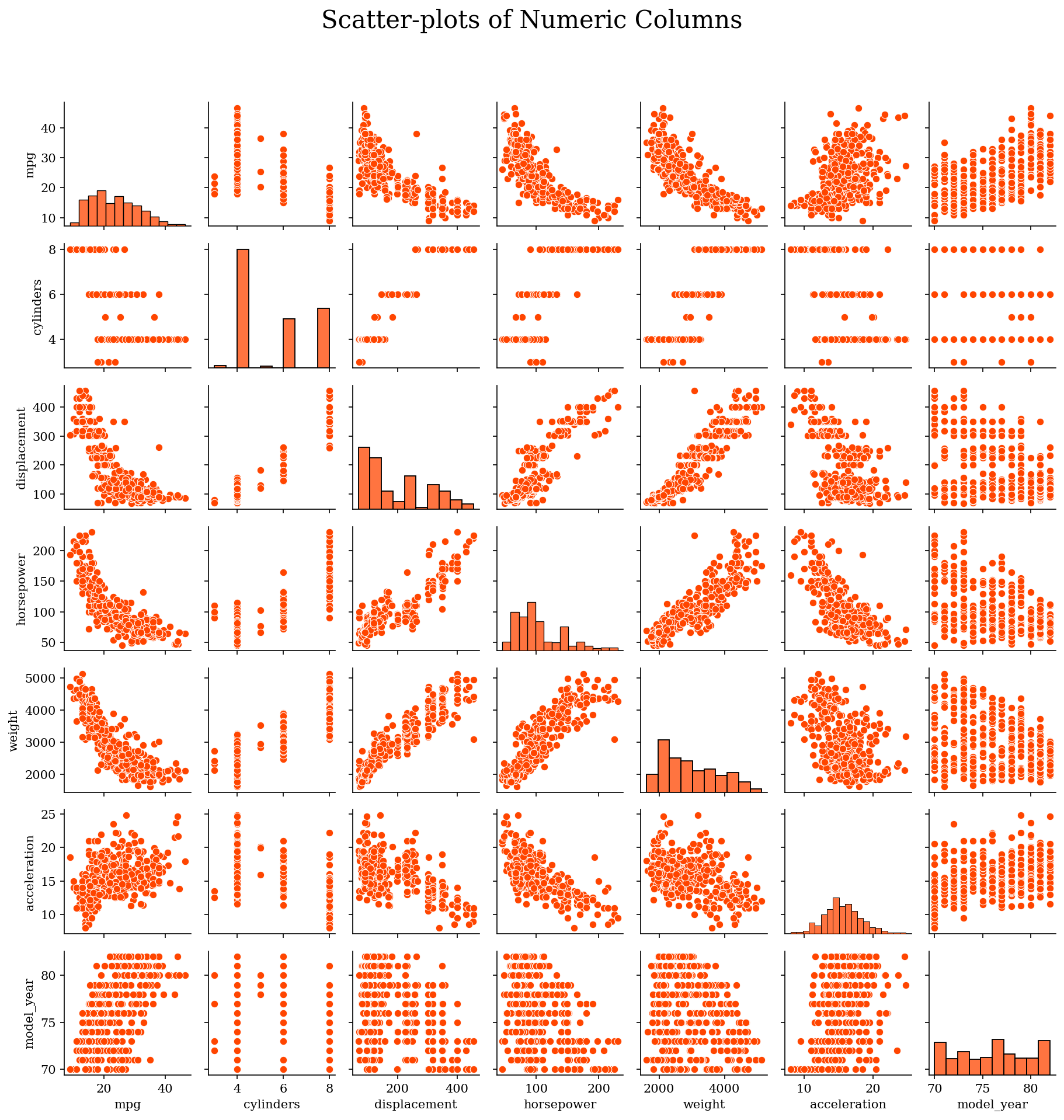
Exploratory Data Analysis Report

The dataset consists of 398 rows (observations) and 9 columns (features), 7 of which are numeric:

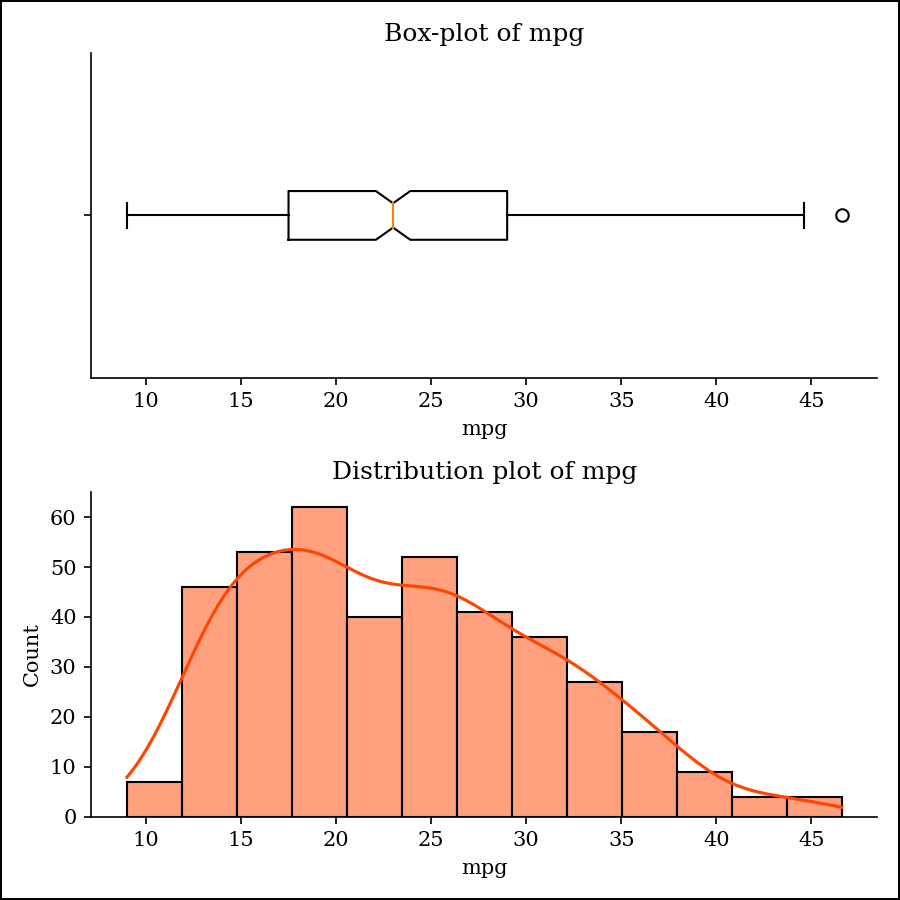


## 1. Mpg

**Mpg** is a numeric variable with 129 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 23.514572864321607 |
| Standard Deviation | 7.815984312565782 |
| Minimum | 9.0 |
| Lower Quartile | 17.5 |
| Median | 23.0 |
| Upper Quartile | 29.0 |
| Maximum | 46.6 |
| Skewness | 0.45706634399491913 |
| Kurtosis | -0.5107812652123154 |

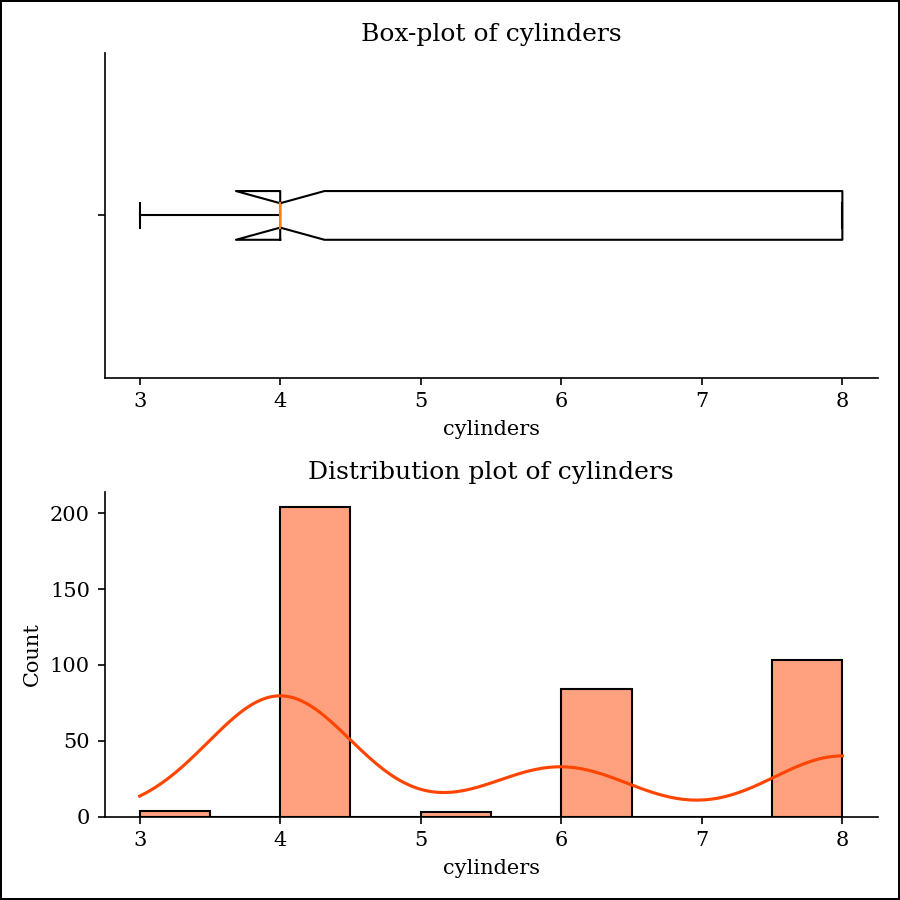


## 2. Cylinders

**Cylinders** is a numeric variable with 5 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 5.454773869346734 |
| Standard Deviation | 1.7010042445332119 |
| Minimum | 3.0 |
| Lower Quartile | 4.0 |
| Median | 4.0 |
| Upper Quartile | 8.0 |
| Maximum | 8.0 |
| Skewness | 0.5269215453528939 |
| Kurtosis | -1.376662176054571 |

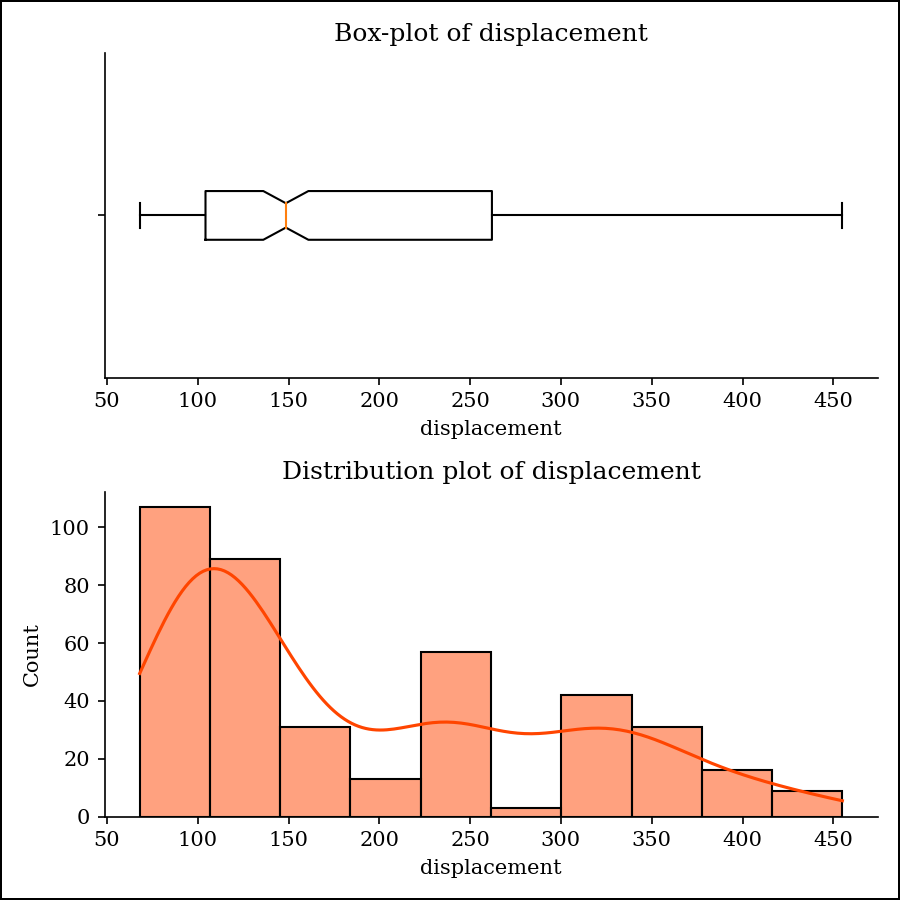


## 3. Displacement

**Displacement** is a numeric variable with 82 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 193.42587939698493 |
| Standard Deviation | 104.26983817119591 |
| Minimum | 68.0 |
| Lower Quartile | 104.25 |
| Median | 148.5 |
| Upper Quartile | 262.0 |
| Maximum | 455.0 |
| Skewness | 0.7196451643005952 |
| Kurtosis | -0.7465966295967594 |

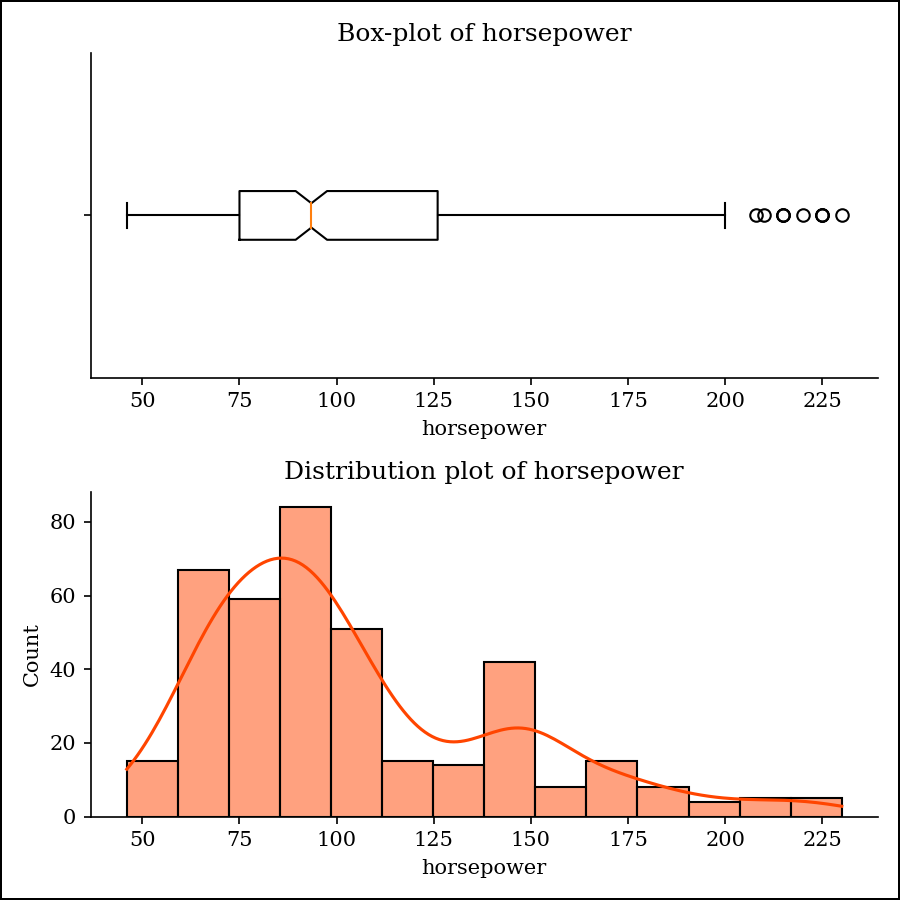


## 4. Horsepower

**Horsepower** is a numeric variable with 93 unique values. 6 (1.51%) of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 392.0 |
| Average | 104.46938775510205 |
| Standard Deviation | 38.49115993282849 |
| Minimum | 46.0 |
| Lower Quartile | 75.0 |
| Median | 93.5 |
| Upper Quartile | 126.0 |
| Maximum | 230.0 |
| Skewness | 1.0873262824048695 |
| Kurtosis | 0.696946999742821 |

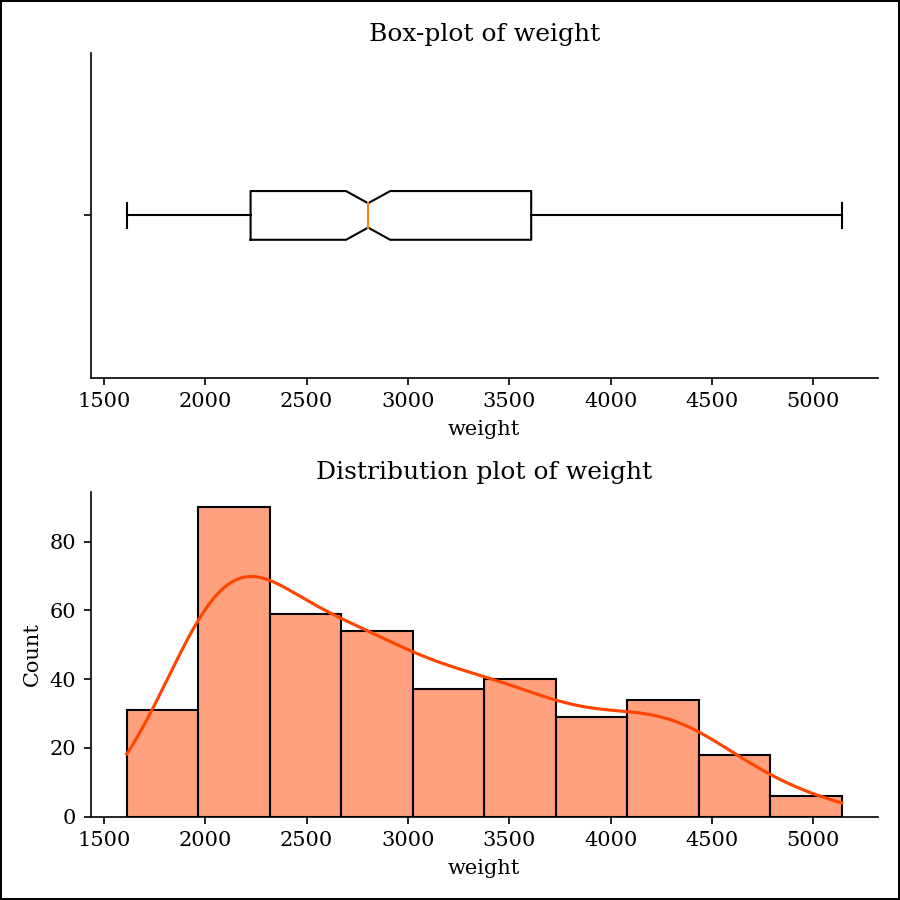


## 5. Weight

**Weight** is a numeric variable with 351 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 2970.424623115578 |
| Standard Deviation | 846.8417741973268 |
| Minimum | 1613.0 |
| Lower Quartile | 2223.75 |
| Median | 2803.5 |
| Upper Quartile | 3608.0 |
| Maximum | 5140.0 |
| Skewness | 0.5310625125994629 |
| Kurtosis | -0.7855289051011454 |

