Assignment 1

Data Audit Report:

1. Metadata Description.
2. Descriptive Statistics.

Metadata Description

My Data Set Name = CARS

It has total 93 observations and consist of

26 variables.

The CONTENTS Procedure

Data Set Name	WORK.CARS	Observations	93
Member Type	DATA	Variables	26
Engine	V9	Indexes	0
Created	01/23/2021 12:11:06	Observation Length	224
Last Modified	01/23/2021 12:11:06	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Total number of variable is 26 the name of variable and the type of variable is displayed in the given picture with the length of each variable.

Variable	Description
Manufacturer	Brand Name of the car manufacturer.
Model	The different models of the particular cars. Different car models can have the same manufacture.
Category	There are different types of category in cars such as sports, van, etc.
Min_Price	This variable shows the minimum price of the particular car that is the lowest variant of the particular car.
Mid_Price	This shows the price of the medium variant of the car.

Variables in Creation Order

#	Variable	Type	Len	Format	Informat	Label
1	Manufacturer	Char	13	$13.	$13.	Manufacturer
2	Model	Char	14	$14.	$14.	Model
3	Category	Char	7	$7.	$7.	Category
4	Min_Price	Num	8	BEST.		Min_Price
5	Mid_Price	Num	8	BEST.		Mid_Price
6	Max_Price	Num	8	BEST.		Max_Price
7	City_Fuel	Num	8	BEST.		City_Fuel
8	Hwy_Fuel	Num	8	BEST.		Hwy_Fuel
9	Air_Bags	Num	8	BEST.		Air_Bags
10	Drive_Train	Num	8	BEST.		Drive_Train
11	Cylinders	Num	8	BEST.		Cylinders
12	Engine_Size	Num	8	BEST.		Engine_Size
13	Max_HP	Num	8	BEST.		Max_HP
14	Max_HP_RPM	Num	8	BEST.		Max_HP_RPM
15	RPM_high	Num	8	BEST.		RPM_high
16	Manual	Num	8	BEST.		Manual
17	Fuel_Tank	Num	8	BEST.		Fuel_Tank
18	Passenger	Num	8	BEST.		Passenger
19	Length	Num	8	BEST.		Length
20	Wheel_Base	Num	8	BEST.		Wheel_Base
21	Width	Num	8	BEST.		Width
22	U_Turn_Diam	Num	8	BEST.		U_Turn_Diam
23	Rear_Room	Num	8	BEST.		Rear_Room
24	Luggage	Num	8	BEST.		Luggage
25	Weight	Num	8	BEST.		Weight
26	Domestic	Num	8	BEST.		Domestic

Variable	Description
Mid_Price	This shows the price of the medium variant of the car.
Max_Price	This shows the price for the top variant of the car.
City_Fuel	Shows the average of the car in city roads.
Hwy_Fuel	Shows the average of the car in the Highway roads.
Air_Bags	The Number of airbags car has.
Drive_Train	Tells us about which car is front wheel drive, which is back wheel drive and which is four wheel drive.
Cylinders	Number of cylinders car consist
Engine_Size	How much CC of engine each car have.

Variable	Description
Max_HP	Shows the max horse power of the cars.
Max_HP_RPM	Max HP based on Rotation per minute.
RPM_high	Highest RPM of the car.
Manual	Differentiate between which car is manual transmission and which car is automatic transmission.
Fuel_Tank	Maximum amount of fuel in each car.
Passenger	Number of Seats in each car.
Length	Length of car in inches.
Wheel_Base	Distance between front axel and rare axel in inches
Width	Width of the car in inches.
U_Turn_Diam	u turn diameter of the car in inches

Variable	Description
Rear_room	Consist of length of rear legroom in the car in inches.
Luggage	Displays the room for luggage.
Weight	Weight of car in pounds.
Domestic	Manufacturer details is it international car or domestic car.

