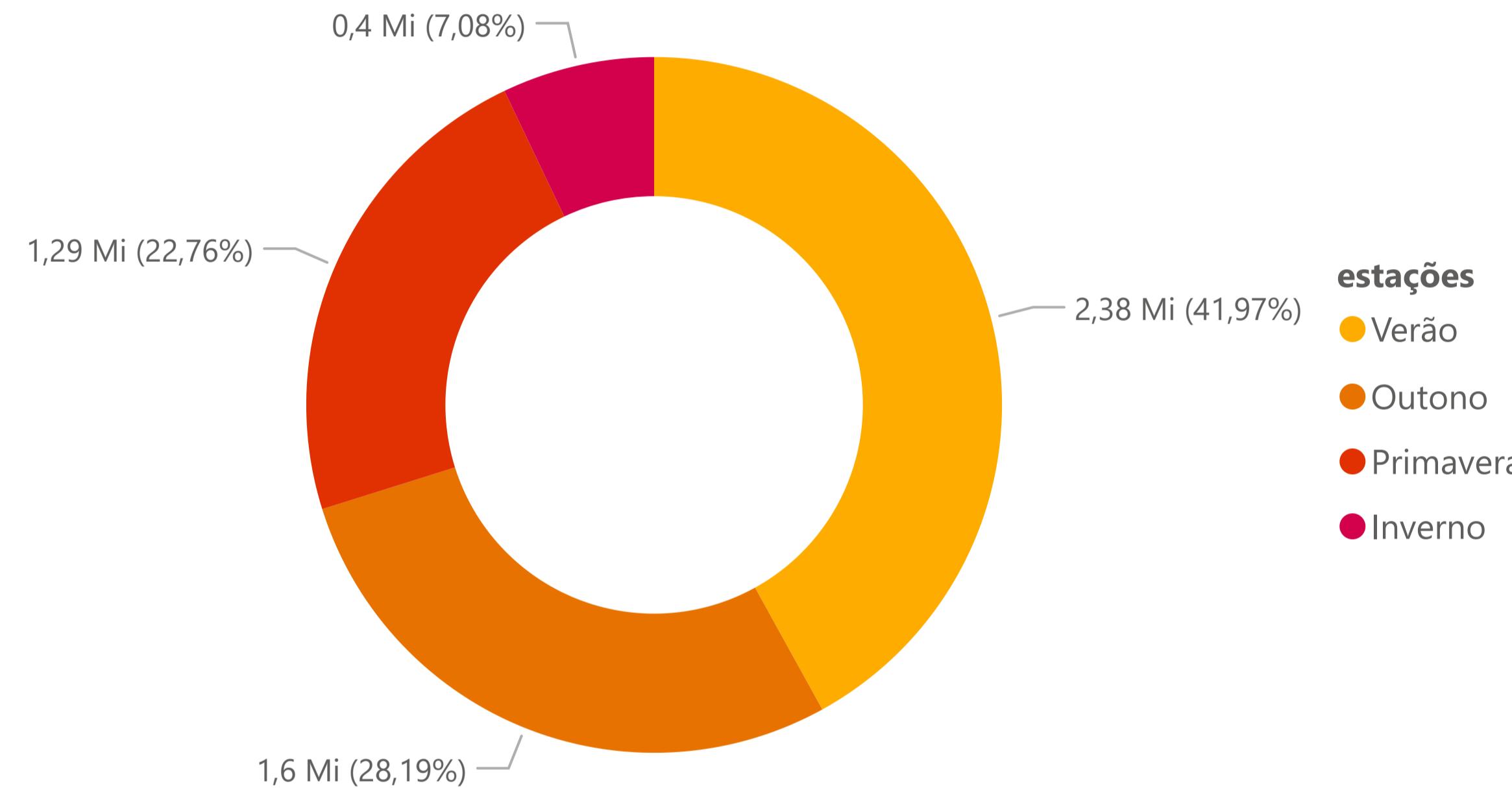


### Usabilidade ao longo dos meses



# QUANTITATIVO DE VIAGENS POR ESTAÇÕES DO ANO DE 2022 EM CHICAGO

Contagem por estações



Em relação ao quantitativo total (assinantes e usuários casuais), as estações **VERÃO** e **OUTONO** tiveram o maior destaque no ano de 2022.

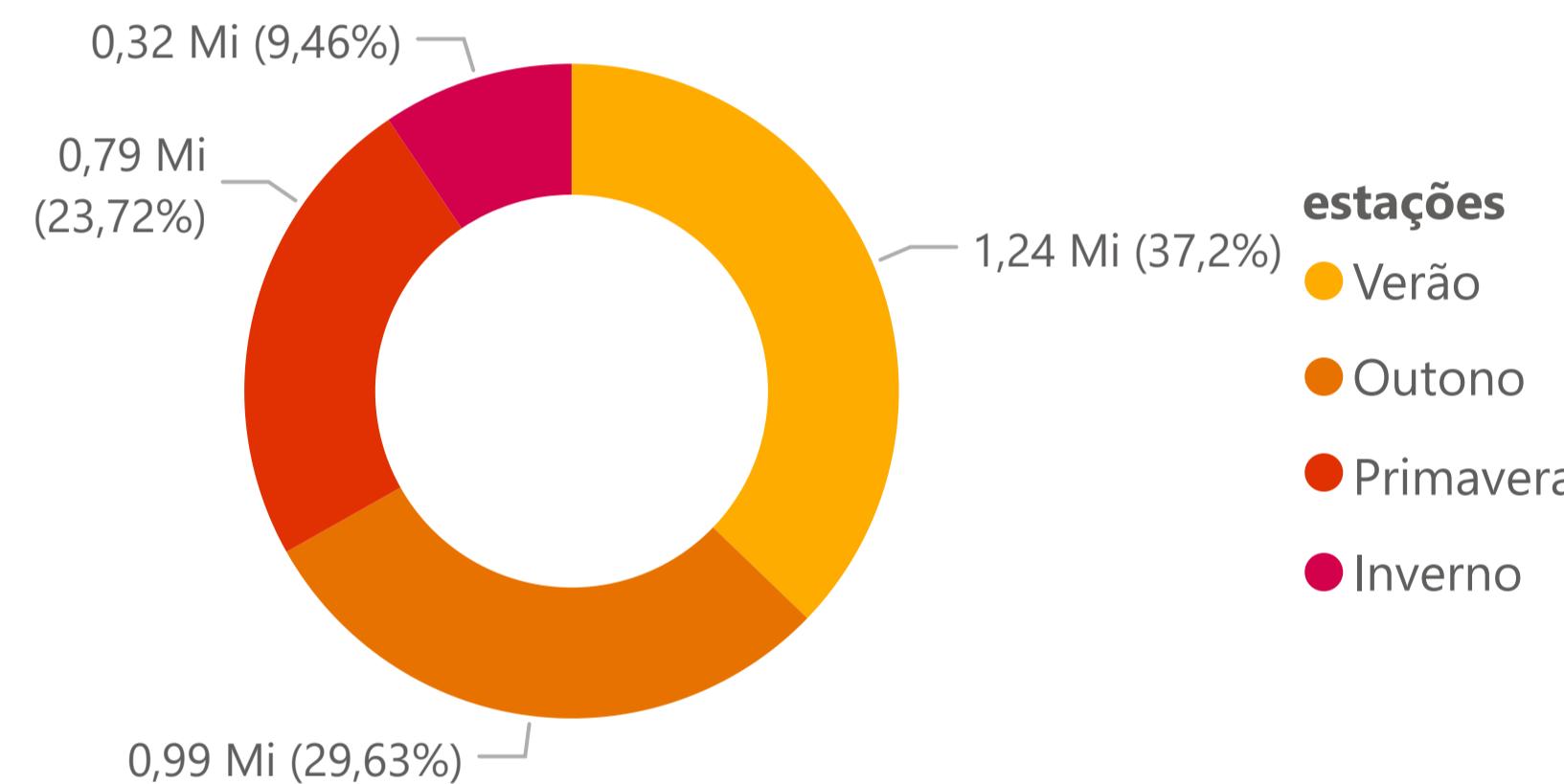
**VERÃO** corresponde a mais de **40% do total de viagens**, sendo uma informação relevante para decisões operacionais e de marketing.

As estações **INVERNO** (somente 7% do total, com **apenas 400 mil viagens**) e **PRIMAVERA** foram as que tiveram o pior desempenho, provavelmente ocasionado pela diminuição da temperatura em Chicago.

## QUANTITATIVO DE VIAGENS POR ESTAÇÕES DO ANO DE 2022 EM CHICAGO

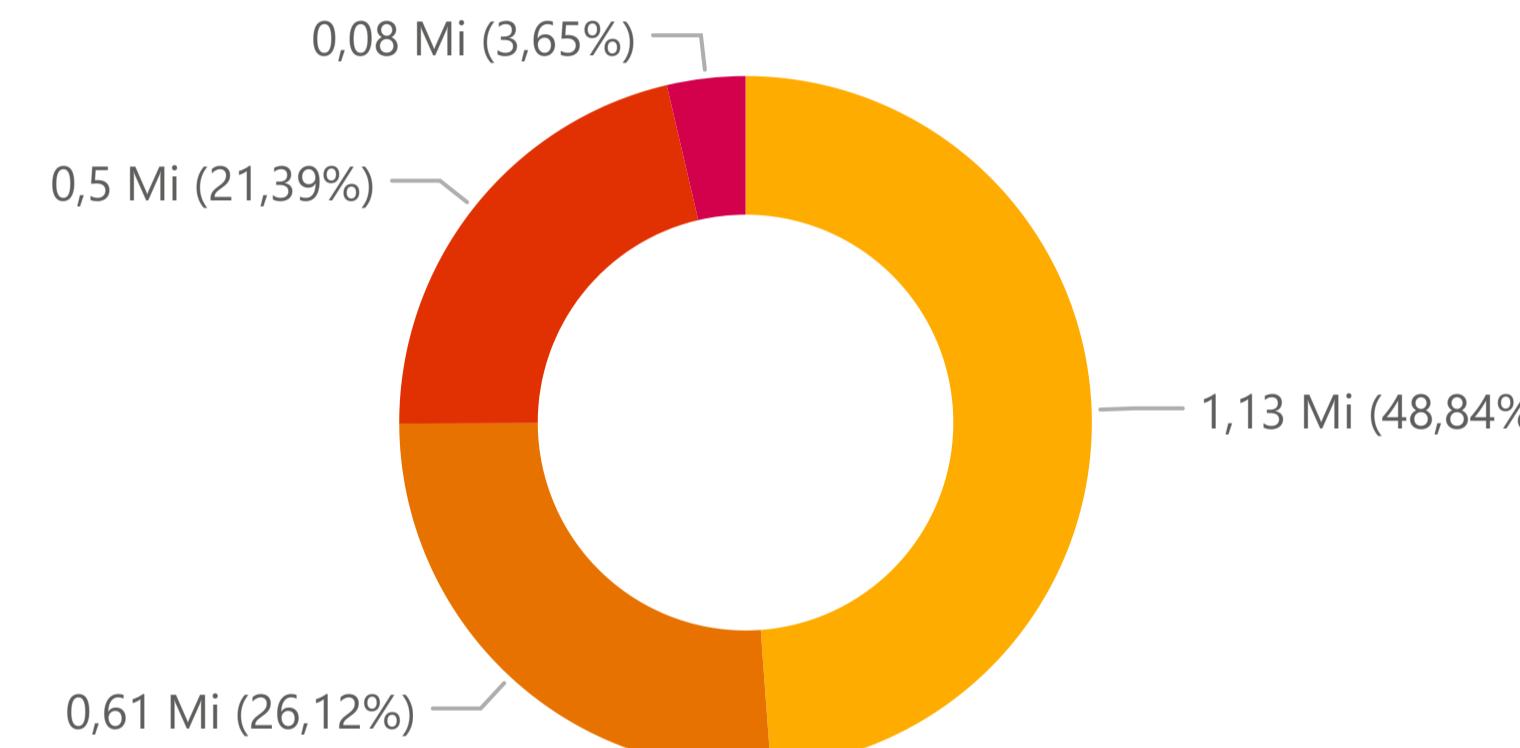
3,35 Mi

Assinantes



2,32 Mi

Usuários casuais



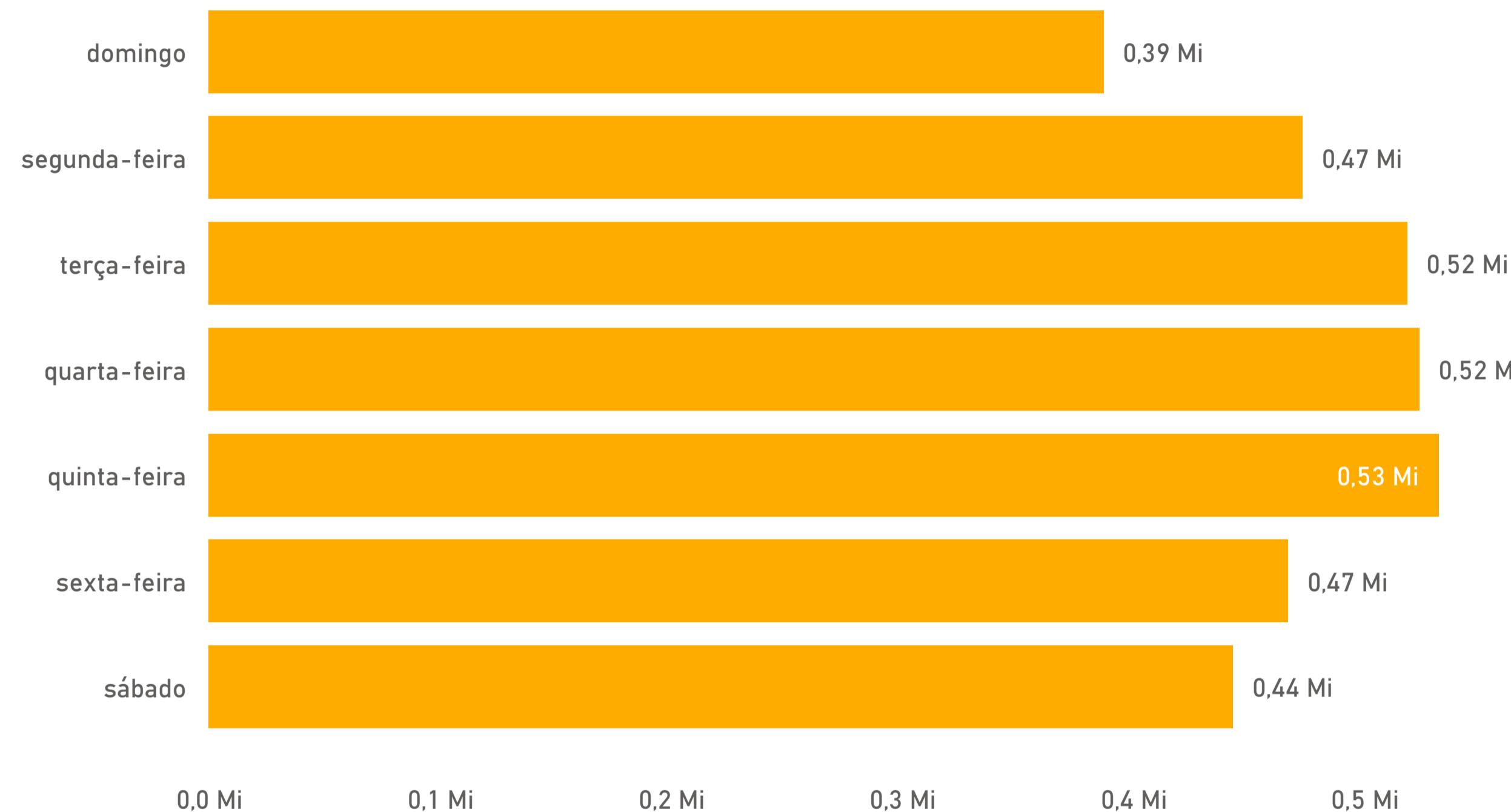
Separando os dados por "Tipo de Usuário", é possível notar o mesmo padrão onde **VERÃO** e **OUTONO** mantém o maior destaque no ano de 2022.

Contudo, em usuários **ASSINANTES** há um distribuição mais uniforme nas estações, mesmo que singela. Isso comprova a seguinte hipótese:

