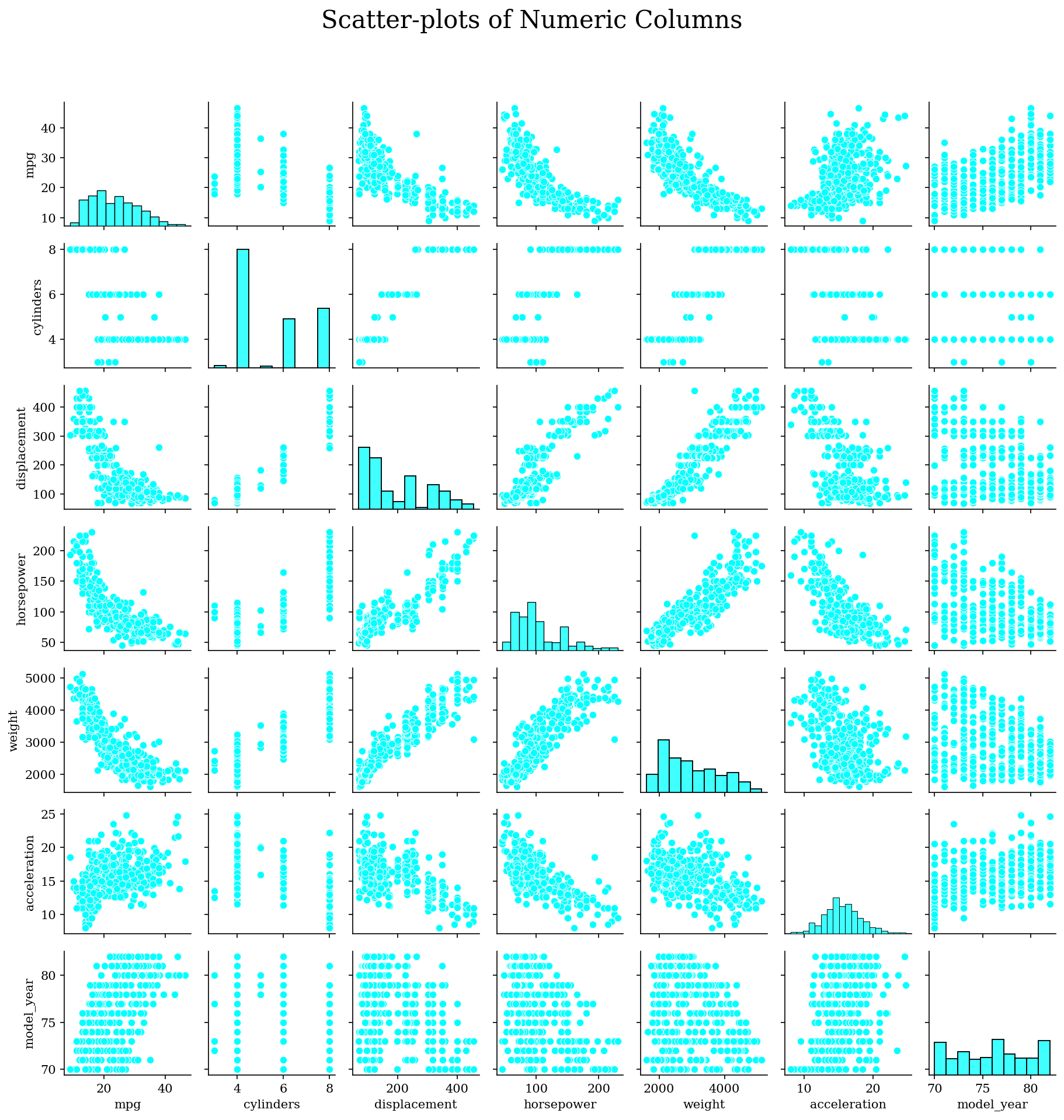
Exploratory Data Analysis Report - Auto MPG Data Set

The data contains 398 observations (rows) and 9 variables (columns), 7 of which are numeric:

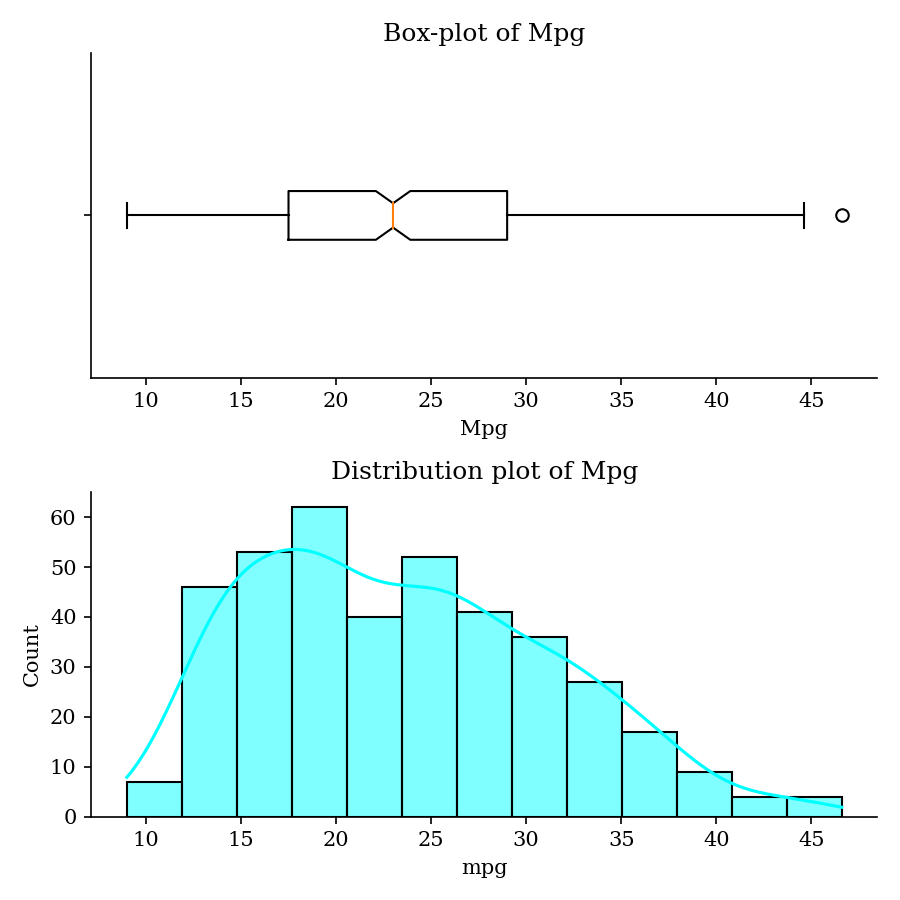


## 1. Mpg

*Mpg* has 129 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 23.5146 |
| Standard Deviation | 7.816 |
| Minimum | 9.0 |
| Lower Quartile | 17.5 |
| Median | 23.0 |
| Upper Quartile | 29.0 |
| Maximum | 46.6 |
| Skewness | 0.4571 |
| Kurtosis | -0.5108 |

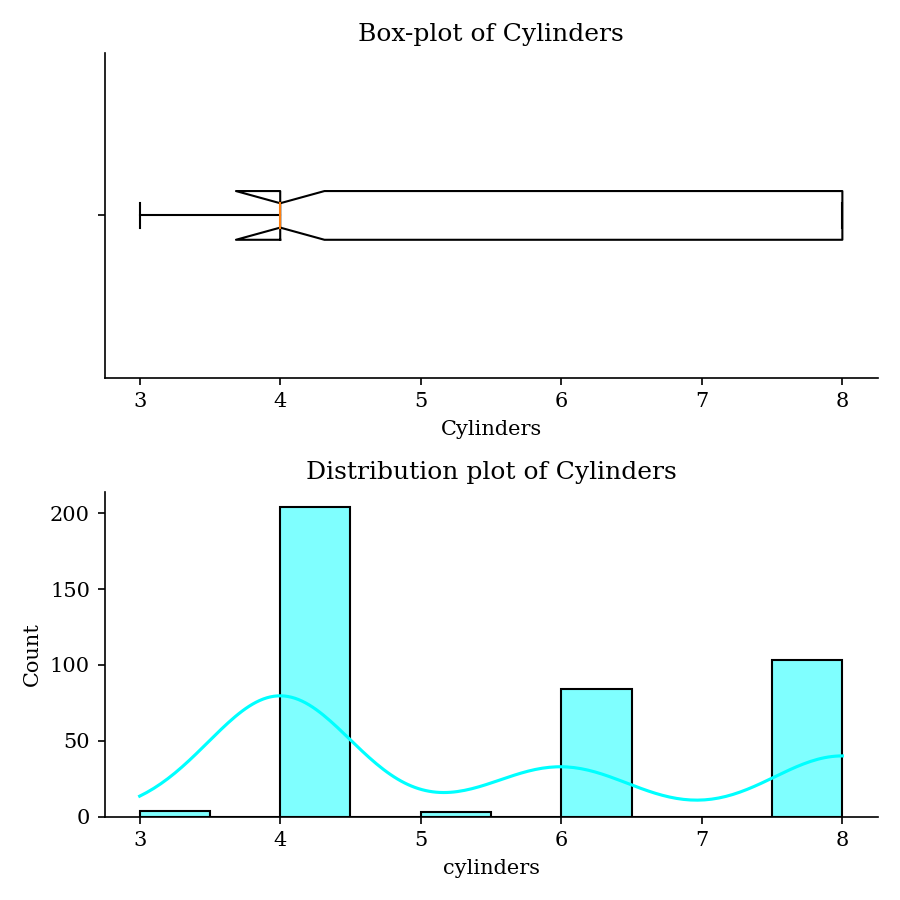


## 2. Cylinders

*Cylinders* has 5 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 5.4548 |
| Standard Deviation | 1.701 |
| Minimum | 3.0 |
| Lower Quartile | 4.0 |
| Median | 4.0 |
| Upper Quartile | 8.0 |
| Maximum | 8.0 |
| Skewness | 0.5269 |
| Kurtosis | -1.3767 |

