In [1]:

```
import pandas as pd
```

In [11]:

```
import plotly
import cufflinks as cf
import pandas as pd
import numpy as np
```

In [12]:

```
# 오프라인 모드에서도 인터렉티브한 그래픽을 가능하도록 하기
# Enabling the offline mode for interactive plotting locally
from plotly.offline import download_plotlyjs,init_notebook_mode,plot,iplot
init_notebook_mode(connected=True)
cf.go_offline()
```

In [15]:

```
#train = pd.read_csv("./titanic/train.csv")
#test = pd.read_csv("./titanic/train.csv")

train = pd.read_csv("train.csv")
test = pd.read_csv("test.csv")
```

In [17]:

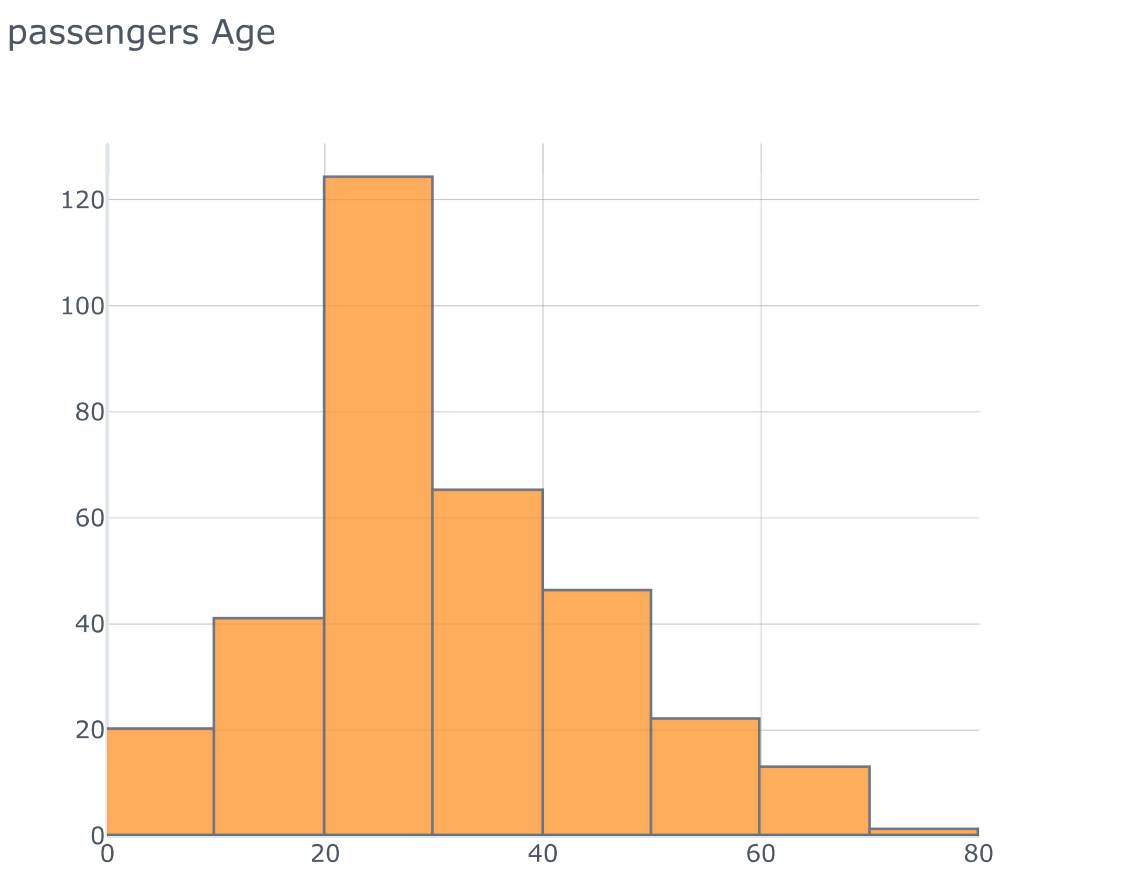
```
test.head()
```

Out[17]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Emba
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	

In [30]:

```
test['Age'].iplot(kind='histogram', bins=15, title = 'passengers Age')
```



Export to plot.ly »

In [24]:

```
test['Fare'].iplot(kind='box')
```

Export to plot.ly »

- 일부 사람들이 512 정도 요금을 지불했다.
- 25~75%의 많은 사람들이 8에서 35사이에 위치하고 있다.

