

Relatório de análise estatística - Old Town Road LTDA

Consultores Responsáveis:

Estatiano 1
Estatiano 2
Estatiano 3

Requerente:
ESTAT

Brasília, 9 de novembro de 2025.



Sumário

	Página
1 Introdução	3
2 Metodologia	4
2.1 Frequência Relativa	4
2.2 Média	4
2.3 Mediana	5
2.4 Quartis	5
2.5 Variância	5
2.5.1 Variância Amostral	6
2.6 Desvio Padrão	6
2.6.1 Desvio Padrão Amostral	6
2.7 Boxplot	6
2.8 Gráfico de Dispersão	7
2.9 Tipos de Variáveis	8
2.9.1 Qualitativas	8
2.9.2 Quantitativas	8
2.10 Coeficiente de Correlação de Pearson	9
2.11 Teste de Correlação de Pearson	9
3 Análises	11
3.1 Receita média das lojas registrada nos anos de 1880 até 1889	11
3.1.1 1880 a 1885	12
3.1.2 1885 a 1889	12
3.1.3 Todo o período	12
3.2 Variação de peso por altura dos clientes.	13
3.2.1 Análises relacionadas à altura	13
3.2.2 Análises relacionadas ao peso	14
3.2.3 Relação entre as variáveis	15
3.3 Idade dos clientes de Âmbar Seco a depender da loja.	15
3.3.1 Idade de todos clientes da amostra	16
3.3.2 Idade dos clientes de Âmbar Seco	17
3.3.3 Idade dos clientes de Âmbar seco a partir da loja	18
3.4 O top 3 produtos mais vendidos nas top 3 lojas com maior receita em 1889.	19
3.4.1 As três lojas com maior receita no ano de 1889.	20
3.4.2 Os três itens mais vendidos nas três lojas com maior receita em 1889.	20
4 Conclusão	22

1 Introdução

O seguinte relatório estatístico foi elaborado a pedido de João Sábio, proprietário da Old Town Road Ltda., uma holding que administra diversas empresas voltadas ao comércio no faroeste. O cliente demonstrou interesse em expandir seus investimentos para uma nova região e, para elaborar suas decisões, solicitou a realização de análises estatísticas descritivas e inferenciais que permitam compreender melhor o comportamento do mercado local, o perfil dos clientes e o desempenho comercial das lojas dessa área.

As análises foram desenvolvidas a partir de um banco de dados disponibilizado e coletado pelo próprio cliente, contendo informações sobre receitas anuais específicas, características físicas dos consumidores, distribuição etária e volume de vendas de produtos.

Todo o processamento e geração dos resultados foram realizados no software R, utilizando a versão mais recente disponível no momento da elaboração, por meio da ferramenta Quarto para integração entre código, texto e formatação automática dos resultados. O nível de significância adotado foi de 5% ($\alpha = 0,05$) para todas as inferências estatísticas realizadas.

Atendendo às solicitações do cliente, o relatório está dividido em quatro análises principais: Receita média das lojas registrada nos anos de 1880 a 1889, Variação de peso por altura dos clientes, Idade dos clientes de Âmbar Seco a depender da loja e Os três produtos mais vendidos nas três lojas com maior receita em 1889.

2 Metodologia

Este relatório é composto por técnicas estatísticas que serão descritas a seguir de acordo com o que foi utilizado em tal estudo:

2.1 Frequência Relativa

A frequência relativa é utilizada para a comparação entre classes de uma variável categórica com c categorias, ou para comparar uma mesma categoria em diferentes estudos.

A frequência relativa da categoria j é dada por:

$$f_j = \frac{n_j}{n}$$

Com:

- $j = 1, \dots, c$
- n_j = número de observações da categoria j
- n = número total de observações

Geralmente, a frequência relativa é utilizada em porcentagem, dada por:

$$100 \times f_j$$

2.2 Média

A média é a soma das observações dividida pelo número total delas, dada pela fórmula:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Com:

- $i = 1, 2, \dots, n$
- n = número total de observações

2.3 Mediana

Sejam as n observações de um conjunto de dados $X = X_{(1)}, X_{(2)}, \dots, X_{(n)}$ de determinada variável ordenadas de forma crescente. A mediana do conjunto de dados X é o valor que deixa metade das observações abaixo dela e metade dos dados acima.

Com isso, pode-se calcular a mediana da seguinte forma:

$$med(X) = \begin{cases} X_{\frac{n+1}{2}}, & \text{para } n \text{ ímpar} \\ \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}, & \text{para } n \text{ par} \end{cases}$$

2.4 Quartis

Os quartis são separatrizes que dividem o conjunto de dados em quatro partes iguais. O primeiro quartil (ou inferior) delimita os 25% menores valores, o segundo representa a mediana, e o terceiro delimita os 25% maiores valores. Inicialmente deve-se calcular a posição do quartil:

- Posição do primeiro quartil P_1 :

$$P_1 = \frac{n+1}{4}$$

- Posição da mediana (segundo quartil) P_2 :

$$P_2 = \frac{n+1}{2}$$

- Posição do terceiro quartil P_3 :

$$P_3 = \frac{3 \times (n+1)}{4}$$

Com n sendo o tamanho da amostra. Dessa forma, $X_{(P_i)}$ é o valor do i -ésimo quartil, onde $X_{(j)}$ representa a j -ésima observação dos dados ordenados.

Se o cálculo da posição resultar em uma fração, deve-se fazer a média entre o valor que está na posição do inteiro anterior e do seguinte ao da posição.

2.5 Variância

A variância é uma medida que avalia o quanto os dados estão dispersos em relação à média, em uma escala ao quadrado da escala dos dados.

2.5.1 Variância Amostral

Para uma amostra, a variância é dada por:

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

Com:

- X_i = i-ésima observação da amostra
- \bar{X} = média amostral
- n = tamanho da amostra

2.6 Desvio Padrão

O desvio padrão é a raiz quadrada da variância. Ele avalia o quanto os dados estão dispersos em relação à média.

2.6.1 Desvio Padrão Amostral

Para uma amostra, o desvio padrão é dado por:

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

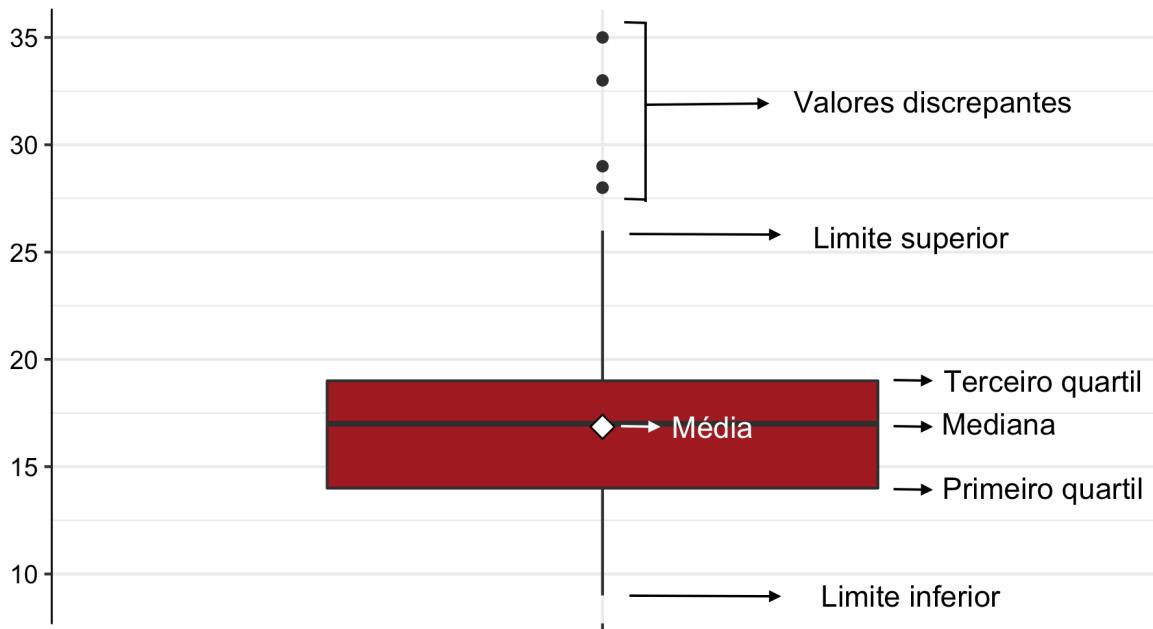
Com:

- X_i = i-ésima observação da amostra
- \bar{X} = média amostral
- n = tamanho da amostra

2.7 Boxplot

O boxplot é uma representação gráfica na qual se pode perceber de forma mais clara como os dados estão distribuídos. A figura abaixo ilustra um exemplo de boxplot.

Figura 1: Exemplo de boxplot

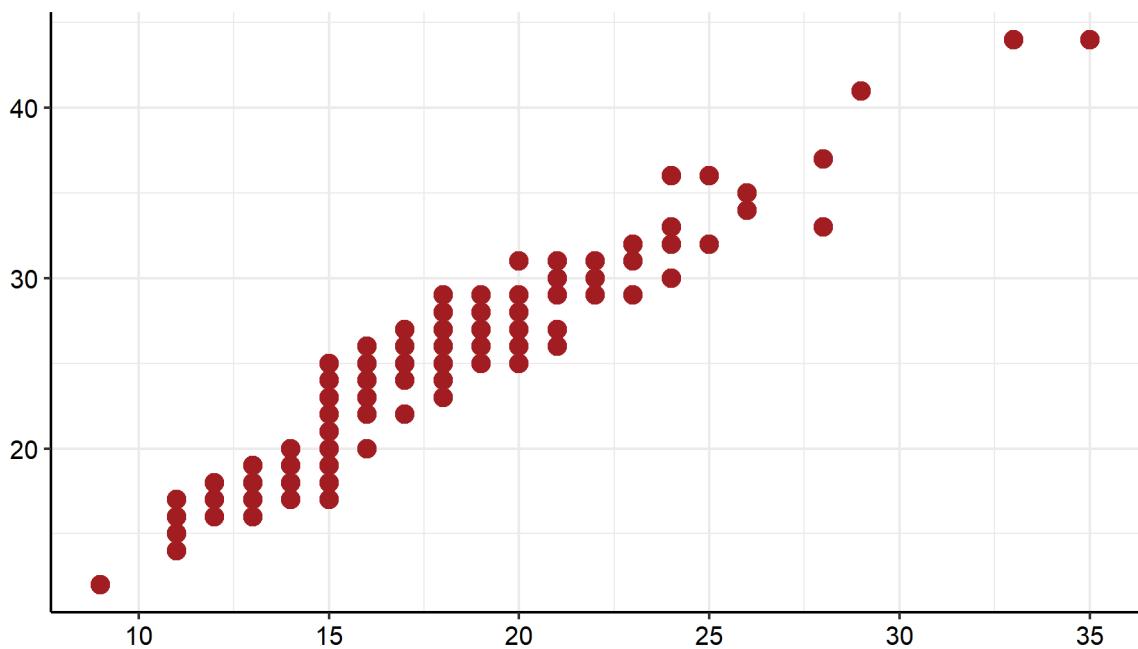


A porção inferior do retângulo diz respeito ao primeiro quartil, enquanto a superior indica o terceiro quartil. Já o traço no interior do retângulo representa a mediana do conjunto de dados, ou seja, o valor em que o conjunto de dados é dividido em dois subconjuntos de mesmo tamanho. A média é representada pelo losango branco e os pontos são *outliers*. Os *outliers* são valores discrepantes da série de dados, ou seja, valores que não demonstram a realidade de um conjunto de dados.

2.8 Gráfico de Dispersão

O gráfico de dispersão é uma representação gráfica utilizada para ilustrar o comportamento conjunto de duas variáveis quantitativas. A figura abaixo ilustra um exemplo de gráfico de dispersão, onde cada ponto representa uma observação do banco de dados.

Figura 2: Exemplo de Gráfico de Dispersão



2.9 Tipos de Variáveis

2.9.1 Qualitativas

As variáveis qualitativas são as variáveis não numéricas, que representam categorias ou características da população. Estas subdividem-se em:

- **Nominais:** quando não existe uma ordem entre as categorias da variável (exemplos: sexo, cor dos olhos, fumante ou não, etc)
- **Ordinais:** quando existe uma ordem entre as categorias da variável (exemplos: nível de escolaridade, mês, estágio de doença, etc)

2.9.2 Quantitativas

As variáveis quantitativas são as variáveis numéricas, que representam características numéricas da população, ou seja, quantidades. Estas subdividem-se em:

- **Discretas:** quando os possíveis valores são enumeráveis (exemplos: número de filhos, número de cigarros fumados, etc)
- **Contínuas:** quando os possíveis valores são resultado de medições (exemplos: massa, altura, tempo, etc)

2.10 Coeficiente de Correlação de Pearson

O coeficiente de correlação de Pearson é uma medida que verifica o grau de relação linear entre duas variáveis quantitativas. Este coeficiente varia entre os valores -1 e 1. O valor zero significa que não há relação linear entre as variáveis. Quando o valor do coeficiente r é negativo, diz-se existir uma relação de grandeza inversamente proporcional entre as variáveis. Analogamente, quando r é positivo, diz-se que as duas variáveis são diretamente proporcionais.