Variables in Creation Order

#	Variable	Type	Len	Format	Informat	Label
1	Manufacturer	Char	13	$13.	$13.	Manufacturer
2	Model	Char	14	$14.	$14.	Model
3	Category	Char	7	$7.	$7.	Category
4	Min_Price	Num	8	BEST.		Min_Price
5	Mid_Price	Num	8	BEST.		Mid_Price
6	Max_Price	Num	8	BEST.		Max_Price
7	City_Fuel	Num	8	BEST.		City_Fuel
8	Hwy_Fuel	Num	8	BEST.		Hwy_Fuel
9	Air_Bags	Num	8	BEST.		Air_Bags
10	Drive_Train	Num	8	BEST.		Drive_Train
11	Cylinders	Num	8	BEST.		Cylinders
12	Engine_Size	Num	8	BEST.		Engine_Size
13	Max_HP	Num	8	BEST.		Max_HP
14	Max_HP_RPM	Num	8	BEST.		Max_HP_RPM
15	RPM_high	Num	8	BEST.		RPM_high
16	Manual	Num	8	BEST.		Manual
17	Fuel_Tank	Num	8	BEST.		Fuel_Tank
18	Passenger	Num	8	BEST.		Passenger
19	Length	Num	8	BEST.		Length
20	Wheel_Base	Num	8	BEST.		Wheel_Base
21	Width	Num	8	BEST.		Width
22	U_Turn_Diam	Num	8	BEST.		U_Turn_Diam
23	Rear_Room	Num	8	BEST.		Rear_Room
24	Luggage	Num	8	BEST.		Luggage
25	Weight	Num	8	BEST.		Weight
26	Domestic	Num	8	BEST.		Domestic