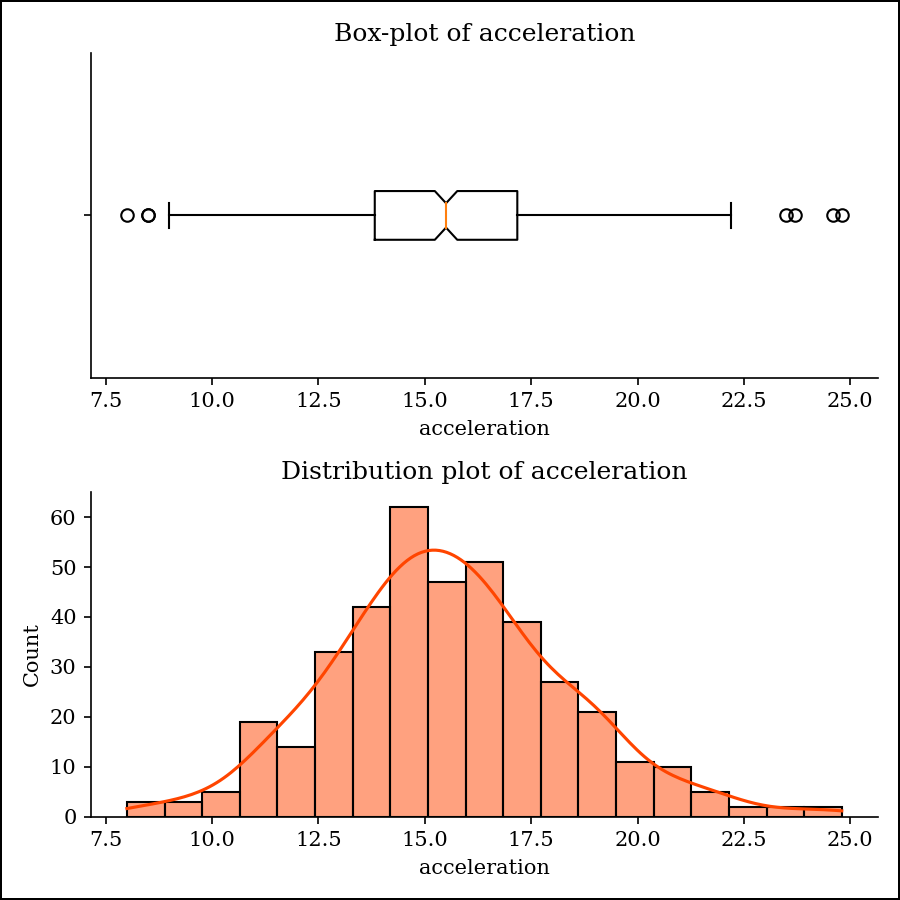


## 6. Acceleration

**Acceleration** is a numeric variable with 95 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 15.568090452261307 |
| Standard Deviation | 2.757688929812676 |
| Minimum | 8.0 |
| Lower Quartile | 13.825000000000001 |
| Median | 15.5 |
| Upper Quartile | 17.175 |
| Maximum | 24.8 |
| Skewness | 0.27877684462588986 |
| Kurtosis | 0.4194968829706842 |

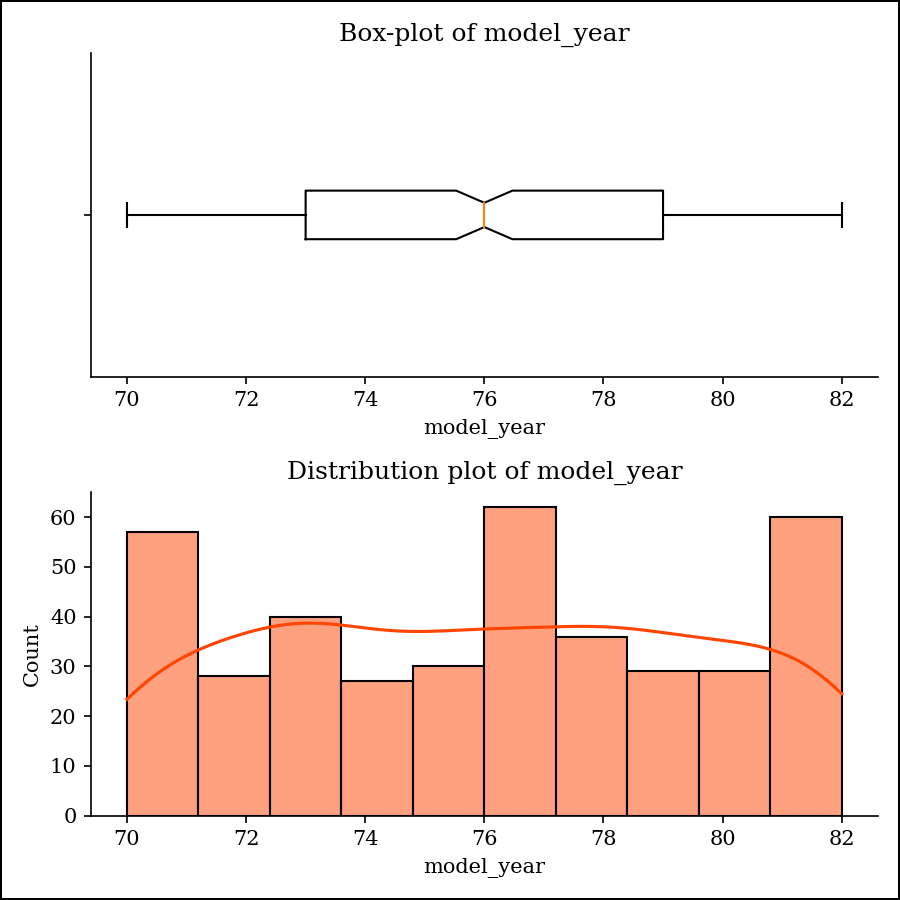


## 7. Model\_Year

**Model\_year** is a numeric variable with 13 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 76.01005025125629 |
| Standard Deviation | 3.697626646732623 |
| Minimum | 70.0 |
| Lower Quartile | 73.0 |
| Median | 76.0 |
| Upper Quartile | 79.0 |
| Maximum | 82.0 |
| Skewness | 0.01153459401509278 |
| Kurtosis | -1.1812317432191861 |



## 8. Origin

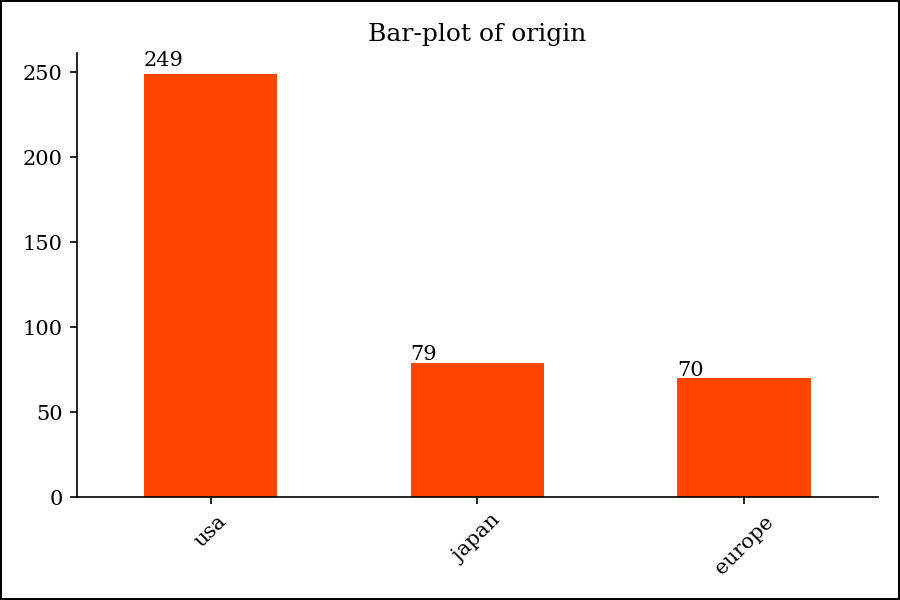
**Origin** is a categorical variable with 3 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398 |
| Unique values | 3 |
| Mode (Highest occurring value) | usa |

#### Most Common Values

|  |  |
| --- | --- |
| usa | 249 (62.56%) |
| japan | 79 (19.85%) |
| europe | 70 (17.59%) |



## 9. Name

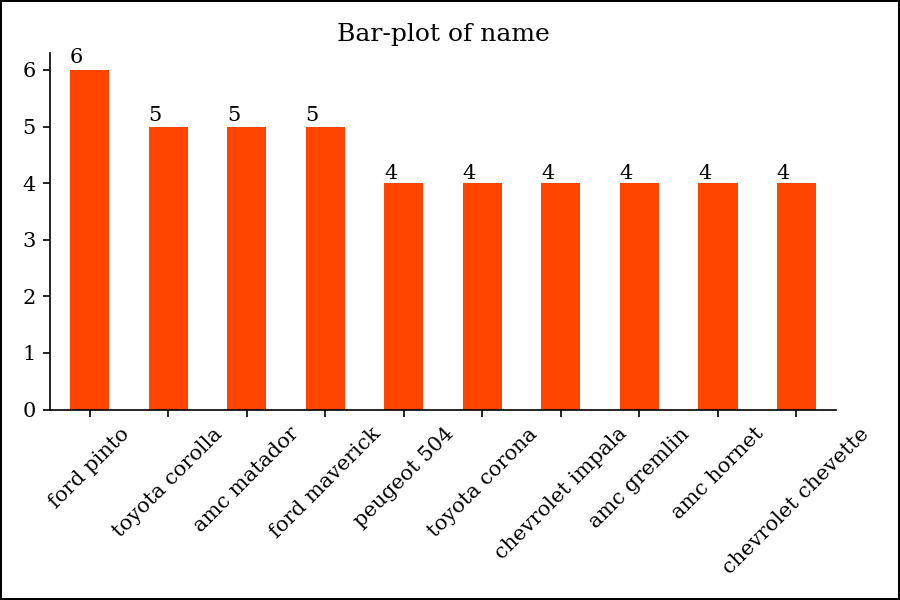
**Name** is a categorical variable with 305 unique values. None of its values are missing.

#### Summary Statistics

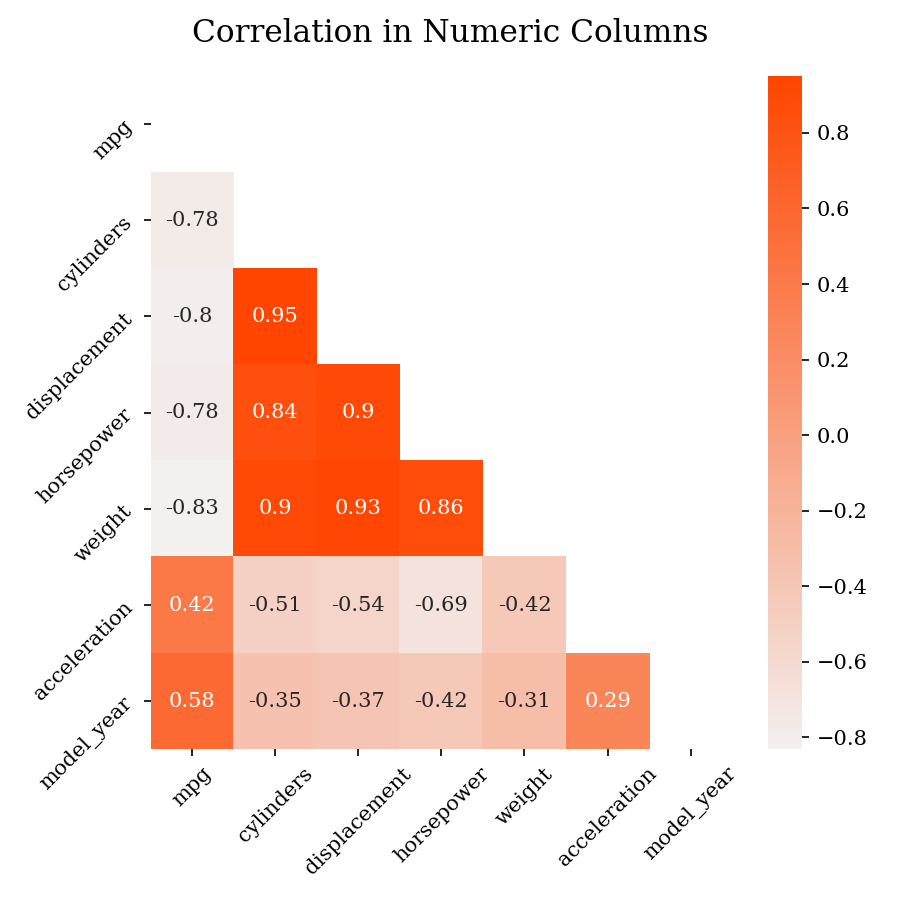
|  |  |
| --- | --- |
| Number of observations | 398 |
| Unique values | 305 |
| Mode (Highest occurring value) | ford pinto |

#### Most Common Values

|  |  |
| --- | --- |
| ford pinto | 6 (1.51%) |
| toyota corolla | 5 (1.26%) |
| amc matador | 5 (1.26%) |
| ford maverick | 5 (1.26%) |
| peugeot 504 | 4 (1.01%) |

