soal5_athlete

April 9, 2019

In [2]:
```
import pandas as pd
athlete = pd.read_csv('athletes.csv')
athlete
```

Out[2]:

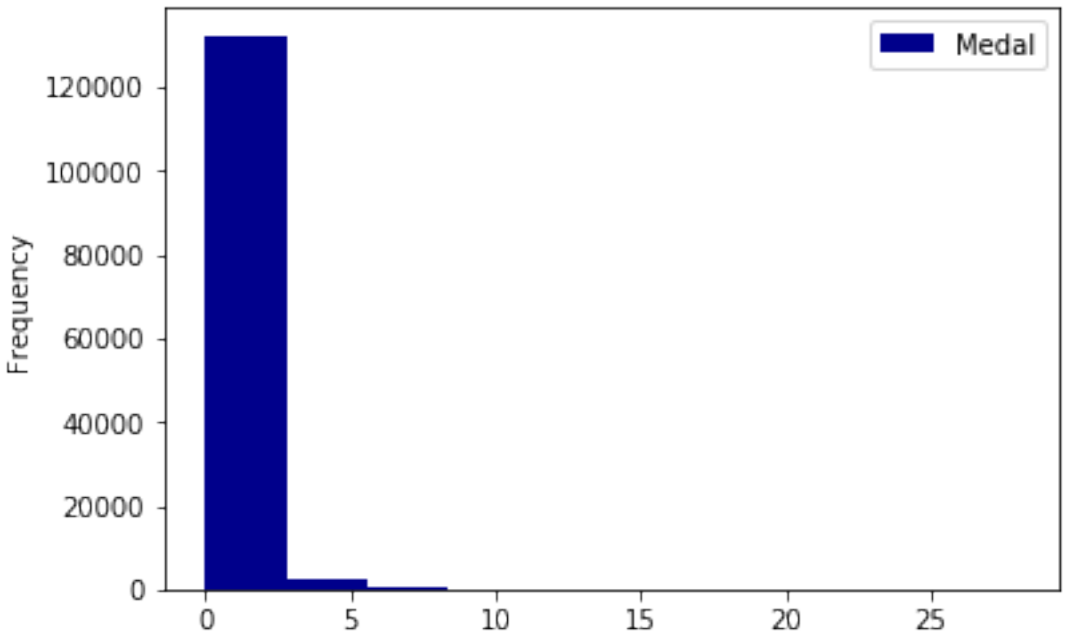
	Name	Medal
0	A Dijiang	0
1	A Lamusi	0
2	A. Aanantha Sambu Mayavo	0
3	A. Abdul Razzak	0
4	A. Brun	0
5	A. Buydens	0
6	A. Charles Six	0
7	A. Christory	0
8	A. Darnis	0
9	A. Duponcheel	0
10	A. E. Page	0
11	A. G. Chagale	0
12	A. Germaine Golding (Regnier-)	0
13	A. Godinat	0
14	A. Hurtado Vargas	0
15	A. Hussain Ahmed	0
16	A. J. J. Fridt	0
17	A. J. Tyronne Benildus "Benny" Fernando	0
18	A. Kordonis	0
19	A. Laffen	0
20	A. Lafontaine	0
21	A. Lambrecht, Jr.	0
22	A. O. Pinner	0
23	A. Papadakis	0
24	A. Porcher	0
25	A. Priftis	0
26	A. Roger	0
27	A. S. Harley	0
28	A. Sattar Basheer	0
29	A. Schmitt	0
...	...	...
134701	Isabelle Regina Werth	10
134702	Marit Bjrgen	10

```
134703                      Polina Hryhorivna Astakhova     10
134704                         Raisa Petrovna Smetanina     10
134705                      Raymond Clarence "Ray" Ewry     10
134706                                Stefania Belmondo     10
134707                   Vitaly Venediktovich Shcherbo     10
134708                                       Yang Yang     10
134709                       gnes Keleti-Srkny (Klein)     10
134710                   Aleksandr Vladimirovich Popov     11
134711                            Carl Townsend Osburn     11
134712                               Mark Andrew Spitz     11
134713                  Matthew Nicholas "Matt" Biondi     11
134714                       Viktor Ivanovych Chukarin     11
134715                          Vra slavsk (-Odloilov)     11
134716                        Aleksey Yuryevich Nemov     12
134717                         Birgit Fischer-Schmidt     12
134718            Dara Grace Torres (-Hoffman, -Minas)     12
134719  Jennifer Elisabeth "Jenny" Thompson (-Cumpelik)     12
134720                  Natalie Anne Coughlin (-Hall)     12
134721                            Paavo Johannes Nurmi     12
134722                          Ryan Steven Lochte     12
134723                                  Sawao Kato     12
134724                    Borys Anfiyanovych Shakhlin     13
134725                           Edoardo Mangiarotti     13
134726                           Ole Einar Bjrndalen     13
134727                                  Takashi Ono     13
134728                  Nikolay Yefimovich Andrianov     15
134729              Larysa Semenivna Latynina (Diriy-)     18
134730                         Michael Fred Phelps, II     28

[134731 rows x 2 columns]
```