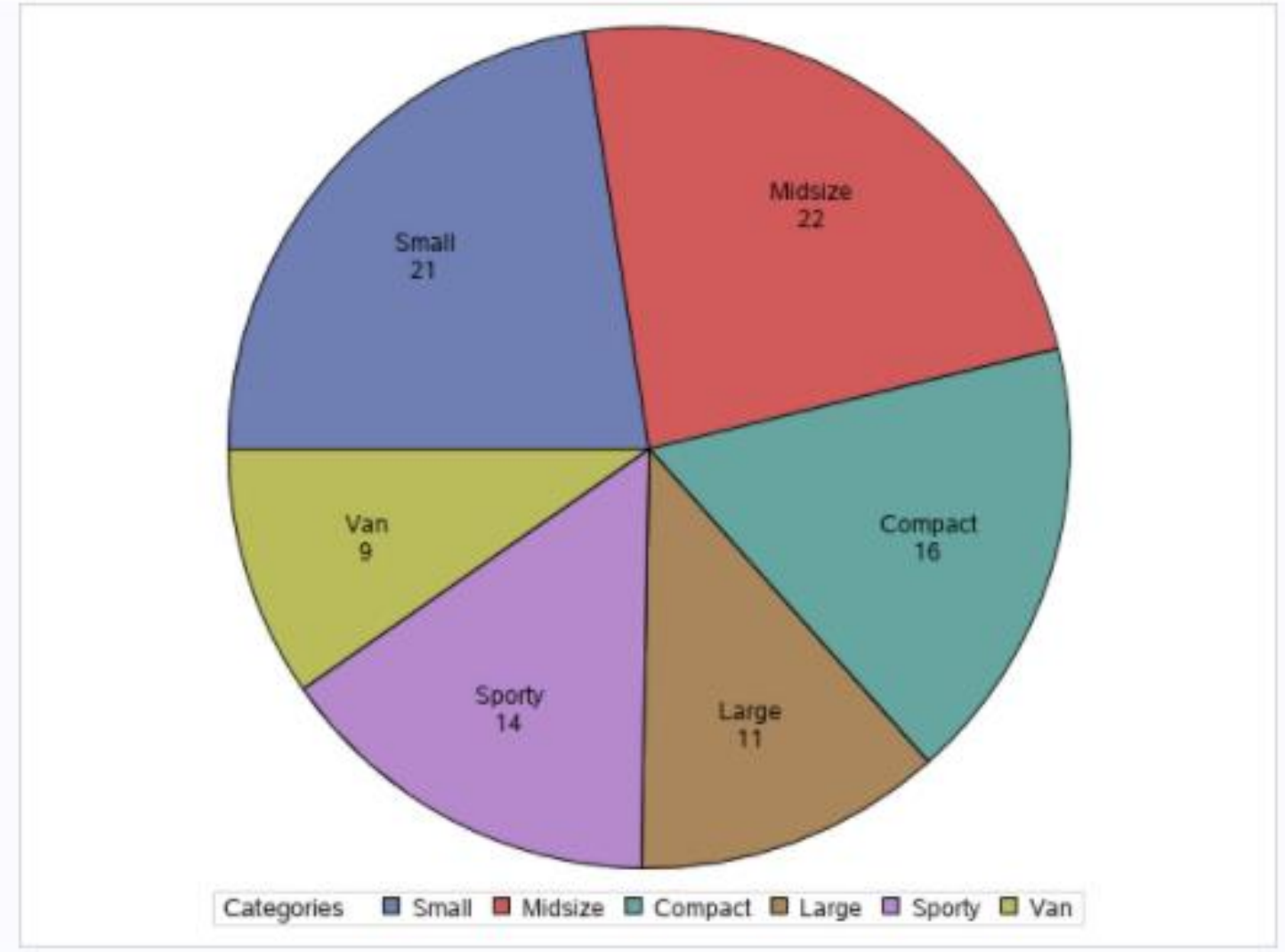
Descriptive Statistics

- I asked some of my friends. How many cylinders should one look for, if he/she thinks of taking a new car? So the common answer to my question was 6 cylinder cars.
- The image displays all the cars and their model which consist of Six cylinders.

Six Cylinder Cars

Manufacturer	Model	Cylinders
Acura	Legend	6
Audi	90	6
Audi	100	6
Buick	LeSabre	6
Buick	Roadmaster	6
Buick	Riviera	6
Chevrolet	Camaro	6
Chevrolet	Lumina APV	6
Chevrolet	Astro	6
Chrysler	Concorde	6
Chrysler	Imperial	6
Dodge	Caravan	6
Dodge	Stealth	6
Eagle	Vision	6
Ford	Aerostar	6
Ford	Taurus	6
Lexus	ES300	6
Lexus	SC300	6
Lincoln	Continental	6
Mazda	MPV	6
Mercedes-Benz	300E	6
Mercury	Cougar	6
Mitsubishi	Diamante	6
Nissan	Quest	6
Nissan	Maxima	6
Oldsmobile	Silhouette	6
Oldsmobile	Eighty-Eight	6
Pontiac	Firebird	6
Pontiac	Grand Prix	6
Pontiac	Bonneville	6
Volkswagen	Corrado	6

The dataset consist of different categories of cars so to find out how many cars in each category we have I made a pie chart using SAS.



Fuel is very important factor in the car.

This raises many questions in the mind while buying a new car.

What will be the range of my car on full tank?
Fuel economy on city roads?
Fuel economy on Highway roads?

So to answer this questions I have divided all the cars into 3 different categories:

1. Low range cars
2. High range cars.
3. Medium range cars.

Range in full fuel tank

Obs	Manufacturer	Model	City_Fuel	Hwy_Fuel	Fuel_Tank	City_Range	City_Range_Category	Hwy_Range	Hwy_Range_Category
1	Acura	Integra	25	31	13.2	330.0	Low	409.2	Mid
2	Acura	Legend	18	25	18	324.0	Low	450.0	Mid
3	Audi	90	20	26	16.9	338.0	Mid	439.4	Mid
4	Audi	100	19	26	21.1	400.9	Mid	548.6	High
5	BMW	535i	22	30	21.1	464.2	High	633.0	High
6	Buick	Century	22	31	16.4	360.8	Mid	508.4	Mid
7	Buick	LeSabre	19	28	18	342.0	Mid	504.0	Mid
8	Buick	Roadmaster	16	25	23	368.0	Mid	575.0	High
9	Buick	Riviera	19	27	18.8	357.2	Mid	507.6	Mid
10	Cadillac	DeVille	16	25	18	288.0	Low	450.0	Mid

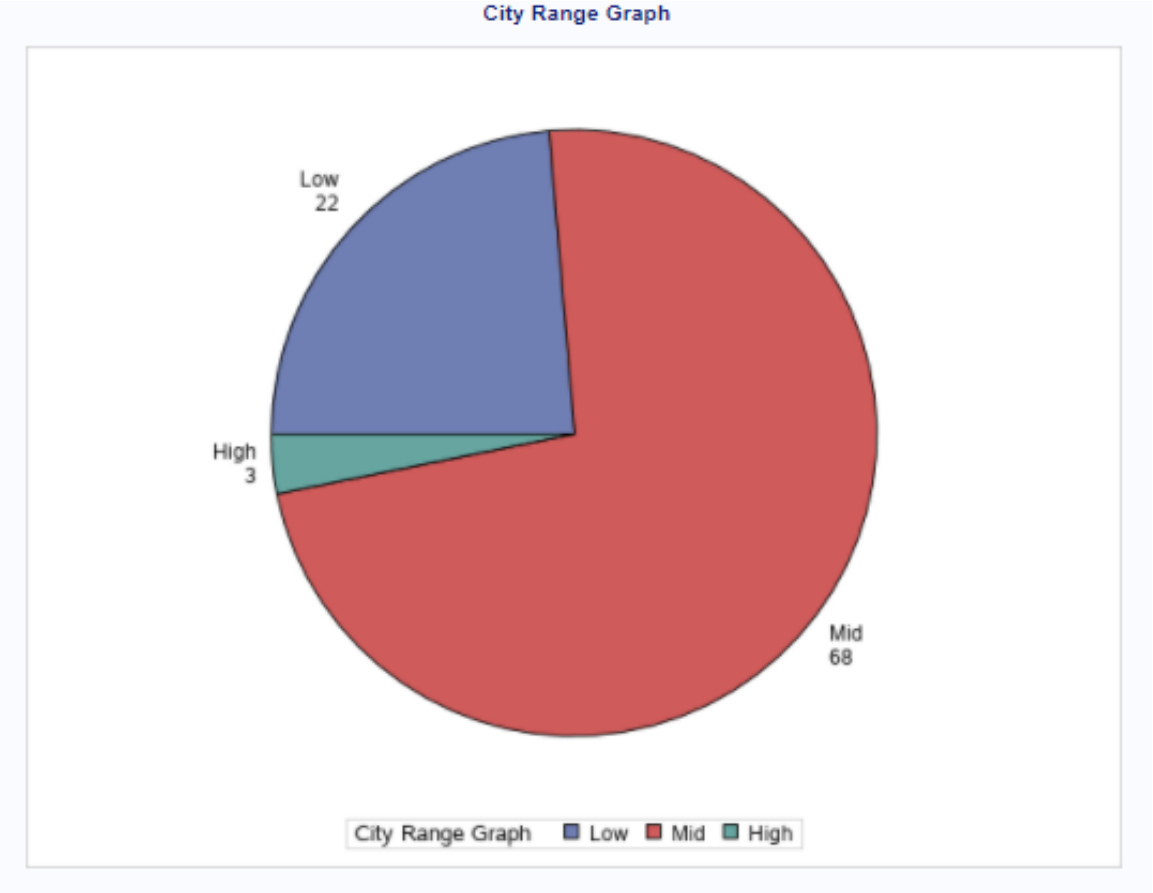
I found out the maximum and the minimum range of the given dataset and divided the entire bandwidth into 3 separate bands:

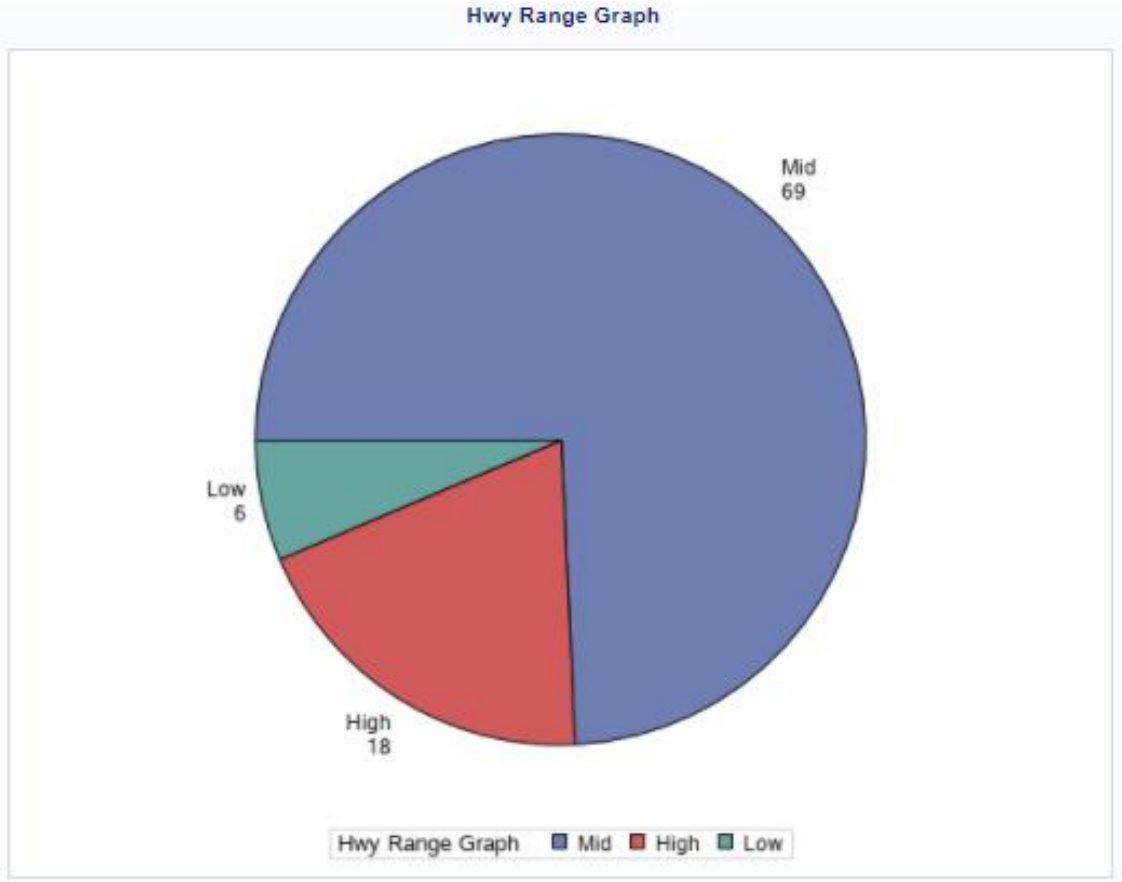
1. Low ranging below 336.
2. Medium ranging between 337 to 417.
3. High ranging above 418.

The above image shows the first 10 observations of the dataset out of 93 observations.

Pie Chart Using SAS

Maximum number of cars in the provided dataset have range from 337 to 417.

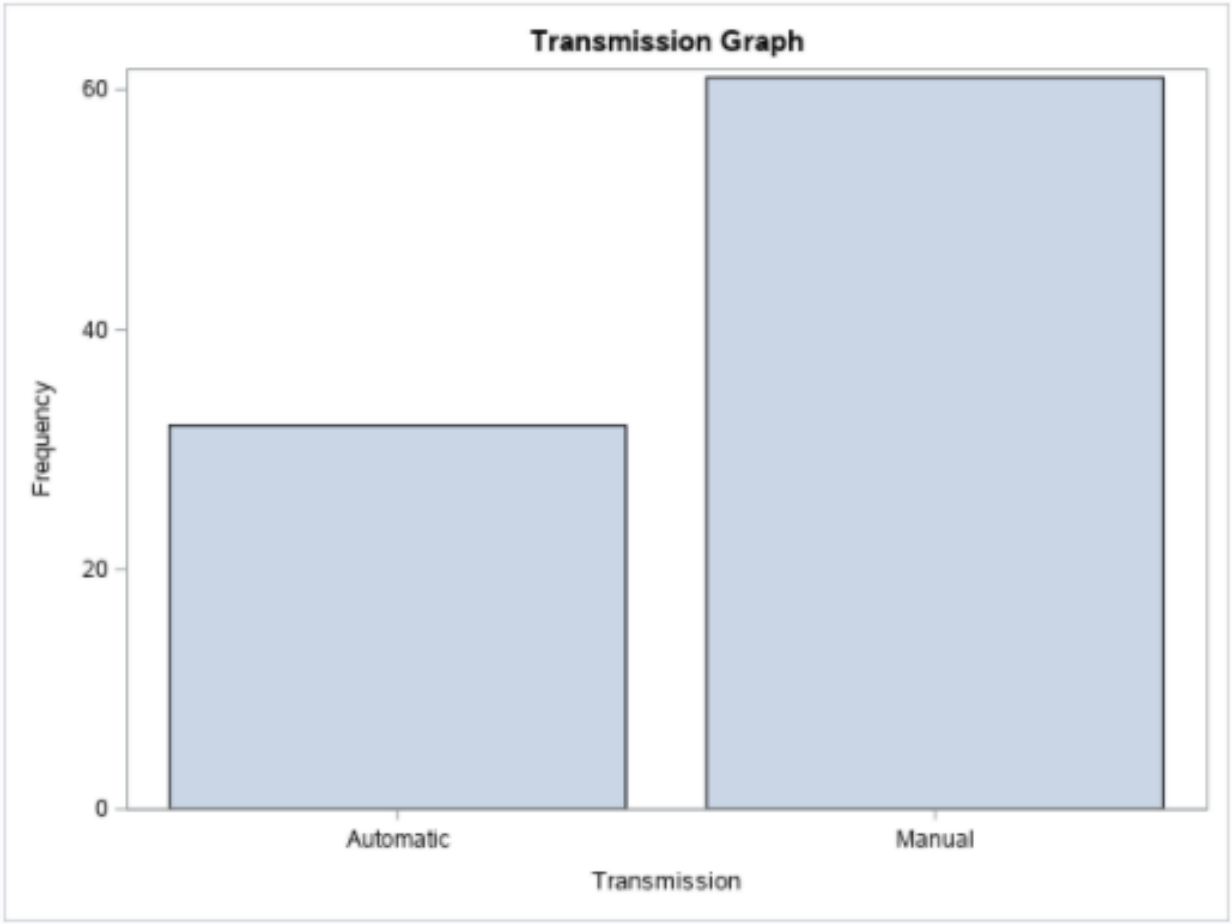




Mean fuel economy by category.

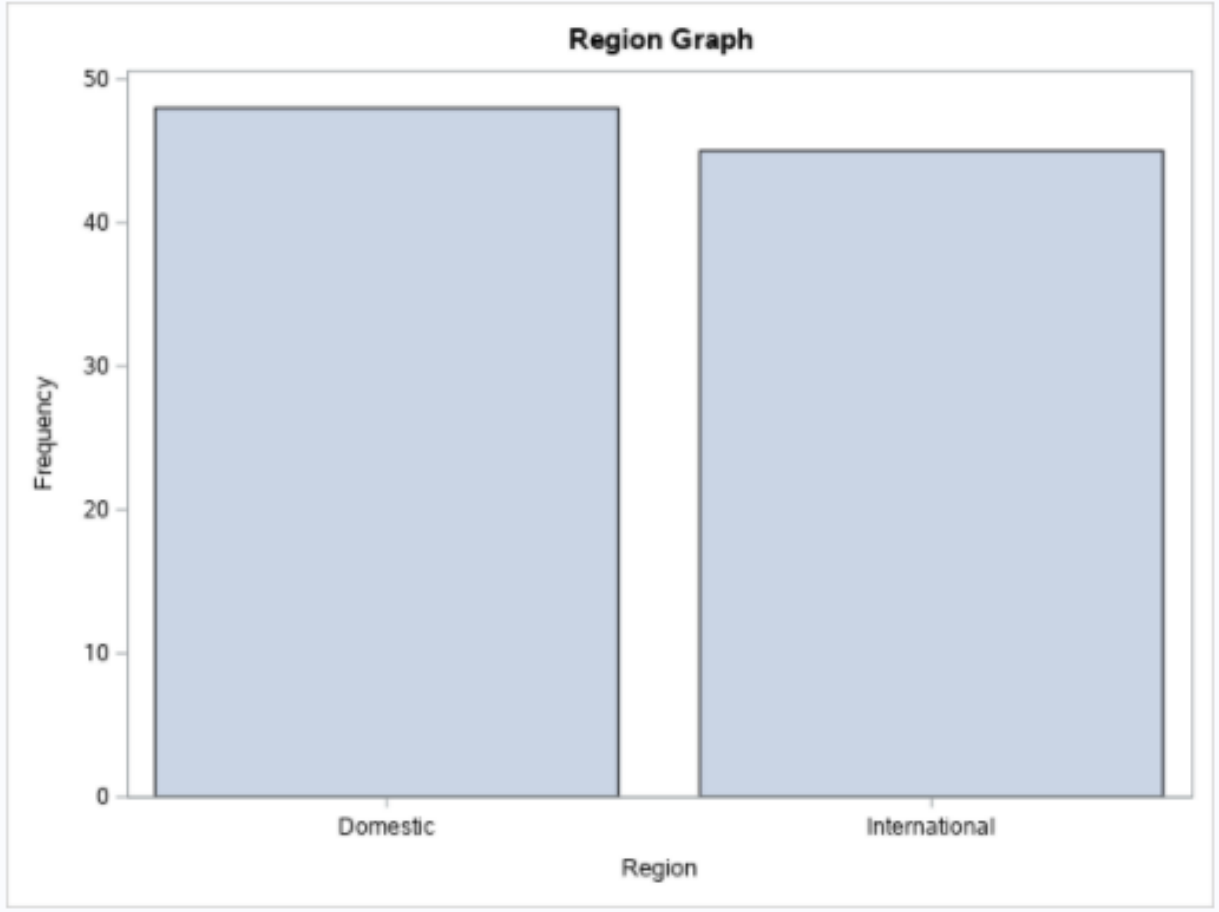
Category	Mean_City_Fuel	Mean_Hwy_Fuel
Compact	23	30
Large	18	27
Midsize	20	27
Small	30	35
Sporty	22	29
Van	17	22

Automatic and Manual Cars in the Dataset:



From 93 cars 61 cars are manual and 32 cars are Automatic.