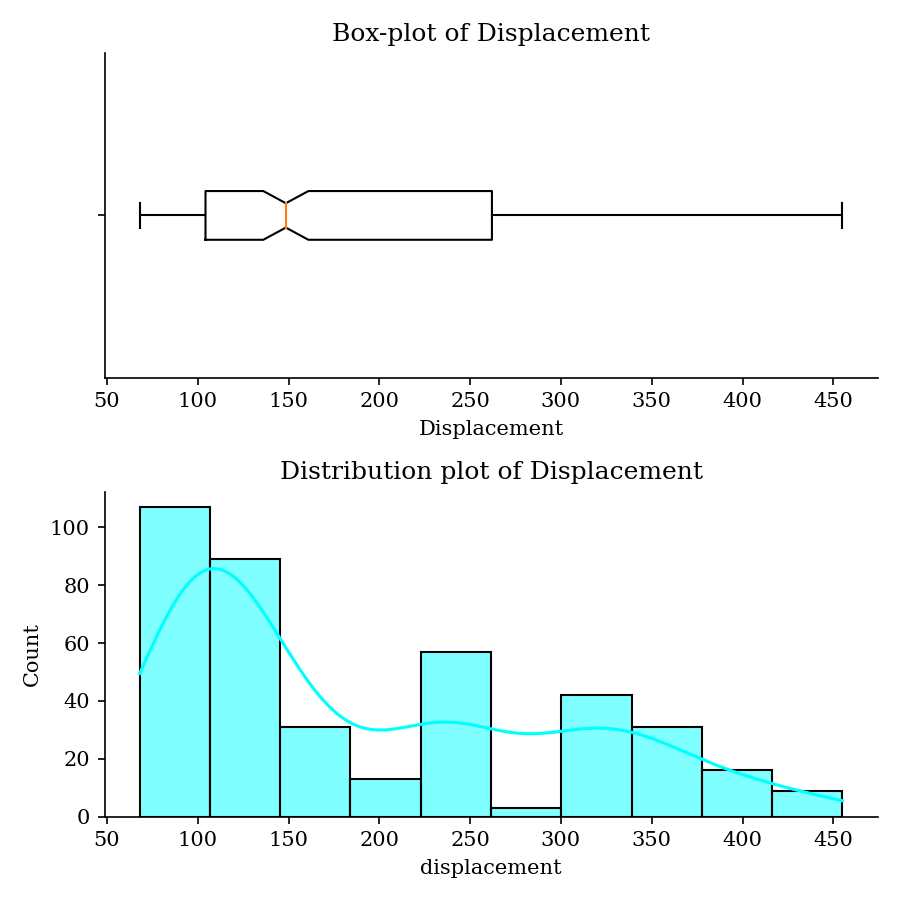


## 3. Displacement

*Displacement* has 82 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 193.4259 |
| Standard Deviation | 104.2698 |
| Minimum | 68.0 |
| Lower Quartile | 104.25 |
| Median | 148.5 |
| Upper Quartile | 262.0 |
| Maximum | 455.0 |
| Skewness | 0.7196 |
| Kurtosis | -0.7466 |

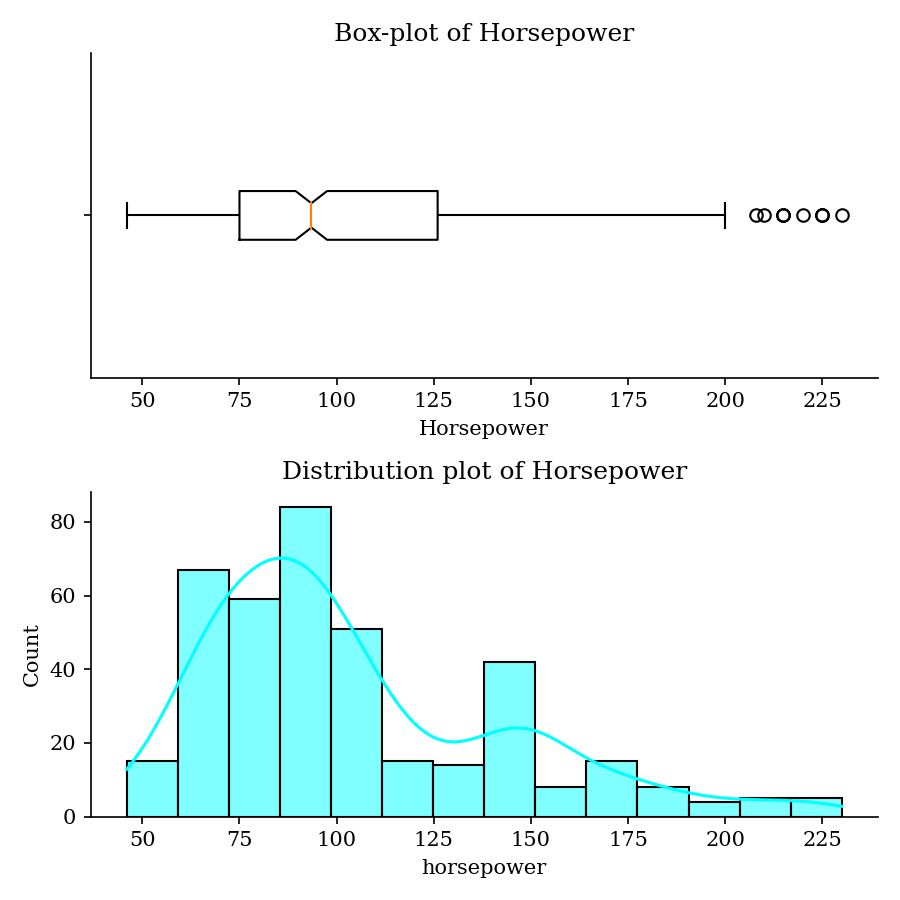


## 4. Horsepower

*Horsepower* has 93 unique values. 6 (1.51%) of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 392.0 |
| Average | 104.4694 |
| Standard Deviation | 38.4912 |
| Minimum | 46.0 |
| Lower Quartile | 75.0 |
| Median | 93.5 |
| Upper Quartile | 126.0 |
| Maximum | 230.0 |
| Skewness | 1.0873 |
| Kurtosis | 0.6969 |

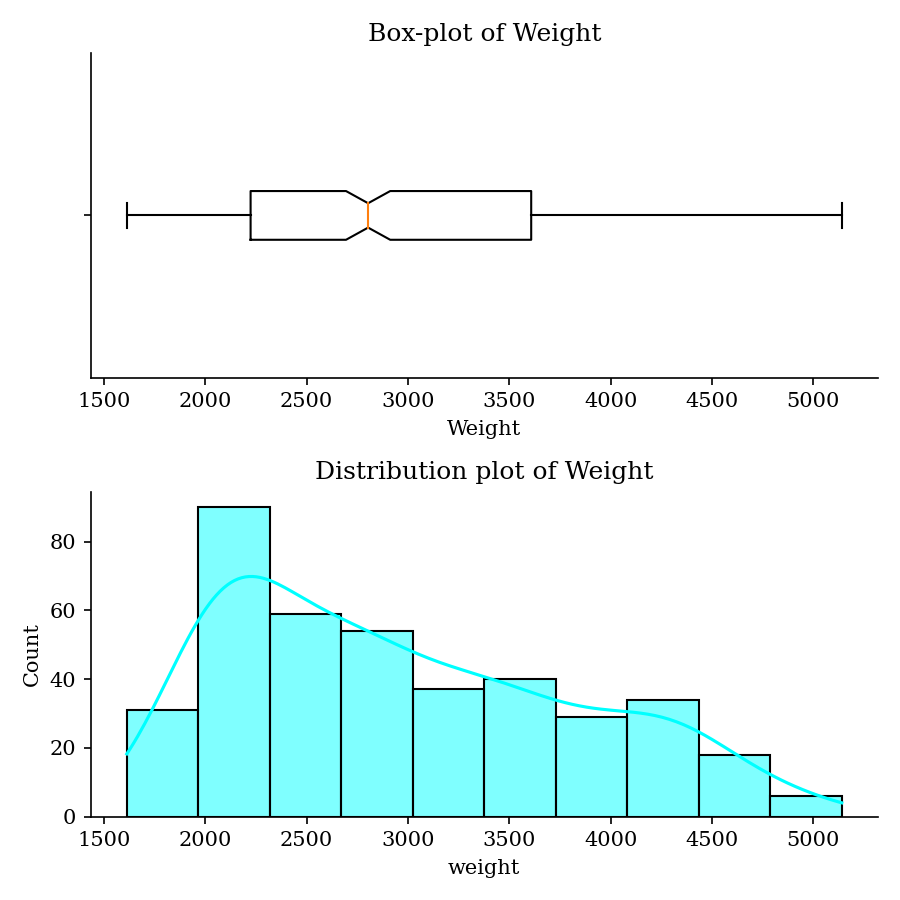


## 5. Weight

*Weight* has 351 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 2970.4246 |
| Standard Deviation | 846.8418 |
| Minimum | 1613.0 |
| Lower Quartile | 2223.75 |
| Median | 2803.5 |
| Upper Quartile | 3608.0 |
| Maximum | 5140.0 |
| Skewness | 0.5311 |
| Kurtosis | -0.7855 |

