

```
In [6]: import pandas as pd
import numpy as np
import plotly
import plotly.graph_objs as go
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.offline as plt
from plotly.offline import init_notebook_mode, download_plotlyjs, iplot, plot
init_notebook_mode(inline=True)
```

```
In [8]: init_notebook_mode(connected=True)
df.go_offline()
```

```
In [9]: df = pd.read_csv('PUBG.csv')
```

```
In [10]: df2 = df.apply(pd.to_numeric, errors='ignore')
df_solo = df2[['player_name','solo_RoundsPlayed','solo_Wins']].head(20)
df_duo = df2[['player_name','duo_RoundsPlayed','duo_Wins']].head(20)
df_squad = df2[['player_name','squad_RoundsPlayed','squad_Wins']].head(20)
wins = df2[['player_name','solo_WinRatio','duo_WinRatio','squad_WinRatio']].head(20)
```

```
In [11]: df.head()
```

	player_name	trucker_id	solo_KillDeathRatio	solo_WinRatio	solo_TimeSurvived	solo_RoundsPlayed	solo_Wins	solo_WinTop5Ratio	solo_Top10s	solo_1
0	BreakNeck	4405	3.14	17.65	18469.14	17	3	0.83	4	
1	Blackwalk	8159	4.41	18.18	33014.86	33	6	0.36	11	
2	mercedes_benz	4454	3.60	0.00	4330.44	5	0	0.00	1	
3	DORA	7729	14.00	50.00	13421.82	8	4	0.67	6	
4	n2star	0	10.50	33.33	9641.04	6	2	0.40	5	

5 rows × 11 columns

```
In [12]: df_solo.head()
```

	player_name	solo_RoundsPlayed	solo_Wins
0	BreakNeck	17	3
1	Blackwalk	33	6
2	mercedes_benz	5	0
3	DORA	8	4
4	n2star	6	2

```
In [13]: df_duo.head()
```

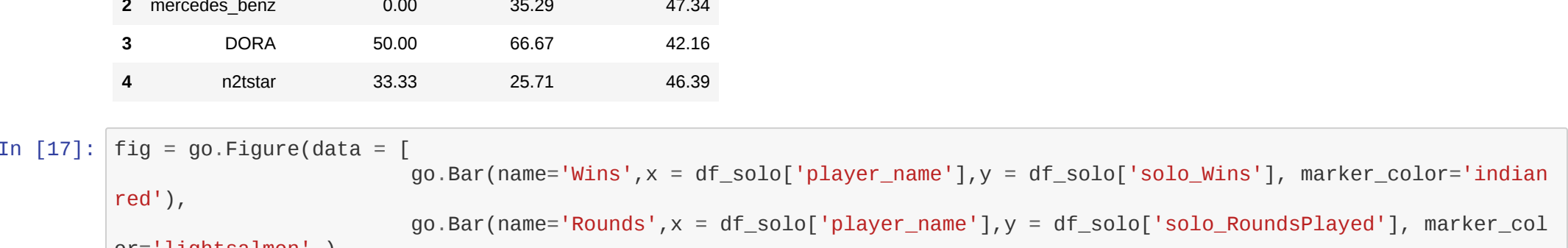
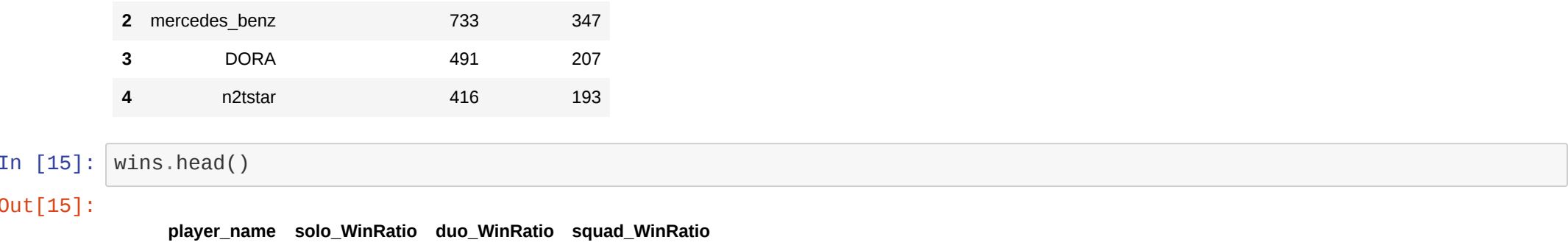
	player_name	duo_RoundsPlayed	duo_Wins
0	BreakNeck	15	5
1	Blackwalk	14	5
2	mercedes_benz	17	6
3	DORA	3	2
4	n2star	105	27

```
In [14]: df_squad.head()
```