O coeficiente de correlação de Pearson é normalmente representado pela letra r e a sua fórmula de cálculo é:

$$r_{Pearson} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \times \sqrt{\sum_{i=1}^n y_i^2 - n\bar{y}^2}}$$

Onde:

- x_i = i-ésimo valor da variável X
- y_i = i-ésimo valor da variável Y
- \bar{x} = média dos valores da variável X
- \bar{y} = média dos valores da variável Y

Vale ressaltar que o coeficiente de Pearson é paramétrico e, portanto, sensível quanto à normalidade (simetria) dos dados.

2.11 Teste de Correlação de Pearson

O coeficiente de correlação linear de Pearson indica a força e a direção do relacionamento linear entre duas variáveis quantitativas. É um índice adimensional com valores situados entre -1 e 1, no qual o valor -1 representa total correlação linear negativa entre as variáveis (quando o valor de uma variável cresce, o valor da outra diminui) e o valor 1 representa total correlação linear positiva entre elas (ambas crescem simultaneamente). Esse coeficiente é obtido por meio da fórmula:

$$r_{Pearson} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \sqrt{\sum_{i=1}^n y_i^2 - n\bar{y}^2}}$$

em que

- x_i = i-ésimo valor da variável X
- y_i = i-ésimo valor da variável Y
- \bar{x} = média dos valores da variável X
- \bar{y} = média dos valores da variável Y
- $r_{Pearson}$ = coeficiente de correlação linear de Pearson amostral

Para o teste de correlação de Pearson, tem-se as seguintes hipóteses:

$$\begin{cases} H_0 : \text{Não há correlação linear entre as variáveis } X \text{ e } Y \\ \quad (\rho_{Pearson} = 0) \\ H_1 : \text{Há correlação linear entre as variáveis } X \text{ e } Y \\ \quad (\rho_{Pearson} \neq 0) \end{cases}$$

em que $\rho_{Pearson}$ é o parâmetro a ser testado: coeficiente de correlação linear populacional.

Se X e Y tem distribuição normal, tem-se que a estatística do teste é dada por:

$$t_{Pearson} = \frac{r_{Pearson} \sqrt{n-2}}{\sqrt{1 - r_{Pearson}^2}} \sim t_{n-2}$$

Assim, sob H_0 , $t_{Pearson}$ segue uma distribuição t -Student com $(n - 2)$ graus de liberdade.

3 Análises

3.1 Receita média das lojas registrada nos anos de 1880 até 1889

Nessa análise, será observada, por ano, a média do faturamento bruto anual de cada uma das lojas descritas no plano amostral, no período de 1880 a 1889, com o objetivo de compreender se houve aumento ou diminuição da receita média ao longo desses anos. Serão analisadas duas variáveis: o ano das vendas, uma variável quantitativa discreta, e a receita média das lojas, que é uma variável quantitativa contínua, expressa em reais. Os valores consideram a cotação atual do dólar a R\$ 5,31 para referência.

Figura 3: Gráfico de linhas da receita média anual de 1880 a 1889

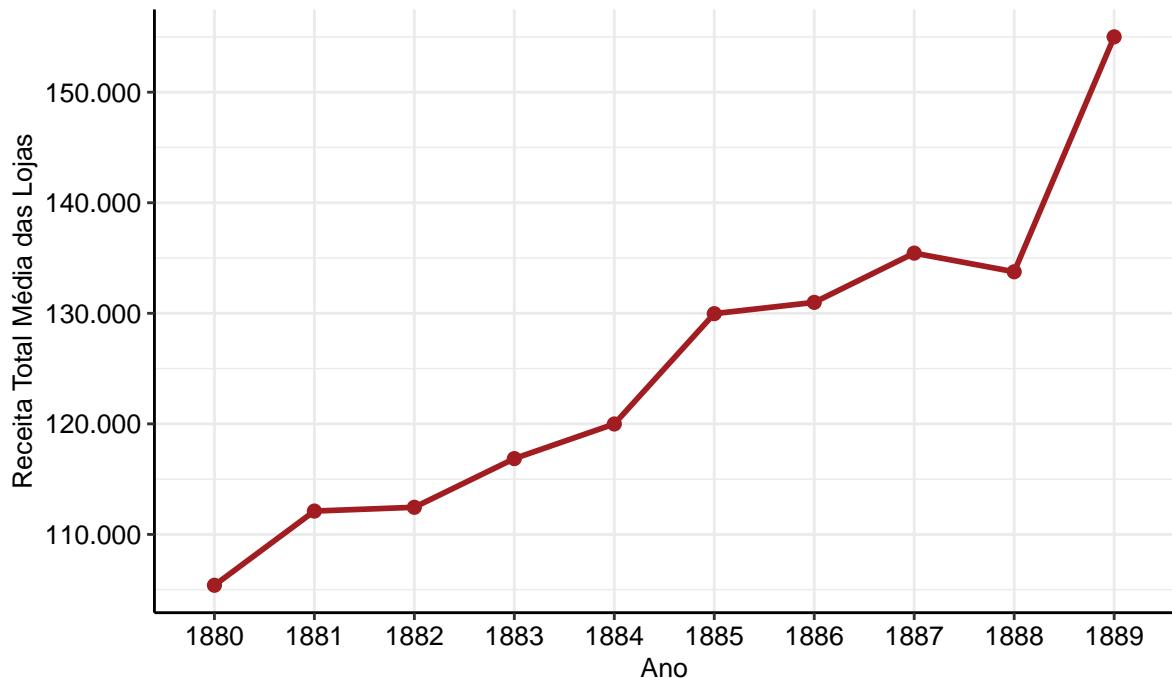


Tabela 1: Média das receitas e variação percentual entre 1880 e 1889

Anos	Média das receitas	Porcentagem de aumento
1880	105.399,0	-
1881	112.110,0	6,37%
1882	112.452,4	0,31%
1883	116.856,9	3,92%
1884	119.989,8	2,68%
1885	129.969,0	8,32%
1886	130.989,2	0,78%
1887	135.444,8	3,40%
1888	133.757,6	-1,25%
1889	155.009,1	15,89%
Total	1.252.977,8	47,07%

De modo geral, ao analisar a tabela acima e a **Figura 3**, é perceptível uma tendência de crescimento contínuo, com algumas flutuações pontuais que merecem destaque. A série indica que, ao final do período, as receitas atingiram o maior valor da década, refletindo um comportamento fortemente ascendente.

3.1.1 1880 a 1885

Nos primeiros 4 anos, o aumento ocorreu de forma gradual, com o acréscimo médio de aproximadamente 3,3% ao ano, o que sugere uma fase inicial estável com crescimento moderado. De 1884 a 1885, verifica-se uma aceleração mais evidente no ritmo de crescimento, quando a receita média chega próxima da faixa de 130 mil reais.

3.1.2 1885 a 1889

O período de 1885 a 1888 apresenta pequenas oscilações, com sutil crescimento até 1887, e uma ligeira redução em 1888. Essa pequena queda, de cerca de 1,25%, não altera a trajetória geral da série, podendo indicar uma correção natural após aumentos sucessivos. Em seguida, o ano de 1889 evidencia o maior salto positivo de todo período, com crescimento estimado em mais de 15% em relação ao ano anterior.

3.1.3 Todo o período

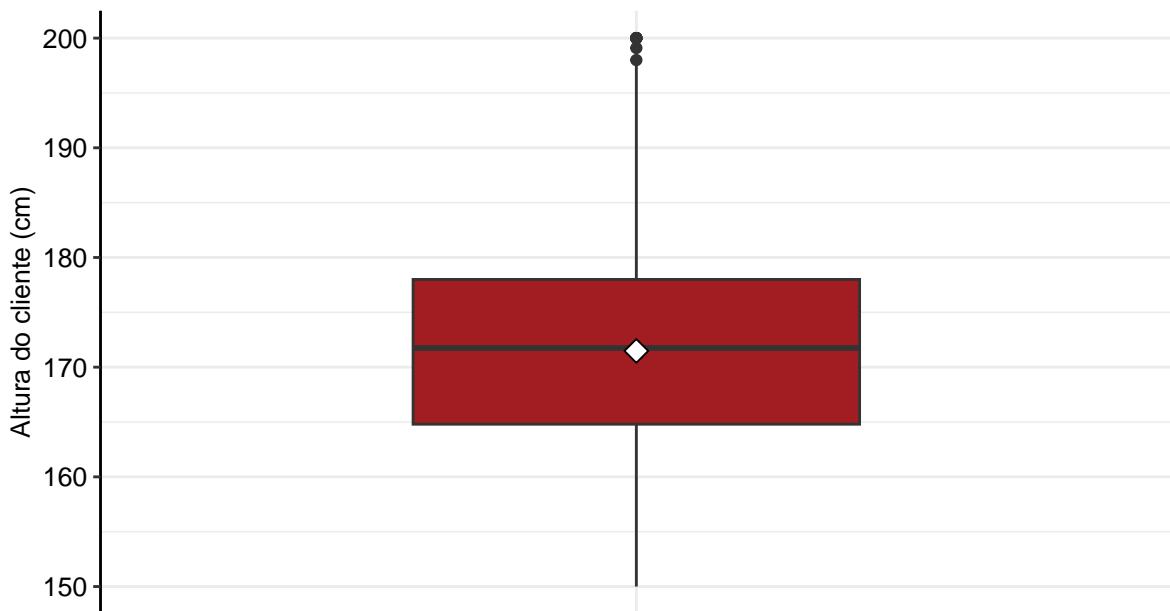
Ao analisar o comportamento geral, constata-se que o crescimento entre 1880 e 1889 foi de aproximadamente 47%, com uma média de crescimento anual próxima de 5%. A partir disso, constata-se que, mesmo diante de pequenas variações, houve uma tendência de expansão ao longo da década.

3.2 Variação de peso por altura dos clientes.

Nessa análise será observada a relação entre duas variáveis fundamentais em estudos corporais: peso (em quilogramas) e altura (em centímetros) dos clientes. A análise busca compreender de que forma essas variáveis se comportam em conjunto, verificando se há uma tendência de aumento do peso conforme cresce a altura, se ocorre o oposto ou se não existe uma relação significativa entre elas. Com base nos dados coletados das duas variáveis quantitativas contínuas, será aplicada uma análise estatística descritiva, permitindo identificar padrões, correlações e possíveis tendências que auxiliem na compreensão do perfil físico dos clientes.

3.2.1 Análises relacionadas à altura

Figura 4: Boxplot da altura dos clientes.



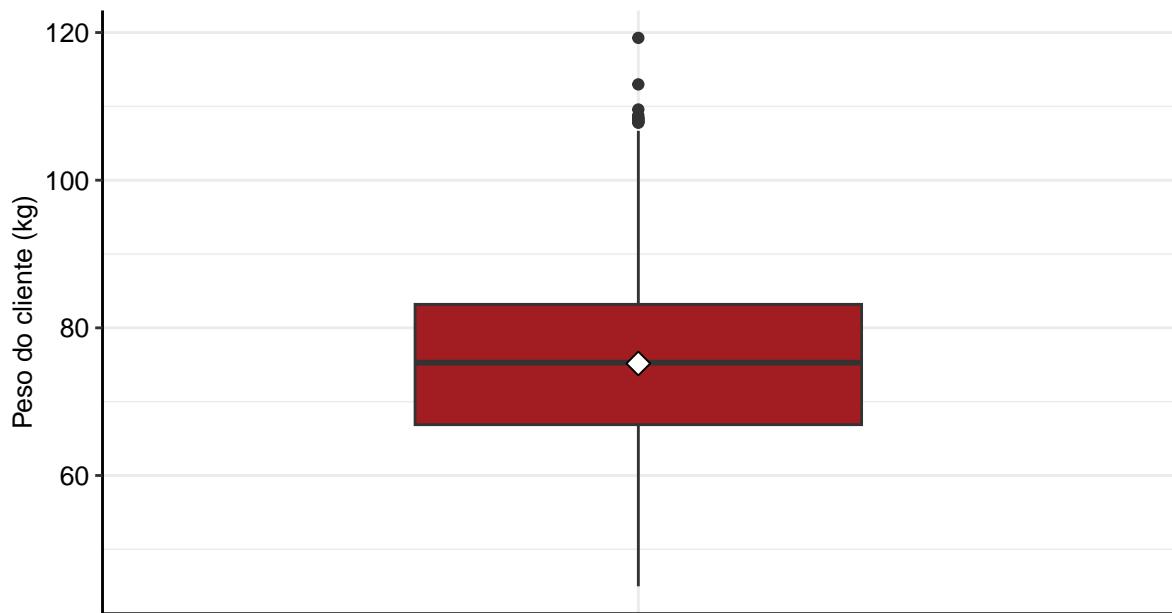
Quadro 1: Medidas resumo da variável

Estatística	Valor
Media	171,48
Desvio Padrao	9,87
Variância	97,38
Mínimo	150,00
1º Quartil	164,80
Mediana	171,75
3º Quartil	178,00
Máximo	200,00

Em termos gerais, de acordo com o que pode ser percebido na tabela acima e na **Figura 4**, a altura dos clientes demonstra-se com uma variação moderada entre os indivíduos, com desvio padrão de 9,9cm, o que indica que a maioria dos clientes possui altura próxima à média, que é 171,48cm. Ainda na **Figura 4**, é possível identificar outliers correspondentes aos clientes mais altos e mais baixos, sendo a maior altura registrada de 200 cm e a menor de 150 cm. Além disso, o valor da mediana (171,8cm), está muito próxima da média, sugerindo uma distribuição bastante simétrica.

3.2.2 Análises relacionadas ao peso

Figura 5: Boxplot do peso dos clientes.



Quadro 2: Medidas resumo da variável

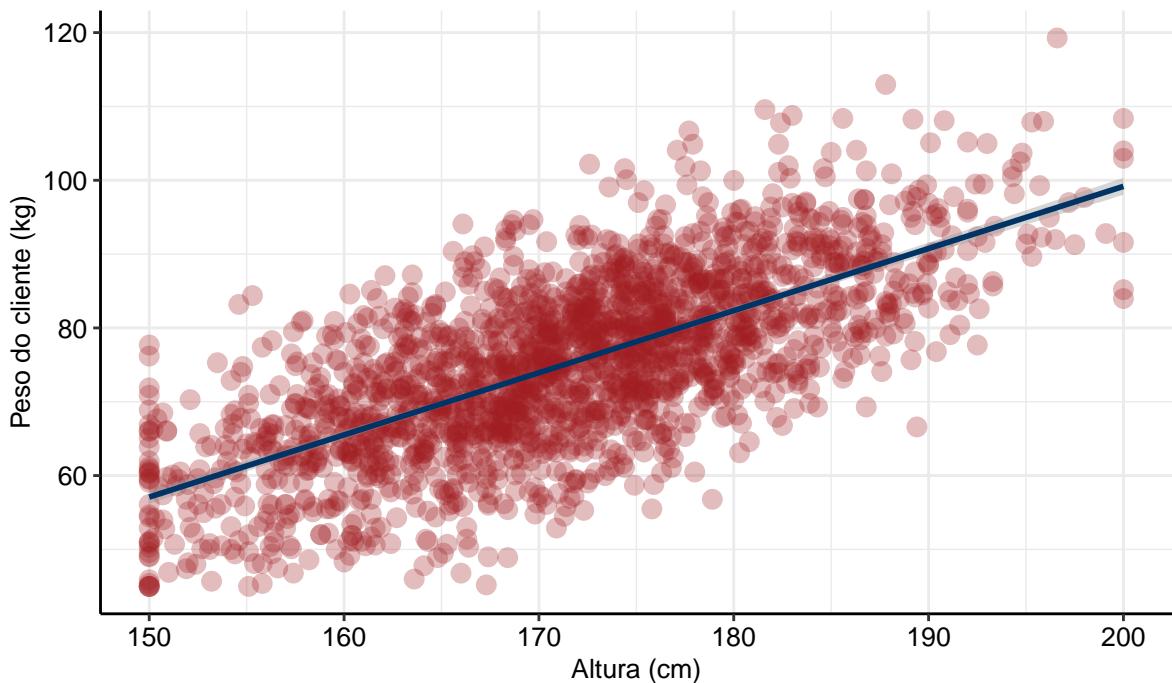
Estatística	Valor
Media	75,17
Desvio Padrao	11,91
Variancia	141,94
Minimo	44,99
1º Quartil	66,89
Mediana	75,28
3º Quartil	83,18
Maximo	119,28

Considerando os dados apresentados na tabela acima e na **Figura 5**, o peso dos clientes também revela uma variação moderada entre os indivíduos, com desvio

padrão de 11,9 kg, o que indica que a maioria apresenta valores próximos à média de 75,17 kg. Ainda na **Figura 5**, é possível identificar outliers que representam os clientes com maiores e menores pesos, sendo o menor valor observado de 44,99 kg e o maior de 119,28 kg. Além disso, assim como na altura, a mediana (75,28 kg) encontra-se praticamente igual à média, sugerindo uma distribuição equilibrada e relativamente simétrica dos dados.

3.2.3 Relação entre as variáveis

Figura 6: Gráfico de dispersão do peso pela altura.



A **Figura 6** apresenta uma tendência de crescimento clara, que significa, de maneira geral, que à medida que a altura aumenta, o peso tende também a crescer. Dessa maneira, com os pontos moderadamente concentrados em torno de uma linha ascendente, é perceptível uma relação linear positiva entre as duas variáveis apesar de existir leve dispersão. Além disso, o coeficiente de correlação de Pearson, que é uma medida estatística que indica a força e a direção da relação linear entre duas variáveis numéricas variando sua força de relação entre -1 (relação negativa) e 1 (relação positiva), de 0,697 também reforça uma correlação forte entre as variáveis.

3.3 Idade dos clientes de Âmbar Seco a depender da loja.

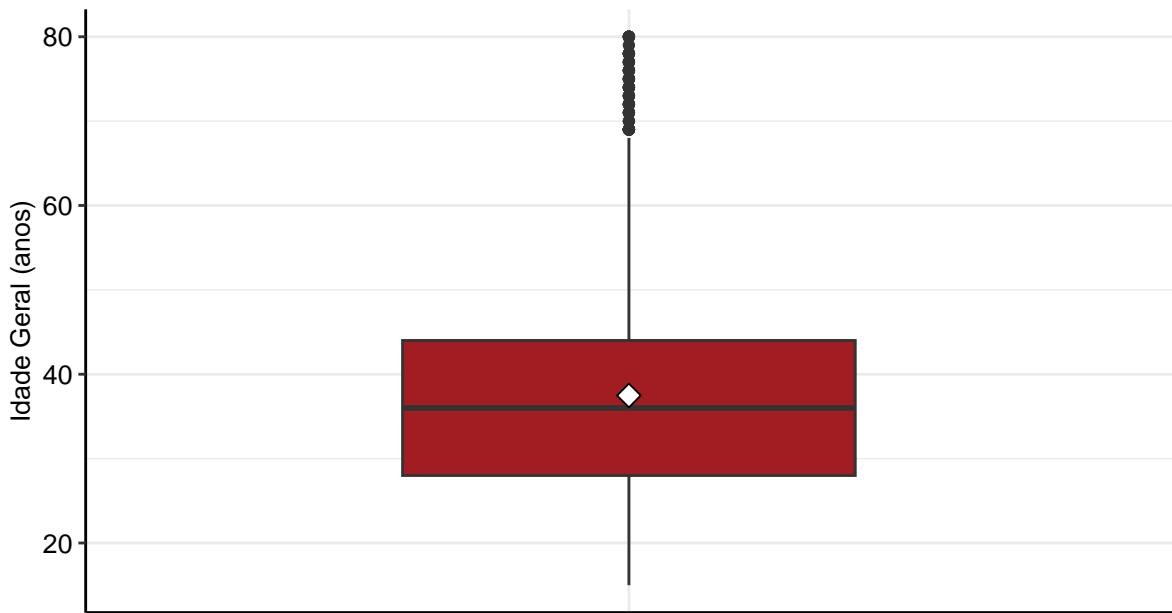
Nessa análise serão observadas medidas de resumo das idades (em anos) dos clientes com o intuito de compreender melhor o perfil etário de tais em três cenários

diferentes: os dados etários gerais da amostra disponibilizada, que contam com todas cidades e lojas, as análises na cidade de Âmbar Seco sem diferenciação por loja, e a análise que leva em conta as lojas de Âmbar Seco individualmente. Tais dados ajudarão a ter um melhor entendimento de quais são os perfis de idades dos clientes em diversos aspectos.

3.3.1 Idade de todos clientes da amostra

Nessa análise, serão consideradas as idades de todos clientes disponibilizadas em todo plano amostral, sem diferenciação por qualquer variável. Com isso, é possível ter uma análise mais ampla e comparativa com cenários mais específicos.

Figura 7: Boxplot da Idade



Quadro 3: Medidas resumo da variável

Estatística	Valor
Media	37,48
Desvio Padrao	12,83
Variância	164,56
Mínimo	15,00
1º Quartil	28,00
Mediana	36,00
3º Quartil	44,00
Máximo	80,00

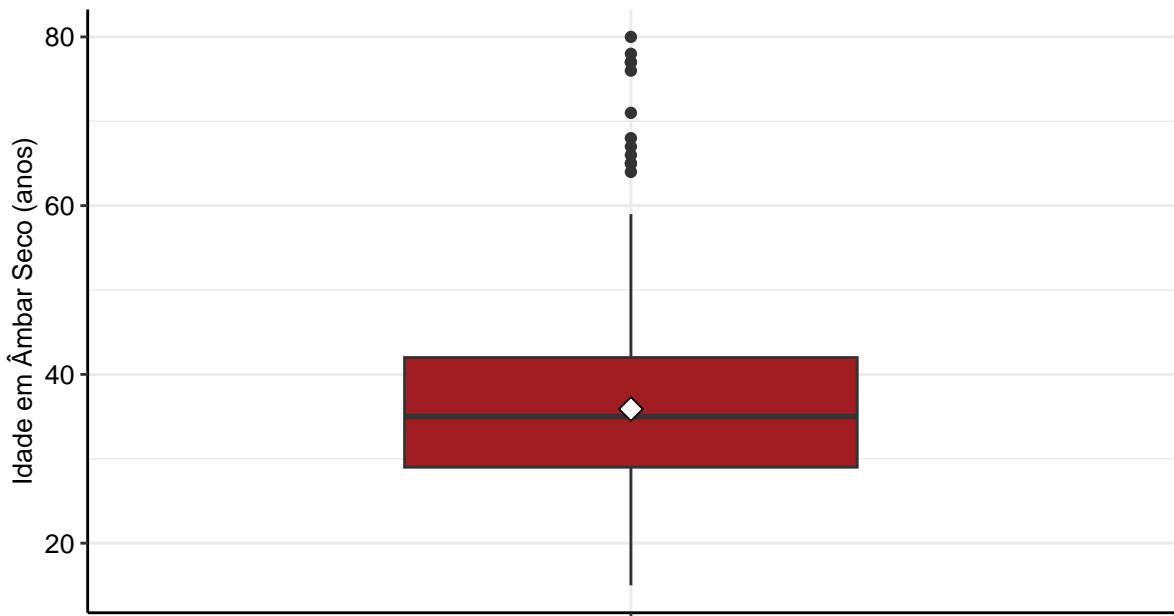
Levando em conta a tabela acima e a **Figura 7**, a distribuição geral das idades dos

clientes apresenta uma média de aproximadamente 37,5 anos, indicando que a maioria dos indivíduos se encontra na faixa etária adulta . O desvio padrão de 12,8 anos evidencia uma variação moderada nas idades observadas, sugerindo que há certa diferença no público analisado. A mediana de 36 anos confirma a concentração dos valores em torno da média, enquanto o intervalo entre o primeiro e o terceiro quartil (28 a 44 anos) mostra que metade dos clientes possui idade dentro dessa faixa. Os valores mínimo e máximo, 15 e 80 anos, indicam a presença de indivíduos tanto jovens quanto idosos na amostra, ainda que estes representem casos menos frequentes (outliers).

3.3.2 Idade dos clientes de Âmbar Seco

Nessa outra análise, é levado em conta somente os dados de idades da pequena cidade de Âmbar Seco, com intuito de aprofundar mais no cenário local antes da análise pedida.

Figura 8: Boxplot da Idade



Quadro 4: Medidas resumo da variável

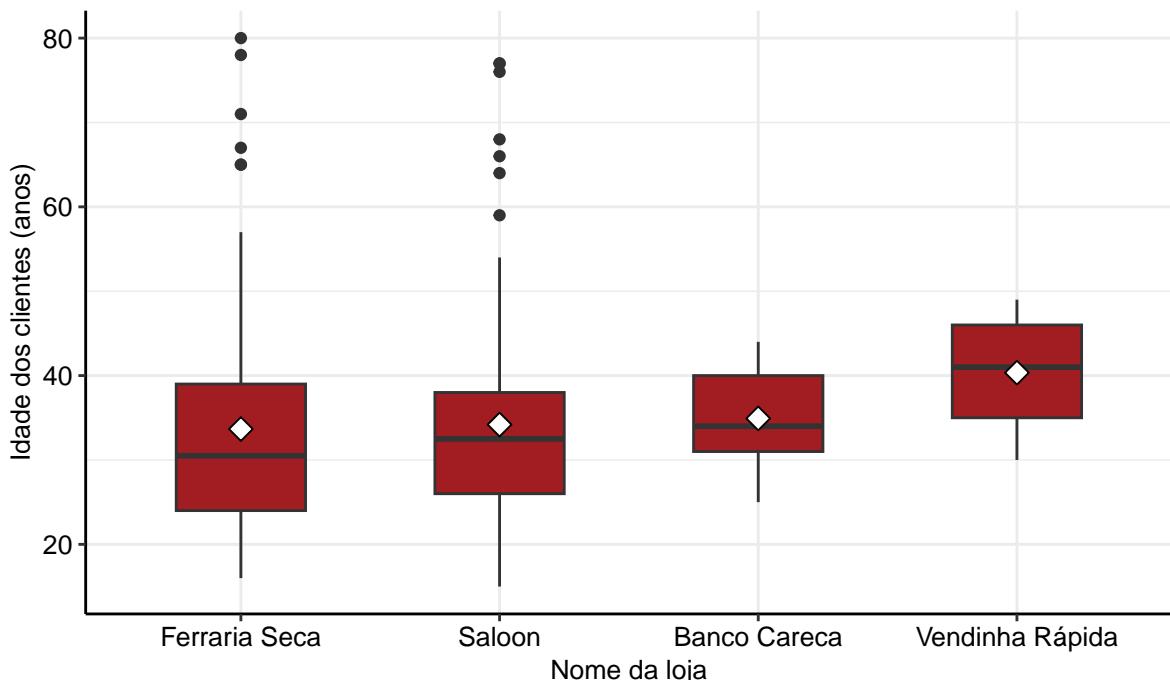
Estatística	Valor
Media	35,89
Desvio Padrão	10,46
Variância	109,42
Mínimo	15,00
1º Quartil	29,00
Mediana	35,00
3º Quartil	42,00
Máximo	80,00

Nessa outra análise, levando em conta a tabela de medidas acima e a **Figura 8**, é perceptível que a média de idade é de 35,9 anos, ligeiramente inferior à média geral da amostra (37,5 anos), indicando um público um pouco mais jovem. O desvio padrão de 10,5 anos e a variância de 109,4 mostraram uma dispersão menor em comparação à amostra total, o que sugere maior homogeneidade nas idades dos clientes locais. A mediana de 35 anos reforça essa concentração em torno da média.

3.3.3 Idade dos clientes de Âmbar seco a partir da loja

Nessa última análise, será possível observar um cenário mais específico, e que foi o proposto desde o início: a relação etária nas lojas de Âmbar Seco.

Figura 9: Boxplot bivariado da idade pelo nome da loja.



Quadro 5: Medidas resumo da idade por cidade

Estatística	Banco Careca	Ferraria Seca	Saloon	Vendinha Rápida
Media	34,92	33,67	34,20	40,35
Desvio Padrao	5,57	13,31	12,70	6,03
Variancia	31,06	177,18	161,23	36,39
Mínimo	25,00	16,00	15,00	30,00
1º Quartil	31,00	24,00	26,00	35,00
Mediana	34,00	30,50	32,50	41,00
3º Quartil	40,00	39,00	38,00	46,00
Máximo	44,00	80,00	77,00	49,00

Considerando a tabela de medidas resumo de cada loja, e a **Figura 9**, é notável diferenças no perfil etário das quatro lojas avaliadas em Âmbar Seco. O boxplot indica que as lojas Ferraria Seca e Saloon concentram clientes mais jovens em comparação às outras duas lojas, com uma mediana próxima a 30 anos. Além disso, ambas apresentam uma maior dispersão entre os dados (desvio padrão mais alto de 13,31 e 12,70 respectivamente), o que indica um atendimento de públicos variados.

A loja Banco Careca apresenta um comportamento mais homogêneo, com menor dispersão (desvio-padrão de 5,57) e idades concentradas entre 25 (mínimo) e 44 anos (máximo). Sua mediana (34 anos) e média (34,92) estão bastante próximas, o que indica uma distribuição mais simétrica. Esse padrão demonstra que o público dessa unidade é mais estável em termos de faixa etária.

Já a Vendinha Rápida se destaca por possuir o público mais velho entre as quatro lojas. A média de idade (40,35 anos) e a mediana (41 anos) são superiores às das demais unidades, com valores concentrados entre 35 e 46 anos e pouca dispersão (desvio-padrão de 6,03). Isso aponta para uma clientela mais velha e com menor variação entre as idades.

Dessa maneira, é possível concluir que as lojas apresentam diferenças consideráveis no perfil etário de seus públicos. Esses resultados podem refletir diferenças no tipo de produto oferecido, na localização ou no posicionamento de cada loja dentro do mercado local.

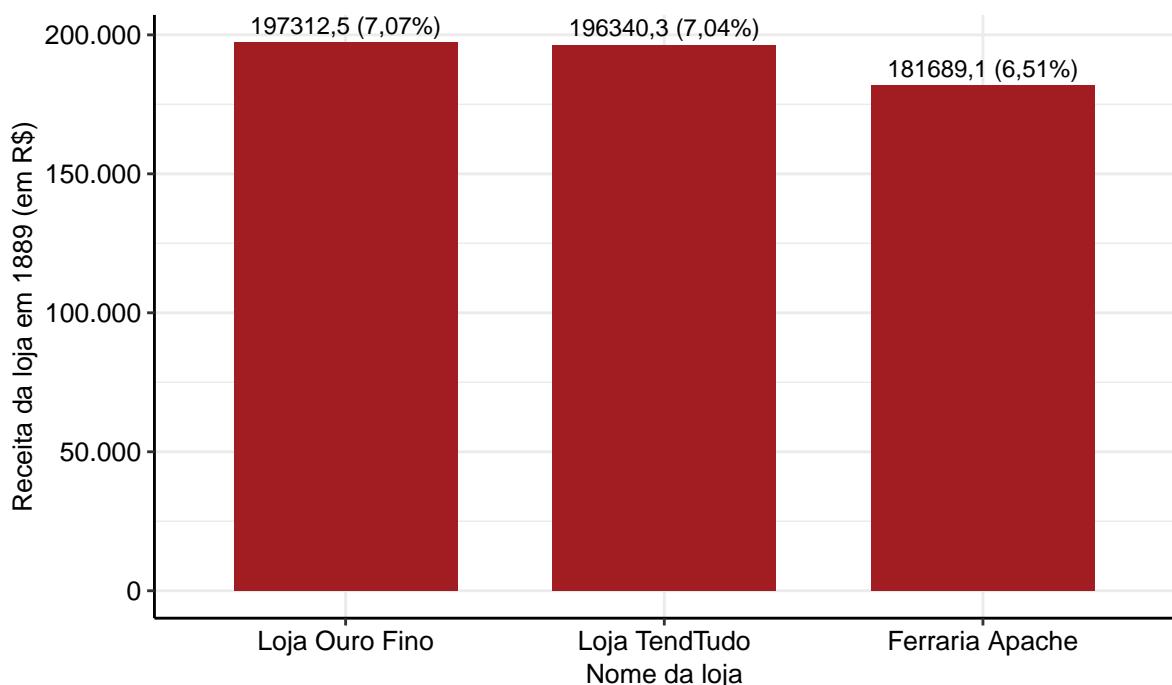
3.4 O top 3 produtos mais vendidos nas top 3 lojas com maior receita em 1889.

