***Aluguer de Bicicletas***

Rafael Palma1, Andrei Truhanov2, Pedro Fernandes 3

1 IG-01, 201600649

2 IG-01, 202001971

3 IG-01, 201600156

**Resumo:** O presente trabalho tem o intuito de realizar uma análise descritiva ao dataset hour. Primeiro foram analisadas e classificadas todas as variáveis do dataset. Das 17 variáveis possíveis as escolhidas para estudar/analisar foram apenas 6: a temperatura, o dia da semana, o casual, a humidade e a temperatura sentida, pois na nossa opinião eram as mais interessantes e relevantes para o estudo. Seguidamente foram criadas funções no software R que possibilitavam o estudo e análise. Ao longo do trabalho notou-se que existiam muitos dados de entrada (cerca de 13400), então foi feita a decisão de dividi-los em dados do dia e dados da noite para observar as diferenças que poderiam ou não existir. Assim, com esta observação, podia-se responder a questões como por exemplo:

* Quando existem mais alugueres? Na parte do dia ou da noite?
* Quais são as horas da parte do dia e da noite onde existem menos utilizadores?
* Qual temperatura média da parte do dia? E da noite?
* Será as diferenças de temperatura e humidade (dia/noite) têm efeito no número de alugueres?

Entre outras.

Por último foi escrito este relatório técnico contendo todos os gráficos e dados da nossa análise (variável a variável) e conclusões.

O resumo deverá consistir numa descrição breve do conteúdo do trabalho. Não deverá ultrapassar as 300 palavras e não deverá conter tabelas, figuras, gráficos, equações ou referências. O trabalho deverá seguir a formatação apresentada neste documento exemplo.

**Palavras-chave**: Temperatura; Dia da Semana; Número de Alugueres; Casual; Humidade

**1. Introdução**

O presente trabalho tem o intuito de realizar uma análise descritiva ao dataset hour, disponível em <https://www.kaggle.com/datasets/marklvl/bike-sharing-dataset>. Este dataset contém 17379 instâncias observadas entre janeiro de 2011 até dezembro de 2012. O dataset é composto oficialmente por 17 variáveis, disponíveis no ficheiro csv. O objetivo principal deste trabalho é escrever em R a análise dos dados do dataset em que iremos descobrir quais os valores com maior frequência e quais os seus resultados finais.

**2. Título da secção**

**Classificação das Variáveis**

**Temp** : Podemos observar que a variável temp(Tempertura) é uma Variável quantitativa continua porque os valores estão casa decimais, e é tratada como Quantitativa Continua normal, visto que, tem grande variedade de variáveis distintas.

**Weekday** : Podemos observar que a variável weekday (Dia da semana) é uma Variável Qualitativa nominal porque os números são tratados como dias da semana, e é tratada como como Qualitativa nominal normal, visto que, não tem grande variedade de variáveis distintas.

**Casual** : Podemos observar que a variável Casual(Guest User) é uma Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

**Atemp :** Podemos observar que a variável atemp (Atemperatura) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

**Hum** : Podemos observar que a variável hum(humidade) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

**Cnt** : Podemos observar que a variável cnt(Count) é uma Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

**2.1. Título da subsecção**

**Tabela de resultados**

No projeto encontra se um conjunto de resultados consoante um determinado intervalo de horas de alugueres de bicicletas.

**Tabela 1** - Horas de aluguer durante o dia e noite

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variável | Temp | Weekday | Casual | Atemp | Hum | cnt |
| Moda | 0.62 | 6 | 0 | 0.6212 | 0.88 | 5 |
| Média | 0.4969872 | 3.003683 | 35.67622 | 0.4757751 | 0.6272288 | 189.4631 |
| Mediana | 0.5 | 3 | 17 | 0.4848 | 0.63 | 142 |
| Quartil | 0% 0.02  25% 0.34  50% 0.50  75% 0.66  100% 1.00 | 0% 0  25% 1  50% 3  75% 5  100% 6 | 0% 0  25% 4  50% 17  75% 48  100% 367 | 0% 0.0000  25% 0.3333  50% 0.4848  75% 0.6212  100% 1.0000 | 0% 0  25% 0.48  50% 0.63  75% 0.78  100% 1 | 0% 1  25% 40 50% 142 75% 281 100% 977 |
| Variância | 0.03707786 | 4.023119 | 2430.986 | 0.0295325 | 0.03722192 | 32901.46 |
| Desvio Padrão | 0.1925561 | 2.005771 | 49.30503 | 0.1718502 | 0.1929298 | 181.3876 |
| Coeficiente de Variação | 38.74469 | 66.77708 | 138.2014 | 36.12005 | 30.75908 | 95.7377 |
| Máximo | 1 | 6 | 367 | 1 | 1 | 977 |
| Mínimo | 0.02 | 0 | 0 | 0 | 0 | 1 |
| Assimetria | -0.006019844 | -0.002997704 | 2.498805 | -0.09041325 | -0.1112679 | 1.277191 |
| Achatamento | -0.9421553 | -1.256181 | 7.567262 | -0.8457618 | -0.8264745 | 1.415942 |

**Temp**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**weekday**

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**Casual**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**atemp**

Podemos observar que a variável atemp (Atemperatura) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**hum**

Podemos observar que a variável hum(humidade) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**cnt**

Podemos observar que a variável cnt(Count) é uma Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

De acordo com o resultado, é assimétrica positiva (ou enviesada para a direita).

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**Tabela 2** - Horas de aluguer durante o dia

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variável | Temp | Weekday | Casual | Atemp | Hum | cnt |
| Moda | 0.72 | 0 | 6 | 0.6212 | 0.46 | 224 |
| Média | 0.5236736 | 2.999886 | 55.84882 | 0.4981943 | 0.572013 | 280.9225 |
| Mediana | 0.52 | 3 | 36 | 0.5 | 0.55 | 237 |
| Quartil | 0% 0.02  25% 0.36  50% 0.52  75% 0.70  100% 1.00 | 0% 0  25% 1  50% 3  75% 5  100% 6 | 0% 0  25% 15  50% 36  75% 74  100% 367 | 0% 0.0000  25% 0.3485  50% 0.5  75% 0.6515  100% 1.0000 | 0% 0  25% 0.42  50% 0.55  75% 0.72  100% 1 | 0% 1  25% 142 50% 237 75% 382 100% 977 |
| Variância | 0.0401759 | 4.020945 | 3436.56 | 0.03187527 | 0.03764531 | 34893.5 |
| Desvio Padrão | 0.2004393 | 2.00523 | 58.62222 | 0.1785365 | 0.194024 | 186.798 |
| Coeficiente de Variação | 38.27561 | 66.84353 | 104.9659 | 35.83671 | 33.91951 | 66.4945 |
| Máximo | 1 | 6 | 367 | 1 | 1 | 977 |
| Mínimo | 0.02 | 0 | 0 | 0 | 0 | 1 |
| Assimetria | -0.03380892 | -0.0001836458 | 1.899737 | -0.1370249 | 0.2254248 | 0.9866874 |
| Achatamento | -0.998132 | -1.255535 | 3.905951 | -0.8923441 | -0.7835361 | 0.6186668 |

**Temp**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**weekday**

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**Casual**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**atemp**

Podemos observar que a variável atemp (Atemperatura) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**hum**

Podemos observar que a variável hum(humidade) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**cnt**

Podemos observar que a variável cnt(Count) é uma Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variaveis distintas.

De acordo com o resultado, é assimétrica positiva (ou enviesada para a direita).