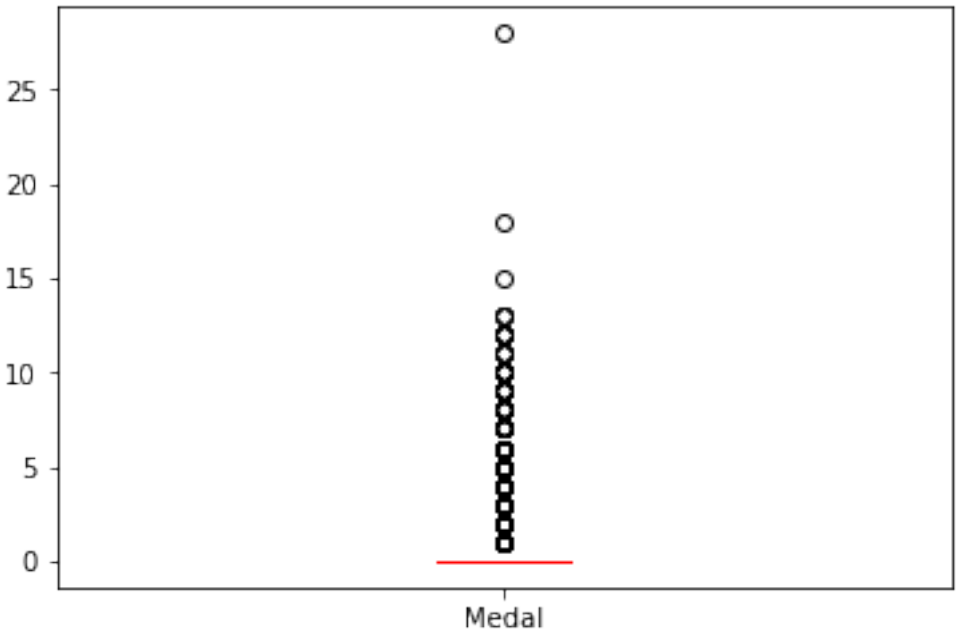
In [6]:
```
athlete.plot(kind = 'hist',color = 'darkblue') #Menampilkan grafik histogram dari atri
```

Out[6]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f4c82e542e8>
```




```
In [7]: athlete.plot(kind = 'box', color = 'red') #Menampilkan grafik box plot dari atribut me
```

```
Out[7]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4c82d56748>
```



```
In [8]: athlete.max() #Menampilkan medal terbanyak

Out[8]: Name     zzet nce
        Medal          28
        dtype: object

In [9]: athlete.min() #Menampilkan medal tersedikit

Out[9]: Name     A Dijiang
        Medal            0
        dtype: object

In [16]: athlete.describe() #Mendapatkan nilai mean dari dataframe athlete

Out[16]:                Medal
         count  134731.000000
         mean        0.295277
         std         0.724716
         min         0.000000
         25%         0.000000
         50%         0.000000
         75%         0.000000
         max        28.000000

In [17]: athlete.median() #mendapatkan nilai tengah/median dari dataframe athlete

Out[17]: Medal    0.0
         dtype: float64

In [19]: athlete.mode(numeric_only = True) #Mendapatkan nilai modus dari dataframe athlete

Out[19]:    Medal
         0      0

In [23]: athlete.std() #mendapatkan standard deviasi dari dataframe athlete

Out[23]: Medal    0.724716
         dtype: float64

In [20]: athlete.var() #mendapatkan variansi dari dataframe athlete

Out[20]: Medal    0.525213
         dtype: float64

In [21]: athlete.skew() #mendapatkan nilai skew dari dataframe athlete (right skew)

Out[21]: Medal    4.819999
         dtype: float64

In [22]: athlete.kurtosis() #Mendapatkan nilai kurtosis dari dataframe athlete
```

```
Out[22]: Medal    51.562526
         dtype: float64
```

#Fungsi ditribusi peluang yang kami pilih pada tugas kali ini adalah distribusi lognormal. Dikarenakan dambar histogram yang menyerupai grafik lognormal

#Soal 1, Peluang meraih tepat 0 medali

#Soal 2, Peluang meraih lebih dari 10 medali

#Soal 3, Peluang meraih tepat 3 medali

#Soal 4, Peluang meraih 1 atau 5 medali

#Soal 5, Confidence Interval 95%

```
In [10]: pd.value_counts(athlete['Medal'].values, sort=False)
```

```
Out[10]: 0     106529
         1      21187
         2       4619
         3       1385
         4        530
         5        214
         6        115
         7         54
         8         39
         9         22
         10        16
         11         6
         12         8
         13         4
         15         1
         18         1
         28         1
         dtype: int64
```