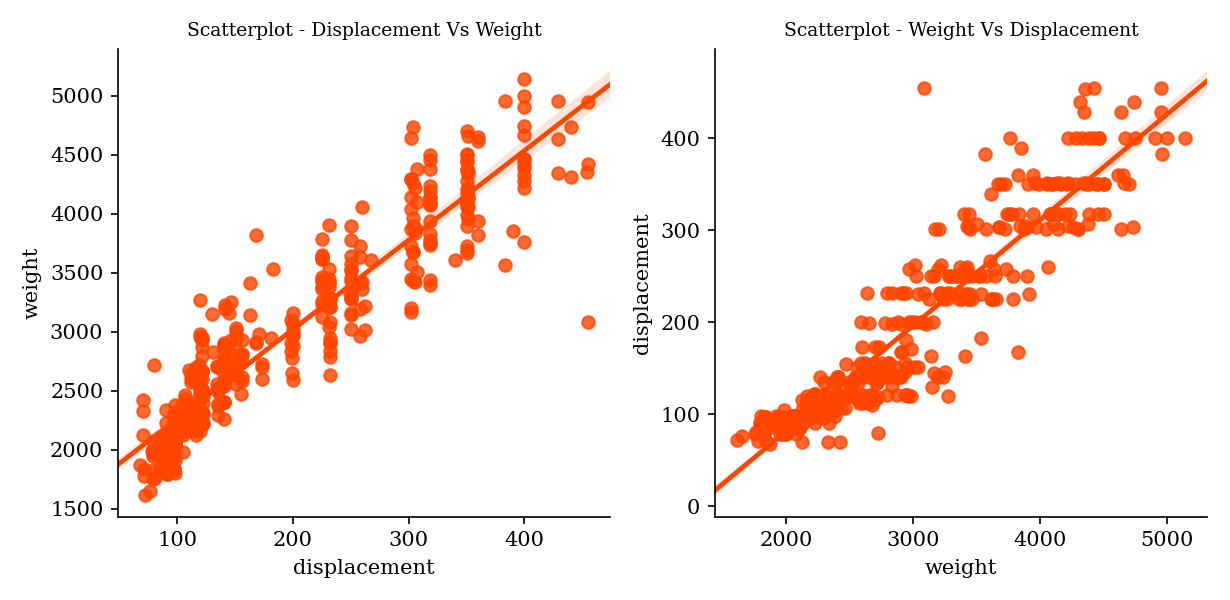


# Bivariate Analysis (Correlation)



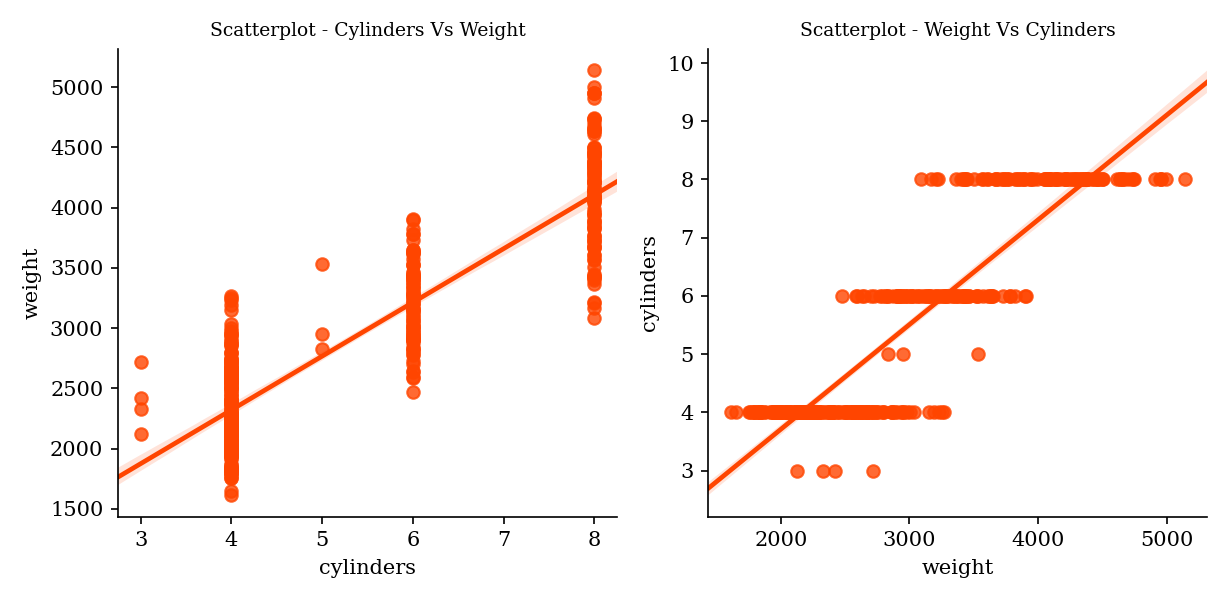
## 1. Displacement vs Weight

**Displacement** and **Weight** have very strong positive correlation (0.93).



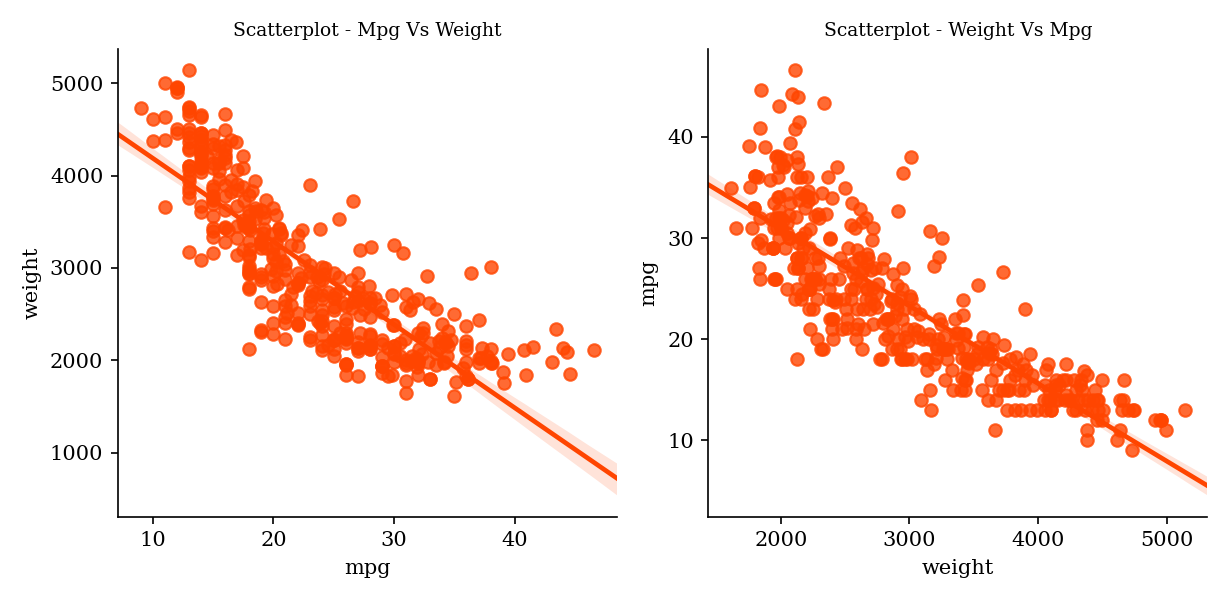
## 2. Cylinders vs Weight

**Cylinders** and **Weight** have strong positive correlation (0.90).



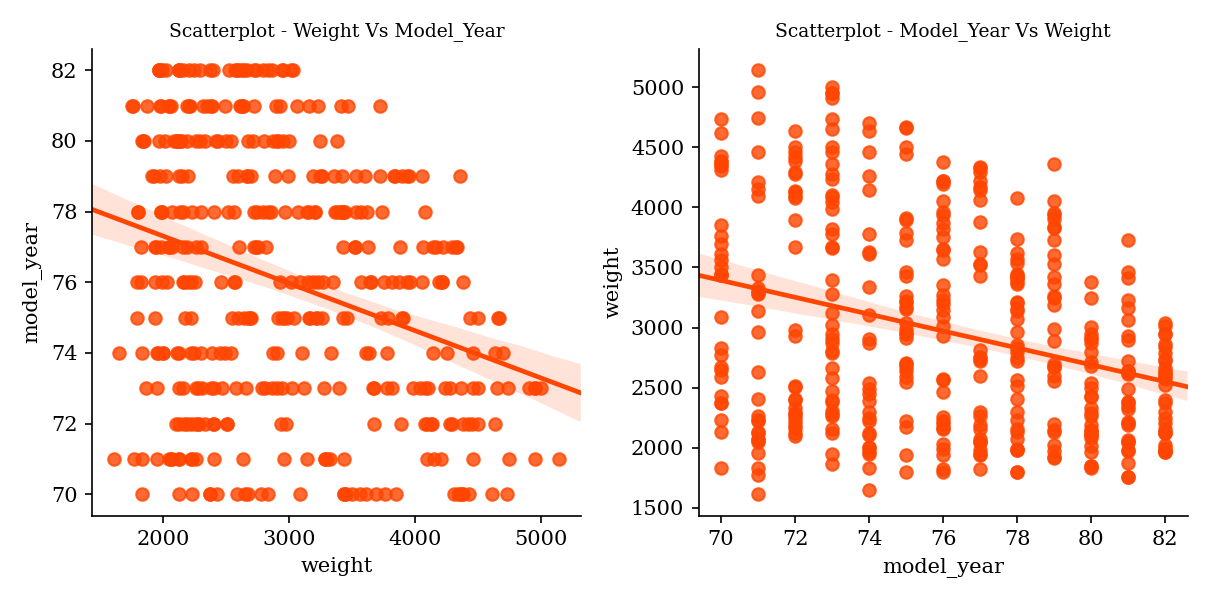
## 3. Mpg vs Weight

**Mpg** and **Weight** have strong negative correlation (-0.83).



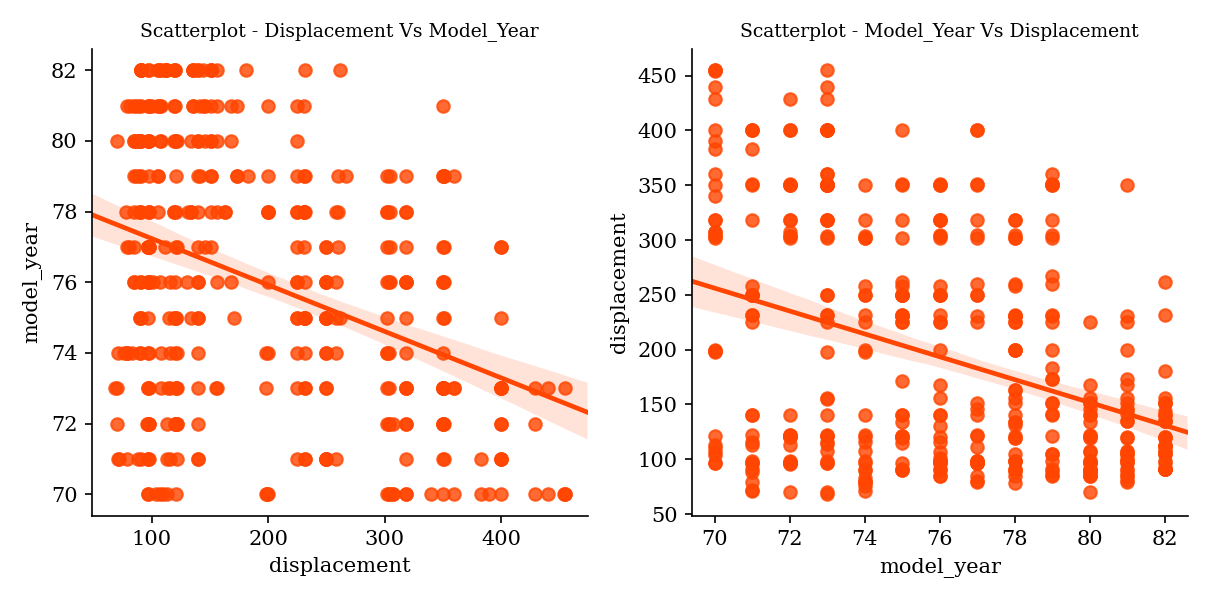
## 4. Weight vs Model\_Year

**Weight** and **Model\_year** have weak negative correlation (-0.31).



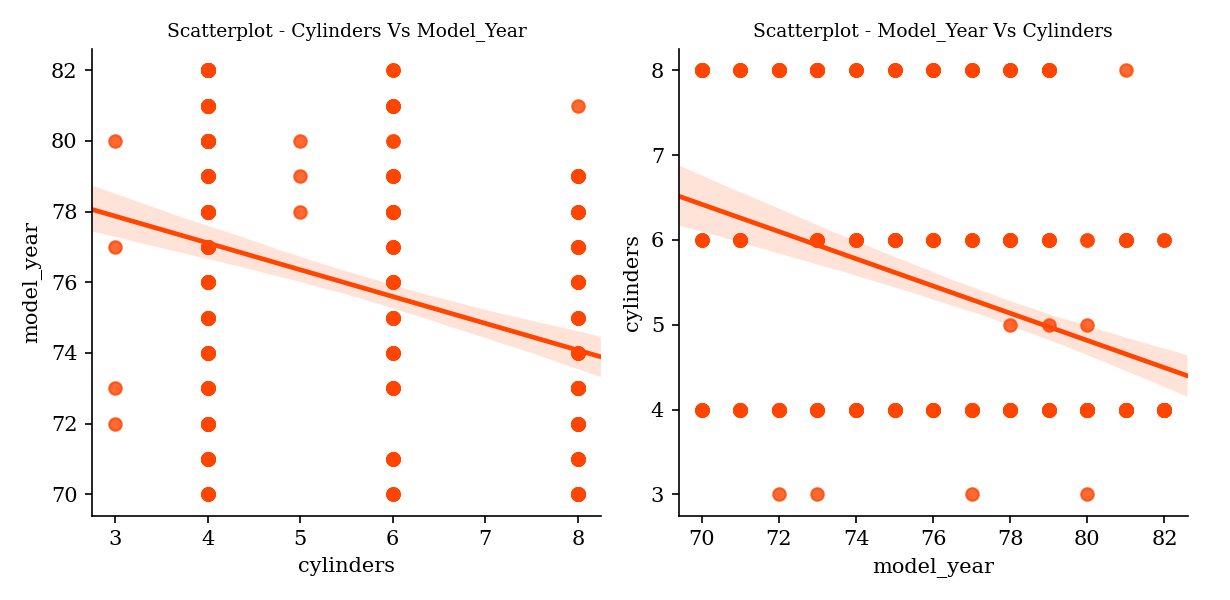
## 5. Displacement vs Model\_Year

**Displacement** and **Model\_year** have weak negative correlation (-0.37).



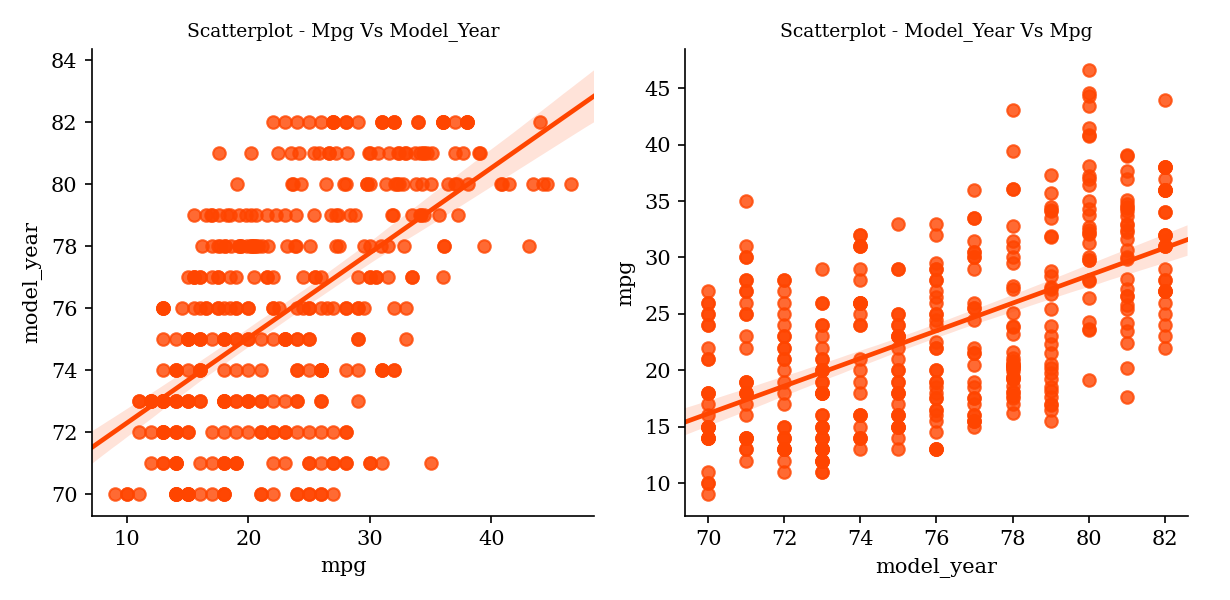