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**Tabela 3** - Horas de aluguer durante a noite

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variável | Temp | Weekday | Casual | Atemp | Hum | cnt |
| Moda | 0.62 | 6 | 0 | 0.6212 | 0.88 | 5 |
| Média | 0.4700012 | 3.007522 | 15.27717 | 0.4531042 | 0.6830645 | 96.97697 |
| Mediana | 0.46 | 3 | 5 | 0.4545 | 0.7 | 48 |
| Quartil | 0% 0.02  25% 0.32  50% 0.46  75% 0.62  100% .92 | 0% 0  25% 1  50% 3  75% 5  100% 6 | 0% 0  25% 1  50% 5  75% 18  100% 237 | 0% 0.0000  25% 0.3182  50% 0.4545  75% 0.5909  100% 0.9242 | 0% 0.00  25% 0.56  50% 0.70  75% 0.83  100% 1.00 | 0% 1 25% 12 50% 48 75% 140 100% 743 |
| Variância | 0.03250075 | 4.025754 | 586.6798 | 0.02614449 | 0.03059674 | 13876.49 |
| Desvio Padrão | 0.1802797 | 2.006428 | 24.22147 | 0.1616926 | 0.1749192 | 117.7985 |
| Coeficiente de Variação | 38.35728 | 66.71366 | 158.5469 | 35.68552 | 25.60801 | 121.4706 |
| Máximo | 0.92 | 6 | 237 | 0.9242 | 1 | 743 |
| Mínimo | 0.02 | 0 | 0 | 0 | 0 | 1 |
| Assimetria | -0.07372828 | -0.005844616 | 2.896634 | -0.1239053 | -0.3915097 | 1.843371 |
| Achatamento | -1.022793 | -1.257213 | 11.42775 | -0.8666668 | -0.4062143 | 3.583108 |

**Temp**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**weekday**

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**Casual**

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**atemp**

Podemos observar que a variável atemp (Atemperatura) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**hum**

Podemos observar que a variável hum(humidade) é uma Variável quantitativa contínua, e é tratada como quantitativa contínua normal, visto que, tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica negativa fraca (ou enviesada para a esquerda).

Curva platicúrtica, achatada, o que significa que os valores estão pouco concentrados em torno da média e, consequentemente, existe uma variação elevada.

**cnt**

Podemos observar que a variável cnt(Count) é uma Variável quantitativa discreta, e é tratada como quantitativa discreta normal, visto que, não tem grande variedade de variáveis distintas.

De acordo com o resultado, é assimétrica positiva (ou enviesada para a direita).

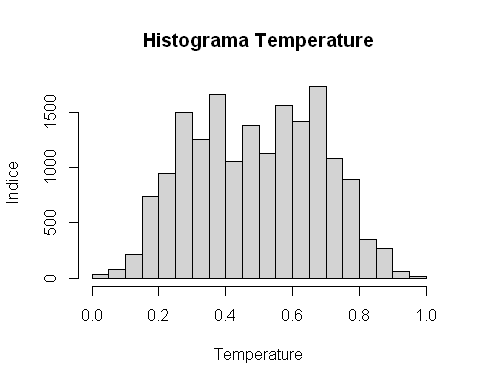
Curva leptocúrtica, alongada, o que significa que os valores estão muito concentrados em torno da média e, consequentemente, existe uma variação é pouco elevada.

**Observação geral dos dados**

Ao verificar todos os dados das tabelas foi encontrado todas com valores semelhantes como é o caso da moda porque 0 que corresponde ao dia da semana Domingo e 6 que corresponde ao dia da semana Sábado e esses dias que são ao fim de semana são um dos dias que as pessoas vão passear e é quando existem mais alugueres de bicicletas porque as pessoas não estão a trabalhar ao fim de semana. Apesar de haver humidade de 88% como moda existiu mais alugueres de bicicleta porque as pessoas ao fim de semana não estão a trabalhar e o cnt fica na mesma porque existe uma necessidade de transporte que as pessoas precisam.

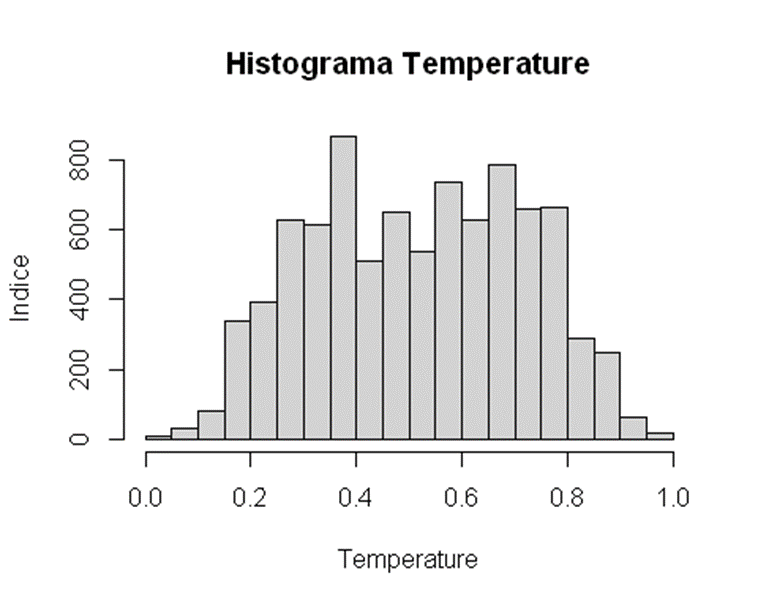
**Os resultados dos histogramas/gráfico de barras**

Os histogramas e gráfico de barras vão mostrar os resultados do número de ocorrências de uma determinada coluna.



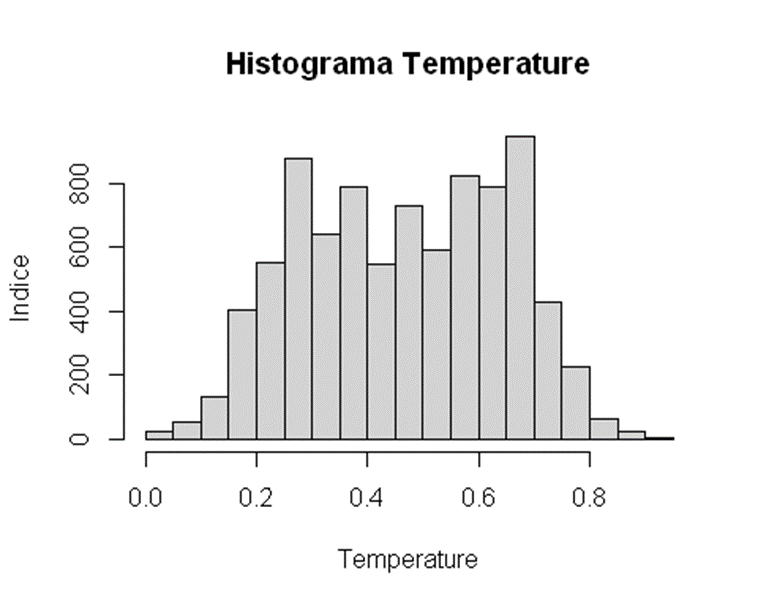
**Gráfico** 1 - Temperature dia e noite

A partir do histograma da temperatura(temp), pode-se observar que o valor do índice aumenta no intervalo 0.02, a 1, tendo um comportamento crescente até atingir o maior índice no intervalo de 0.608 a 0.673.



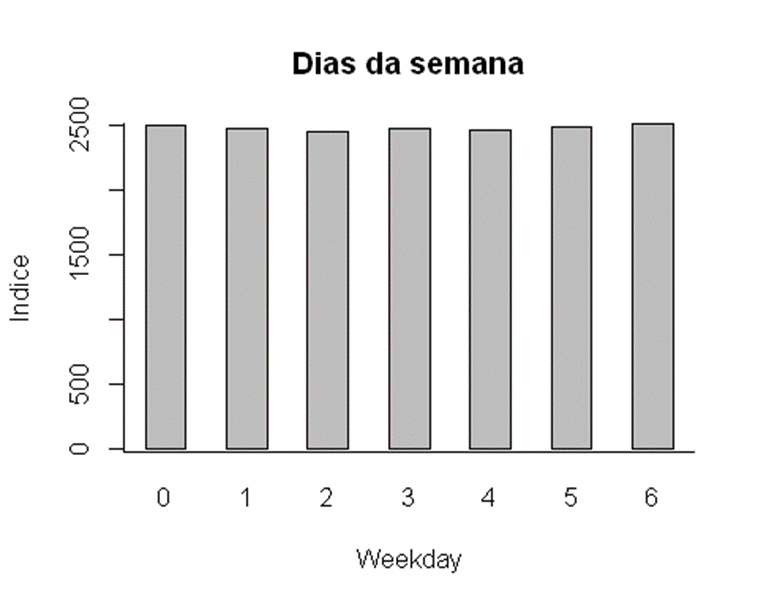
**Gráfico 2 -** Temperature dia

A partir do histograma da temperatura(temp), pode-se observar que o valor do índice aumenta no intervalo 0.02, a 0.72 e começa a diminuir de 0.72 a 1 tendo um comportamento crescente até atingir o maior índice no intervalo de 0.65 a 0.72.