Number of Domestic and International Cars:



From 93 Cars 45 cars are international and 48 cars are domestic

Highest weighing Car and Lowest weighing Car:

Highest weighing car and Lowest weighing car

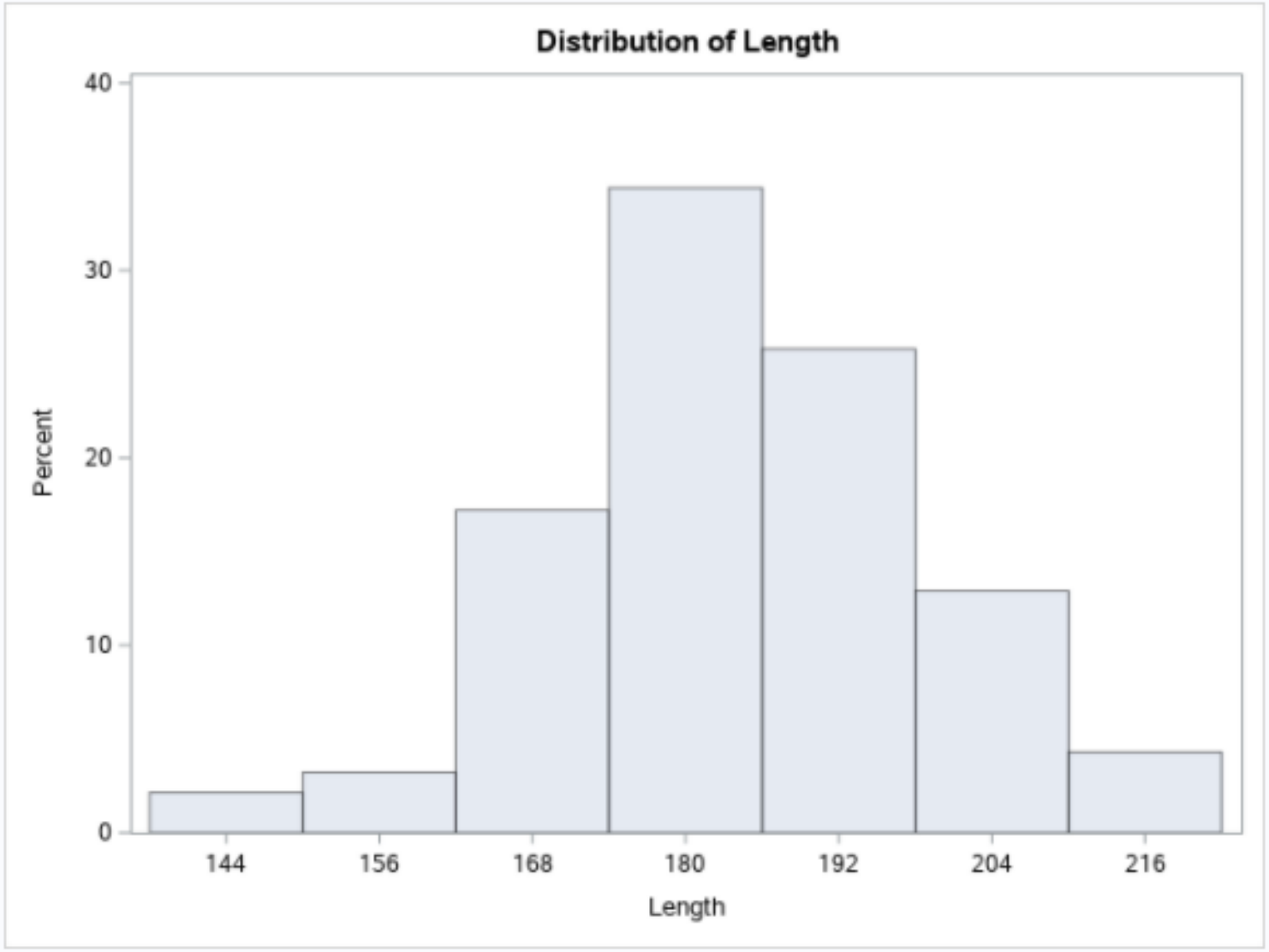
The MEANS Procedure

Analysis Variable : Weight Weight				
N	Mean	Std Dev	Minimum	Maximum
93	3072.9	589.9	1695.0	4105.0

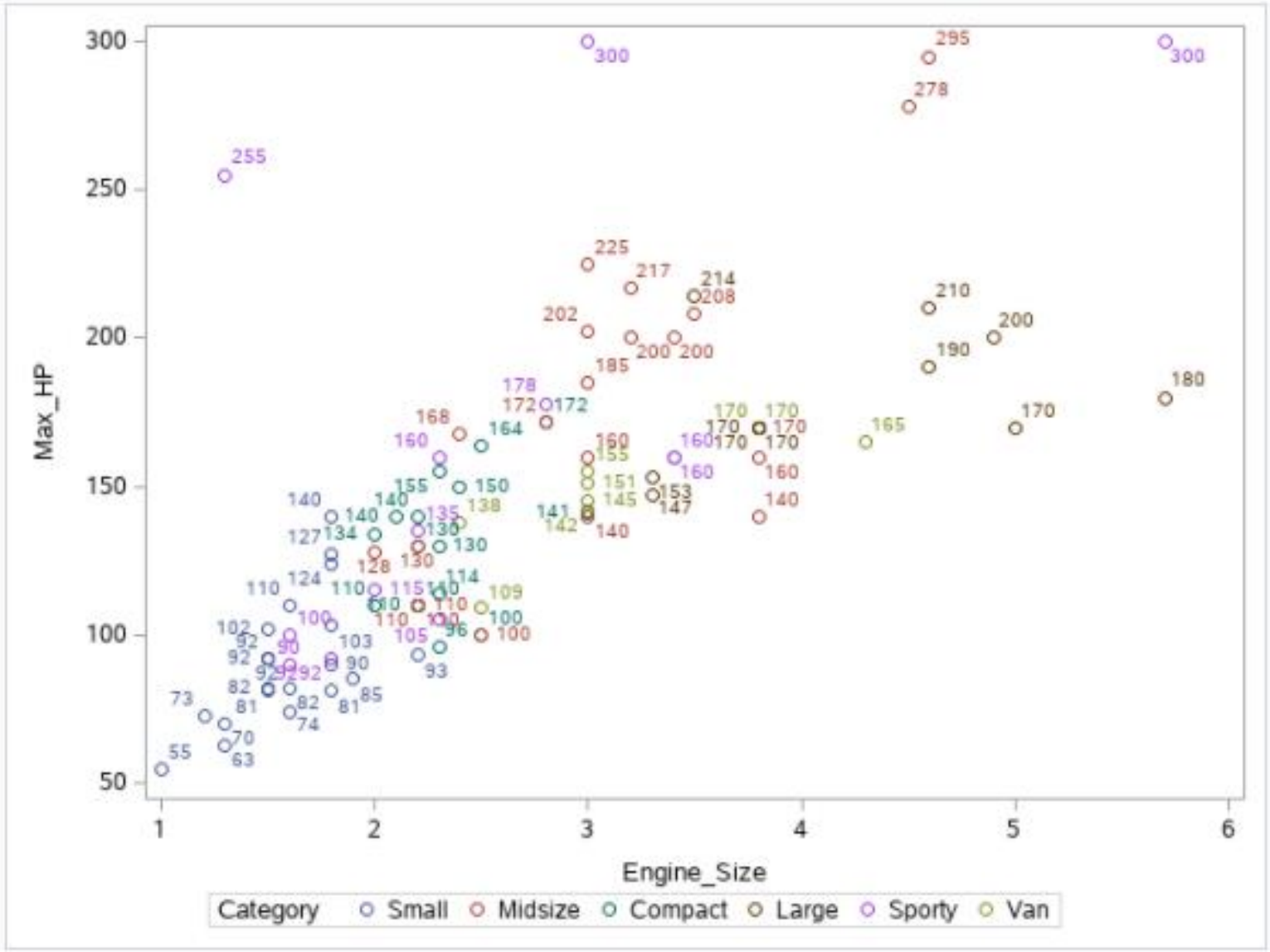
Out of 93 Cars the Heaviest car was of 4105 pounds and lightest car was of 1695 pounds.
The average Weight of all 93 cars is 3072.9 pounds.

Summary of length of cars in percentage:

The graph states that 35% of cars from the Dataset have length around 180 inches.



Scatter plot of Max_HP*Engine_Size grouped by Category:



Looking at the plot we can see that Max_HP and Engine_Size are some what related to each other.

My Observations, Commentary about the Data:

1. Inconsistent Data because there are null values.
2. Units of measurement are absent for all variables.
3. Very small dataset but includes lot of information in terms of column but less data in terms of row.