## 6. Cylinders vs Model\_Year

**Cylinders** and **Model\_year** have weak negative correlation (-0.35).



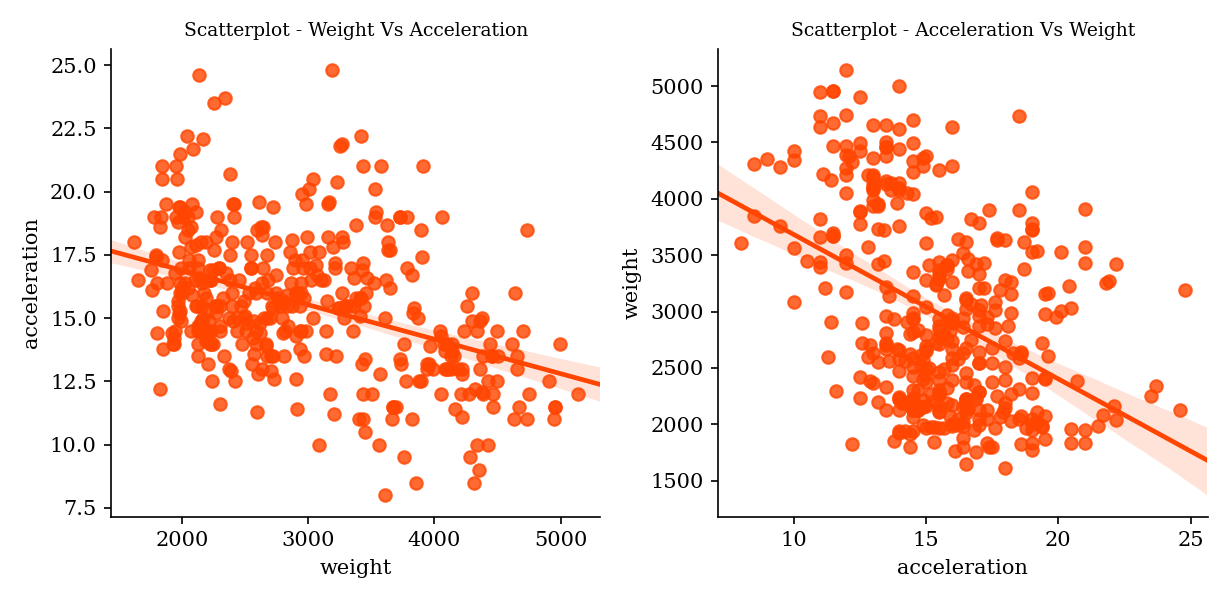
## 7. Mpg vs Model\_Year

**Mpg** and **Model\_year** have moderate positive correlation (0.58).



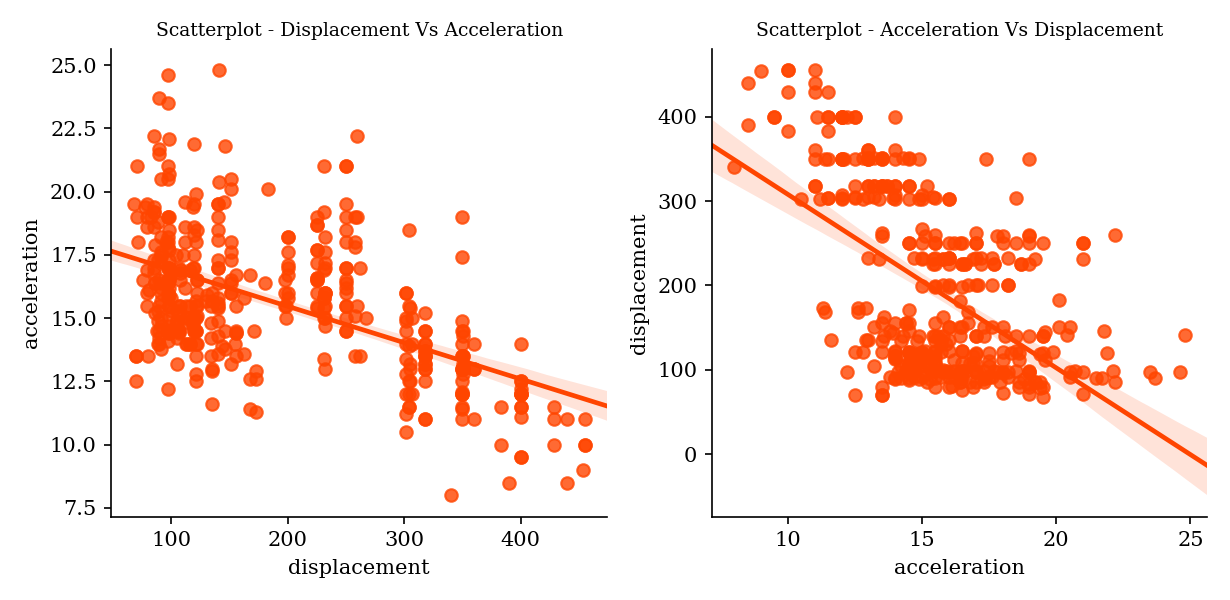
## 8. Weight vs Acceleration

**Weight** and **Acceleration** have weak negative correlation (-0.42).



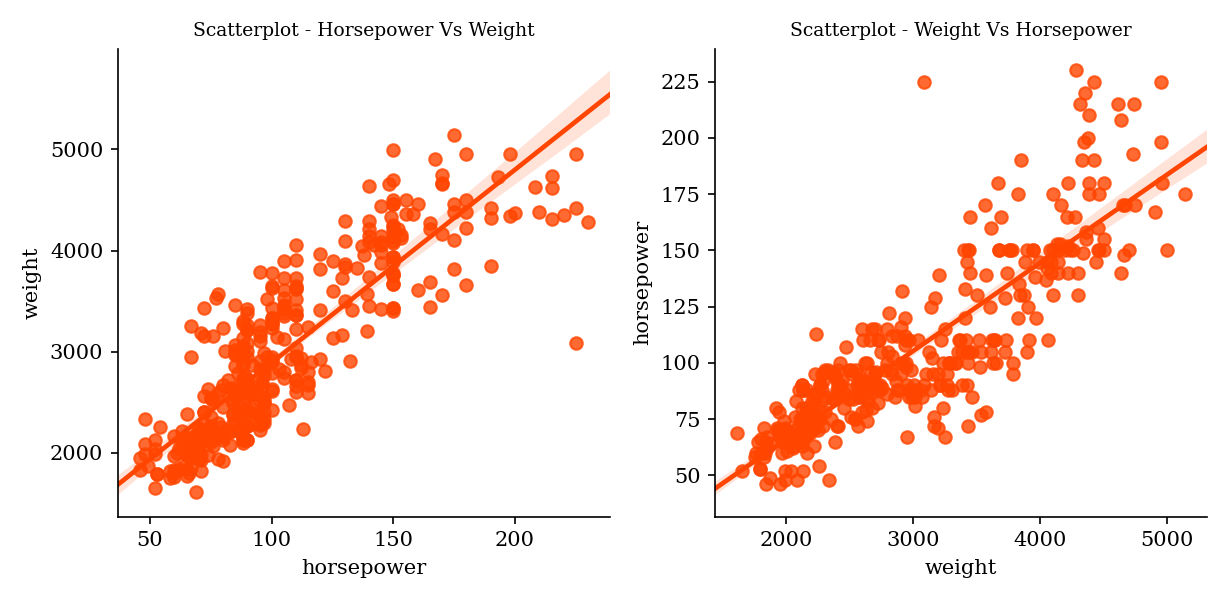
## 9. Displacement vs Acceleration

**Displacement** and **Acceleration** have moderate negative correlation (-0.54).



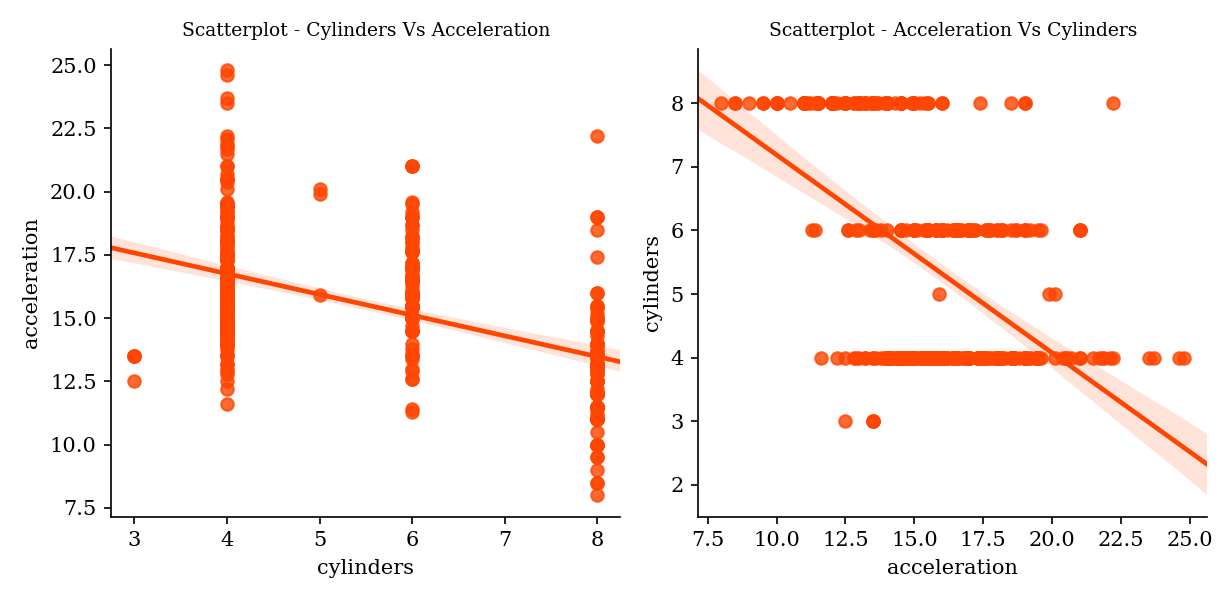
## 10. Horsepower vs Weight

**Horsepower** and **Weight** have strong positive correlation (0.86).



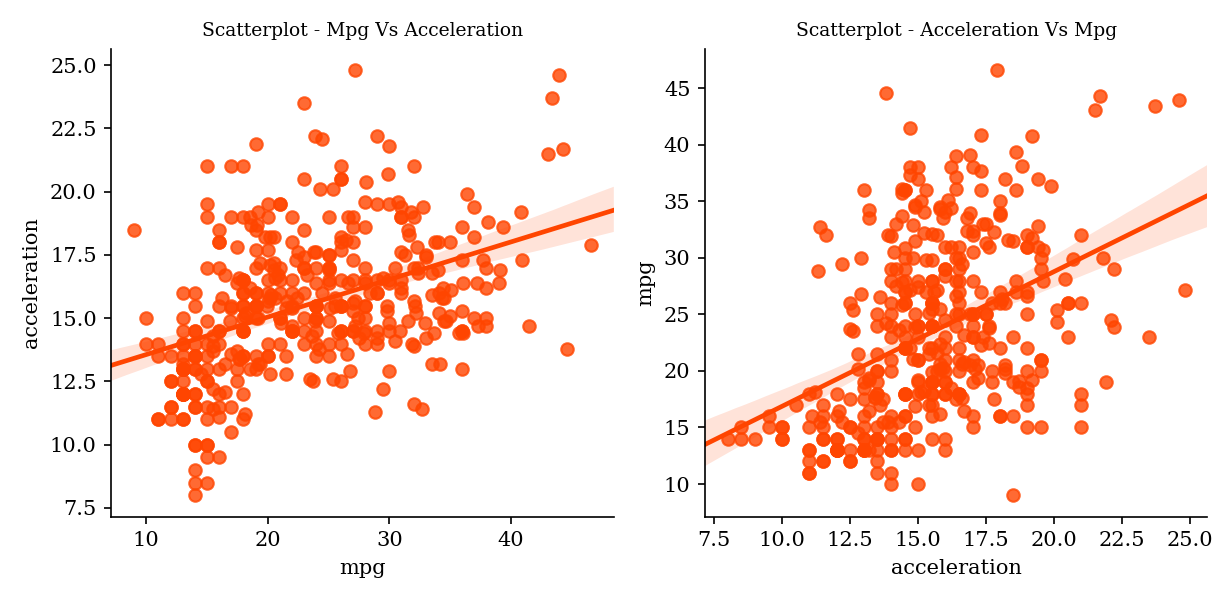
## 11. Cylinders vs Acceleration

**Cylinders** and **Acceleration** have moderate negative correlation (-0.51).



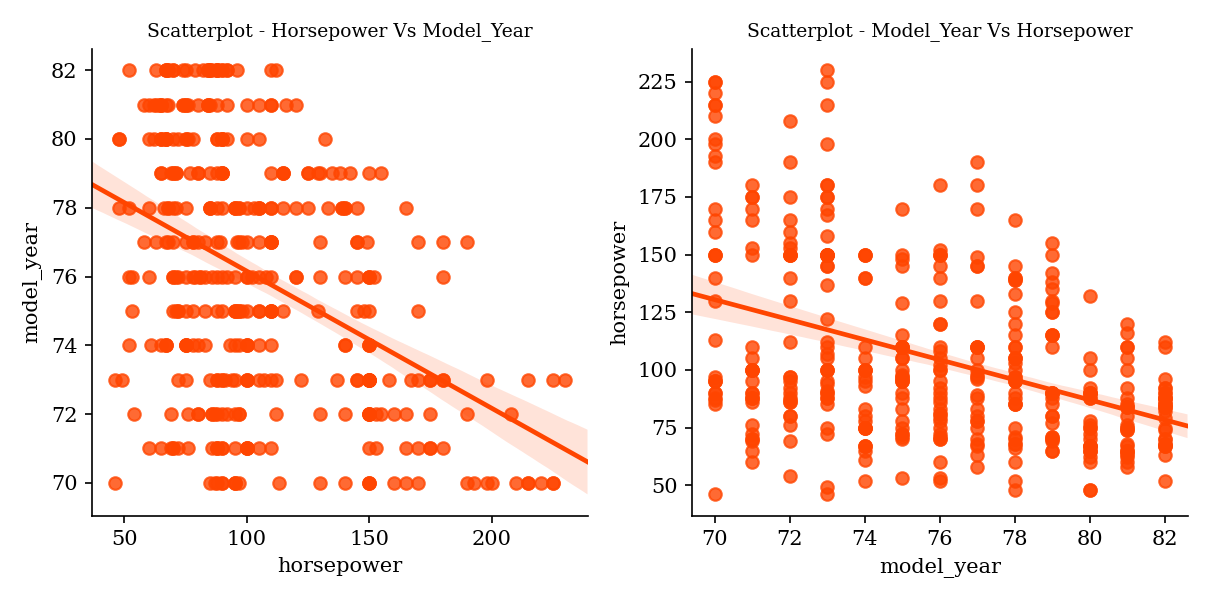
## 12. Mpg vs Acceleration

**Mpg** and **Acceleration** have weak positive correlation (0.42).



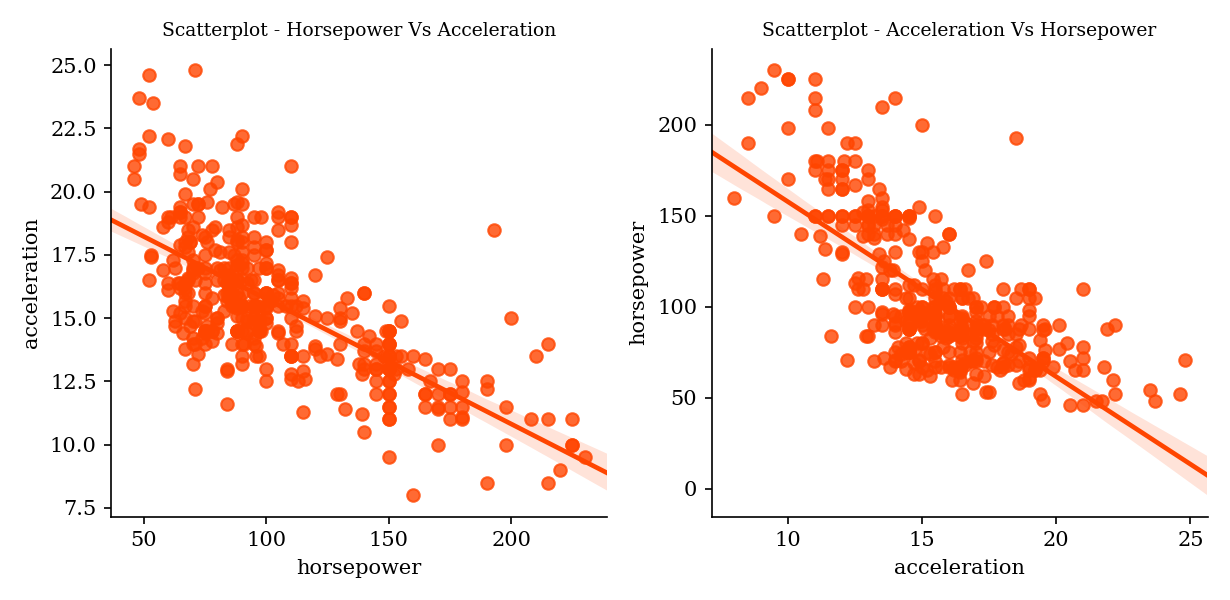
## 13. Horsepower vs Model\_Year

**Horsepower** and **Model\_year** have weak negative correlation (-0.42).



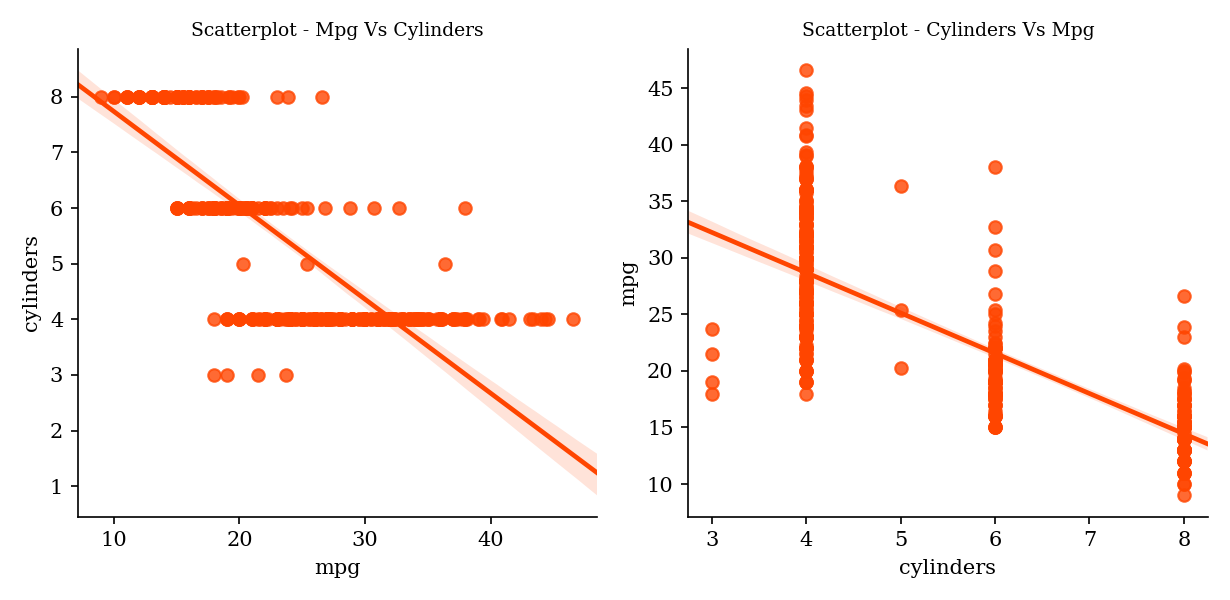
## 14. Horsepower vs Acceleration

**Horsepower** and **Acceleration** have moderate negative correlation (-0.69).



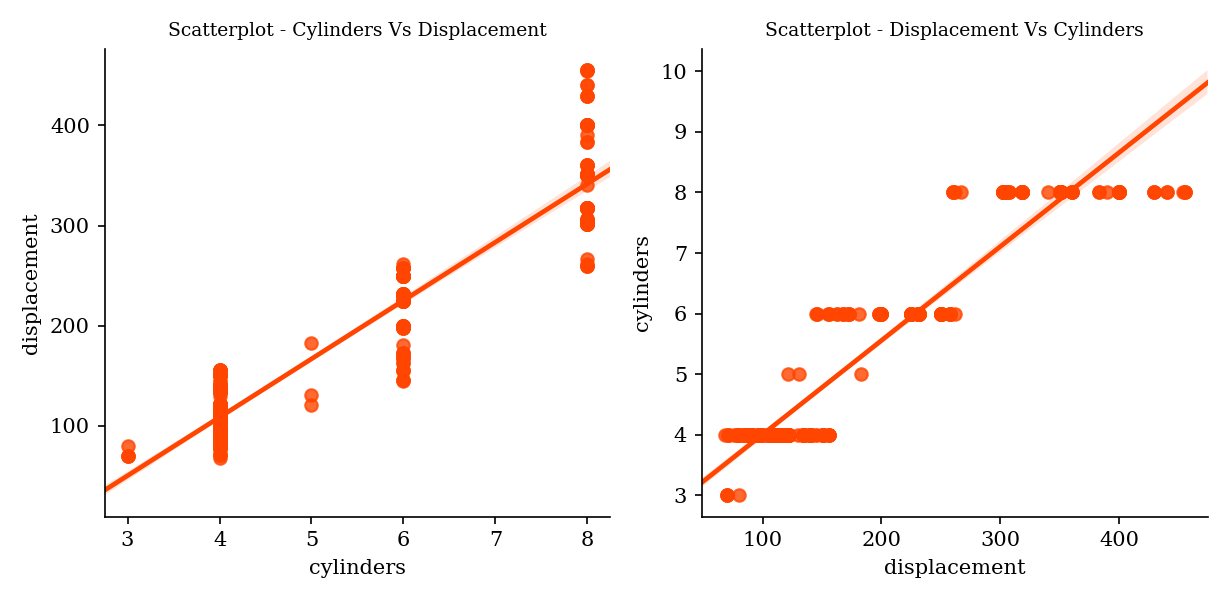
## 15. Mpg vs Cylinders

**Mpg** and **Cylinders** have strong negative correlation (-0.78).



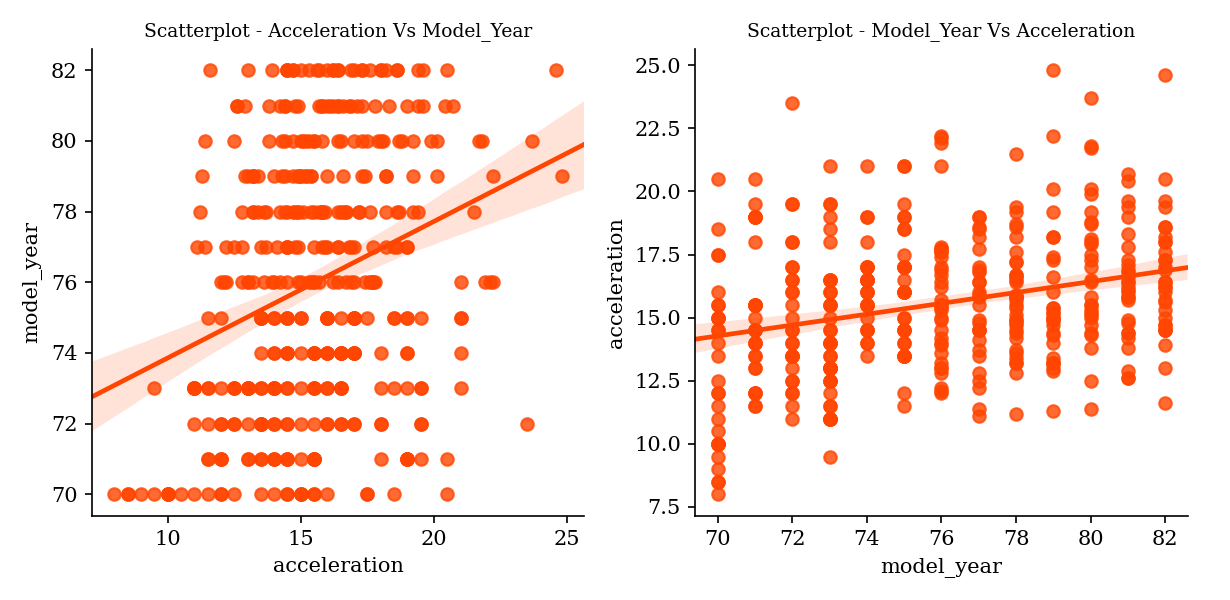