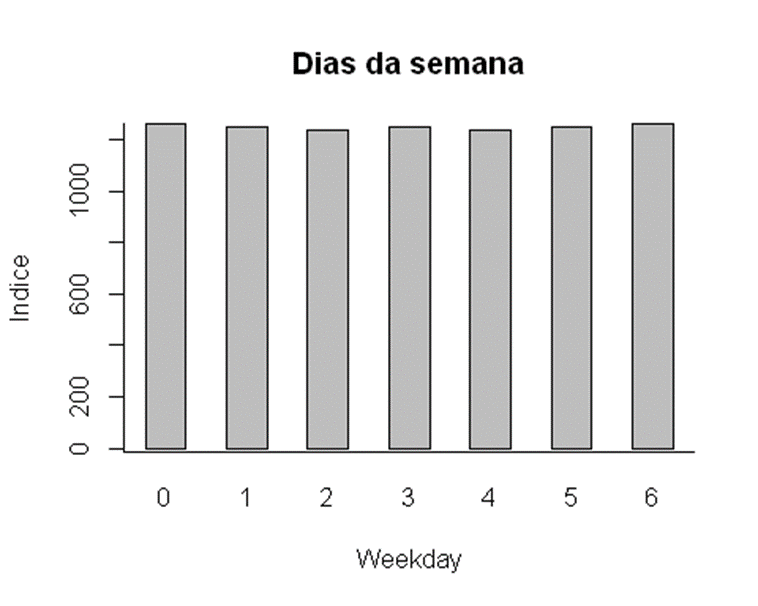
**Gráfico 3 -** Temperature noite

A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0.02 a 0.663 e começa a diminuir de 0.663 até 0.92 tendo um comportamento crescente até atingir o maior índice no intervalo de 0.599 a 0.663.



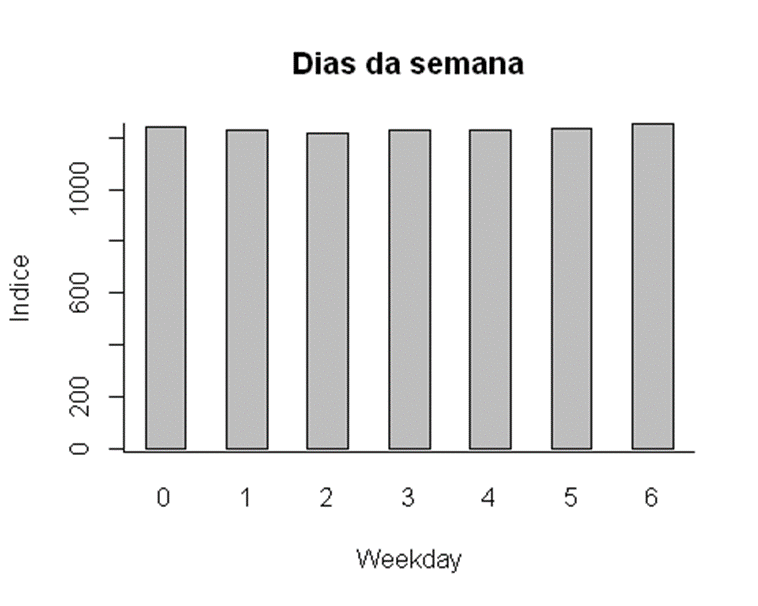
**Gráfico 4 -** Weekday dia e noite

A partir do Gráfico de barras do weekday(dia da semana), pode-se observar que o valor do índice mantém se no intervalo 0 a 6, tendo um comportamento crescente até atingir o maior índice que é 6.



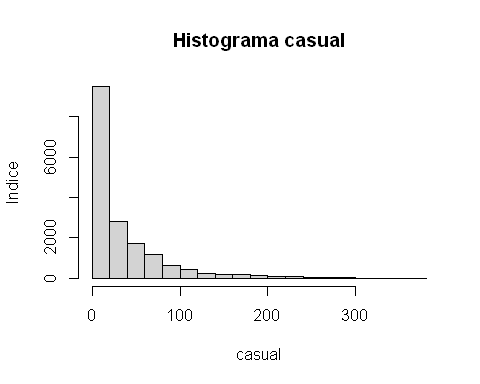
**Gráfico 5 -** Weekday dia

A partir do Gráfico de barras do weekday(dia da semana), pode-se observar que o valor do índice mantém se no intervalo 0 a 6, tendo um comportamento crescente até atingir o maior índice que é 0.



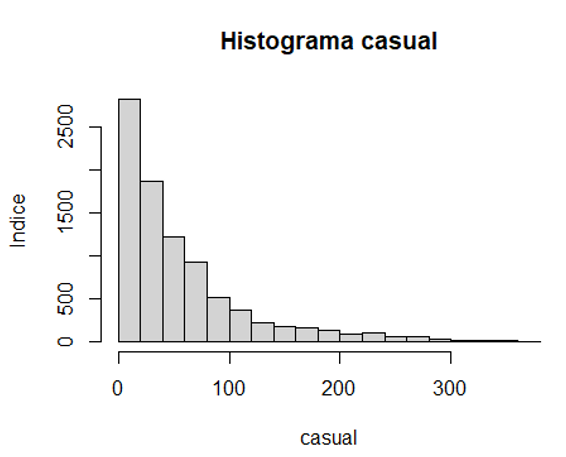
**Gráfico 6 -** Weekday noite

A partir do Gráfico de barras do weekday(dia da semana), pode-se observar que o valor do índice mantém se no intervalo 0 a 6, tendo um comportamento crescente até atingir o maior índice que é 6.



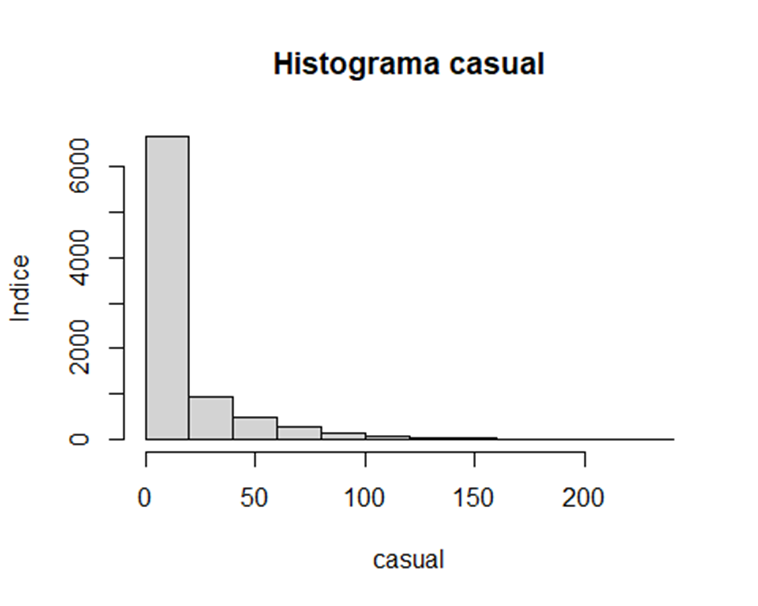
**Gráfico 7 -** Casual dia e noite

A partir do histograma de utilizadores casuais, pode-se observar que o valor do índice diminui no intervalo 0 a 367, tendo um comportamento decrescente e encontro o maior índice no intervalo de 0 a 24.5.



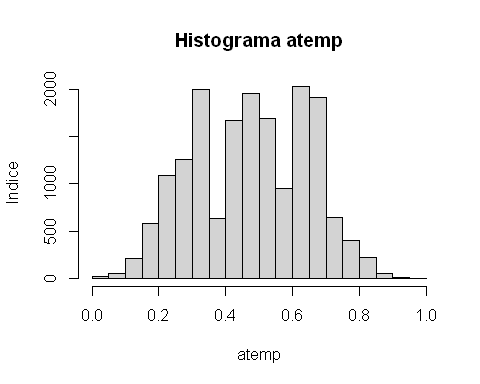
**Gráfico 8 -** Casual dia

A partir do histograma, pode-se observar que o valor do índice diminui no intervalo 0, a 367, tendo um comportamento decrescente até atingir o maior índice no intervalo de 0 a 24.5.



