In [2]:
```
###### Question 1 - Import Data
import pandas as pd
import os

path = "D:\\Documents\\DAAN862"
os.chdir(path)

cars = pd.read_csv("mtcars.csv")

cars
```

Out[2]:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
5	Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
6	Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
7	Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
8	Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
9	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
10	Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
11	Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
12	Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
13	Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
14	Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
15	Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4
16	Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
17	Fiat 128	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
18	Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
19	Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.90	1	1	4	1
20	Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
21	Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
22	AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2
23	Camaro Z28	13.3	8	350.0	245	3.73	3.840	15.41	0	0	3	4
24	Pontiac Firebird	19.2	8	400.0	175	3.08	3.845	17.05	0	0	3	2
25	Fiat X1-9	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
26	Porsche 914-2	26.0	4	120.3	91	4.43	2.140	16.70	0	1	5	2
27	Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
28	Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.50	0	1	5	4
29	Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.50	0	1	5	6
30	Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	5	8
31	Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.60	1	1	4	2

In [3]:
```
###### Question 2 - Perform Statistical Analysis
cars.describe()

# From the produced statistical profile of the data set, we can see the count, average, min, max,
# and standard deviations of the data set. This profile is useful in understanding where the
# data falls amongst this set.
```

Out[3]:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
count	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.0000
mean	20.090625	6.187500	230.721875	146.687500	3.596563	3.217250	17.848750	0.437500	0.406250	3.687500	2.8125
std	6.026948	1.785922	123.938694	68.562868	0.534679	0.978457	1.786943	0.504016	0.498991	0.737804	1.6152
min	10.400000	4.000000	71.100000	52.000000	2.760000	1.513000	14.500000	0.000000	0.000000	3.000000	1.0000
25%	15.425000	4.000000	120.825000	96.500000	3.080000	2.581250	16.892500	0.000000	0.000000	3.000000	2.0000
50%	19.200000	6.000000	196.300000	123.000000	3.695000	3.325000	17.710000	0.000000	0.000000	4.000000	2.0000
75%	22.800000	8.000000	326.000000	180.000000	3.920000	3.610000	18.900000	1.000000	1.000000	4.000000	4.0000
max	33.900000	8.000000	472.000000	335.000000	4.930000	5.424000	22.900000	1.000000	1.000000	5.000000	8.0000

In [4]:
```
###### Question 3
cars_gear = cars.iloc[:, [0, 1, 10]]
cars_gear.sort_values("mpg", axis=0, ascending=False)

# From the produced table, we can see that in general, cars with more gears tend to get
# better MPG but the it is not a clear comparison as there are outliers such as car ID 28
# (Ford Pantera) and car ID 20 (Toyota Corona). Additionally, since the range of gears is
# only 3-5, we also see plenty of middle values (gear = 4) at the top of the MPG list.
```

Out[4]:

	model	mpg	gear
19	Toyota Corolla	33.9	4
17	Fiat 128	32.4	4
27	Lotus Europa	30.4	5
18	Honda Civic	30.4	4
25	Fiat X1-9	27.3	4
26	Porsche 914-2	26.0	5
7	Merc 240D	24.4	4
2	Datsun 710	22.8	4
8	Merc 230	22.8	4
20	Toyota Corona	21.5	3
31	Volvo 142E	21.4	4
3	Hornet 4 Drive	21.4	3
1	Mazda RX4 Wag	21.0	4
0	Mazda RX4	21.0	4
29	Ferrari Dino	19.7	5
9	Merc 280	19.2	4
24	Pontiac Firebird	19.2	3
4	Hornet Sportabout	18.7	3
5	Valiant	18.1	3
10	Merc 280C	17.8	4
12	Merc 450SL	17.3	3
11	Merc 450SE	16.4	3
28	Ford Pantera L	15.8	5
21	Dodge Challenger	15.5	3
13	Merc 450SLC	15.2	3
22	AMC Javelin	15.2	3
30	Maserati Bora	15.0	5
16	Chrysler Imperial	14.7	3
6	Duster 360	14.3	3
23	Camaro Z28	13.3	3
15	Lincoln Continental	10.4	3
14	Cadillac Fleetwood	10.4	3

In [5]:
```
cars_gear.groupby('gear', as_index=False)['mpg'].describe()

# When averaging, the 4-gear value actually has the best MPG results, which could be related
# to the difference of other factors related to the 4-gear cars.
```

Out[5]:

	gear	count	mean	std	min	25%	50%	75%	max
0	3	15.0	16.106667	3.371618	10.4	14.5	15.5	18.400	21.5
1	4	12.0	24.533333	5.276764	17.8	21.0	22.8	28.075	33.9
2	5	5.0	21.380000	6.658979	15.0	15.8	19.7	26.000	30.4

In [6]:
```
###### Question 4
cars_carb = cars.iloc[:, [0, 1, 11]]
cars_carb.sort_values("mpg", axis=0, ascending=False)

# From the produced table, we can see a similar relation where a lower Carb value
# trends to better MPG. All cars with an MPG of 21.4 or higher have only 1-2 Carb values
# and generally cars with or 4 carbs are produce a lower MPG.
```

Out[6]:

	model	mpg	carb
19	Toyota Corolla	33.9	1
17	Fiat 128	32.4	1
27	Lotus Europa	30.4	2
18	Honda Civic	30.4	2
25	Fiat X1-9	27.3	1
26	Porsche 914-2	26.0	2
7	Merc 240D	24.4	2
2	Datsun 710	22.8	1
8	Merc 230	22.8	2
20	Toyota Corona	21.5	1
31	Volvo 142E	21.4	2
3	Hornet 4 Drive	21.4	1
1	Mazda RX4 Wag	21.0	4
0	Mazda RX4	21.0	4
29	Ferrari Dino	19.7	6
9	Merc 280	19.2	4
24	Pontiac Firebird	19.2	2
4	Hornet Sportabout	18.7	2
5	Valiant	18.1	1
10	Merc 280C	17.8	4
12	Merc 450SL	17.3	3
11	Merc 450SE	16.4	3
28	Ford Pantera L	15.8	4
21	Dodge Challenger	15.5	2
13	Merc 450SLC	15.2	3
22	AMC Javelin	15.2	2
30	Maserati Bora	15.0	8
16	Chrysler Imperial	14.7	4
6	Duster 360	14.3	4
23	Camaro Z28	13.3	4
15	Lincoln Continental	10.4	4
14	Cadillac Fleetwood	10.4	4

In [7]:
```
cars_carb.groupby('carb', as_index=False)['mpg'].describe()

# In the produced averages, the carb values seems to indicate lower carb equates
# to better MPG as well where not enough data is available to determine
# the results of cars with 6 or 8 carbs
```

Out[7]:

	carb	count	mean	std	min	25%	50%	75%	max
0	1	7.0	25.342857	6.001349	18.1	21.450	22.80	29.85	33.9
1	2	10.0	22.400000	5.472152	15.2	18.825	22.10	25.60	30.4
2	3	3.0	16.300000	1.053565	15.2	15.800	16.40	16.85	17.3
3	4	10.0	15.790000	3.911081	10.4	13.550	15.25	18.85	21.0
4	6	1.0	19.700000	NaN	19.7	19.700	19.70	19.70	19.7
5	8	1.0	15.000000	NaN	15.0	15.000	15.00	15.00	15.0

In [8]:
```
###### Question 5 - Determine Correlation to MPG
cars_pruned = cars.iloc[:, 1:11]
cars_pruned.corrwith(cars.mpg)
```

Out[8]:
```
mpg     1.000000
cyl    -0.852162
disp   -0.847551
hp     -0.776168
drat    0.681172
wt     -0.867659
qsec    0.418684
vs      0.664039
am      0.599832
gear    0.480285
dtype: float64
```

In [9]:
```
cars_pruned.cov().iloc[0, :]
```

Out[9]:
```
mpg      36.324103
cyl      -9.172379
disp   -633.097208
hp     -320.732056
drat      2.195064
wt       -5.116685
qsec      4.509149
vs        2.017137
am        1.803931
gear      2.135685
Name: mpg, dtype: float64
```

In [10]:
```
cars_pruned.cov()
```

Out[10]:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear
mpg	36.324103	-9.172379	-633.097208	-320.732056	2.195064	-5.116685	4.509149	2.017137	1.803931	2.135685
cyl	-9.172379	3.189516	199.660282	101.931452	-0.668367	1.367371	-1.886855	-0.729839	-0.465726	-0.649194
disp	-633.097208	199.660282	15360.799829	6721.158669	-47.064019	107.684204	-96.051681	-44.377621	-36.564012	-50.802621
hp	-320.732056	101.931452	6721.158669	4700.866935	-16.451109	44.192661	-86.770081	-24.987903	-8.320565	-6.358871
drat	2.195064	-0.668367	-47.064019	-16.451109	0.285881	-0.372721	0.087141	0.118649	0.190151	0.275988
wt	-5.116685	1.367371	107.684204	44.192661	-0.372721	0.957379	-0.305482	-0.273661	-0.338105	-0.421081
qsec	4.509149	-1.886855	-96.051681	-86.770081	0.087141	-0.305482	3.193166	0.670565	-0.204960	-0.280403
vs	2.017137	-0.729839	-44.377621	-24.987903	0.118649	-0.273661	0.670565	0.254032	0.042339	0.076613
am	1.803931	-0.465726	-36.564012	-8.320565	0.190151	-0.338105	-0.204960	0.042339	0.248992	0.292339
gear	2.135685	-0.649194	-50.802621	-6.358871	0.275988	-0.421081	-0.280403	0.076613	0.292339	0.544355

In [ ]: