

Creating Descriptive Statistics



Matthew Renze

DATA SCIENCE CONSULTANT

@matthewrenze www.matthewrenze.com



Overview



Creating Descriptive Statistics

Demo



Descriptive Statistics

Describe data

Numerical analysis

Summary statistics

Fuel Economy	
Statistic	Value (mpg)
Minimum	10.40
1 st Quartile	15.42
Median	19.20
Mean	20.09
3 rd Quartile	22.80
Maximum	33.90



Statistical Terms

Observations

Variables

Categorical variable

Numeric variable

Cars			
Name	Transmission	Cylinders	Fuel Economy
Mazda RX4	Manual	6	21.0
Mazda Wag	Manual	6	21.0
Datsun 710	Manual	4	22.8
Hornet 4 Drive	Automatic	6	21.4
Sportabout	Automatic	8	18.7
Valiant	Automatic	6	18.1
Duster 360	Automatic	8	14.3
Merc 240D	Automatic	4	24.4



Categorical Variables

Qualitative values

Frequency of observations

Frequency table

Frequency of Transmission Type	
Transmission	Frequency
Automatic	19
Manual	13



Numeric Variables

Quantitative values

Measures of center

Measures of spread

Fuel Economy	
Statistic	Value (mpg)
Minimum	10.40
1 st Quartile	15.42
Mean	20.09
Median	19.20
3 rd Quartile	22.80
Maximum	33.90
Std. Dev.	6.02



Correlation

Strength of relationship

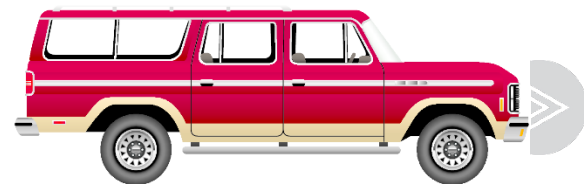
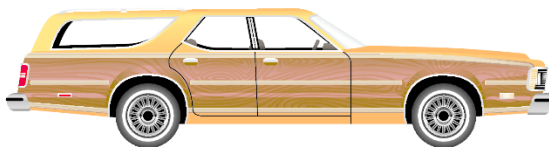
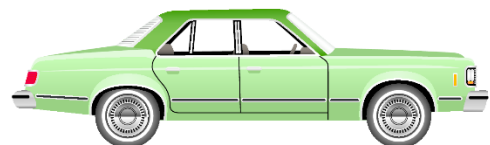
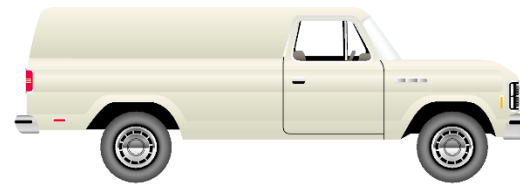
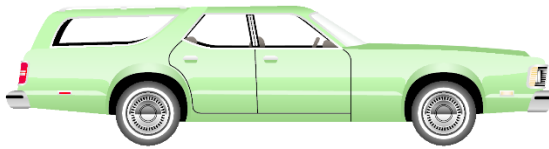
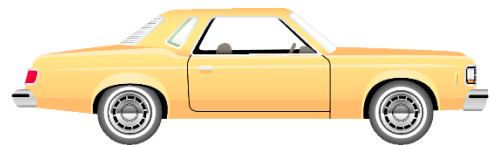
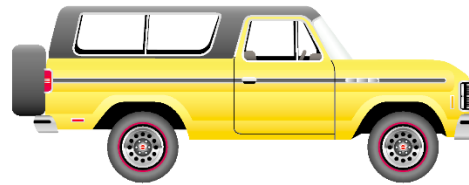
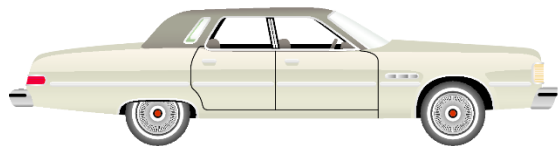
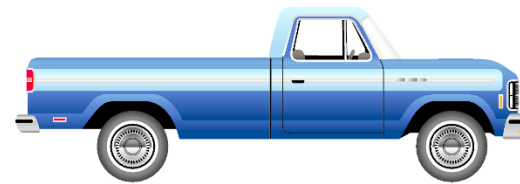
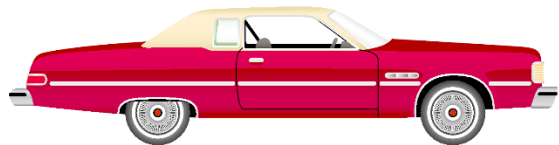
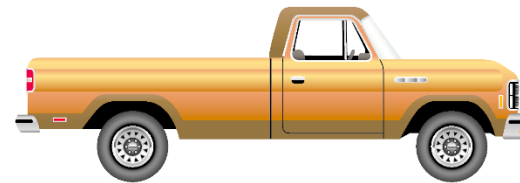
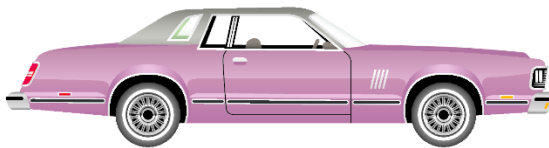
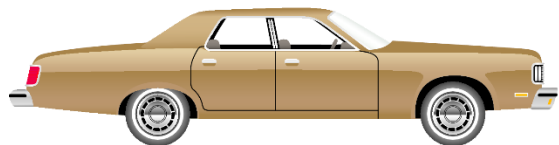
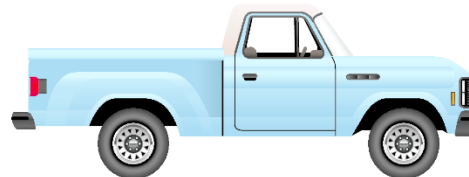
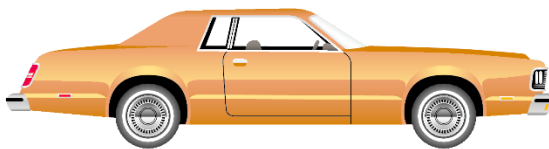
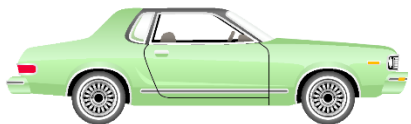
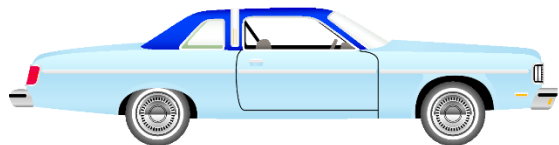
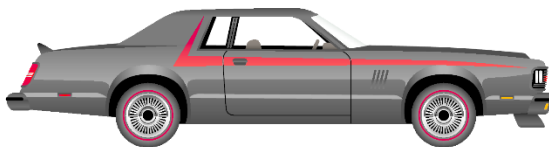
Correlation coefficient