In [ ]:

```

```

```

```

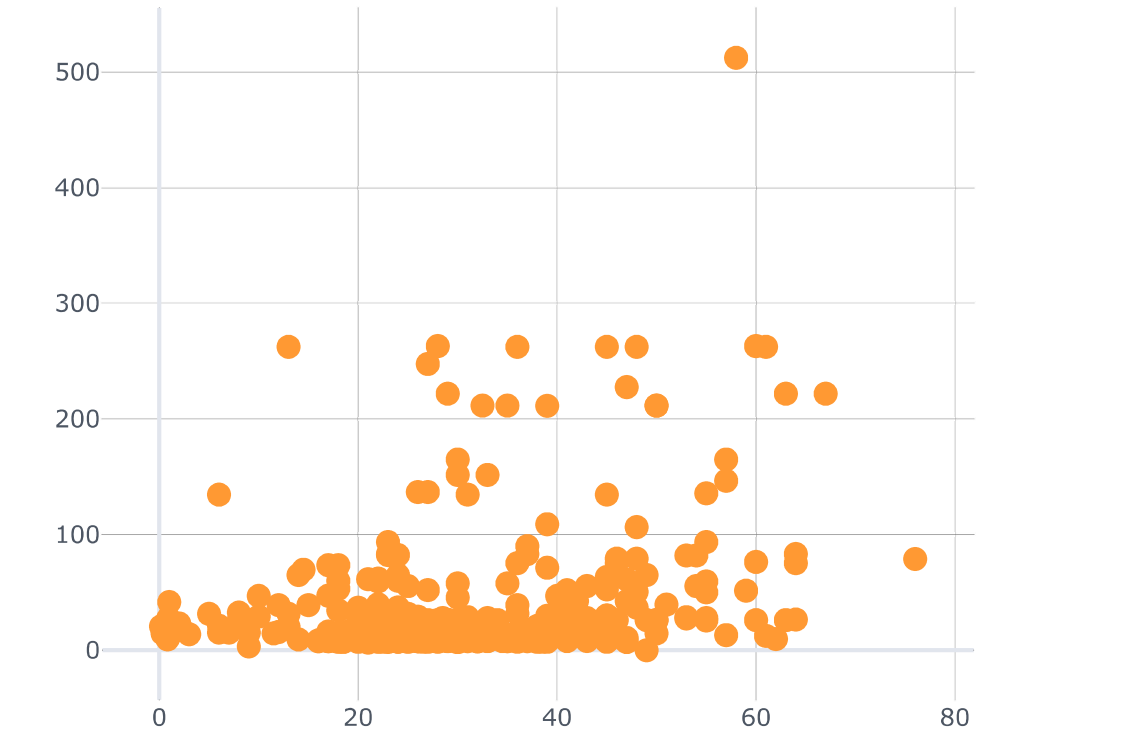
In [34]:

```
test['Sex'].iplot(kind='bar', x='male', y='female')
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
<ipython-input-34-0a5b35bdd8f3> in <module>
----> 1 test['Sex'].iplot(kind='bar', x='male', y='female')

C:\ProgramData\Anaconda3\lib\site-packages\cufflinks\plotlytools.py in _iplot(self, kind, data, layout, filename, sharing, title, xTitle, yTitle, zTitle, theme, colors, colorscale, fill, width, dash, mode, interpolation, symbol, size, barmode, sortbars, bargap, bargroupgap, bins, histnorm, histfunc, orientation, boxpoints, annotations, keys, bestfit, bestfit_colors, mean, mean_colors, categories, x, y, z, text, gridcolor, zerolinecolor, margin, labels, values, secondary_y, secondary_y_title, subplots, shape, error_x, error_y, error_type, locations, lon, lat, asFrame, asDates, asFigure, asImage, dimensions, asPlot, asUrl, online, **kwargs)
    824                                     df=pd.DataFrame({df.name:df})
    825                                 if x:
--> 826                                     df=df.set_index(x)
    827                                 if y and secondary_y:
    828                                     _y = [y] if not isinstance(y, list) else y

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\frame.py in set_index(self, keys, drop, append, inplace, verify_integrity)
   4301
   4302         if missing:
-> 4303             raise KeyError(f"None of {missing} are in the columns")
   4304
   4305         if inplace:

KeyError: "None of ['male'] are in the columns"
```