**ASSINANTES** usam mais as bicicletas independente da estação e isso não se da ao fato de uma proporcionalidade em relação ao quantitativo de usuários casuais, mas sim ao fator **NECESSIDADE**, como precisar usar a bicicleta para trabalhar e estudar até mesmo nos meses **MAIS FRIOS**.

# USO NOS DIAS DA SEMANA POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Usabilidade ao longo da semana (Por assinantes)



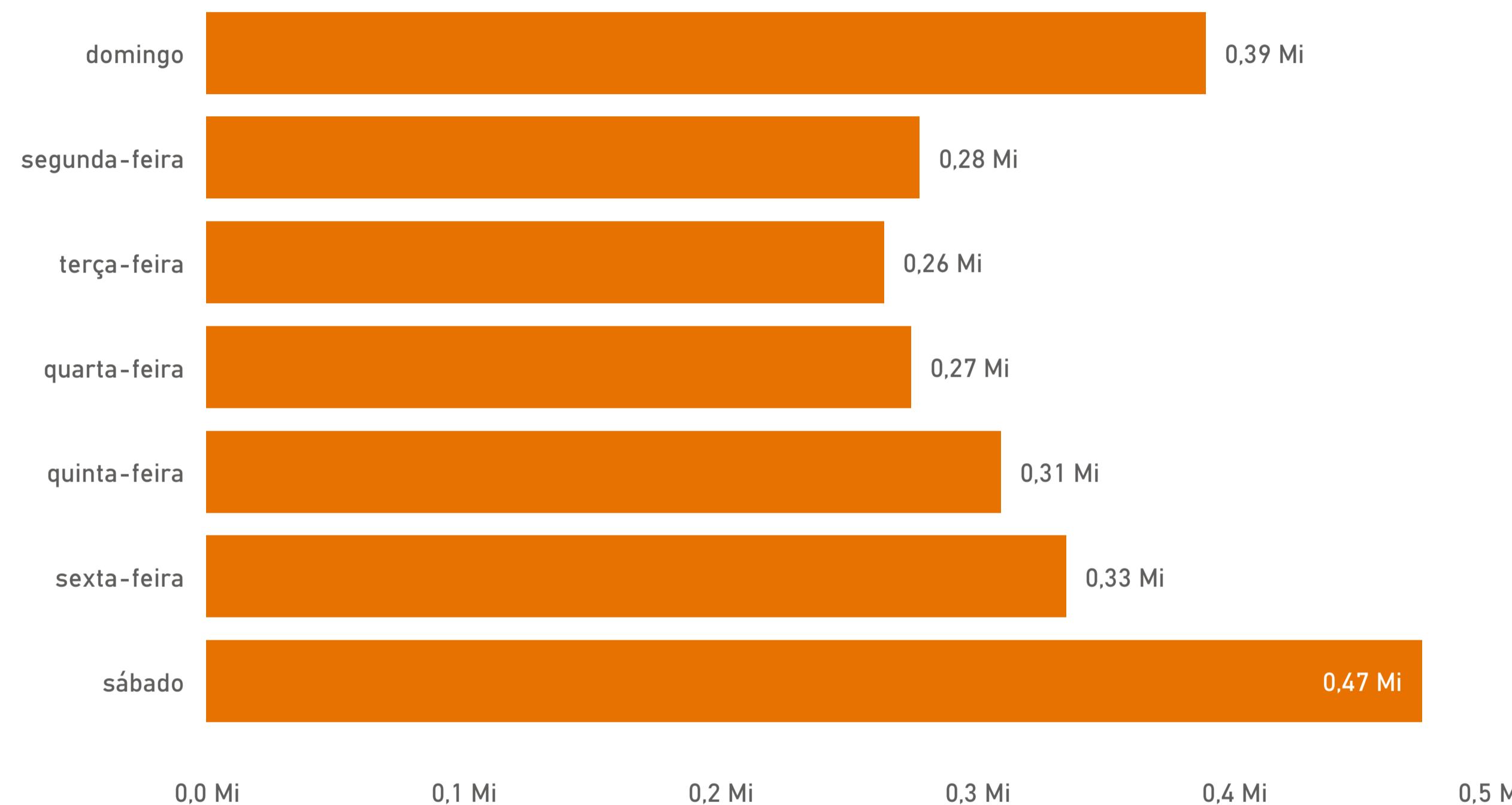
Quando se trata do uso das bicicletas ao longo da semana notamos uma clara diferença.

**ASSINANTES** usam com maior frequência durante os dias da semana, sendo o pico nas QUINTAS-FEIRAS, tendo os finais de semana com a menor usabilidade.

Isso provavelmente se dá ao motivo de que **assinantes usam mais por motivos profissionais** quando comparados aos usuários casuais.