	player_name	squad_RoundsPlayed	squad_Wins
0	BreakNeck	642	305
1	Blackwalk	722	338
2	mercedes_benz	733	347
3	DORA	481	287
4	n2star	416	193

```
In [15]: wins.head()
```

	player_name	solo_WinRatio	duo_WinRatio	squad_WinRatio
0	BreakNeck	17.65	33.33	47.51
1	Blackwalk	18.18	35.71	46.81
2	mercedes_benz	0.00	35.29	47.34
3	DORA	50.00	66.67	42.16
4	n2star	33.33	25.71	46.39



```
In [13]: from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler
```

```
In [14]: wins.head()
```

	player_name	solo_WinRatio	duo_WinRatio	squad_WinRatio
0	BreakNeck	17.65	33.33	47.51
1	Blackwalk	18.18	35.71	46.81
2	mercedes_benz	0.00	35.29	47.34
3	DORA	50.00	66.67	42.16
4	n2star	33.33	25.71	46.39

```
In [15]: wins_x = wins.iloc[:,1:4]
wins_x.head()
```

	solo_WinRatio	duo_WinRatio	squad_WinRatio
0	17.65	33.33	47.51
1	18.18	35.71	46.81
2	0.00	35.29	47.34
3	50.00	66.67	42.16
4	33.33	25.71	46.39

```
In [16]: x_array = np.array(wins_x)
```

```
Out[16]: array([[17.65, 33.33, 47.51],
 [18.18, 35.71, 46.81],
 [ 0. , 35.29, 47.34],
 [50. , 66.67, 42.16],
 [33.33, 25.71, 46.39],
 [31.25, 14.67, 32.11],
 [ 8.45, 23.31, 33. ],
 [29.32, 18.52, 38.32],
 [31.64, 13.64, 32.89],
 [25. , 17.65, 34.83],
 [ 6.83, 11.46, 27.44],
 [ 1.19, 3.19, 18.5 ],
 [10.34, 24.49, 27.42],
 [ 5.56, 12.82, 27.52],
 [13.84, 21.48, 39.78],
 [33.33, 28.57, 44.41],
 [23.03, 33.33, 41. ],
 [18.46, 32.47, 35.21],
 [37.5 , 21.5 , 37.84],
 [ 4.44, 3.51, 17.91]])
```

```
In [17]: scaler = MinMaxScaler()
x_scaled = scaler.fit_transform(x_array)
x_scaled
```

```
Out[17]: array([[0.353 , 0.47479521, 1. ],
 [0.3638 , 0.51228733, 0.97635135],
 [0. , 0.38567108, 0.38422678],
 [1. , 1. , 0.81925676],
 [0.6566 , 0.3547574 , 0.36216216],
 [0.625 , 0.18684436, 0.47938189],
 [0.169 , 0.31695822, 0.5897973 ],
 [0.5264 , 0.24149338, 0.41925676],
 [0.2728 , 0.15461878, 0.56688188],
 [0. , 0.22778828, 0.57162162],
 [0.1286 , 0.13527725, 0.32188846],
 [0.0238 , 0. , 0.01993243],
 [0.2668 , 0.33533915, 0.32128378],
 [0.1112 , 0.15178132, 0.32468216],
 [0.2668 , 0.28812254, 0.73885135],
 [0.6566 , 0.38581086, 0.89217021],
 [0.4668 , 0.47479521, 0.78986757],
 [0.2692 , 0.40214764, 0.58445446],
 [0.75 , 0.2854373 , 0.67668919],
 [0.0888 , 0.08504966, 0. ]])
```

```
In [18]: kmeans = KMeans(n_clusters = 3, random_state = 100)
kmeans.fit(x_scaled)
```

```
Out[18]: KMeans(n_clusters=3, random_state=100)
```

```
In [19]: kmeans.cluster_centers_
```

```
Out[19]: array([[0.189925 , 0.16828258, 0.38287162],
 [0.45594545, 0.37895973, 0.78719287],
 [1. , 1. , 0.81925678]])
```

```
In [20]: print(kmeans.labels_)
```

```
Out[20]: [1 1 2 1 1 0 0 0 1 0 0 0 0 1 1 1 1 1 0]
```

```
In [21]: wins['clusters'] = kmeans.labels_
```

```
Out[21]: wins
```

```
In [22]: df.head()
```