Nessa análise será observada a relação entre duas variáveis fundamentais para o desempenho comercial das lojas: os produtos mais vendidos e as lojas com maior receita no ano de 1889. A análise busca compreender de que forma essas variáveis se comportam em conjunto, verificando quais produtos concentram as maiores vendas nas principais lojas, se há predominância de determinados itens em alguma loja es-

pecífica ou se a distribuição de vendas é equilibrada entre os produtos. Com base nos dados coletados dessas variáveis quantitativas, será aplicada uma análise estatística descritiva, permitindo identificar padrões, tendências de consumo e possíveis preferências dos clientes, auxiliando na compreensão do desempenho e perfil comercial das lojas.

3.4.1 As três lojas com maior receita no ano de 1889.

Figura 10: Gráfico de colunas da receita das lojas.

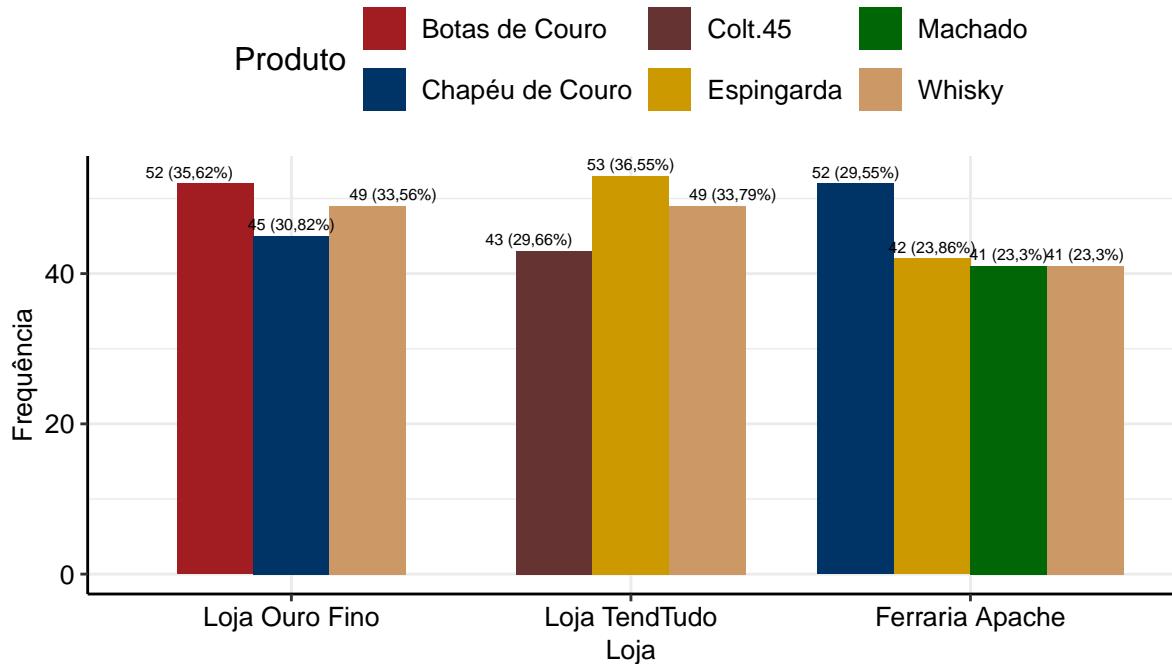


A partir do que está descrito na **Figura 10**, é possível concluir que as três lojas com maior receita no ano de 1889 foram: Loja “Ouro Fino” com aproximadamente 7,1% de todo faturamento do ano, Loja “TendTudo” ligeiramente abaixo da anterior com 7% da receita total, e “Ferraria Apache”, com 6,5% da receita. A partir disso será feita a análise individual de cada uma dessas lojas para que seja analisado quais produtos foram predominantes nas vendas desse ano.

3.4.2 Os três itens mais vendidos nas três lojas com maior receita em 1889.

Nessa análise, será possível observar a distribuição dos produtos mais vendidos em cada uma das três lojas com maior faturamento em 1889.

Figura 11: Gráfico de colunas da loja pela frequência de vendas do produto.



A partir dos valores apresentados na **Figura 11**, pode-se concluir que, na loja “Ouro Fino” em 1889, o produto mais vendido foi: “Botas de Couro”, seguido pelo “Whisky”, e em terceiro “Chapéu de Couro”. Na loja “TendTudo”, o mais vendido foi “Espinarda” seguido por: “Whisky” e “Colt. 45” respectivamente. Por último, na “Ferraria Apache”, o campeão de vendas foi: ” Chapéu de couro” seguido por “Espinarda” e em terceiro seguem dois itens empataos com 41 vendas, “Machado” e “Whisky”.

Dessa maneira, a partir da análise da **Figura 11** das três lojas em 1889, observa-se que cada estabelecimento apresenta um padrão de vendas distinto, com produtos líderes diferentes em cada caso. Apesar disso, certos itens, como “Whisky” e “Chapéu de Couro”, aparecem em mais de uma loja entre os mais vendidos, indicando uma demanda relativamente alta desses produtos no período. Além disso, a variação percentual entre os produtos mais vendidos mostra que, embora cada loja tenha seu produto mais vendido, a composição do top 3 evidencia a diversidade nas preferências dos clientes e na estratégia de vendas de cada loja.

4 Conclusão

Com base nas análises realizadas a partir do banco de dados disponibilizado por João Sábio, proprietário da Old Town Road, foi possível obter uma visão abrangente sobre o desempenho das lojas e o perfil dos clientes da região estudada. Os resultados apresentados permitem identificar tendências relevantes para o planejamento estratégico e a tomada de decisão do cliente no contexto comercial do faroeste.

De maneira geral, considerando todas análises solicitadas, observou-se um crescimento consistente da receita entre 1880 e 1889, totalizando cerca de 47% no período e média anual próxima de 5%, o que evidencia uma tendência de expansão ao longo da década. As análises de correlação apontaram relação linear positiva moderadamente forte, entre as variáveis “altura” e “peso”, indicando que o aumento de uma tende a acompanhar o da outra. Verificaram-se também diferenças marcantes no perfil etário das lojas avaliadas: unidades como Ferraria Seca e Saloon concentram públicos mais jovens e heterogêneos, enquanto Banco Careca e Vendinha Rápida atendem clientelas mais homogêneas e de faixas etárias mais elevadas. Por fim, em 1889, as três lojas de maior receita apresentaram padrões de venda distintos, com produtos líderes variados — ainda que itens como Whisky e Chapéu de Couro se destaquem em mais de um estabelecimento —, refletindo diversidade nas preferências dos consumidores e nas estratégias comerciais de cada unidade.

Destarte, os resultados obtidos indicam que o mercado regional analisado apresenta crescimento consistente, consumidores com perfis bem definidos e tendências comerciais diversificadas entre as unidades de negócio. As informações geradas a partir deste estudo fornecem dados concretos para o planejamento e expansão das atividades da Old Town Road, apoiando decisões futuras de investimento, segmentação de público e otimização de portfólio de produtos.