# USO NOS DIAS DA SEMANA POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Usabilidade ao longo da semana (Por casuais)



Analizando os dados dos clientes **CASUAIS** notamos a diferença na usabilidade.

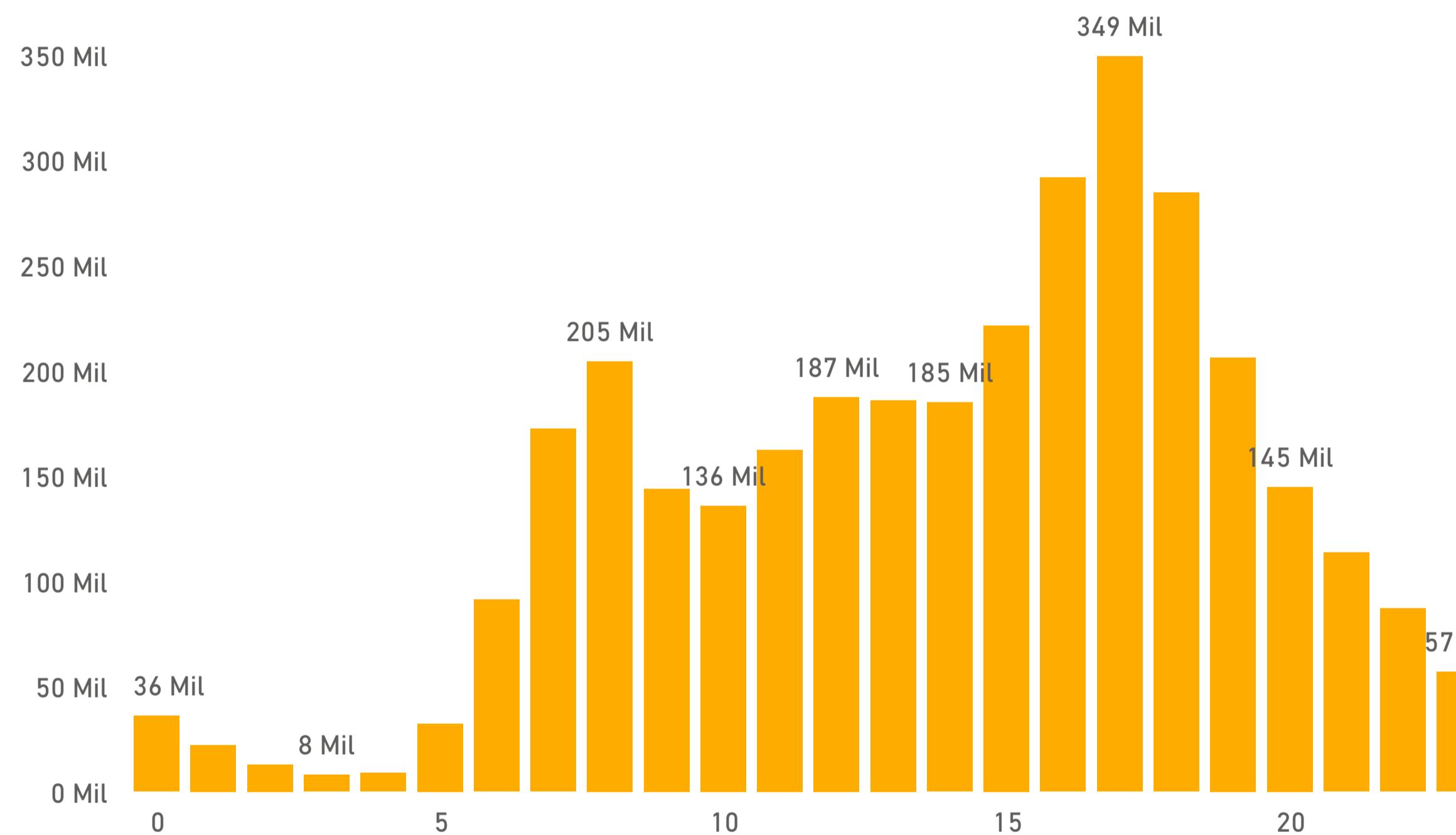
Enquanto **ASSINANTES** usam com maior frequência durante os dias da semana, os usuários **CASUAIS** usam muito mais aos finais de semana, tendo o maior pico aos **SÁBADOS**.

Esse fator clarifica que os usuários **não assinantes preferencialmente usam fora dos dias comerciais, provavelmente priorizando passeios e viagens.**

# HORÁRIOS MAIS MOVIMENTADOS POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Contagem de horas registradas em milhares (Por assinantes)

Horários das 00h00 até 23h00



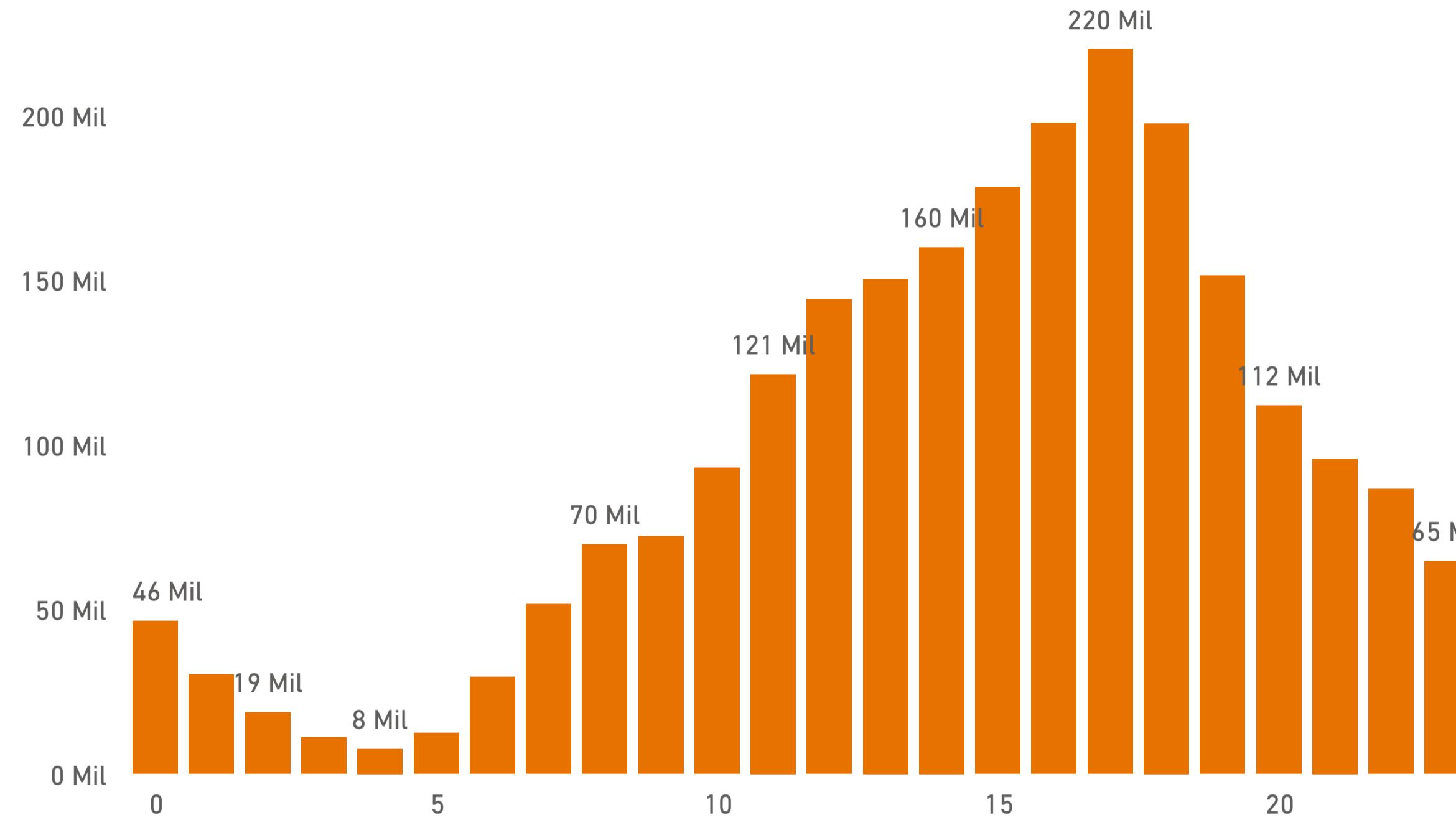
Uma informação relevante é como os clientes divergem quanto aos horários de utilização. **ASSINANTES** tendem a ter **picos de utilização** nos horários das **7h às 10h** e nos horários das **15h às 19h**.

Isso talvez se dá ao motivo de que **ASSINANTES** possuem a necessidade de usar as bicicletas para tarefas como **cursos, faculdades e serviços** nos horários comerciais ou até mesmo para realização de entregas em serviços como **iFood**, sendo interessante se possível, fazer uma **relação futura com dados do iFood** e nossos clientes para saber quantos trabalham como **ENTREGADORES**.

# HORÁRIOS MAIS MOVIMENTADOS POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Contagem de horas registradas em milhares (Por clientes casuais)

Horários das 00h00 até 23h00



Os clientes que usam casualmente nossas bicicletas usam de maneira constante **a partir do horário das 6h**, com elevação até o horário das 17h quando começam a diminuir o uso em horários noturnos.

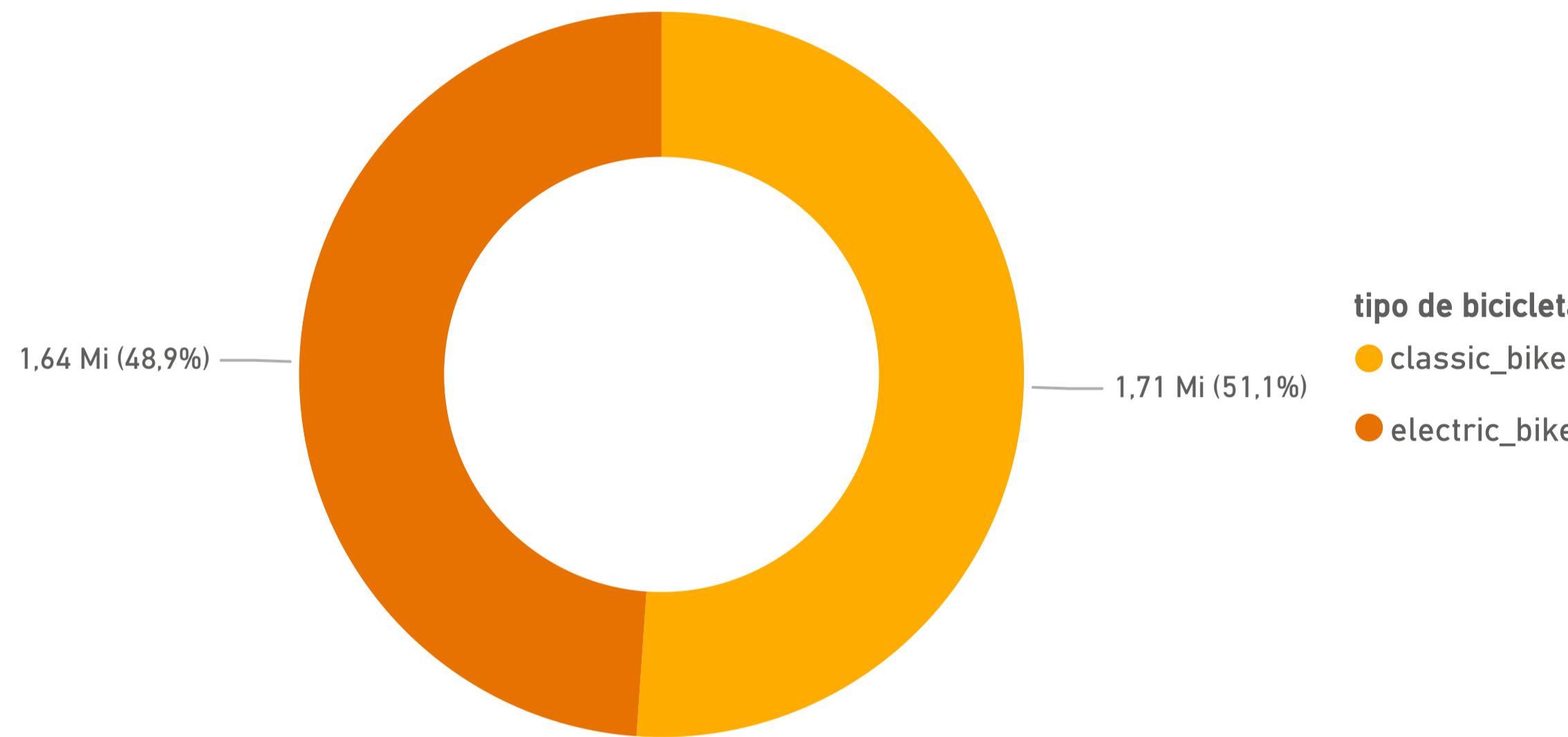
Como usuários casuais **não usam tanto pela manhã** como assinantes, as estratégias de marketing poderiam ser direcionadas para os horários a partir das **12h até as 18h**.

Como nesta análise não possuímos dados específicos dos clientes como idade e profissão e número de filhos ou dos locais que eles frequentam só podemos fazer presunções e partir deste gráfico.

Talvez usem esporadicamente para **passeios com familiares e amigos ou para frequentar parques e academias**.

# ESCOLHA DE MODELO DE BICICLETA POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Uso de cada modelo de bicicleta (Por clientes assinantes)



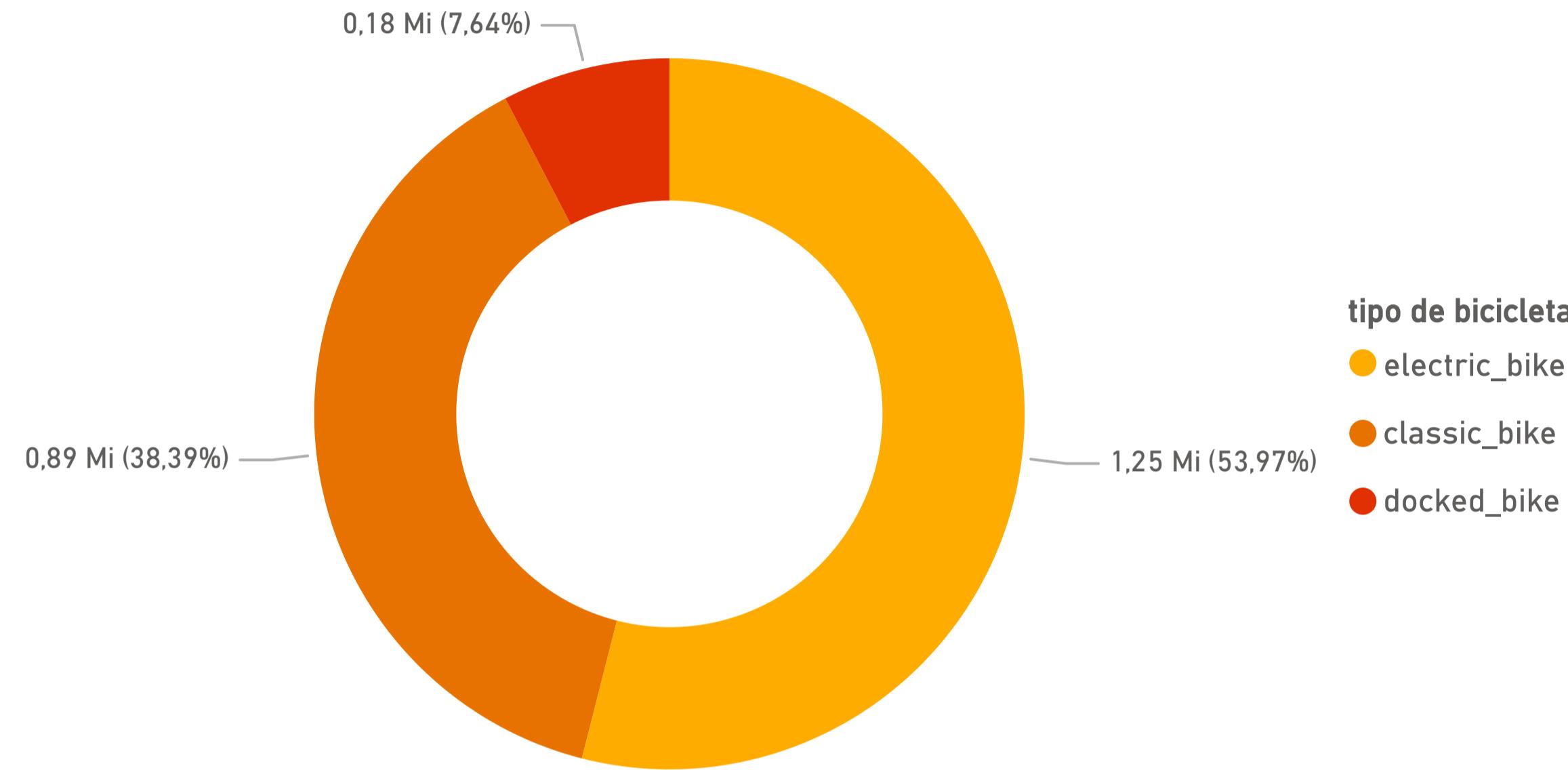
Cientes **ASSINANTES** usam de maneira bem distribuída os modelos de bicicletas clássicas e elétricas, tendo uma leve preferência pelo modelo **CLÁSSICO**.

É possível notar que não há a categoria de **BICICLETAS EM DOCAS** quando se trata dos **ASSINANTES**, tendo em vista uma **vantagem do plano de assinatura o fato de não ter que se preocupar em ter que encontrar uma doca para deixar a bike**.

Esse é um fator relevante para nossa análise, tendo em vista que a categoria de assinantes prioriza apenas o modelo da bicicleta, **quanto mais clientes assinantes, menor é a quantidade de docas espalhadas pelo Estado, sendo uma vantagem financeira** (mas levando em consideração que há a necessidade de um polos para a coleta de bicicletas).

# ESCOLHA DE MODELO DE BICICLETA POR TIPO DE CLIENTE DO ANO DE 2022 EM CHICAGO

Uso de cada modelo de bicicleta (Por clientes casuais)



Cientes **CASUAIS** usam as bicicletas de maneira mais dispersa, sendo usado preferencialmente as bicicletas **ELÉTRICAS**.

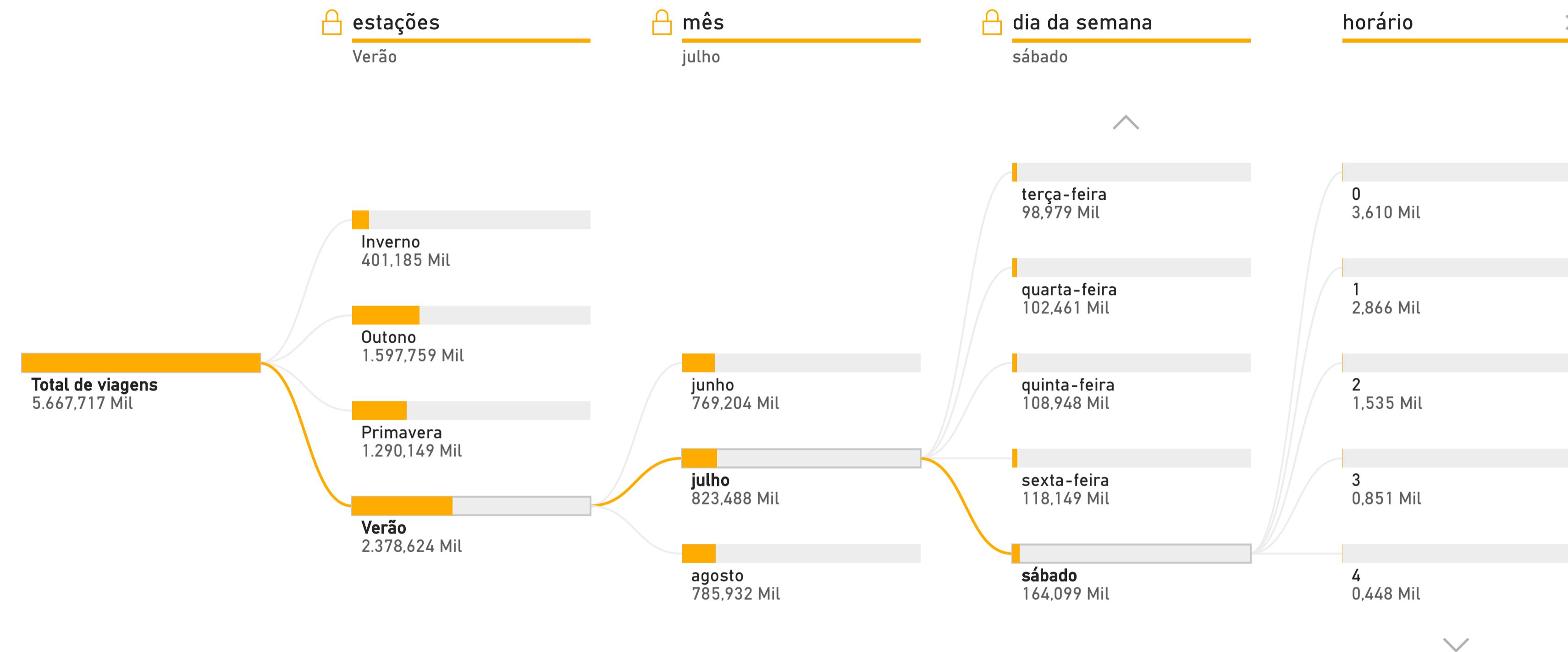
Esse aumento do uso das bicicletas elétricas é um fator **extremamente relevante**, talvez pela curiosidade dos clientes novos em conhecer esse modelo naturalmente mais caro do que modelos clássicos e também para a empresa se preparar para futuros projetos com a equipe de marketing para adquirir mais bicicletas elétricas e promover seu uso.

Também é possível notar a presença da categoria de **BICICLETAS EM DOCAS**, haja visto a necessidade de coleta e entrega de bicicletas em horários específicos.

Como os usuários casuais divergem entre **passes de viagem única** e **passes de viagens de dia inteiro** há a grande necessidade de ter diversas docas espalhadas, tanto pela conveniência do cliente (que caso perca o horário de entrega deve pagar o valor de acréscimo) e pela necessidade da empresa para ter controle das bicicletas.

# PERÍODO DE MAIOR MOVIMENTAÇÃO DE VIAGENS DO ANO DE 2022 EM CHICAGO

Estação, mês, dia da semana e horário mais movimentado



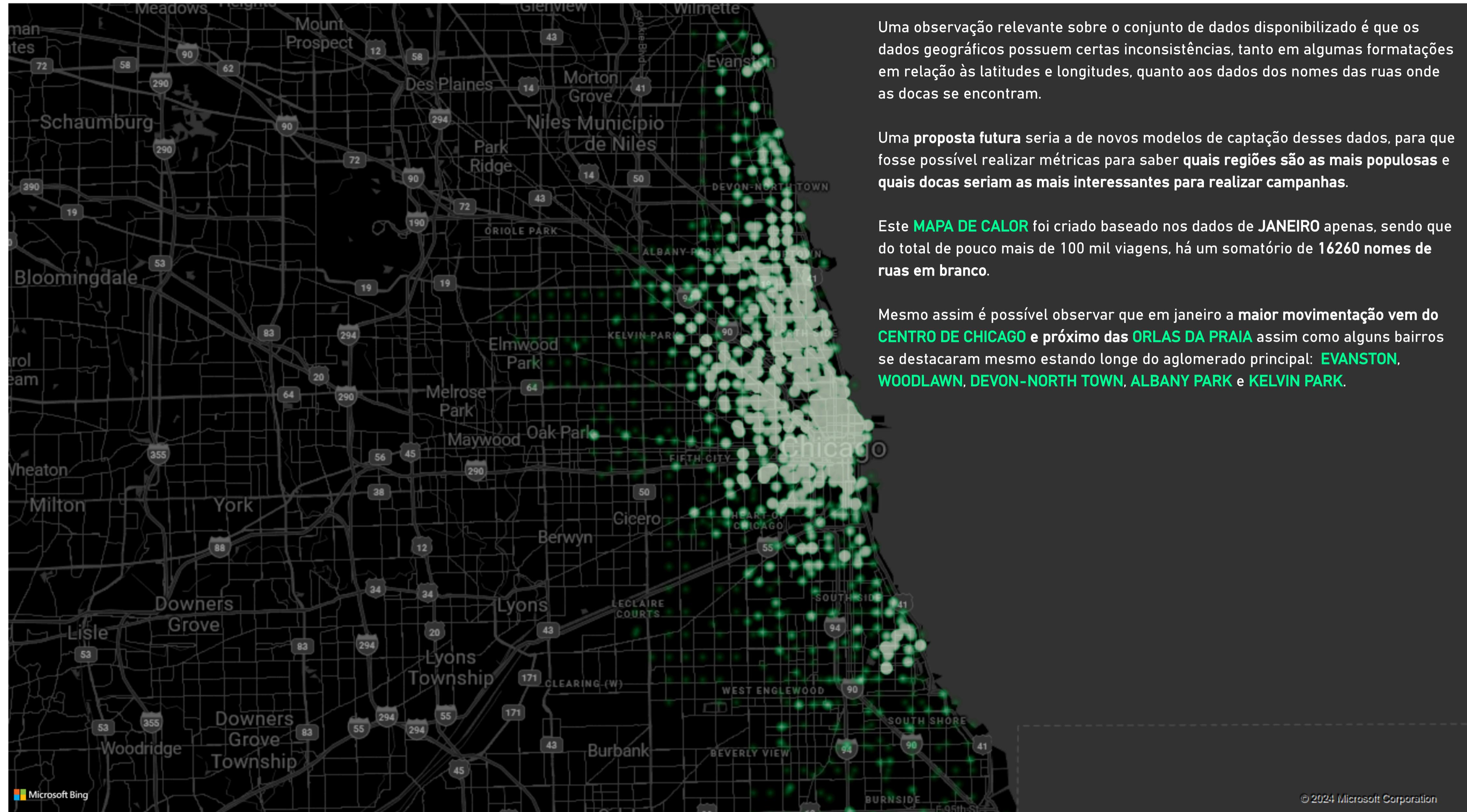
Em uma visão panorâmica, podemos observar que o período do **VERÃO** corresponde a **mais de 40% da viagens**.

O mês de **JULHO** vence dentre os meses quentes de verão, tendo **quase 15% da viagens totais**.

Os **SÁBADOS** são os dias vencedores, com **mais de 150 MIL VIAGENS** cada.

E o **horário que se destaca** é o período da tarde, por volta das **15H** tendo uma contagem de **quase 13 MIL VIAGENS** registradas nesse período.

## Contagem de viagens por doca inicial



## Questões de Negócio

Como assinantes e usuários casuais **divergem** em relação ao uso das bicicletas?

Porque usuários casuais **pagariam** por uma assinatura?

Como a empresa pode usar mídia social para **influenciar** usuários casuais a se tornarem assinantes?

# Como assinantes e usuários casuais divergem em relação ao uso das bicicletas?

Levando em considerações os insights levantados ao longo dos gráficos, é possível notar alguns pontos diferenciais entre os clientes:

**ASSINANTES** usam mais as bicicletas independente da estação e isso não se da ao fato de uma proporcionalidade em relação ao quantitativo de usuários casuais, mas sim ao fator **NECESSIDADE**, como precisar usar a bicicleta para trabalhar e estudar até mesmo nos meses **MAIS FRIOS**.

**ASSINANTES** usam com maior frequência **durante os dias da semana**, tendo os finais de semana com a menor usabilidade. Isso provavelmente se dá ao motivo de que **assinantes usam mais por motivos profissionais** quando comparados aos usuários casuais.

Enquanto **ASSINANTES** usam com maior frequência durante os dias da semana, os usuários **CASUAIS** usam muito mais aos finais de semana, tendo o maior pico aos **SÁBADOS**. Esse fator clarifica que os usuários **não assinantes preferencialmente usam fora dos dias comerciais, provavelmente priorizando passeios e viagens**.

Uma informação relevante é como os clientes divergem quanto aos horários de utilização. **ASSINANTES** tendem a ter **picos de utilização** nos horários das **7h às 10h** e nos horários das **15h às 19h**. Isso talvez se dá ao motivo de que **ASSINANTES** possuem a necessidade de usar as bicicletas para tarefas como **cursos, faculdades e serviços** nos horários comerciais ou até mesmo para realização de entregas em serviços como **iFood**, sendo interessante se possível, fazer uma **relação futura com dados do iFood** e nossos clientes para saber quantos trabalham como **ENTREGADORES**.

Os clientes que usam casualmente nossas bicicletas usam de maneira constante a partir do horário das **6h**, com elevação até o horário das **17h** quando começam a diminuir o uso em horários noturnos. Como usuários casuais **não usam tanto pela manhã** como assinantes, as estratégias de marketing poderiam ser direcionadas para os horários a partir das **12h até as 18h**. Como nesta análise não possuímos dados específicos dos clientes como idade e profissão e número de filhos ou dos locais que eles frequentam só podemos fazer presunções. Talvez usem esporadicamente para **passeios com familiares e amigos ou para frequentar parques e academias**.

Clientes **ASSINANTES** usam de maneira bem distribuída os modelos de bicicletas clássicas e elétricas, tendo uma leve preferência pelo modelo **CLÁSSICO**. É possível notar que não há a categoria de **BICICLETAS EM DOCAS** quando se trata dos **ASSINANTES**, tendo em vista uma **vantagem do plano de assinatura o fato de não ter que se preocupar em ter que encontrar uma doca para deixar a bike**. Esse é um fator relevante para nossa análise, tendo em vista que a categoria de assinantes prioriza apenas o modelo da bicicleta, **quanto mais clientes assinantes, menor é a quantidade de docas espalhadas pelo Estado, sendo uma vantagem financeira** (mas levando em consideração que há a necessidade de um polos para a coleta de bicicletas).

Clientes **CASUAIS** usam as bicicletas de maneira mais dispersa, sendo usado preferencialmente as bicicletas **ELÉTRICAS**. Esse aumento do uso das bicicletas elétricas é um fator **extremamente relevante**, talvez pela curiosidade dos clientes novos em conhecer esse modelo naturalmente mais caro do que modelos clássicos e também para a empresa se preparar para futuros projetos com a equipe de marketing para adquirir mais bicicletas elétricas e promover seu uso. Também é possível notar a presença da categoria de **BICICLETAS EM DOCAS**, haja visto a necessidade de coleta e entrega de bicicletas em horários específicos. Como os usuários casuais divergem entre **passes de viagem única e passes de viagens de dia inteiro** há a grande necessidade de ter diversas docas espalhadas, tanto pela conveniência do cliente (que caso perca o horário de entrega deve pagar o valor de acréscimo) e pela necessidade da empresa para ter controle das bicicletas.

Em uma visão panorâmica, podemos observar que o período do **VERÃO** corresponde a **mais de 40% da viagens**. O mês de **JULHO** vence dentre os meses quentes de verão, tendo **quase 15% da viagens totais**. Os **SÁBADOS** são os dias vencedores, com **mais de 150 MIL VIAGENS** cada. E o horário que se destaca é o período da tarde, por volta das **15H** tendo uma contagem de **quase 13 MIL VIAGENS** registradas nesse período.

# Porque usuários casuais pagariam por uma assinatura?

Antes de responder a esta pergunta devo salientar algumas sugestões quanto a coleta de dados e possíveis melhorias futuras para que fosse possível melhores insights e respostas mais precisas:

Sobre os clientes, seria interessante ter dados como **faixa etária, profissão e quantos filhos têm** para sabermos como são nossos clientes, quantas crianças utilizam as bicicletas para brincar em parques e quantos adultos trabalham como entregadores por exemplo.

A respeito dos próprios planos de assinaturas, poderíamos ter os **dados de preço entre os planos** para que fosse possível ter uma noção de métricas de quanto é movimentado, assim como dados de gasto para termos em média uma perspectiva de ganho e possíveis perdas em meses de menores movimentações, como JANEIRO, FEVEREIRO e DEZEMBRO.

Com base nos dados, poderíamos ter uma mudança na aquisição dos dados, tendo em vista que 127.793 viagens contabilizaram “0” minutos de viagem e na coluna *end\_station\_name* há um total de 892.742 linhas vazias e para observações mais precisas seria apropriado ter uma quantidade menor de inconsistências.

Concluindo com a pergunta de negócio: **Porque um usuário casual pagaria uma assinatura anual?**

Para que ele obtenha uma assinatura anual ele precisa de sentir que **será vantajoso essa mudança de planos**.

Pelos gráficos foi possível observar que os usuários casuais preferem as bicicletas **ELÉTRICAS**, talvez pela curiosidade de utilizar uma ou pelo conforto disponibilizado por elas, sendo um ponto interessante para a **equipe de marketing investir em campanhas destacando novos modelos para USUÁRIOS ANUAIS** (e também para as equipes do financeiro e operacional se preparesm antecipadamente);

Usuários casuais preferem usar aos **FINAIS DE SEMANA**, se soubéssemos quantos filhos cada cliente têm, poderíamos oferecer pacotes de família, que incentivasse a prática esportiva assim como o aumento das vendas do plano anual.

E como o **VERÃO** é o período mais movimentado, toda o preparo de marketing poderia ser realizado poucos antes dessa Estação.

# Como a empresa pode usar mídia social para influenciar usuários casuais a se tornarem assinantes?

Tendo em mente as preferências dos usuários casuais a equipe de marketing pode realizar as seguintes campanhas:

Campanhas destacando o uso de bicicletas **ELÉTRICAS** e suas vantagens;

Campanhas **influenciando o uso das bicicletas para atividades físicas em parques e passeios com a família**, já sabendo que estes usuários tendem a usar mais aos fins de semana, ficando implícito que aos sábados e domingos são dias preferidos para passeios como esses.

Durante as estações de **VERÃO** e **OUTONO** realizar parcerias com celebridades para cativar novos clientes a se tornarem assinantes, sabendo que essas duas estações são as mais movimentadas.

Sabendo que usuários casuais usam de forma exponencial a partir das **6h da manhã com o pico às 17h** seria interessante saber quais eventos e atividades em Chicago geralmente acontece nesse horário matinal e vespertino para que fosse mais uma estratégia de marketing.

# **FINALIZAÇÃO DA APRESENTAÇÃO**

E validação do Certificado Google Data Analytics Certificate

Saliento que os dados utilizados na apresentação estarão em meu repositório no GitHub caso você não esteja visualizando pelo PowerBi.

De antemão agradeço e estarei disponível no LinkedIn e GitHub para contato.

Para validação do meu certificado clique na imagem a seguir:



Ou acesse o link: <https://www.credly.com/badges/5d5a939c-c674-462d-be30-ac194c9f262b>