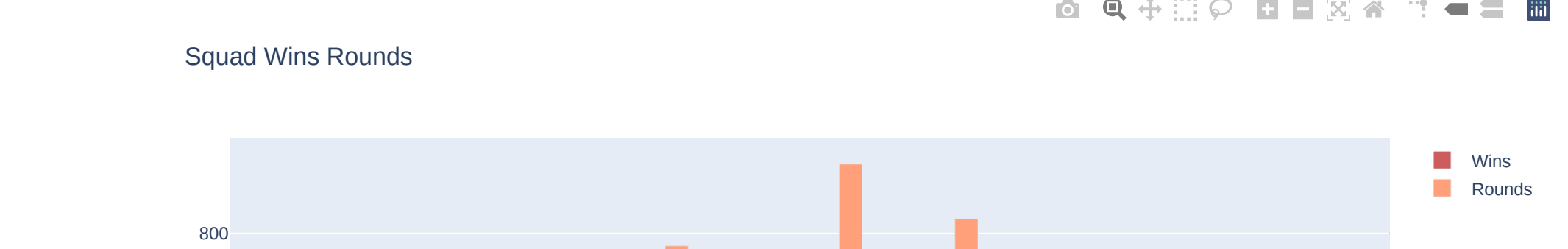
	player_name	solo_WinRatio	duo_WinRatio	squad_WinRatio	clusters
0	BreakNeck	17.65	33.33	47.51	1
1	Blackwalk	18.18	35.71	46.81	1
2	mercedes_benz	0.00	35.29	47.34	1
3	DORA	50.00	66.67	42.16	2
4	n2star	33.33	25.71	46.39	1
5	coldkorygen	31.25	14.67	32.10	1
6	Giken	8.45	23.31	33.00	0
7	KoreaKoi	26.32	18.52	30.32	0
8	undor	13.64	13.64	32.89	0
9	Fortune	25.00	17.65	34.83	1
10	PandaTV_Tongk	6.03	11.46	27.44	0
11	Berry_	1.19	3.18	18.50	0
12	PanDaTV_Karl	10.34	24.49	27.42	0
13	homeloye	5.56	12.82	27.52	0
14	YUPPIEE	13.04	21.48	39.78	1
15	Doomhammer	33.33	28.57	44.41	1
16	Dinghokhan	23.03	33.33	41.00	1
17	YoonRoot	18.46	32.47	35.21	1
18	3NIGMA	37.50	21.50	37.84	1
19	synthral	4.44	3.51	17.91	0

```
In [23]: output = plt.scatter(x_scaled[:,0],
x_scaled[:,1],
c = wins_clusters,
marker = 'o',
alpha = 1)
```

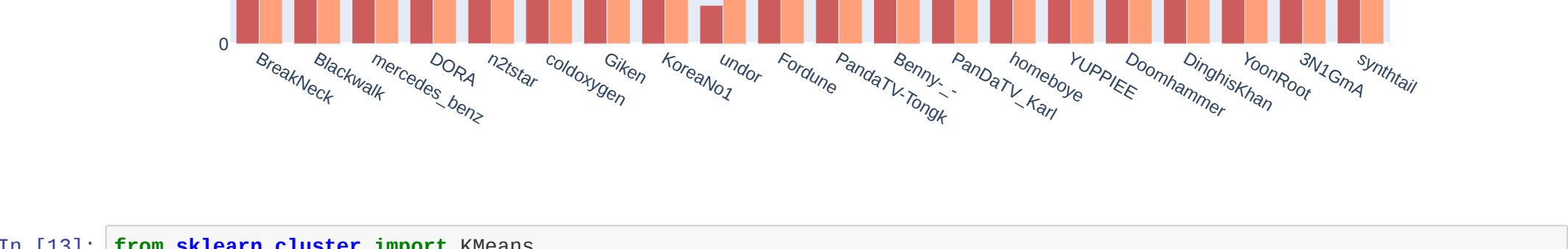
```
centers = kmeans.cluster_centers_
plt.scatter(centers[:,0],
centers[:,1],
c = 'red',
s = 100,
alpha = 0,
marker = 's')
```

```
plt.title('Hasil Clustering')
plt.colorbar(output)
```

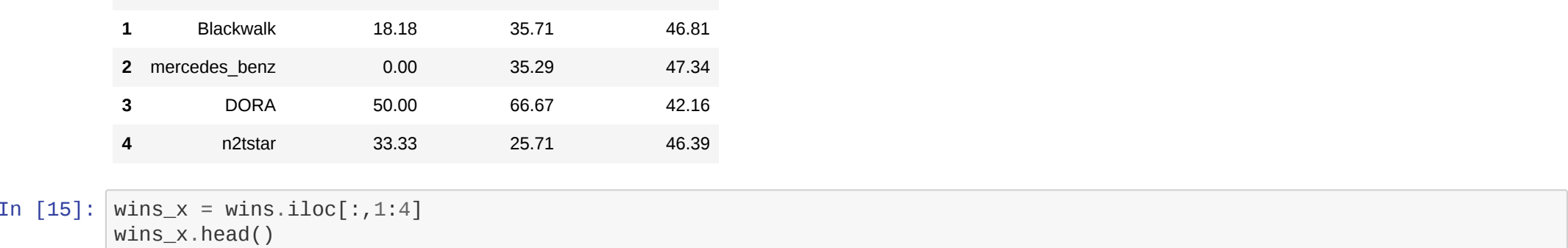
```
Out[23]: <matplotlib.colorbar.Colorbar at 6x172967388c>
```



```
In [24]: fig = go.Figure()
fig.add_trace(go.Scatter(
    x = wins['clusters'] == 0,
    y = wins['player_name'],
    marker=dict(color='purple',size=13),
    mode='markers',
    name='cluster 0'
))
fig.add_trace(go.Scatter(
    x = wins['clusters'] == 1,
    y = wins['player_name'],
    marker=dict(color='darkseagreen', size = 12),
    mode='markers',
    name='cluster 1'
))
fig.add_trace(go.Scatter(
    x = wins['clusters'] == 2,
    y = wins['player_name'],
    marker=dict(color='gold', size = 12),
    mode='markers',
    name='cluster 2'
))
```



```
In [25]: fig = go.Figure()
fig.add_trace(go.Scatter(
    x = wins['solo_WinRatio'],
    y = wins['player_name'],
    marker=dict(color='EC5E64', size = 13),
    mode = 'markers',
    name = 'Solo'
))
fig.add_trace(go.Scatter(
    x = wins['duo_WinRatio'],
    y = wins['player_name'],
    marker=dict(color='985E5B', size = 13),
    mode = 'markers',
    name = 'Duo'
))
fig.add_trace(go.Scatter(
    x = wins['squad_WinRatio'],
    y = wins['player_name'],
    marker=dict(color='FFFA4B', size = 13),
    mode = 'markers',
    name = 'Squad'
))
fig.show()
```



```
In [7]: pwd
```

```
Out[7]: C:\Users\Aneke1
```

```
In [9]: cd \
```

```
C:\
```

```
In [10]: cd D:
```

```
D:\
```

```
In [11]: D:\penulisan iliniah
```

```
File "c:\python-input-11-8a834bab1e24", line 1
D:\penulisan iliniah
^
SyntaxError: unexpected character after line continuation character
```

```
In [12]: pwd
```

```
Out[12]: C:\Users\Aneke1
```

```
In [5]: cd D:\penulisan iliniah
```

```
D:\penulisan iliniah
```

```
In [20]: df
```

	player_name	trucker_id	solo_KillDeathRatio	solo_WinRatio	solo_TimeSurvived	solo_RoundsPlayed	solo_Wins	solo_WinTop5Ratio	solo_Top10s	solo_1
0	BreakNeck	4405	3.14	17.65	18469.14	17	3	0.83	4	
1	Blackwalk	8159	4.41	18.18	33014.86	33	6	0.36	11	
2	mercedes_benz	4454	3.60	0.00	4330.44	5	0	0.00	1	
3	DORA	7729	14.00	50.00	13421.82	8	4	0.67	6	
4	n2star	0	10.50	33.33	9641.04	6	2	0.40	5	

87898 rows × 11 columns

```
In [24]: df_solo
```

	player_name	solo_RoundsPlayed	solo_Wins
0	BreakNeck	17	3
1	Blackwalk	33	6
2	mercedes_benz	5	0
3	DORA	8	4
4	n2star	6	2
5	coldkorygen	16	5
6	Giken	142	12
7	KoreaKoi	19	5
8	undor	44	6
9	Fortune	8	2
10	PandaTV_Tongk	116	7
11	Berry_	335	4
12	PanDaTV_Karl	87	9
13	homeloye	72	4
14	YUPPIEE	23	3
15	Doomhammer	3	1
16	Dinghokhan	165	38
17	YoonRoot	65	12
18	3NIGMA	16	6
19	synthral	45	2

```
In [ ]:
```