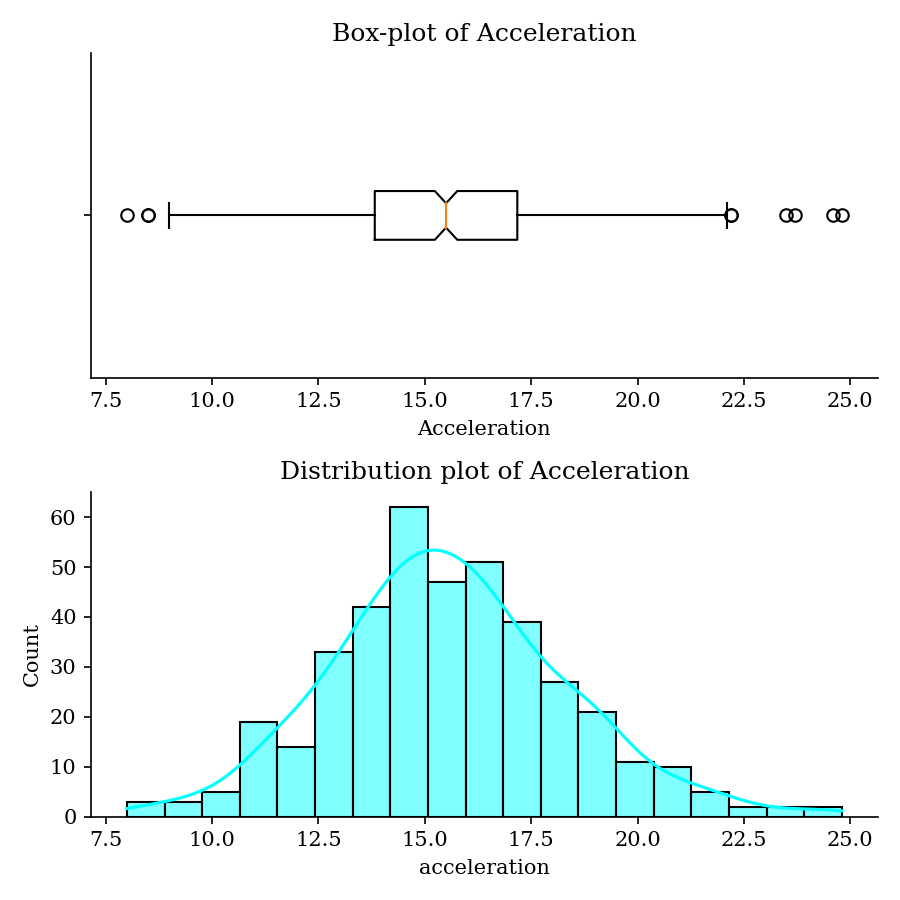


## 6. Acceleration

*Acceleration* has 95 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 15.5681 |
| Standard Deviation | 2.7577 |
| Minimum | 8.0 |
| Lower Quartile | 13.825 |
| Median | 15.5 |
| Upper Quartile | 17.175 |
| Maximum | 24.8 |
| Skewness | 0.2788 |
| Kurtosis | 0.4195 |

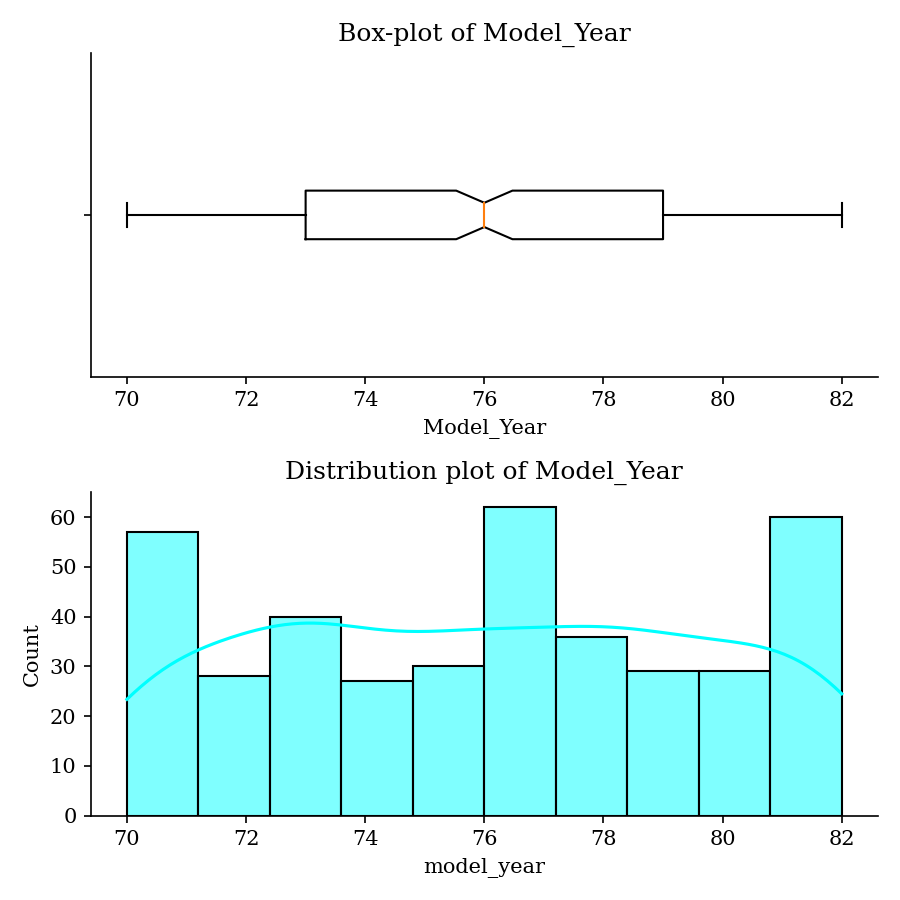


## 7. Model\_Year

*Model\_year* has 13 unique values. None of its values are missing.

#### Summary Statistics

|  |  |
| --- | --- |
| Number of observations | 398.0 |
| Average | 76.0101 |
| Standard Deviation | 3.6976 |
| Minimum | 70.0 |
| Lower Quartile | 73.0 |
| Median | 76.0 |
| Upper Quartile | 79.0 |
| Maximum | 82.0 |
| Skewness | 0.0115 |
| Kurtosis | -1.1812 |



## 8. Origin

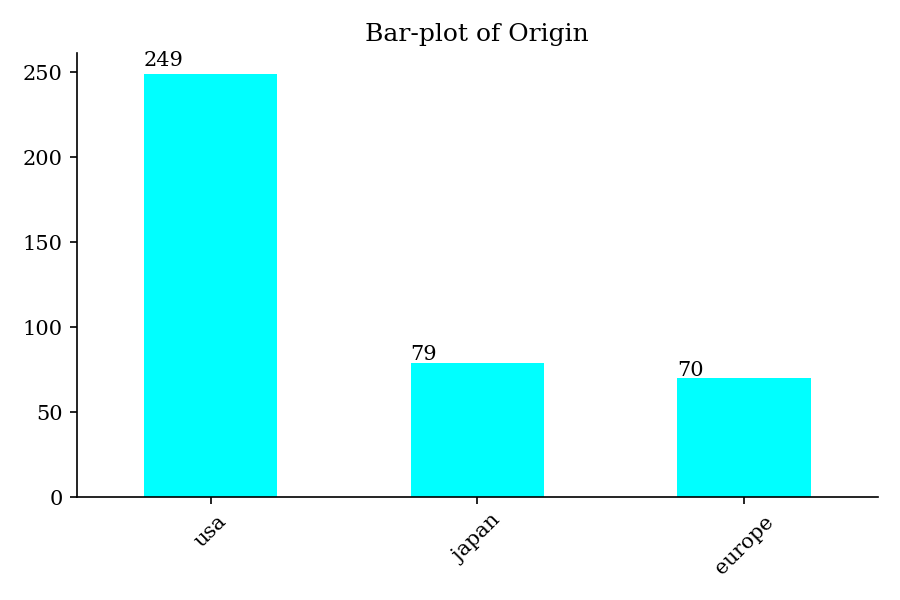
*Origin* has 3 unique values. None of its values are missing.

#### Summary

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

#### Most Frequent

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



## 9. Name

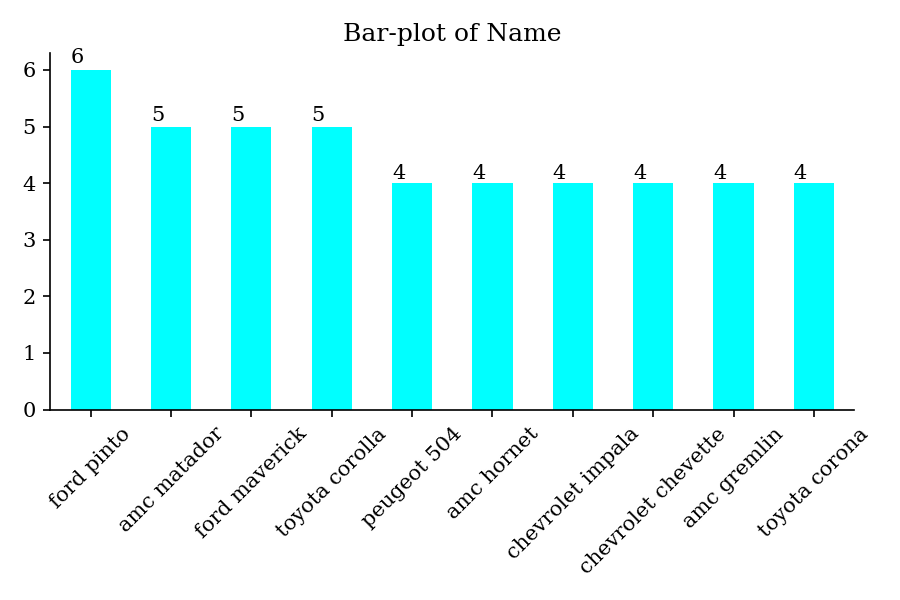
*Name* has 305 unique values. None of its values are missing.