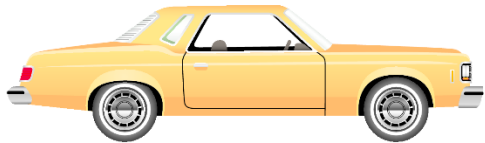
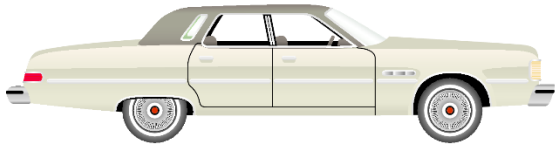
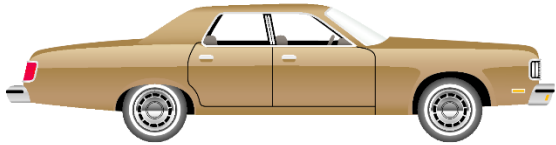
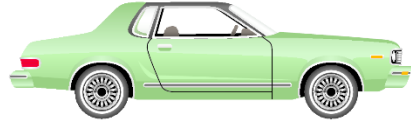
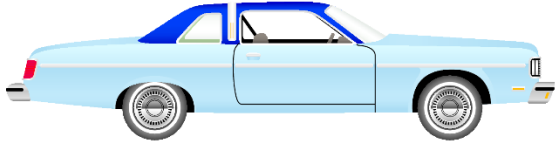
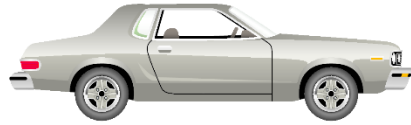
Standardized scale

Correlation	
	Fuel Economy
Cylinders	-0.85

Fuel Efficiency	
Cylinders	Consumption
4	9.73
6	8.13
8	6.40







1. How many cars of each type of transmission?
2. What does fuel economy look like?
3. How are cylinders and fuel economy related?



Load the Data



Create a frequency table



Create Numeric Statistics



Get Correlation Coefficient



Summarize Table



Transmission	
Transmission	Frequency
Automatic	19
Manual	13

Correlation	
	Fuel Economy
Cylinders	-0.85

Fuel Economy	
Statistic	Value (mpg)
Minimum	10.40
1 st Quartile	15.42
Median	19.20
Mean	20.09
3 rd Quartile	22.80
Maximum	33.90



Summary



Descriptive Statistics

Demo

