

soal1_fifa

April 9, 2019

```
In [2]: import pandas as pd #Menampilkan isi dari file fifa.csv disimpan dalam dataset fifa
fifa = pd.read_csv('fifa.csv')
fifa
```

```
Out [2]:
```

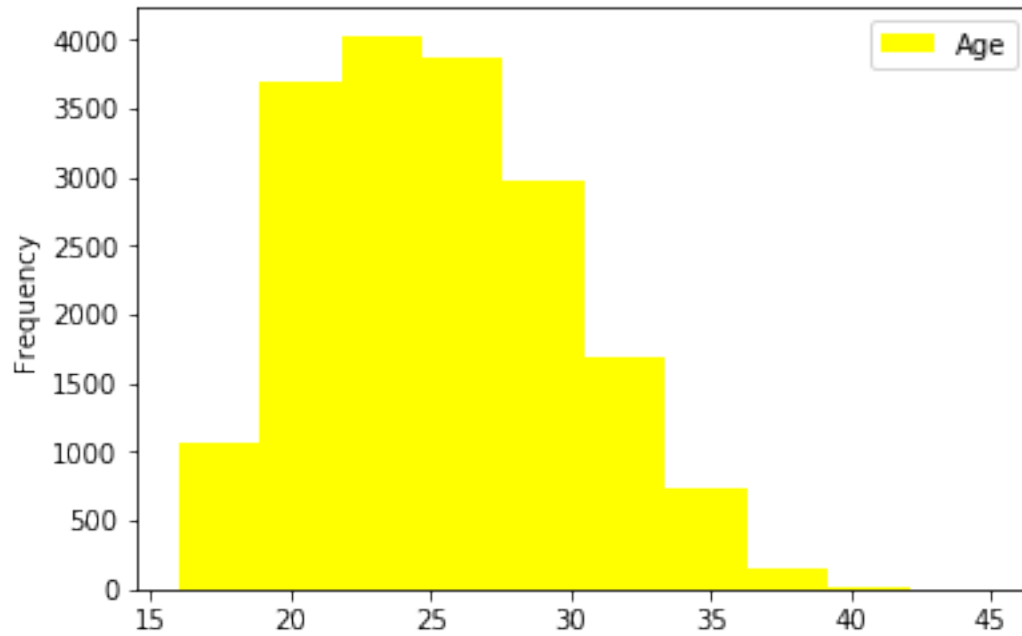
	Name	Age	Nationality
0	L. Messi	31	Argentina
1	Cristiano Ronaldo	33	Portugal
2	Neymar Jr	26	Brazil
3	De Gea	27	Spain
4	K. De Bruyne	27	Belgium
5	E. Hazard	27	Belgium
6	L. Modri	32	Croatia
7	L. Suárez	31	Uruguay
8	Sergio Ramos	32	Spain
9	J. Oblak	25	Slovenia
10	R. Lewandowski	29	Poland
11	T. Kroos	28	Germany
12	D. Godín	32	Uruguay
13	David Silva	32	Spain
14	N. Kanté	27	France
15	P. Dybala	24	Argentina
16	H. Kane	24	England
17	A. Griezmann	27	France
18	M. ter Stegen	26	Germany
19	T. Courtois	26	Belgium
20	Sergio Busquets	29	Spain
21	E. Cavani	31	Uruguay
22	M. Neuer	32	Germany
23	S. Agüero	30	Argentina
24	G. Chiellini	33	Italy
25	K. Mbappé	19	France
26	M. Salah	26	Egypt
27	Casemiro	26	Brazil
28	J. Rodríguez	26	Colombia
29	L. Insigne	27	Italy
...
18177	R. Roache	18	Republic of Ireland
18178	L. Wahlstedt	18	Sweden

18179	J. Williams	17	England
18180	M. Hurst	22	Scotland
18181	C. Maher	17	Republic of Ireland
18182	Y. Góez	18	Colombia
18183	K. Pilkington	44	England
18184	D. Horton	18	England
18185	E. Tweed	19	Republic of Ireland
18186	Zhang Yufeng	20	China PR
18187	C. Ehlich	19	Germany
18188	L. Collins	17	Wales
18189	A. Kaltner	18	Germany
18190	L. Watkins	18	England
18191	J. Norville-Williams	18	England
18192	S. Squire	18	England
18193	N. Fuentes	18	Chile
18194	J. Milli	18	Italy
18195	S. Griffin	18	Republic of Ireland
18196	K. Fujikawa	19	Japan
18197	D. Holland	18	Republic of Ireland
18198	J. Livesey	18	England
18199	M. Baldisimo	18	Canada
18200	J. Young	18	Scotland
18201	D. Walsh	18	Republic of Ireland
18202	J. Lundstram	19	England
18203	N. Christoffersson	19	Sweden
18204	B. Worman	16	England
18205	D. Walker-Rice	17	England
18206	G. Nugent	16	England

[18207 rows x 3 columns]

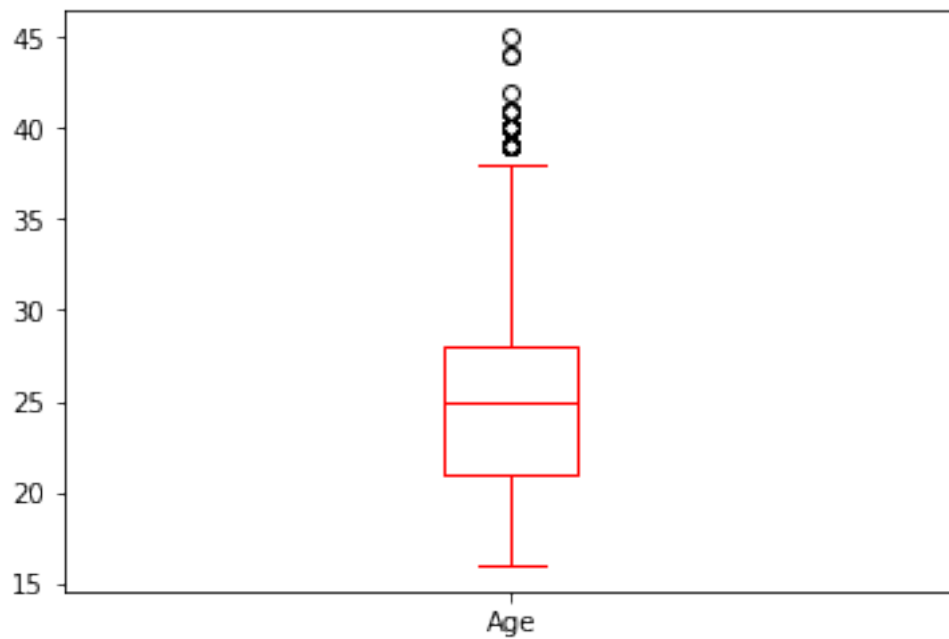
In [4]: `fifa.plot(kind = 'hist',color = 'yellow')` *#Menampilkan grafik histogram dari atribut w*

Out[4]: `<matplotlib.axes._subplots.AxesSubplot at 0x7fb6ff3d7208>`



In [5]: `fifa.plot(kind = 'box',color = 'red')` *#Menampilkan grafik box plot dari atribut umur p*