#### Summary

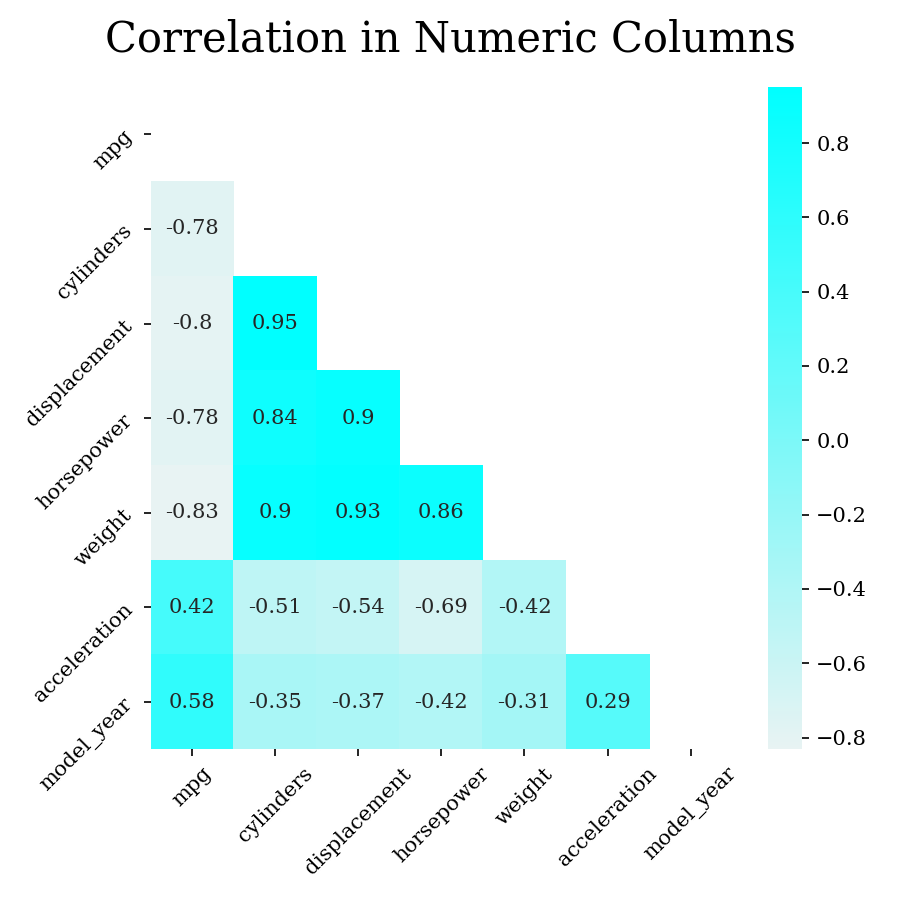
|  |  |
| --- | --- |
| Number of observations | 398 |
| Unique values | 305 |
| Mode (Highest occurring value) | ford pinto |

#### Most Frequent

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

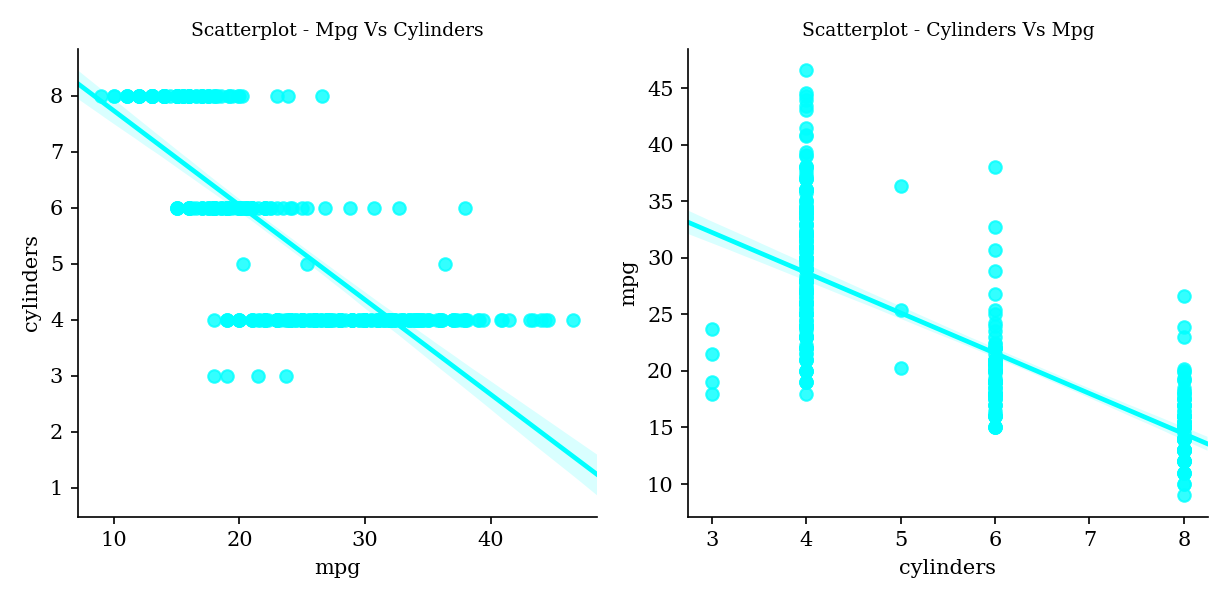


# Bivariate Analysis (Correlation)



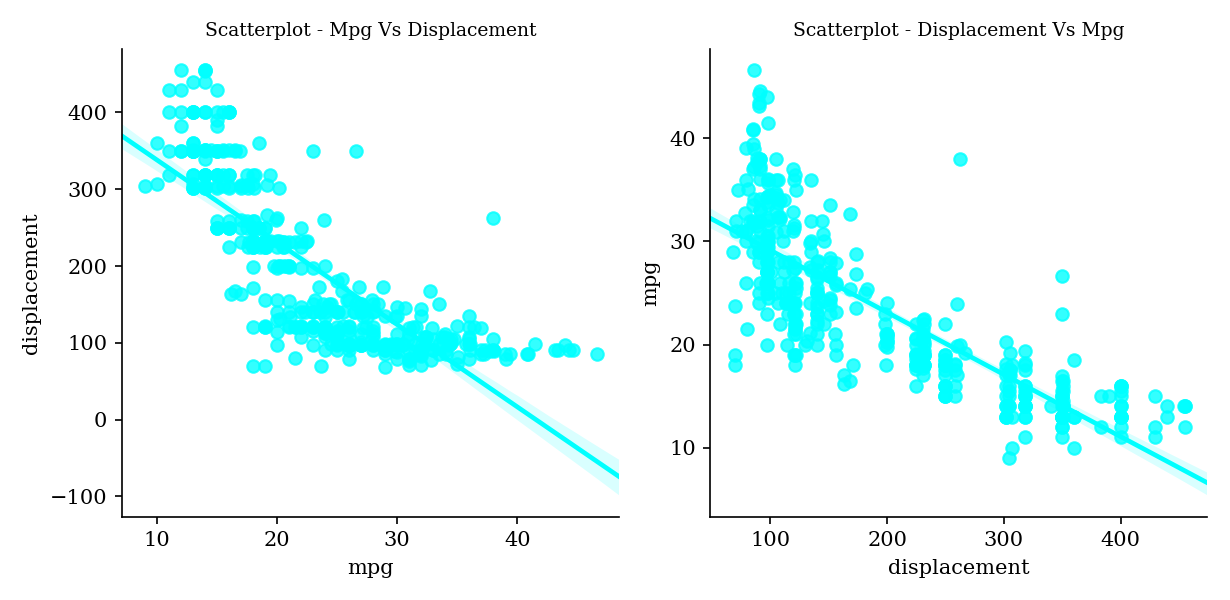
## 1. Mpg vs Cylinders

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



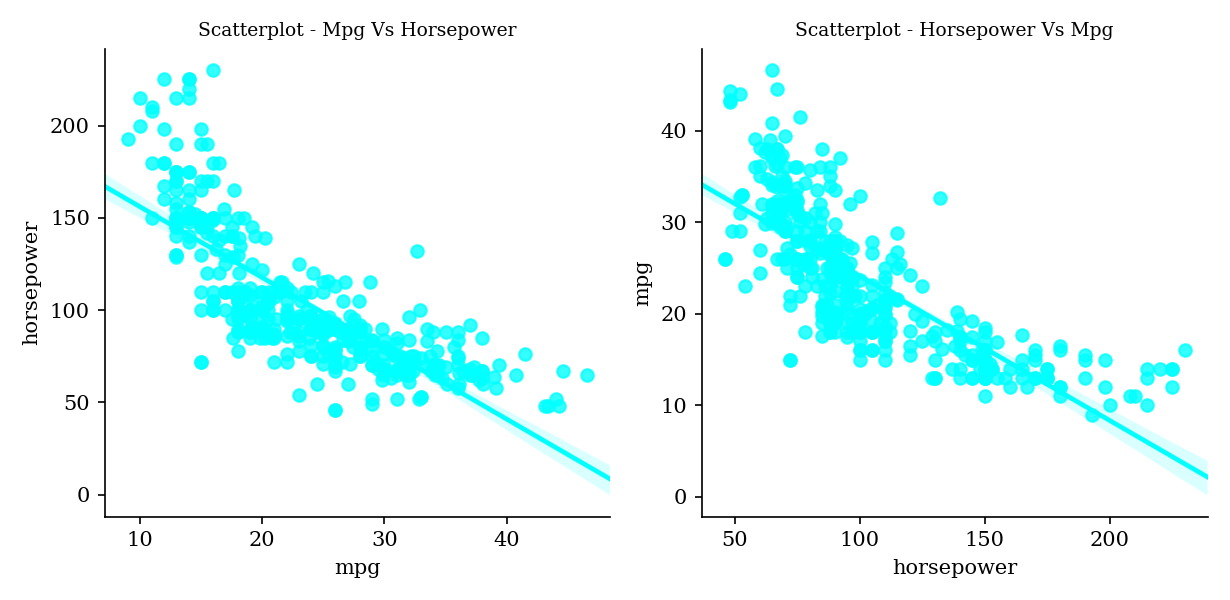
## 2. Mpg vs Displacement

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



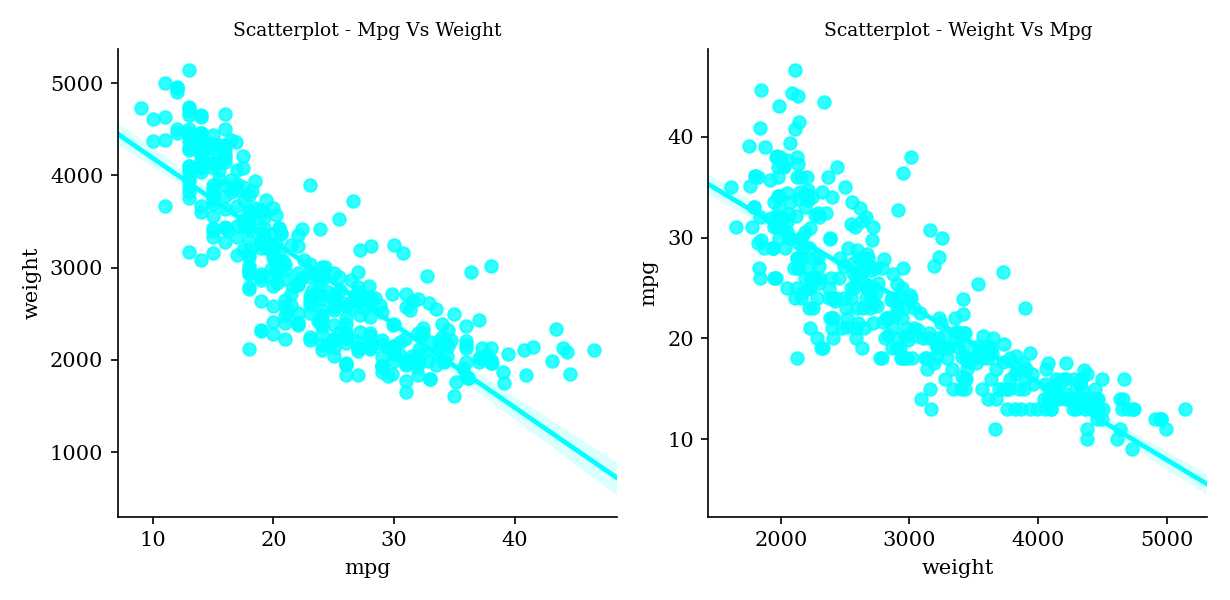
## 3. Mpg vs Horsepower

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



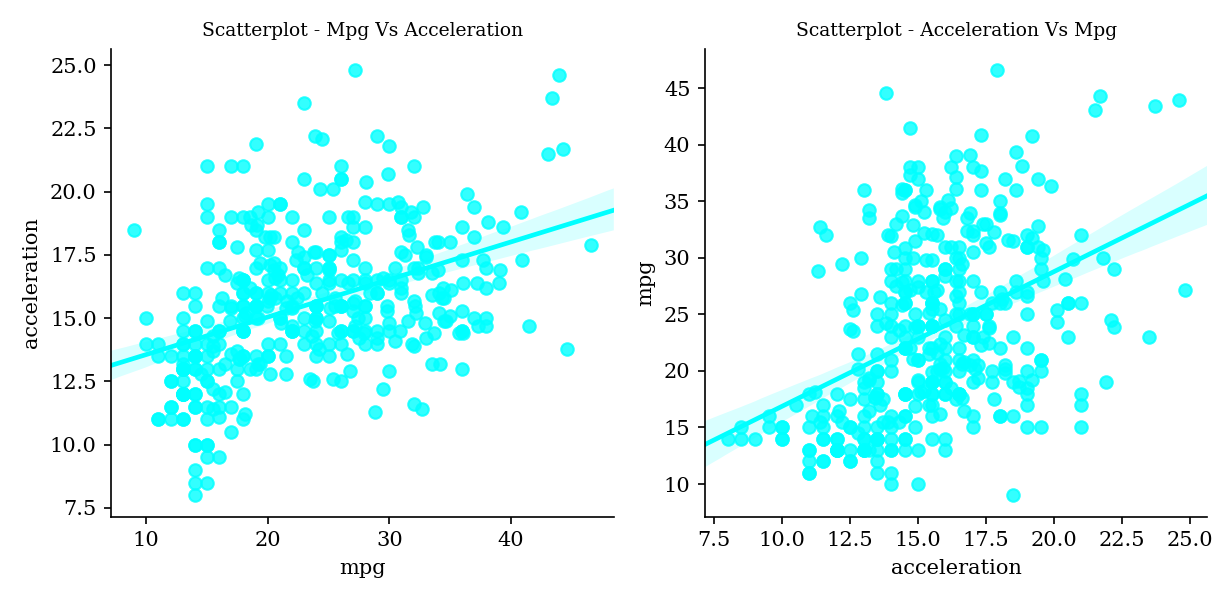
## 4. Mpg vs Weight

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