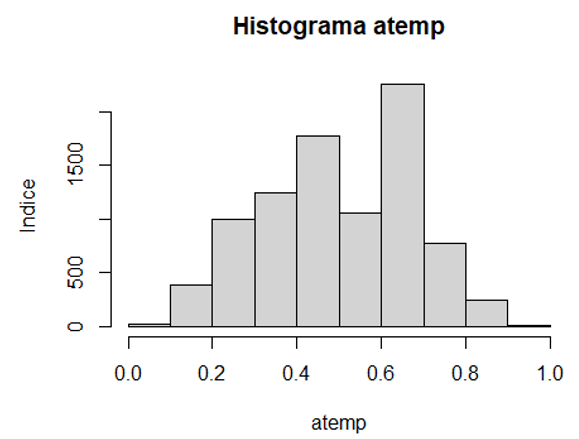
**Gráfico 9 -**  Casual noite

A partir do histograma de utilizadores casuais do weekday(dia da semana), pode-se observar que o valor do índice diminui no intervalo 0 a 367, tendo um comportamento decrescente e atingi o maior índice no intervalo de 0 a 16.9.



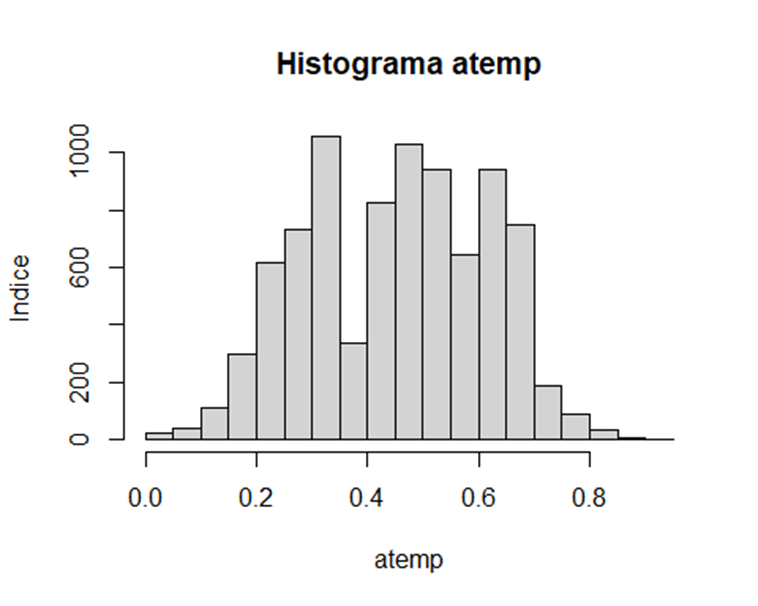
**Gráfico 10 -** atemp dia e noite

A partir do histograma de temperatura sentida do atemp(temperatura sentida), pode-se observar que o valor do índice aumenta no intervalo 0 a 0.5337 e diminui no intervalo 0.5337 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.467 a 0.533.



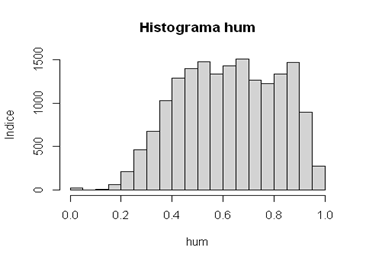
**Gráfico 11 -** atemp dia

A partir do histograma de temperatura sentida do atemp dia(temperatura sentida), pode-se observar que o valor do índice aumenta no intervalo 0 a 0.643, diminui no intervalo 0.643 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.571 a 0.643.



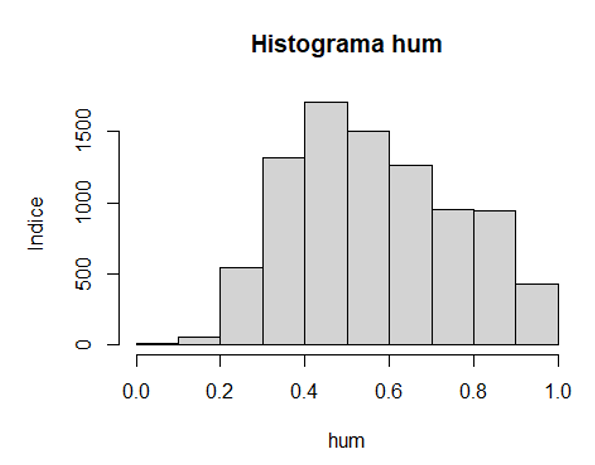
**Gráfico 12 -** atemp noite

A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0 a 0.594 e diminui no intervalo 0.594 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.528 a 0.594.



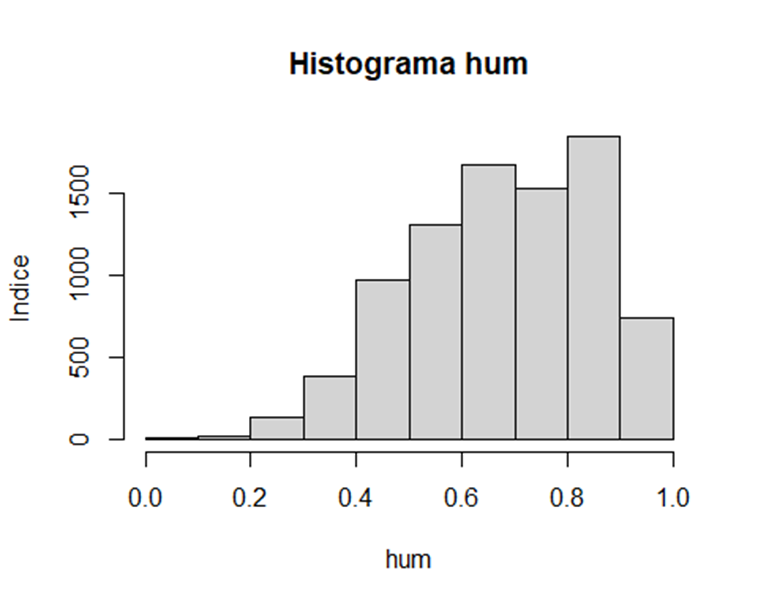
**Gráfico 13 -** hum dia e noite

A partir do histograma Hum(Humidade), pode-se observar que o valor do índice aumenta no intervalo 0 a 0.6 e diminui no intervalo 0.6 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.533 a 0.6.



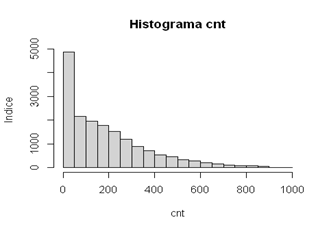
**Gráfico 14 -** hum dia

A partir do histograma Hum(Humidade), pode-se observar que o valor do índice aumenta no intervalo 0 a 0.5 e diminui no intervalo 0.5 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.429,0.5.



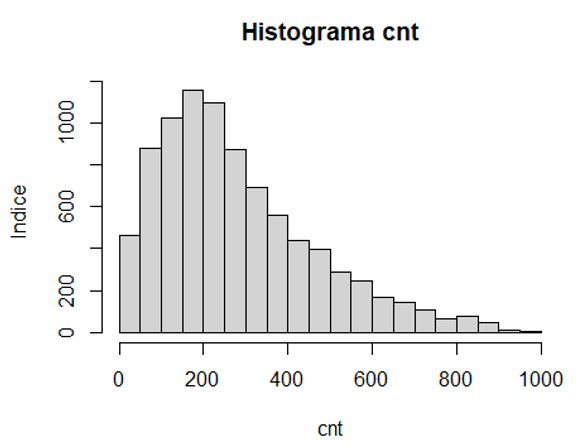
**Gráfico 15 -** hum noite

A partir do histograma, pode-se observar que o valor do índice aumenta no intervalo 0 a 0.714 e diminui no intervalo 0.714 a 1, tendo um comportamento crescente e decrescente até atingir o maior índice no intervalo de 0.643 a 0.714.



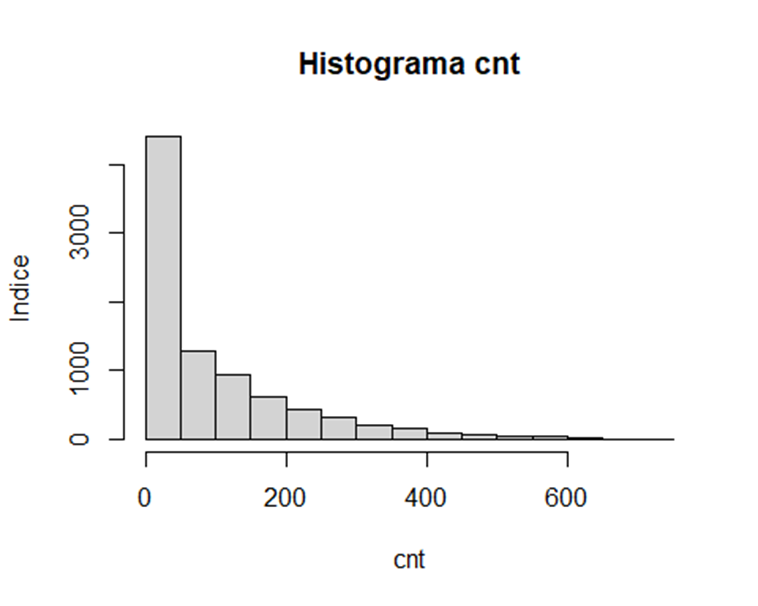
**Gráfico 16 -** cnt dia e noite

A partir do histograma da cnt(contagem total de alugueres de bicicletas), pode-se observar que o valor do índice diminui no intervalo 0, a 977, tendo um comportamento decrescente até atingir o maior índice no intervalo de 1 a 66.1.



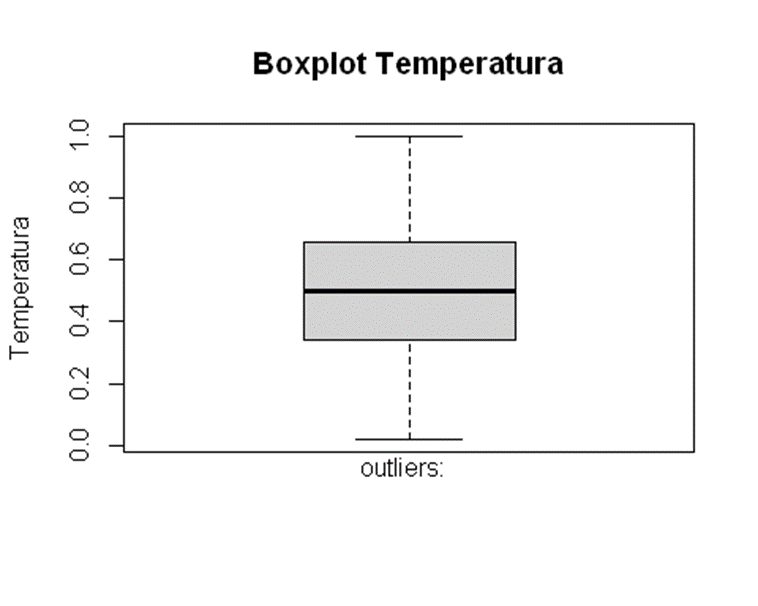
**Gráfico 17 -** cnt dia

A partir do histograma da cnt(contagem total de alugueres de bicicletas), pode-se observar que o valor do índice aumenta no intervalo 0, a 210, atingindo o seu máximo e posteriormente tem um comportamento decrescente até atingir o menor indíce a 977.



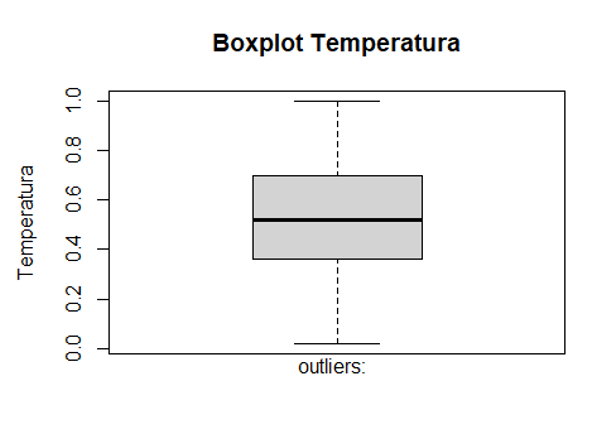
**Gráfico 18 -** cnt noite

A partir do histograma da cnt(contagem total de alugueres de bicicletas), pode-se observar que o valor do índice diminui no intervalo 0, a 743, tendo um comportamento decrescente atingido o maior índice no intervalo de 1 a 54.



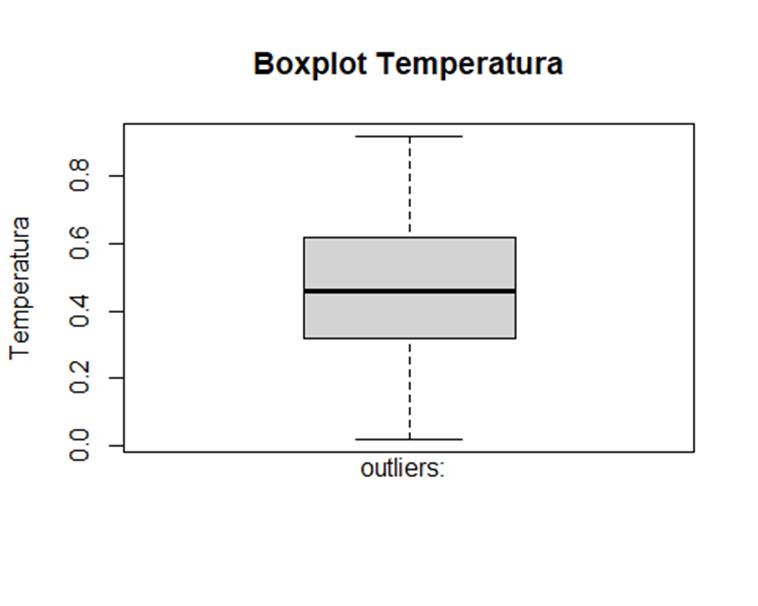
**Gráfico 19 -** temp outliers dia e noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



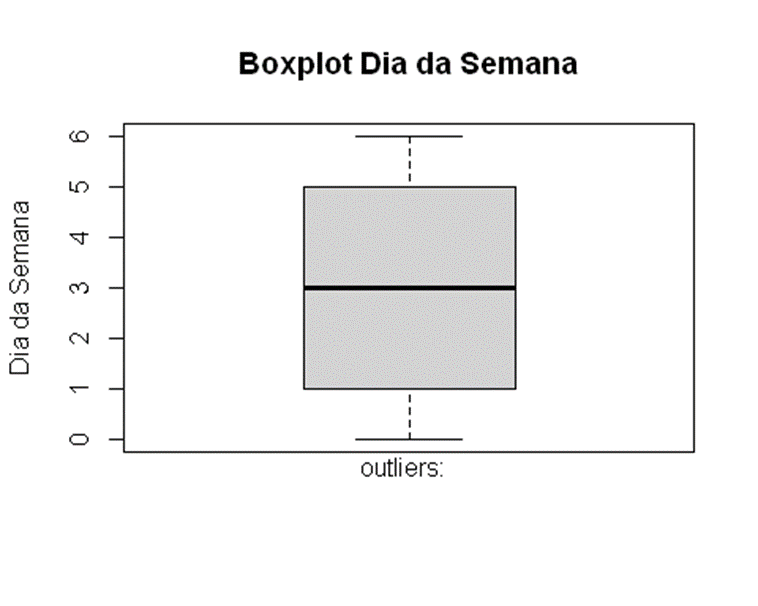
**Gráfico 20 -** temp outliers dia

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



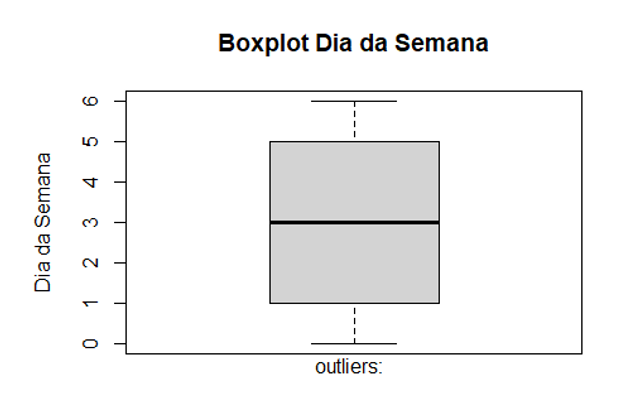
**Gráfico 21 -** temp outliers noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



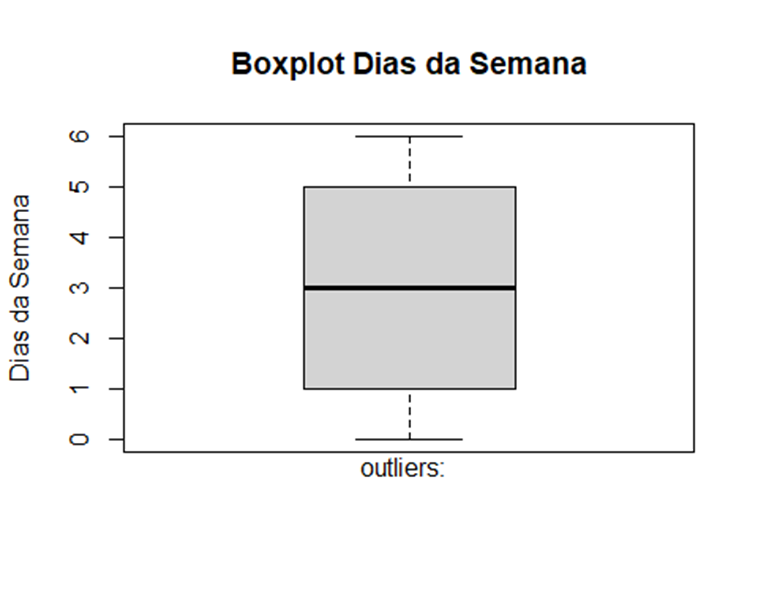
**Gráfico 22 -** weekday outliers dia e noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



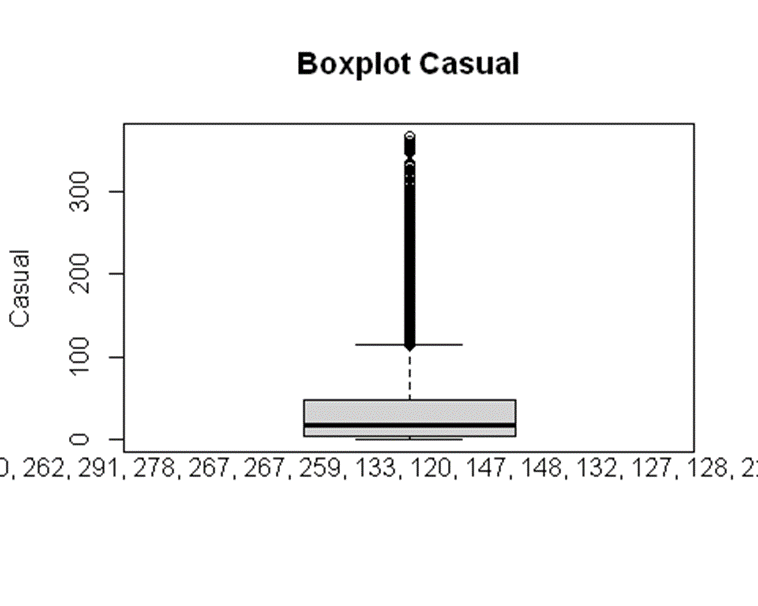
**Gráfico 23 -** weekday outlier dia

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



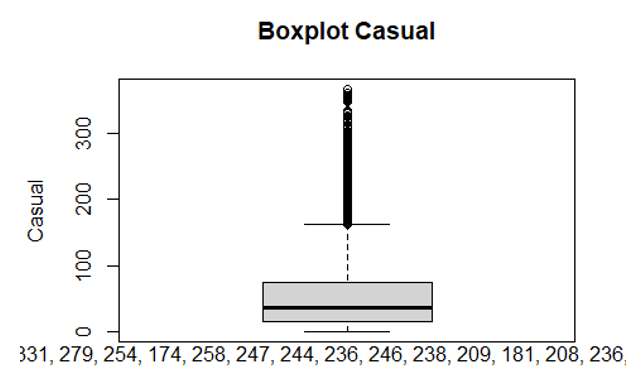
**Gráfico 24 -** weekday outliers noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



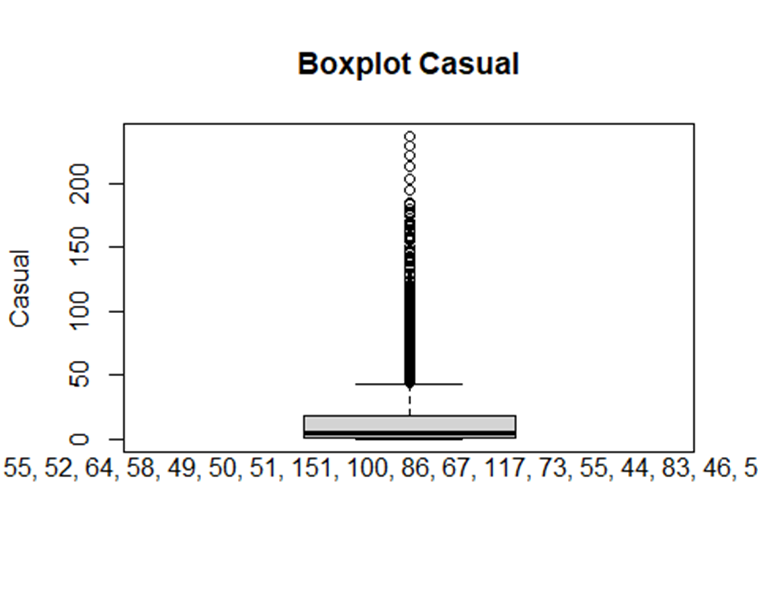
**Gráfico 25 -** casual outliers dia e noite

Verificou-se que existem outliers presentes no gráfico.



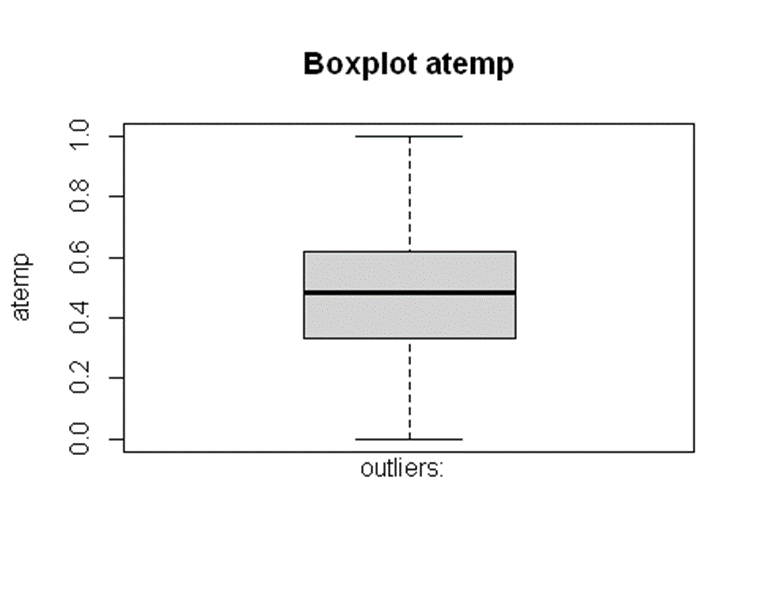
**Gráfico 26 -** casual outliers dia

Verificou-se que existem outliers presentes no gráfico.



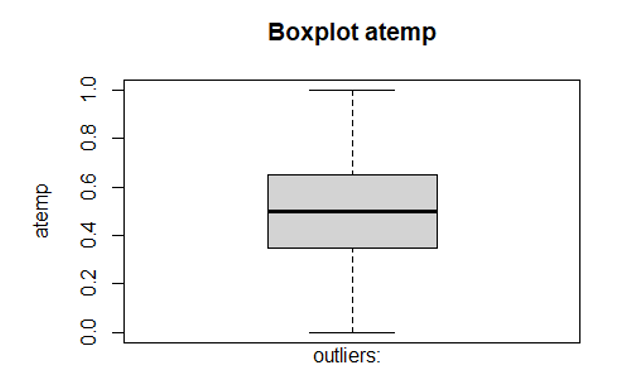
**Gráfico 27 -** casual outliers noite

Verificou-se que existem outliers presentes no gráfico.



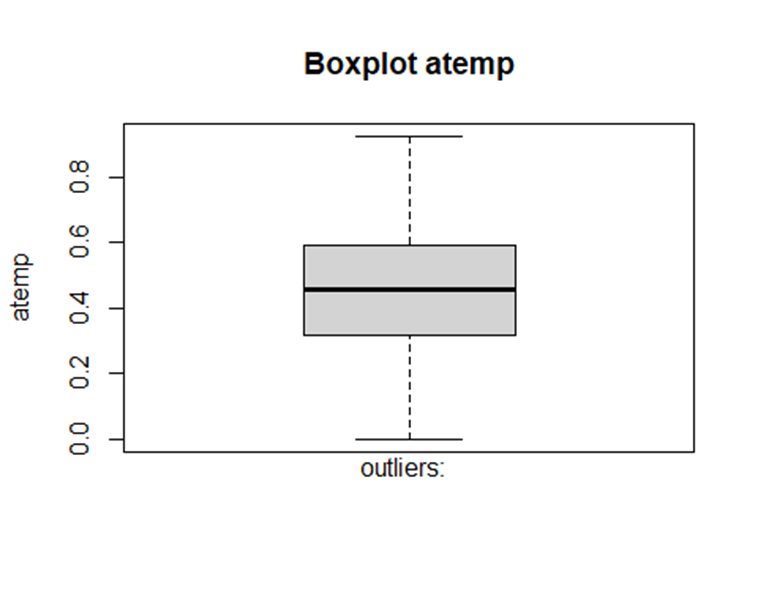
**Gráfico 28 -** atemp outliers dia e noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



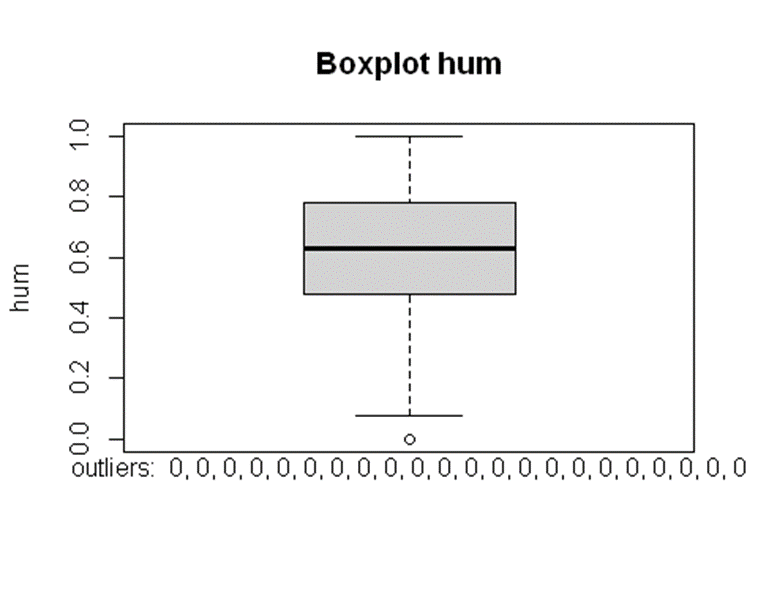
**Gráfico 29 -** atemp outliers dia

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



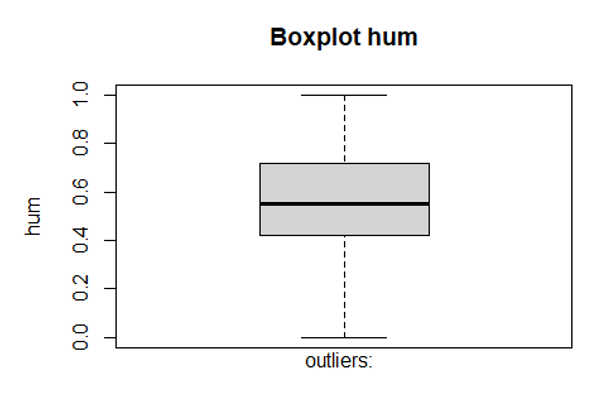
**Gráfico 30 -** atemp outliers noite

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



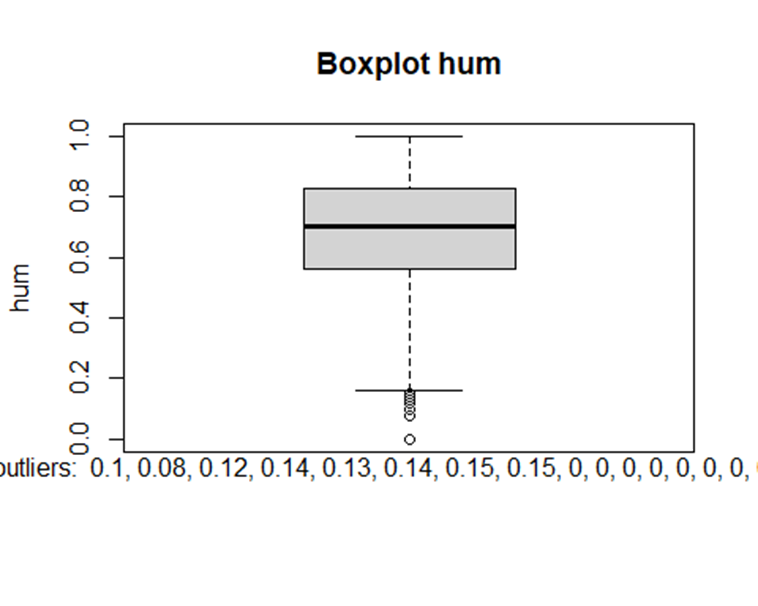
**Gráfico 31 -** hum outliers dia e noite

Verificou-se que existem outliers presentes no gráfico no ponto 0.



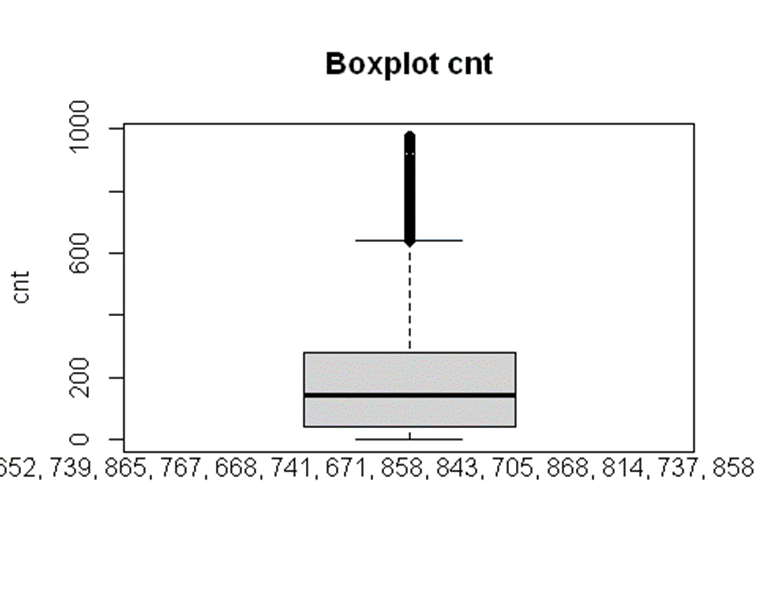
**Gráfico 32 -** hum outliers dia

Verificou-se que não existem outliers presentes no gráfico porque os resultados da coluna não estão fora do diagrama de extremos e quartis.



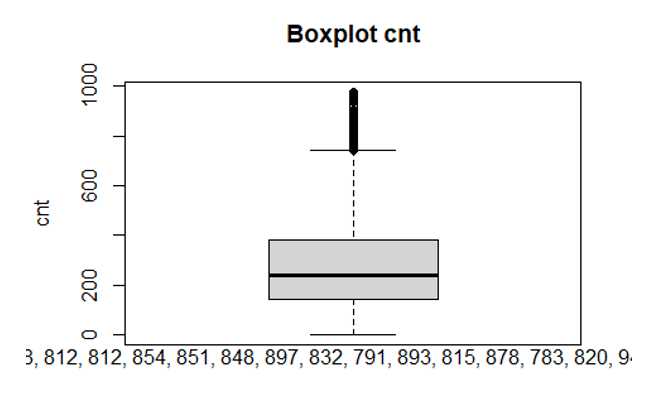
**Gráfico 33 -** hum outliers noite

Verificou-se que existem outliers presentes no gráfico no ponto 0 e perto dele.



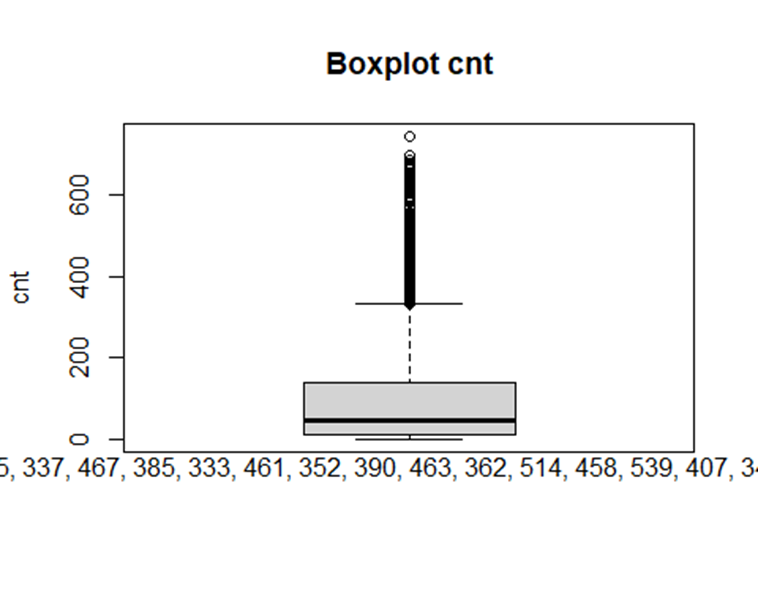
**Gráfico 34 -** cnt outliers dia e noite

Verificou-se que existem bastantes outliers presentes no gráfico.



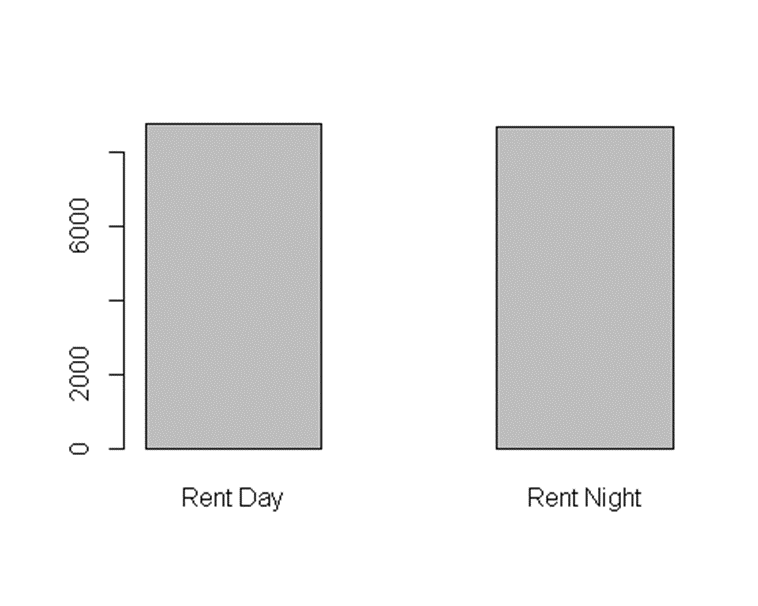
**Gráfico 35 -** cnt outliers dia

Verificou-se que existem bastantes outliers presentes no gráfico.



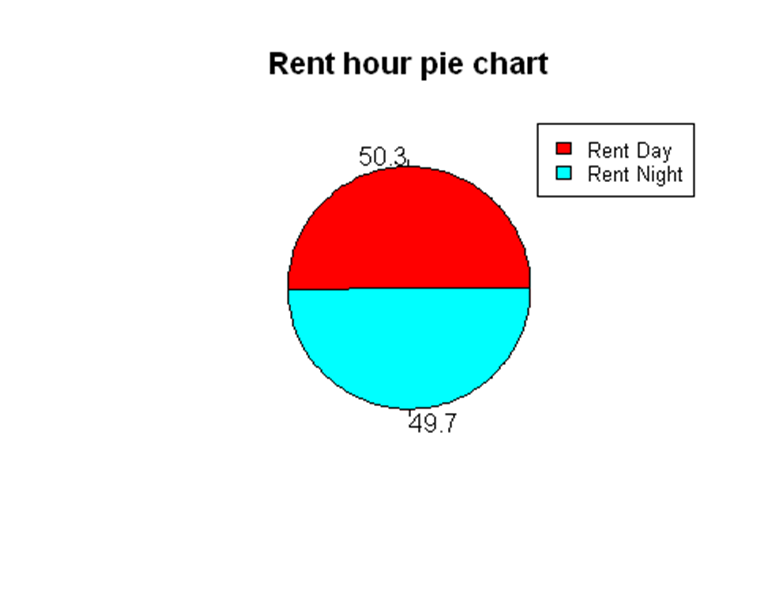
**Gráfico 36 -** cnt outliers noite

Verificou-se que existem bastantes outliers presentes no gráfico.



**Gráfico 37 -** BarPlot com o número de algueres de dia e noite

Podemos ver no gráfico que o Rent Day 8738 e Rent Night 8641, ou seja, existiu mais alugueres durante o dia do que durante a noite



**Gráfico 38 -** PieChart com a percentagem de algueres de dia e noite

Podemos ver no gráfico em tarte podemos ver com mais detalhe que Rent Day tem 50.3% dos alugueres de bicicletas e o Rent Night tem 49.7% dos alugueres de bicicletas.

**Equações/funções utilizadas no código**

Durante o projeto foram utilizadas algumas funções para fazer os cálculos dos dados de forma eficaz e de forma que se entenda o que se está fazer, baseando –se com os formulários da matéria Estatística descritiva. Como por exemplo fazer a moda, sturges, amplitude, coeficiente de variação, a assimetria, o achantamento e os diversos outliers. Todas as fórmulas que foram utilizadas no R foram baseadas através dos slides de Estatística Descritiva.

**5. Conclusões**

No fim de análise de todos os dados dos alugueres de bicicletas reparou se que os alugueres de dia e noite tem resultados diferentes quando se analisa só por uma parte do dia. Reparou se que todos os dados que foram analisados, nos alugueres de dia e noite tem algumas semelhanças em relação às modas e valores extremos e histogramas.

Ao observar os dados também foi bastante importante fazer uma análise dos dias da semana onde ocorreram mais alugueres e menos alugueres.

**Referências Bibliográficas**

Documentação Não Eletrónia

1. O’Reilly, Bruce Bruce - Pratical Statistics for Data scientists.(2022)

2. O’Reilly, Hadley Wickham & Garrett Glomund - R for Data Science.(2022)

3. ANJOS, Adilson dos - Estatística Básica com uso do Software R: Departamento de Estatística UFPR.(2022)

Documentação Eletrónica

1. <https://www.tutorialspoint.com/r/r_pie_charts.htm> - 02/04/2022
2. <https://moodle.ips.pt/2122/mod/resource/view.php?id=3386> - 26/03/2022
3. <https://moodle.ips.pt/2122/mod/resource/view.php?id=3408> - 08/03/2022
4. <https://moodle.ips.pt/2122/mod/resource/view.php?id=3410> - 15/03/2022