Out[5]: `<matplotlib.axes._subplots.AxesSubplot at 0x7fb6ff3bc550>`



```

In [8]: fifa.max() #Nilai maximum dari dataframe fifa

Out[8]: Name          Óscar Whalley
        Age            45
        Nationality    Zimbabwe
        dtype: object

In [9]: fifa.min() #Nilai minimum dari dataframe fifa

Out[9]: Name          A. Abang
        Age            16
        Nationality    Afghanistan
        dtype: object

In [6]: fifa.mean() #Nilai rata-rata dari dataframe fifa

Out[6]: Age      25.122206
        dtype: float64

In [11]: fifa.mode(numeric_only = True) #Modus yang diperoleh dari dataframe fifa

Out[11]:    Age
         0    21

In [19]: fifa.median() #Nilai Median dari dataframe fifa

Out[19]: Age      25.0
        dtype: float64

In [20]: fifa.var() #Nilai variansi dari dataframe fifa

Out[20]: Age      21.808365
        dtype: float64

In [21]: fifa.std() #Nilai standard deviasi dari dataframe fifa

Out[21]: Age      4.669943
        dtype: float64

In [23]: fifa.skew() #Mendapatkan nilai skew (karena skew >0 maka skew right)

Out[23]: Age      0.391764
        dtype: float64

In [24]: fifa.kurtosis() # Mendapatkan nilai kurtosis

Out[24]: Age     -0.459514
        dtype: float64

```

#Fungsi distribusi peluang yang kami pilih pada tugas kali ini adalah distribusi normal. Salah satu pertimbangan yang kami lihat adalah nilai skew = 0.39 yang mendekati nilai 0. Serta gambar histogram yang menyerupai grafik normal

#Soal 1, Ekspektasi untuk 1000 pemain yang berumur kurang dari 22 tahun, Menggunakan distribusi normal $P(X < 22)$

```
In [63]: mu = fifa.mean()
```

```
In [64]: sigma = fifa.std()
```

```
In [53]: x = 22
```

```
In [54]: z = (x-mu)/sigma
          z #Hasil transformasi
```

```
Out[54]: Age    -0.668575
          dtype: float64
```

```
In [55]: #P(X < 22) = P(Z < -0.66) = 0.2546 (Dengan tabel normal)
          p = 0.2546
          ekspektasi = p*1000
          ekspektasi
```

```
Out[55]: 254.6
```

#Jadi ekspektasi umur pemain bola yang kurang dari 22 tahun untuk 1000 pemain bola baru adalah sekitar 254 orang

#Soal 2, Ekspektasi untuk 1000 pemain yang berumur lebih dari 40 tahun, Menggunakan distribusi normal $P(X > 40)$

```
In [58]: x = 40
```

```
In [59]: z = (x-mu)/sigma
          z #Hasil transformasi
```

```
Out[59]: Age     3.185862
          dtype: float64
```

```
In [65]: #P(X > 40) = P(Z > 3.18) = 1-0.9993 (Dengan tabel normal)
          p = 1-0.9993
          ekspektasi = p*1000
          ekspektasi
```

```
Out[65]: 0.70000000000000339
```

#Jadi ekspektasi umur pemain bola yang lebih dari 40 tahun untuk 1000 pemain bola baru adalah 0 orang

===== #bagian
bawah hanya test

```
In [ ]: from scipy import stats
```

```
In [31]: from scipy.stats import norm  
norm.pdf(10,2,4)
```

```
Out[31]: 0.013497741628297016
```

```
In [17]: fifa['Age'].unique() #Jangan ljupa dihapus
```

```
Out[17]: array([31, 33, 26, 27, 32, 25, 29, 28, 24, 30, 19, 40, 22, 23, 34, 35, 36,  
               37, 21, 18, 20, 39, 41, 17, 38, 45, 42, 16, 44])
```

```
In [18]: pd.value_counts(fifa['Age'].values, sort=False) #Dihapus nanti
```

```
Out[18]: 16      42  
        32     574  
        17     289  
        33     408  
        18     732  
        34     404  
        19    1024  
        35     196  
        20    1240  
        36     127  
        21    1423  
        37      82  
        22    1340  
        38      37  
        23    1332  
        39      25  
        24    1358  
        40      13  
        25    1319  
        41       5  
        26    1387  
        42       1  
        27    1162  
        28    1101  
        44       2  
        29     959  
        45       1  
        30     917  
        31     707  
dtype: int64
```