## 5. Mpg vs Acceleration

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



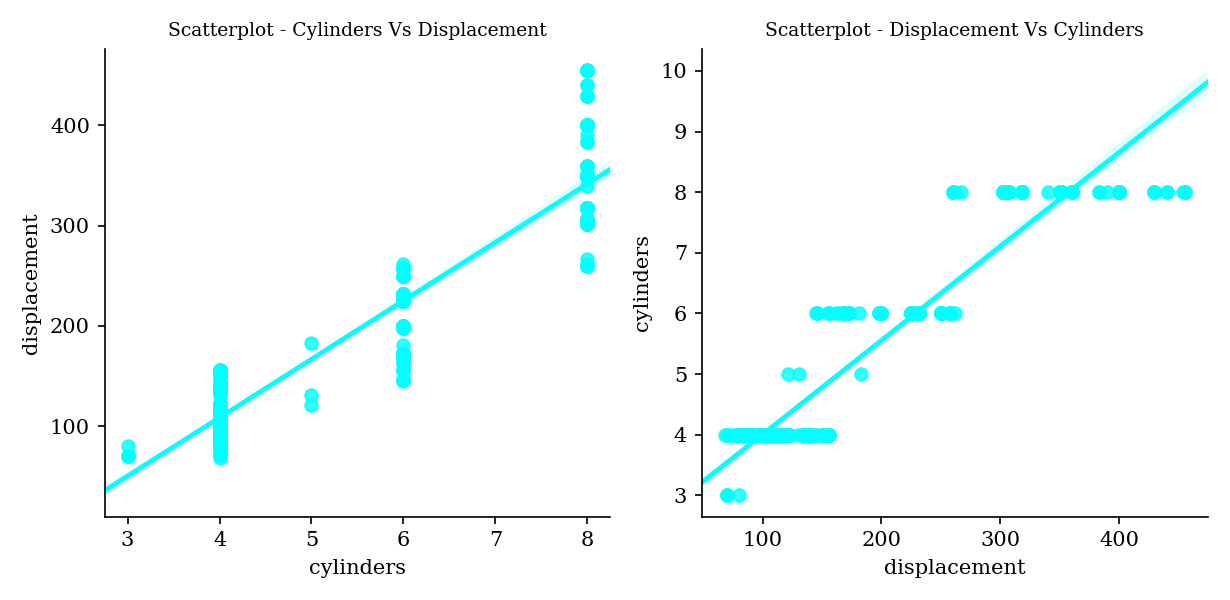
## 6. Mpg vs Model\_Year

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



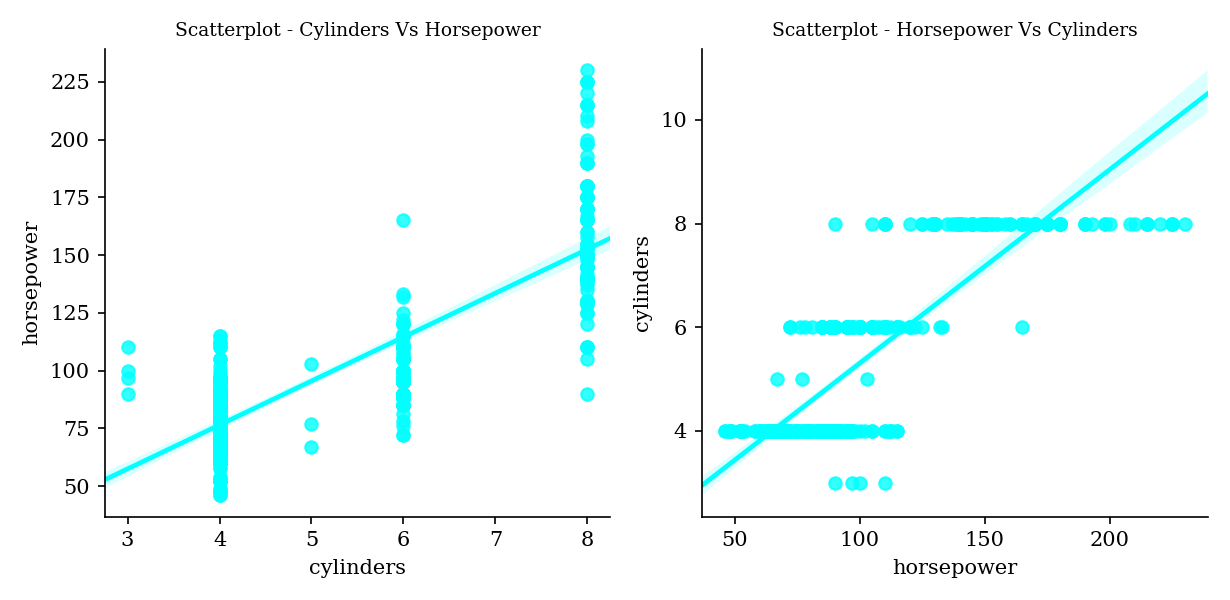
## 7. Cylinders vs Displacement

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



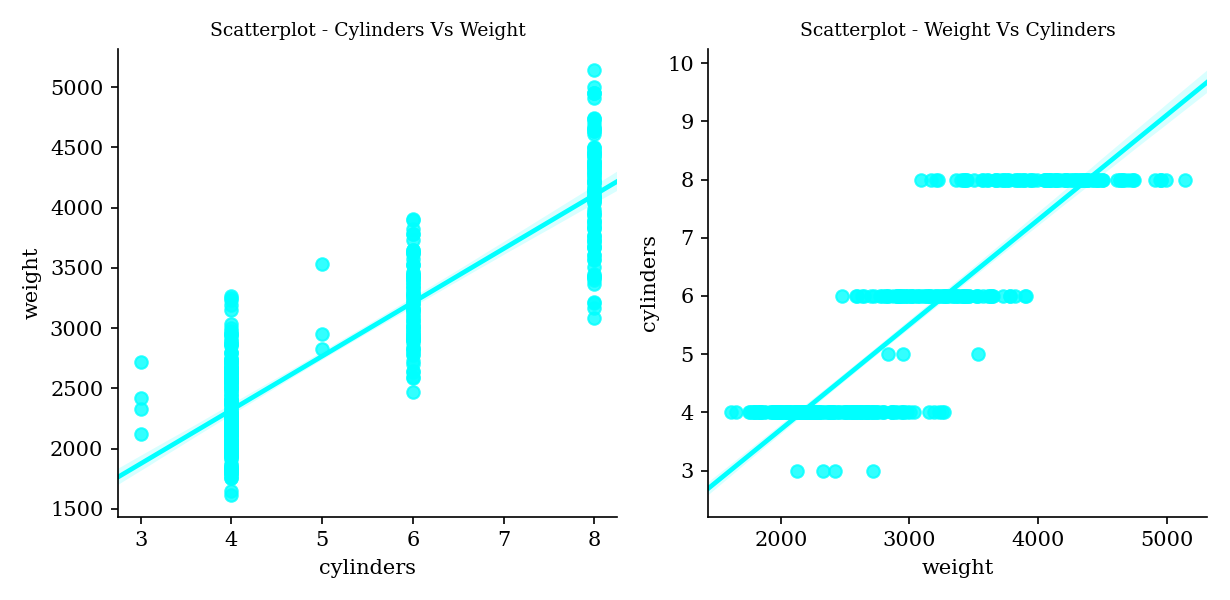
## 8. Cylinders vs Horsepower

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



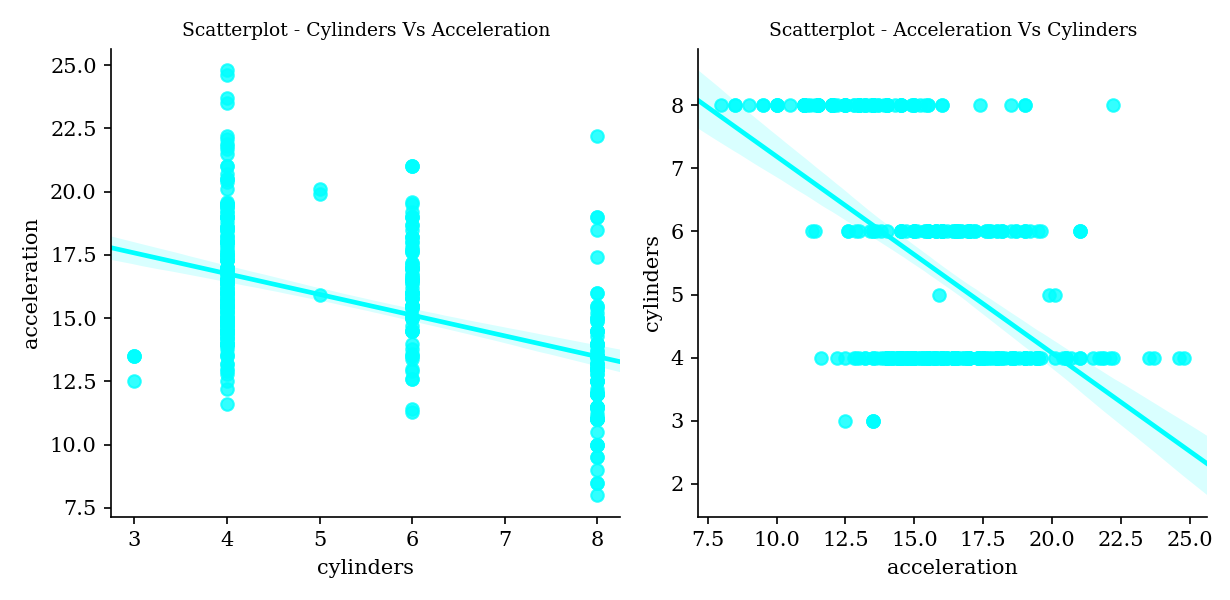
## 9. Cylinders vs Weight

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



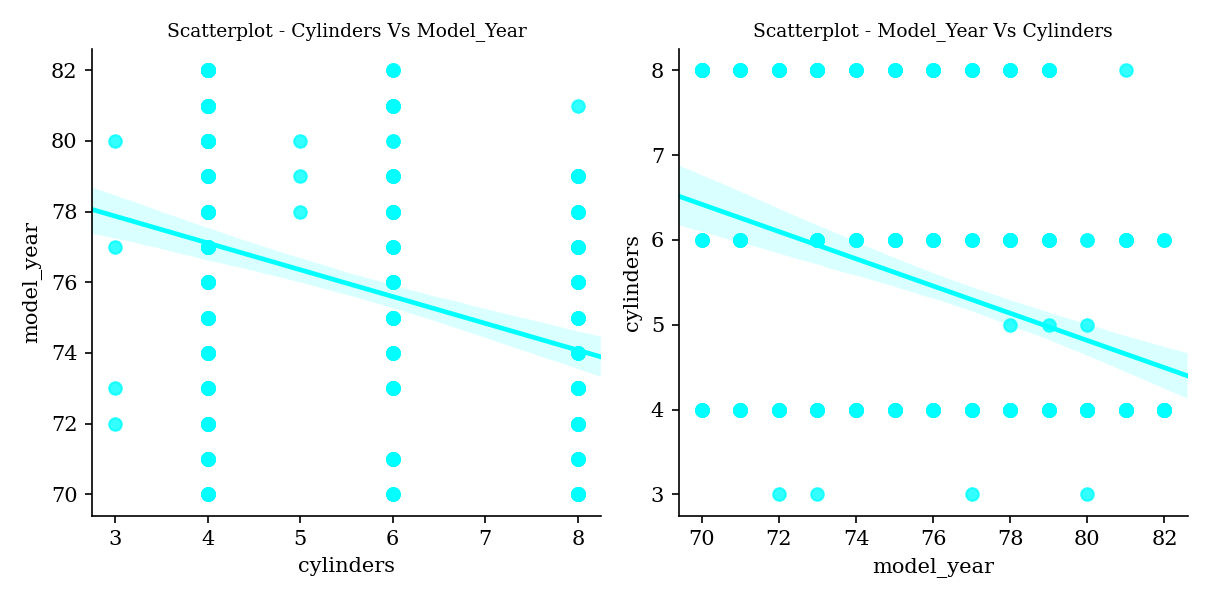
## 10. Cylinders vs Acceleration

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