## 16. Cylinders vs Displacement

**Cylinders** and **Displacement** have very strong positive correlation (0.95).



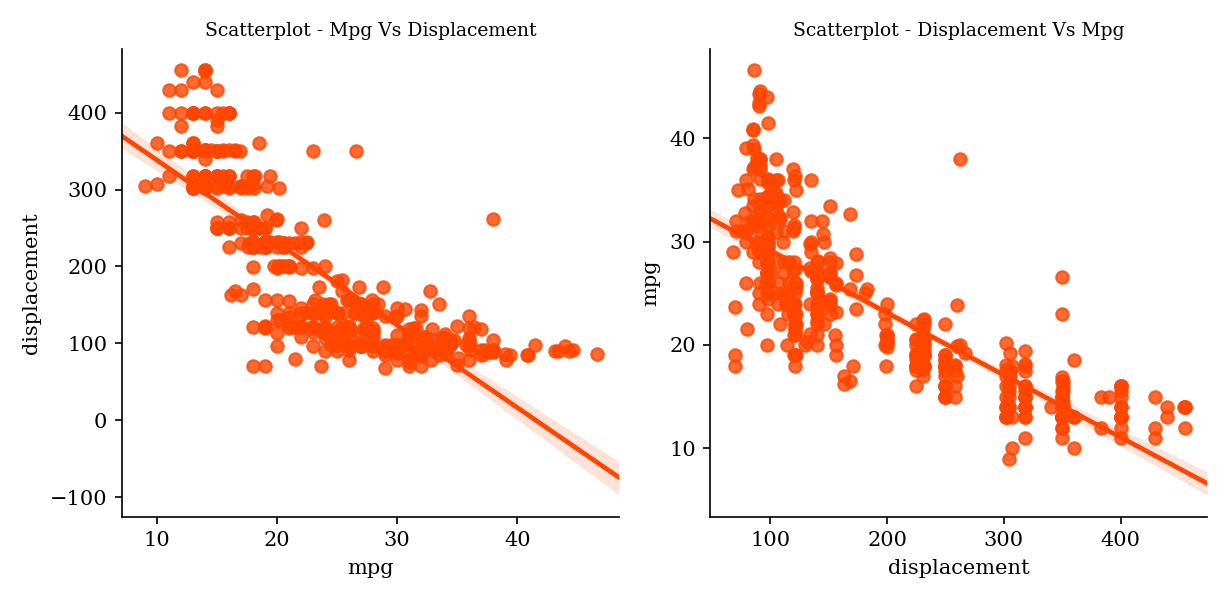
## 17. Acceleration vs Model\_Year

**Acceleration** and **Model\_year** have very weak positive correlation (0.29).



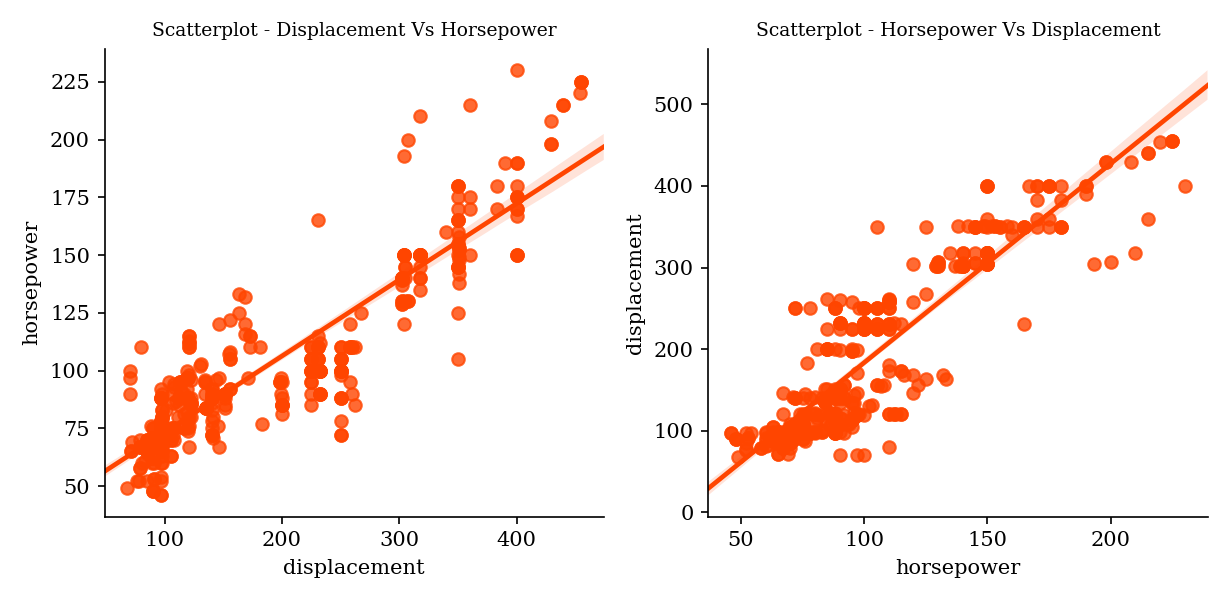
## 18. Mpg vs Displacement

**Mpg** and **Displacement** have strong negative correlation (-0.80).



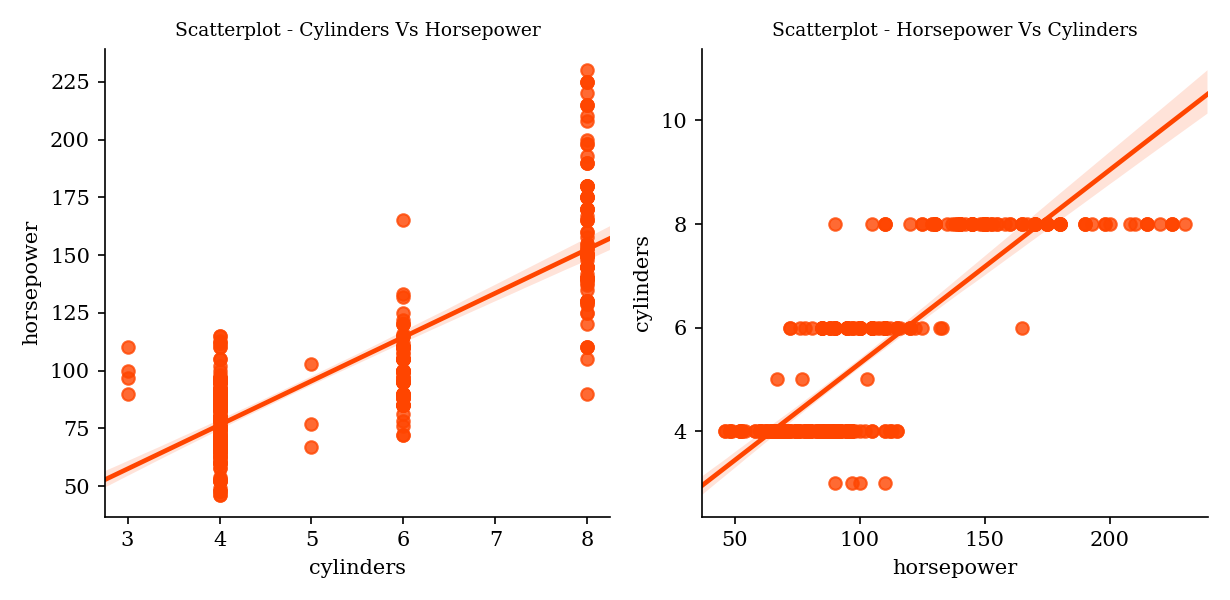
## 19. Displacement vs Horsepower

**Displacement** and **Horsepower** have strong positive correlation (0.90).



## 20. Cylinders vs Horsepower

**Cylinders** and **Horsepower** have strong positive correlation (0.84).



## 21. Mpg vs Horsepower

**Mpg** and **Horsepower** have strong negative correlation (-0.78).

