

Estatística Básica e Introdução ao R.

Aluno: Emerson Adam - Profa. Natália Giordani - Exercício CDA – Aula 6

Dataset: <https://www.kaggle.com/datasets/debajyotipodder/co2-emission-by-vehicles>

Pergunta 1: Entender melhor estes dados contextualizando.

Este dataset contém os detalhes de como a emissão de CO₂ dos veículos pode variar com diferentes características, como número de cilindros, tamanho do motor, tipo de combustível, a fabricante e o modelo, e algumas variáveis quantitativas de consumo. Os dados compilados deste dataset foram extraídos do Governo do Canada e refletem um período de 7 anos de coleta. Referencia do site original dos dados: [LINK](#)

Existem 7385 linhas com 12 diferentes atributos, 42 marcas, 2053 veículos, com 16 classificações de chassi, 27 tipos de transmissão e 5 tipos de combustível.

Realizando a correlação linear entre tamanho do motor e emissão de CO₂, nota-se que há uma correlação de 0,87 utilizando-se o método *spearman* devido aos outliers.

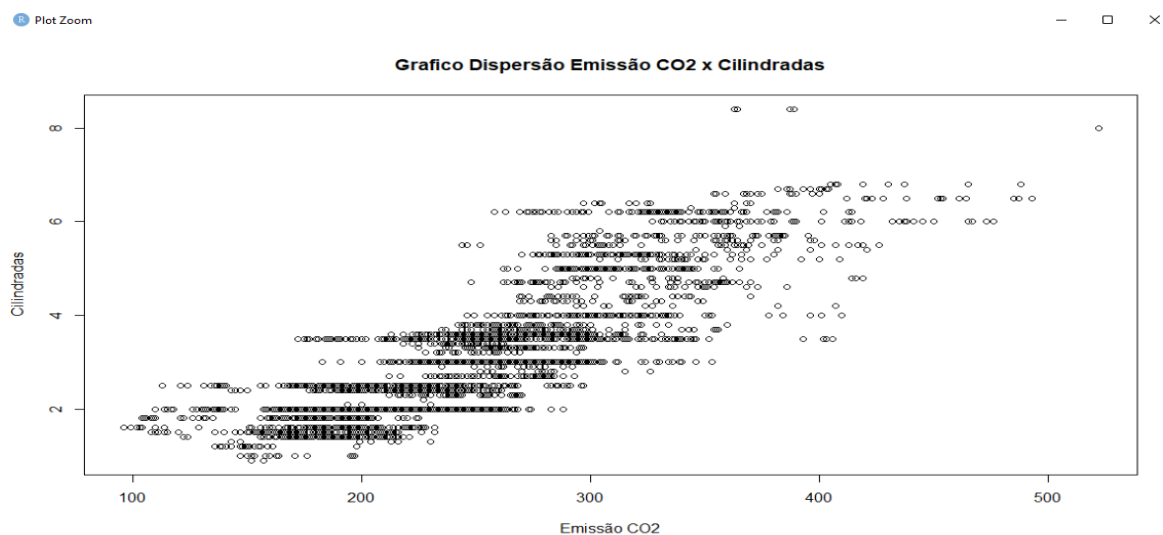
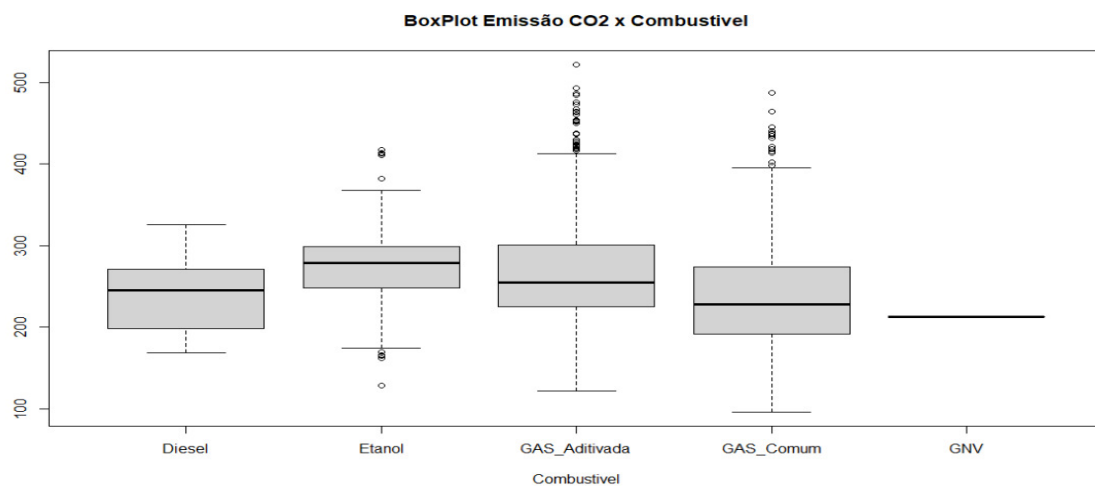


Gráfico de dispersão entre emissão de CO₂ e Tamanho do motor

Ao analisarmos a variável quantitativa Emissão Co₂ e a variável qualitativa combustível, nota-se que o etanol é a mais poluente neste dataset com a mediana em 275,1 gramas/KM contra 237,5 gramas/KM do Diesel.



BoxPlot de entre combustível e emissão de CO₂

Pergunta 2: Existe diferença na proporção de carros automáticos entre Ford e Chevrolet?

#H0 a proporção é = entre as 2 marcas

#H1 a proporção é != entre as 2 marcas

Para resolver esta questão, os 27 tipos de transmissão foram reclassificados em somente 2, manual ou automático, para o universo filtrado Chevrolet e Ford. Resultando em 1216 linhas, onde somente 174 são manuais e os 1042 restantes sendo automáticos num total de 230 modelos únicos de veículos.

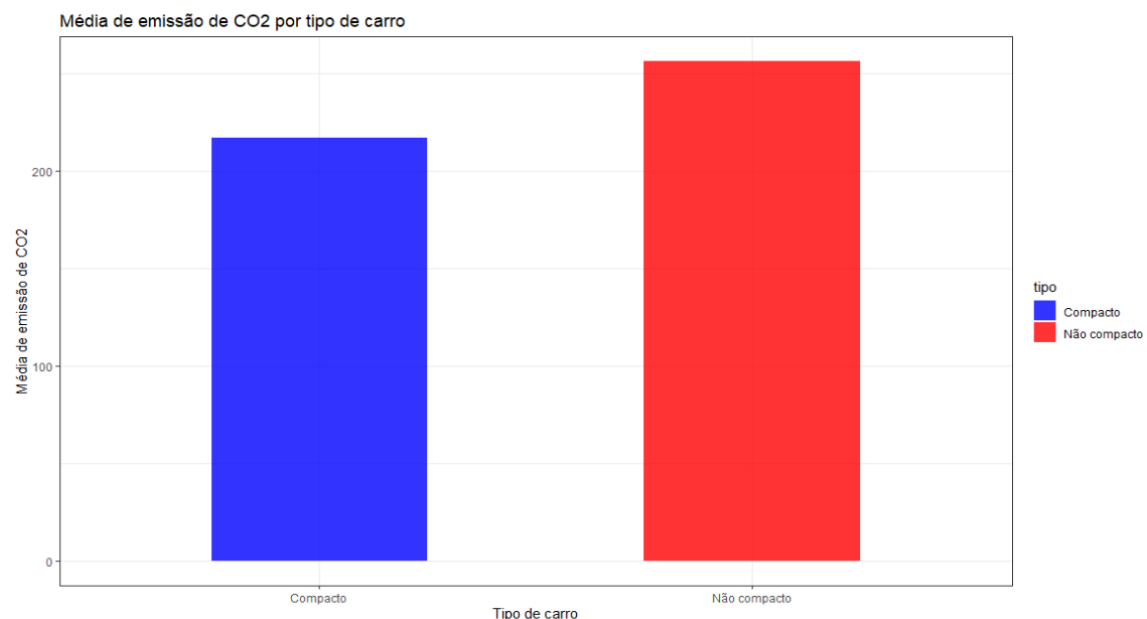
Utilizando o teste de hipótese entre os carros automáticos das 2 marcas, encontramos um valor de “p” igual a 0.98, o que indica que não há evidência estatística significativa para rejeitar a H0 de que as proporções de carros automáticos entre FORD e Chevrolet são iguais. **Isso significa que não há diferença significativa na proporção de carros automáticos entre as duas marcas e que H1 é válida.**

Pergunta 3: Existe correlação entre consumo de combustível (rodovia) e emissão de CO2 considerando os veículos do tipo COMPACT?