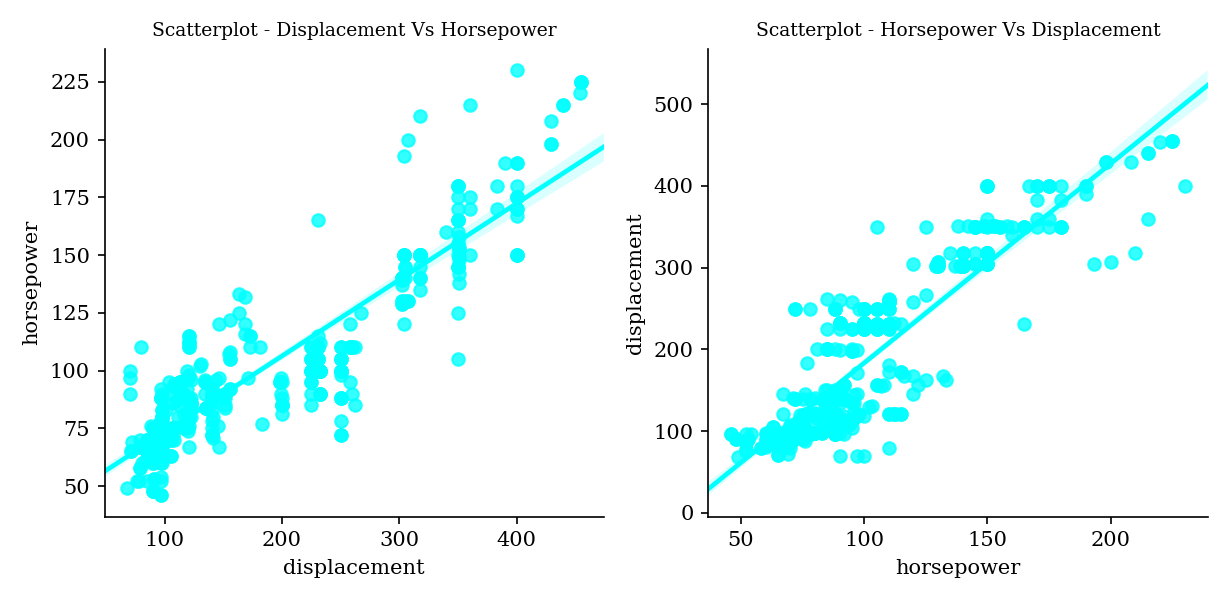
## 11. Cylinders vs Model\_Year

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



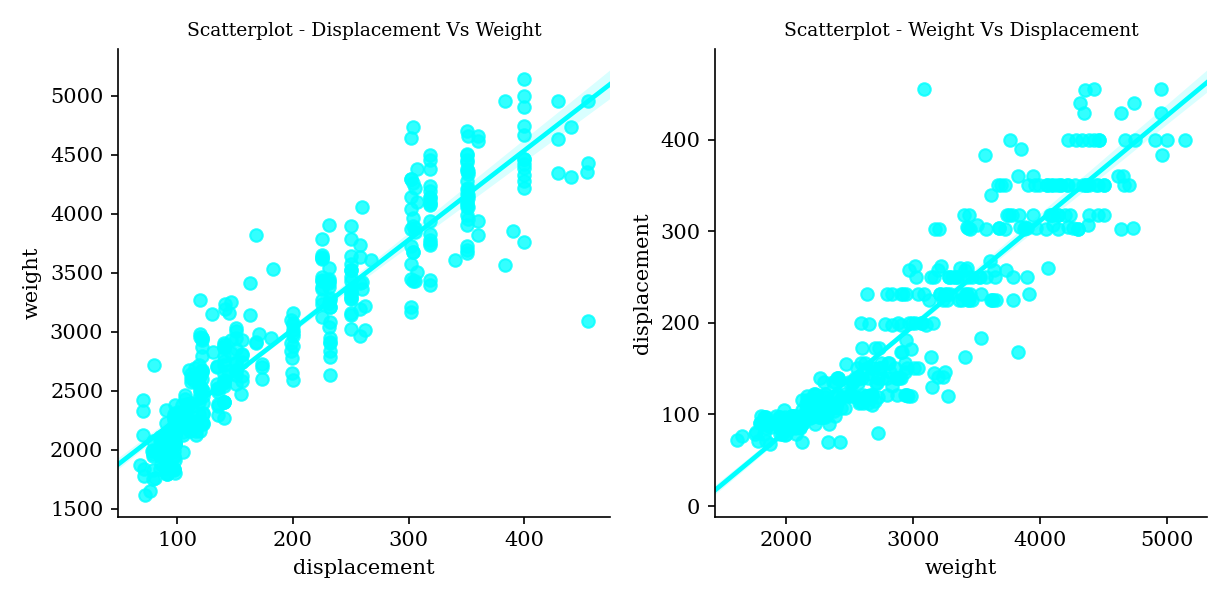
## 12. Displacement vs Horsepower

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



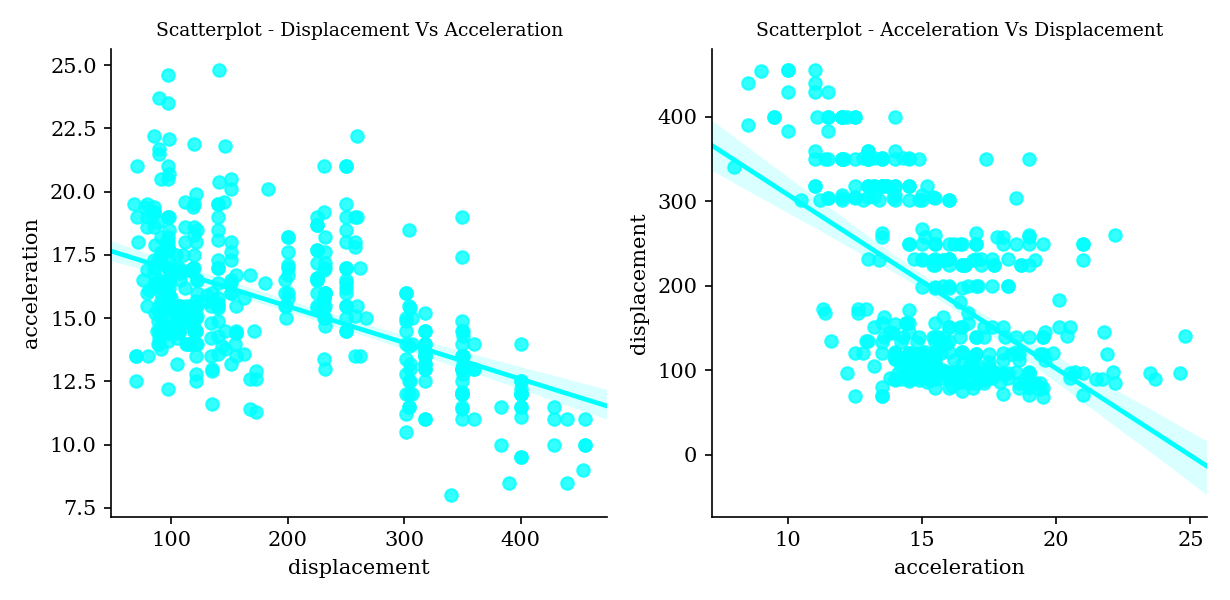
## 13. Displacement vs Weight

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



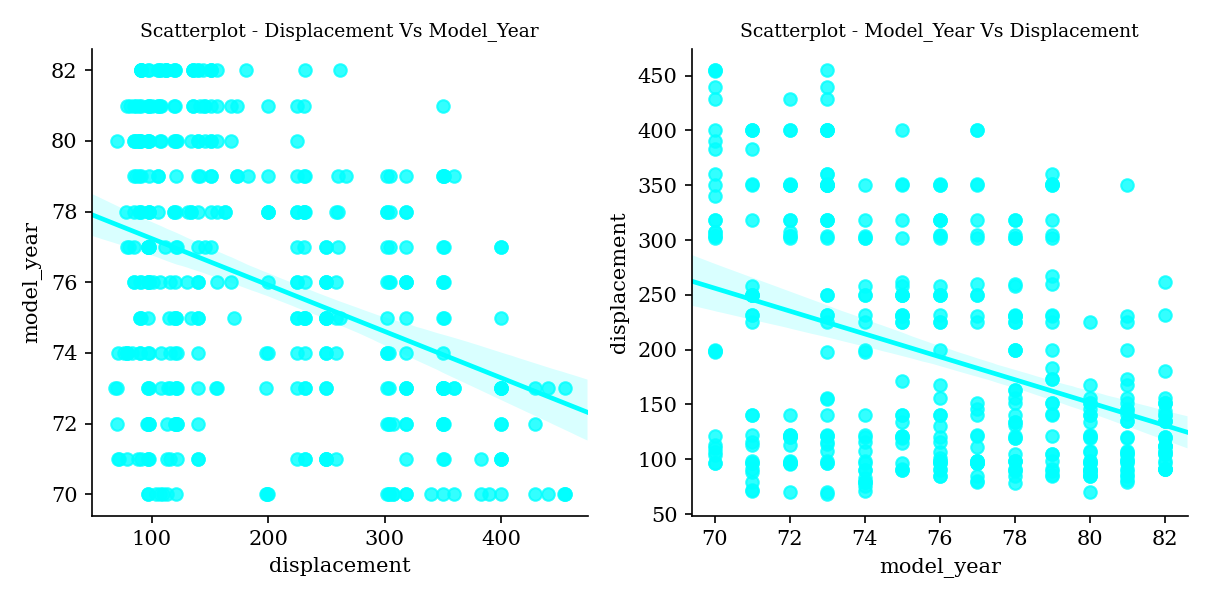
## 14. Displacement vs Acceleration

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



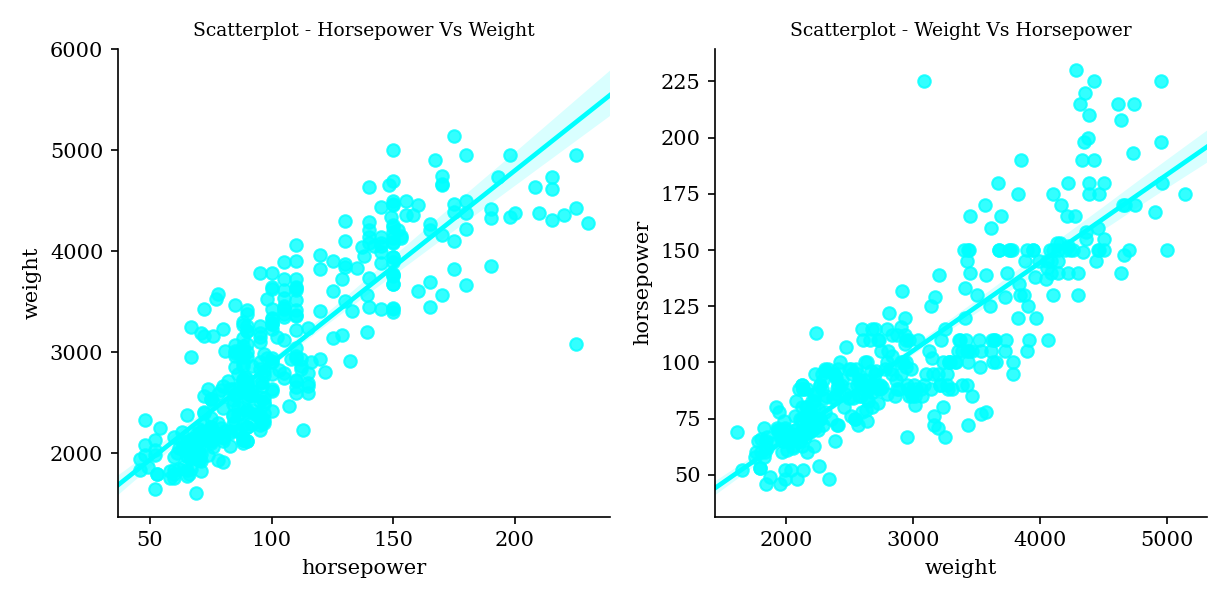
## 15. Displacement vs Model\_Year

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



