Date 05-07-2021

Bubble Plots

In [1]:
```
import pandas as pd
import plotly.offline as pyo
import plotly.graph_objs as go
```

In [2]:
```
mpg_data_csv = pd.read_csv("mpg.csv")
mpg_data_csv
```

Out[2]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model_year	origin	name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino
...	...	...	...	...	...	...	...	...	...
393	27.0	4	140.0	86	2790	15.6	82	1	ford mustang gl
394	44.0	4	97.0	52	2130	24.6	82	2	vw pickup
395	32.0	4	135.0	84	2295	11.6	82	1	dodge rampage
396	28.0	4	120.0	79	2625	18.6	82	1	ford ranger
397	31.0	4	119.0	82	2720	19.4	82	1	chevy s-10

398 rows × 9 columns

In [3]:
```
mpg_data_csv.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 398 entries, 0 to 397
Data columns (total 9 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   mpg           398 non-null    float64
 1   cylinders     398 non-null    int64  
 2   displacement  398 non-null    float64
 3   horsepower    398 non-null    object 
 4   weight        398 non-null    int64  
 5   acceleration  398 non-null    float64
 6   model_year    398 non-null    int64  
 7   origin        398 non-null    int64  
 8   name          398 non-null    object 
dtypes: float64(3), int64(4), object(2)
memory usage: 28.1+ KB
```

In [4]:
```
mpg_data_csv.describe()
```

Out[4]:

	mpg	cylinders	displacement	weight	acceleration	model_year	origin
count	398.000000	398.000000	398.000000	398.000000	398.000000	398.000000	398.000000
mean	23.514573	5.454774	193.425879	2970.424623	15.568090	76.010050	1.572864
std	7.815984	1.701004	104.269838	846.841774	2.757689	3.697627	0.802055
min	9.000000	3.000000	68.000000	1613.000000	8.000000	70.000000	1.000000
25%	17.500000	4.000000	104.250000	2223.750000	13.825000	73.000000	1.000000
50%	23.000000	4.000000	148.500000	2803.500000	15.500000	76.000000	1.000000
75%	29.000000	8.000000	262.000000	3608.000000	17.175000	79.000000	2.000000
max	46.600000	8.000000	455.000000	5140.000000	24.800000	82.000000	3.000000

In [5]:
```
mpg_data_csv.columns
```

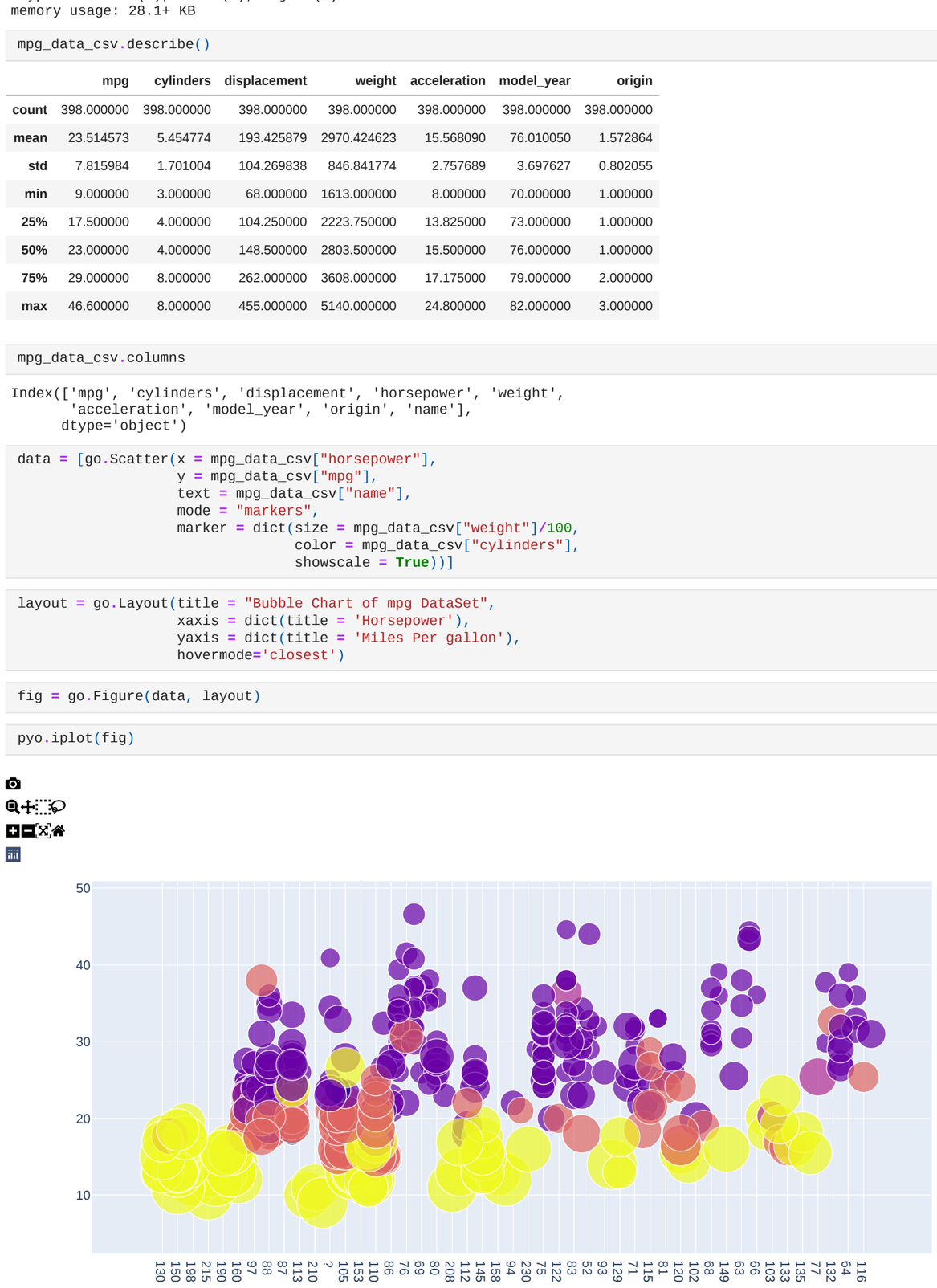
Out[5]:
```
Index(['mpg', 'cylinders', 'displacement', 'horsepower', 'weight',
       'acceleration', 'model_year', 'origin', 'name'],
      dtype='object')
```

In [6]:
```
data = [go.Scatter(x = mpg_data_csv["horsepower"],
                   y = mpg_data_csv["mpg"],
                   text = mpg_data_csv["name"],
                   mode = "markers",
                   marker = dict(size = mpg_data_csv["weight"]/100,
                                 color = mpg_data_csv["cylinders"],
                                 showscale = True))]
```

In [7]:
```
layout = go.Layout(title = "Bubble Chart of mpg DataSet",
                   xaxis = dict(title = 'Horsepower'),
                   yaxis = dict(title = 'Miles Per gallon'),
                   hovermode='closest')
```

In [8]:
```
fig = go.Figure(data, layout)
```

In [9]:
```
pyo.iplot(fig)
```



In [10]:
```
pyo.plot(fig,
         filename = "tutorial_10 (Bubble Plots).html",
         image_width=1600,
         image_height=900,)
```

Out[10]:
```
'tutorial_10 (Bubble Plots).html'
```