H0 não existe relação, hipótese nula sem valor e H1 existe relação

Para realizar esta análise de correlação, precisamos filtrar o dataset somente os veículos do tipo compactos. Após isto, analisar se existe realmente a correlação. Temos um valor de correção de 0,95, o que significa que existe uma alta correlação entre o carro compacto e a emissão de CO2.

Porém analisar somente este grupo está errado visto que TODOS emitem CO2. Portanto é preciso comparar estatisticamente o grupo compacto com os demais grupos e neste caso encontramos um valor de P menor que 0.05, o que significa que estatisticamente existe uma diferença significativa ao verificar a média do grupo compacto com os demais, rejeitando assim H0. Há uma diferença de 39gr/KM de CO2 emitido.

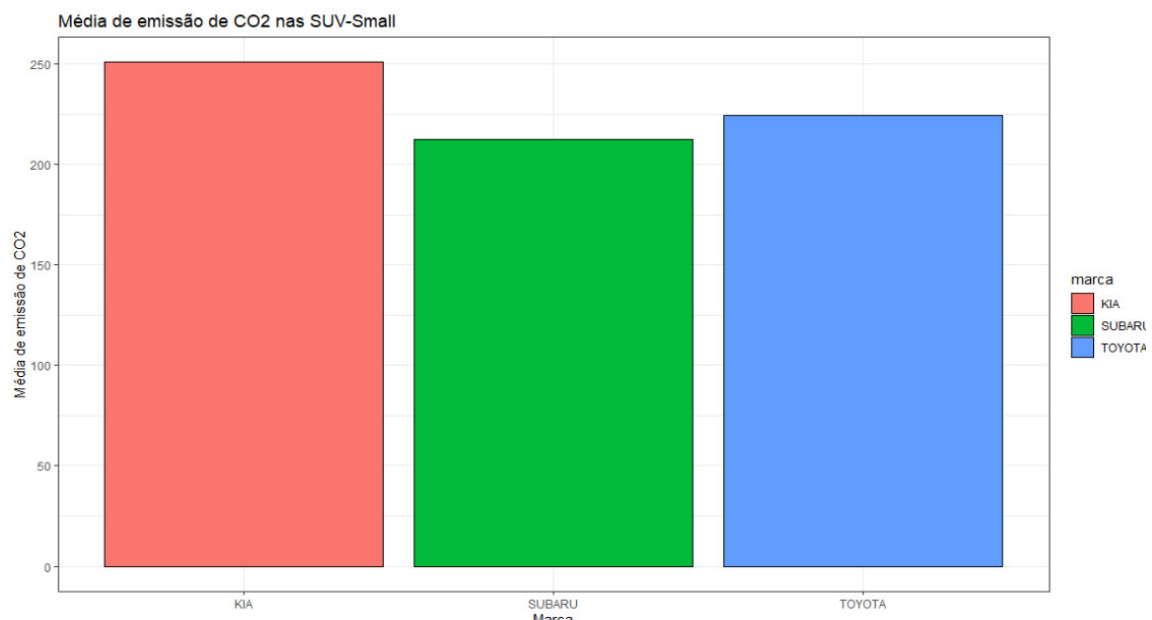


Pergunta 4: Existe diferença na quantidade média de emissão de CO2 nas SUV-Small das marcas KIA, SUBARU e TOYOTA?

H0 não há diferença, hipótese nula sem valor e H1 há diferença

Primeiro precisamos filtrar o DataSet somente nas marcas Kia, Subaru e Toyota e na categoria SUV-SMALL. Após isto realizar o teste de hipótese ANOVA(3 ou mais variáveis) para verificar se há diferença significativa entre as médias de emissão de cada grupo.

Os resultados obtidos de P são menores do que 0,05 e que podemos rejeitar H0, logo há diferença significativa entre as médias de emissão de cada marca, sendo que a ordem do menor para o maior seria, Kia, Toyota e Subaru.



No teste Tukey, encontramos P de 0,051 somente para o grupo Toyota/Subaru, o que significa que H0 é válido e que não há diferença, nos demais grupos, Subaru/Kia e Toyota/Kia a diferença é menor que 0,05, logo rejeitamos H0 e validamos H1, significando que há diferença entre as médias de emissão.