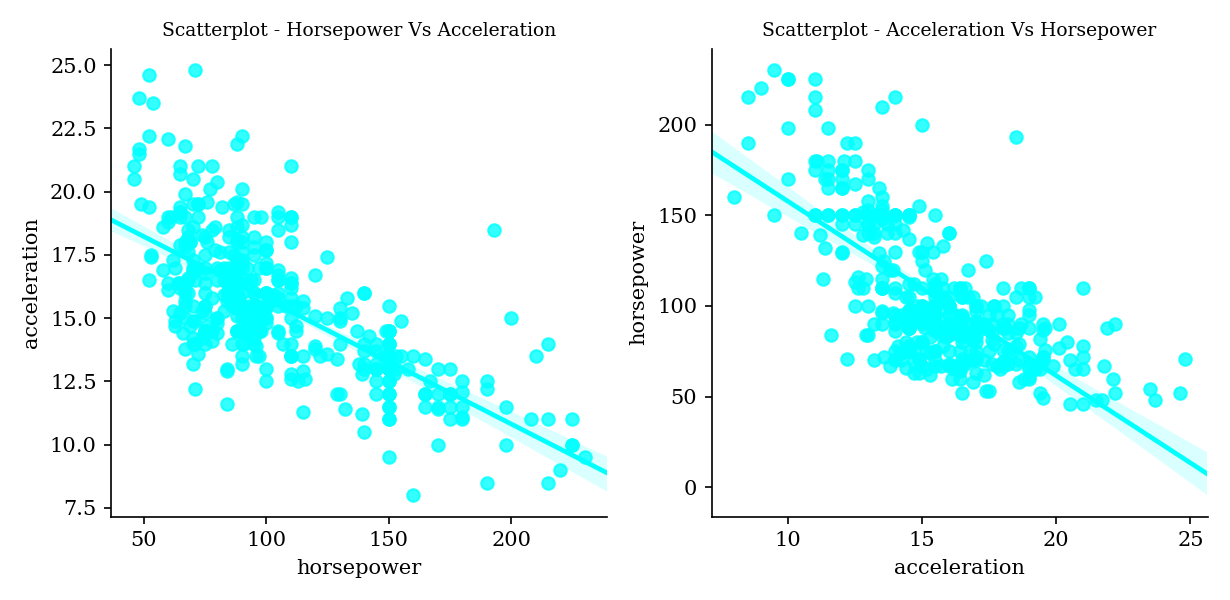
## 16. Horsepower vs Weight

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



## 17. Horsepower vs Acceleration

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



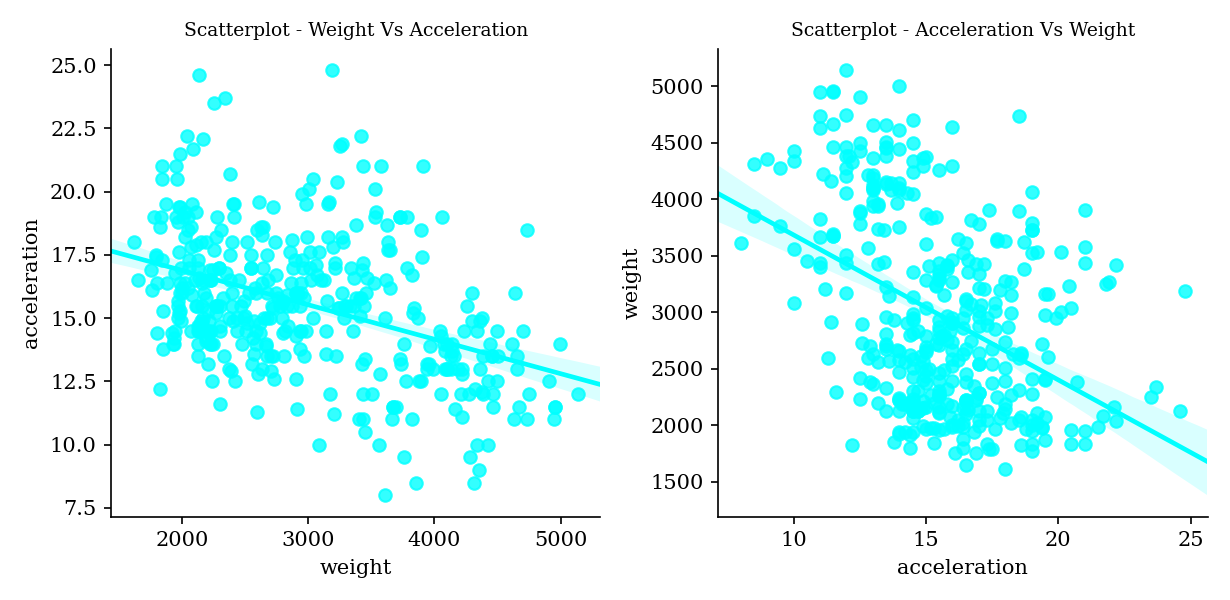
## 18. Horsepower vs Model\_Year

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



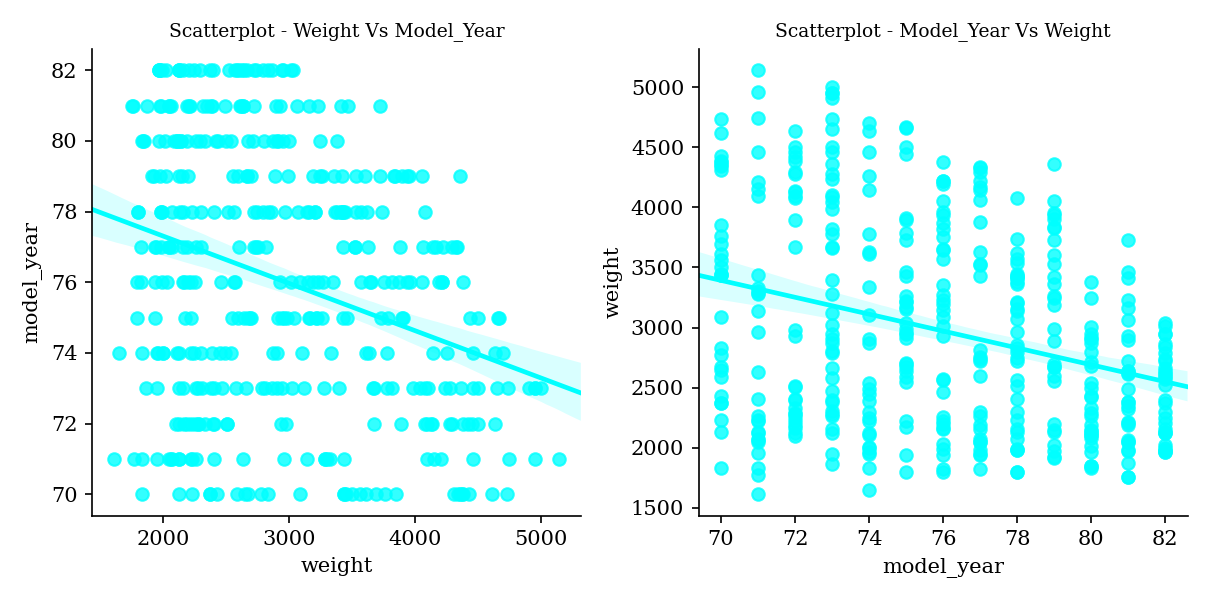
## 19. Weight vs Acceleration

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



## 20. Weight vs Model\_Year

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



## 21. Acceleration vs Model\_Year

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