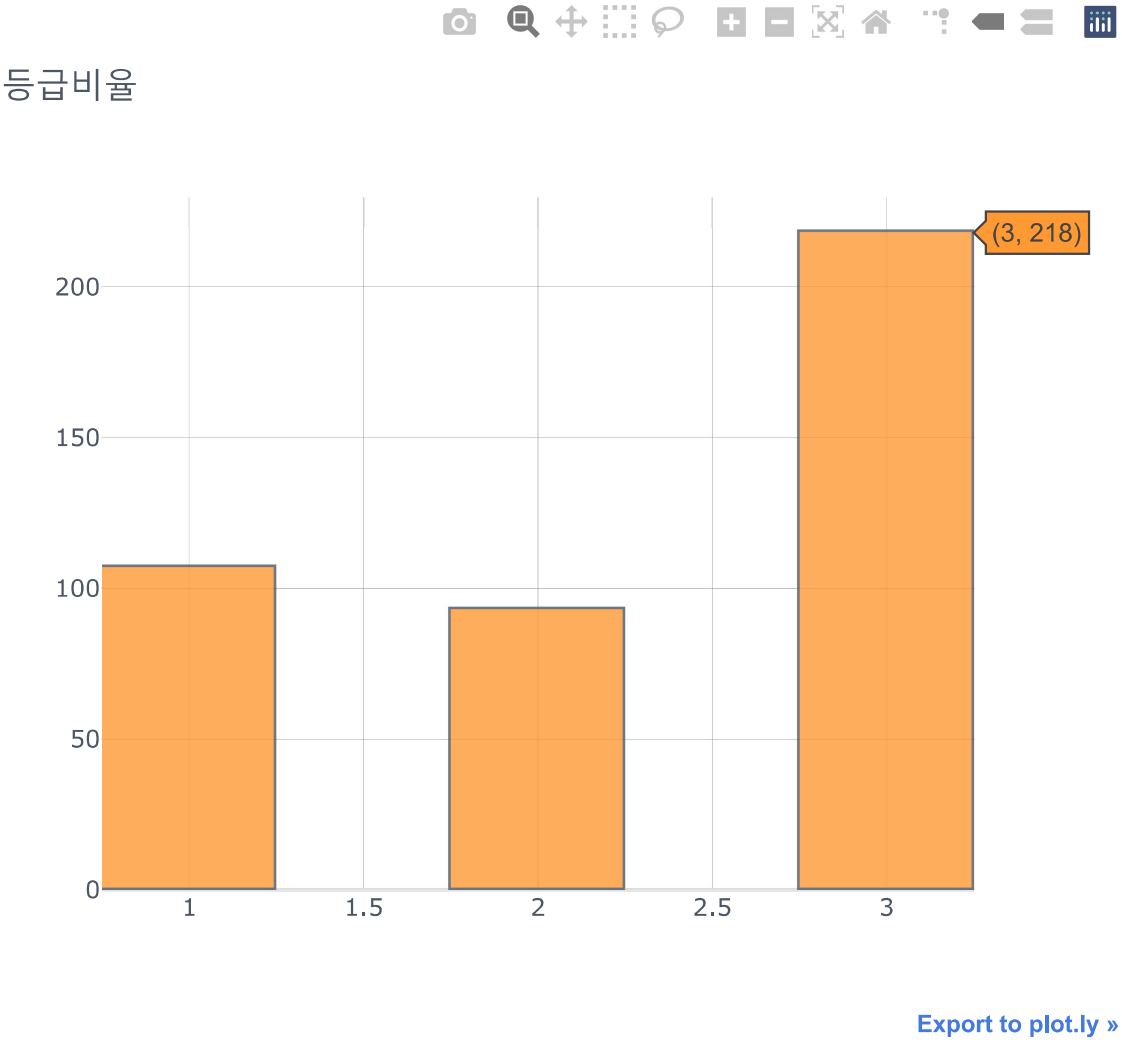
In [44]:

```
test.iplot(kind='scatter', x='Age', y='Fare', mode = 'markers', title = '연령별 요금')
```



In [48]:

```
test['Pclass'].iplot(kind='histogram', bins=10, title = '등급비율')
```



In [ ]:

```

```