In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
# to view the plot in jupiter notebook
%matplotlib inline
```

In [28]:
```
df_auto_dataset = pd.read_csv(r"C:\Users\ctoqu\Desktop\auto_data.csv")
```

In [29]:
```
df_auto_dataset.head(5)
```

Out[29]:

	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

In [31]:
```
# User define function for origin
#1-USA, 2-Europe, 3-Asia

def origin(num):
    if num==1:
        return 'USA'
    elif num==2:
        return 'Europe'
    else:
        return 'Asia'

#use apply function
df_auto_dataset['origin'] = df_auto_dataset['origin'].apply(origin)
```

In [32]:
```
df_auto_dataset.head(30)
```

Out[32]:

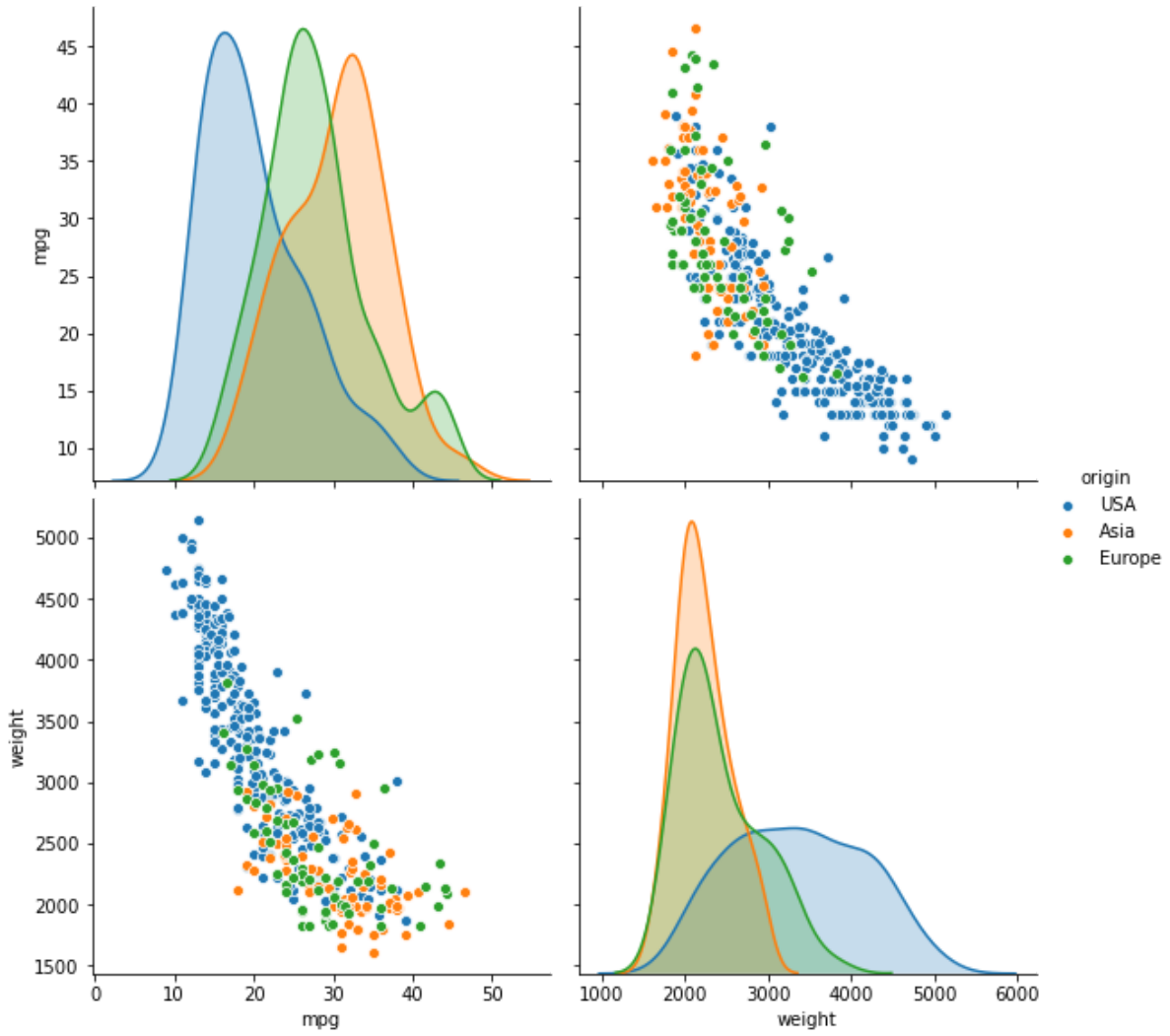
	mpg	cylinders	displacement	horsepower	weight	acceleration	model_year	origin	
0	18.0	8	307.0	130	3504	12.0	70	USA	che ch n
1	15.0	8	350.0	165	3693	11.5	70	USA	skylar
2	18.0	8	318.0	150	3436	11.0	70	USA	plyr sa
3	16.0	8	304.0	150	3433	12.0	70	USA	amc
4	17.0	8	302.0	140	3449	10.5	70	USA	ford
5	15.0	8	429.0	198	4341	10.0	70	USA	ford g
6	14.0	8	454.0	220	4354	9.0	70	USA	che ir
7	14.0	8	440.0	215	4312	8.5	70	USA	plyr
8	14.0	8	455.0	225	4425	10.0	70	USA	p ca
9	15.0	8	390.0	190	3850	8.5	70	USA	ambas
10	15.0	8	383.0	170	3563	10.0	70	USA	chall
11	14.0	8	340.0	160	3609	8.0	70	USA	plyr 'cud
12	15.0	8	400.0	150	3761	9.5	70	USA	che monte
13	14.0	8	455.0	225	3086	10.0	70	USA	buick e wagor
14	24.0	4	113.0	95	2372	15.0	70	Asia	t corona
15	22.0	6	198.0	95	2833	15.5	70	USA	plyr
16	18.0	6	199.0	97	2774	15.5	70	USA	amc h
17	21.0	6	200.0	85	2587	16.0	70	USA	ma
18	27.0	4	97.0	88	2130	14.5	70	Asia	d
19	26.0	4	97.0	46	1835	20.5	70	Europe	volksw 1131 d
20	25.0	4	110.0	87	2672	17.5	70	Europe	peuge
21	24.0	4	107.0	90	2430	14.5	70	Europe	audi
22	25.0	4	104.0	95	2375	17.5	70	Europe	saa

	mpg	cylinders	displacement	horsepower	weight	acceleration	model_year	origin	
23	26.0	4	121.0	113	2234	12.5	70	Europe	bmw
24	21.0	6	199.0	90	2648	15.0	70	USA	amc gr
25	10.0	8	360.0	215	4615	14.0	70	USA	forc
26	10.0	8	307.0	200	4376	15.0	70	USA	chev
27	11.0	8	318.0	210	4382	13.5	70	USA	dodge
28	9.0	8	304.0	193	4732	18.5	70	USA	hi
29	27.0	4	97.0	88	2130	14.5	71	Asia	d

In [34]:
```
# pair plot using sns for mpg, weight, orign and with hue origin, set the size to 4
#note: hue is variable in data set to map plot aspects to different colors

sns.pairplot(df_auto_dataset[['mpg','weight','origin']],hue='origin',height=4)
```

Out[34]: <seaborn.axisgrid.PairGrid at 0x1f343ca0190>



In [ ]: