```
import pandas as pd

from google.colab import drive
drive.mount('/content/gdrive')
```

```
Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive.mou
```

Load the data from wine dataset.

```
wine_data=pd.read_csv('/content/gdrive/MyDrive/wine.csv')
print(wine_data)
```

```
     Wine  Alcohol  Malic.acid   Ash  ...  Color.int   Hue    OD  Proline
0       1    14.23        1.71  2.43  ...       5.64  1.04  3.92     1065
1       1    13.20        1.78  2.14  ...       4.38  1.05  3.40     1050
2       1    13.16        2.36  2.67  ...       5.68  1.03  3.17     1185
3       1    14.37        1.95  2.50  ...       7.80  0.86  3.45     1480
4       1    13.24        2.59  2.87  ...       4.32  1.04  2.93      735
..    ...      ...         ...   ...  ...        ...   ...   ...      ...
173     3    13.71        5.65  2.45  ...       7.70  0.64  1.74      740
174     3    13.40        3.91  2.48  ...       7.30  0.70  1.56      750
175     3    13.27        4.28  2.26  ...      10.20  0.59  1.56      835
176     3    13.17        2.59  2.37  ...       9.30  0.60  1.62      840
177     3    14.13        4.10  2.74  ...       9.20  0.61  1.60      560

[178 rows x 14 columns]
```

Check whether all attributes are standardized or not (mean is 0 and standard deviation is 1). If not, standardize the attributes.

```
Mean_result=list()
for i in range(1,wine_data.shape[1]):
  column=wine_data.columns[i]
  Mean_result.append(wine_data[column].mean())
print(Mean_result)
print(len(Mean_result))
```

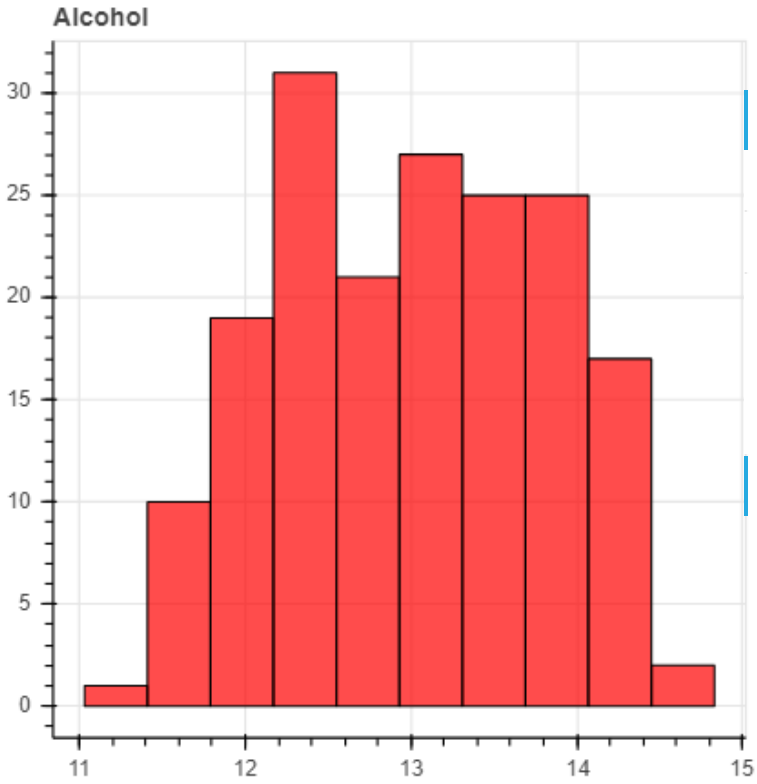
```
[13.000617977528083, 2.336348314606741, 2.3665168539325854, 19.49494382022472, 99.74157
13
```

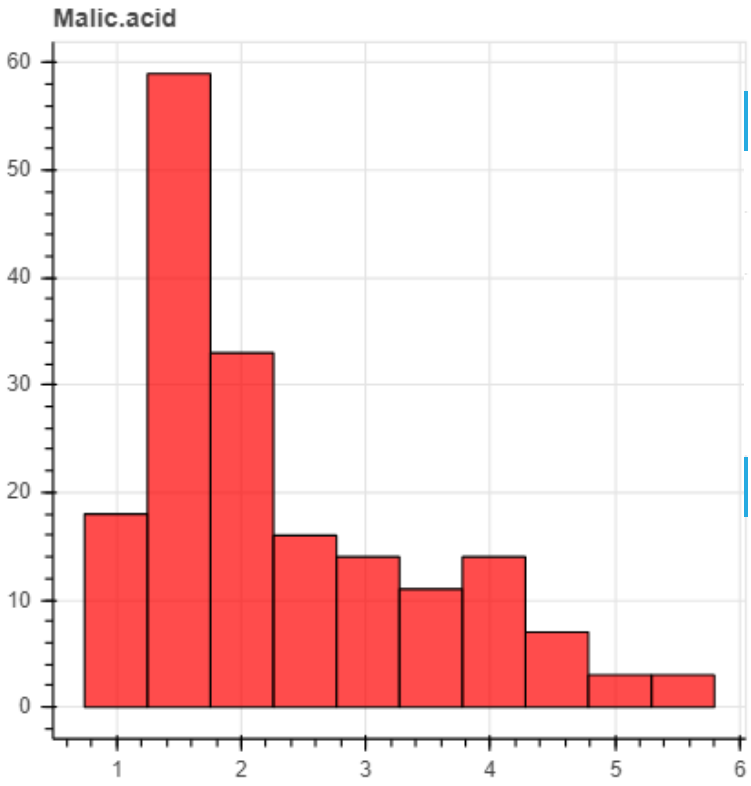
```
SD_result=list()
for i in range(1,wine_data.shape[1]):
  column=wine_data.columns[i]
  SD_result.append(wine_data[column].std())
print(SD_result)
```

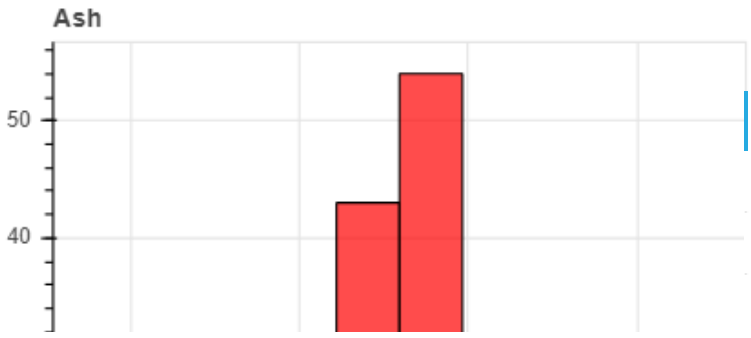
```
    [0.8118265380058577, 1.1171460976144627, 0.2743440090608148, 3.3395637671735052, 14.2824
```

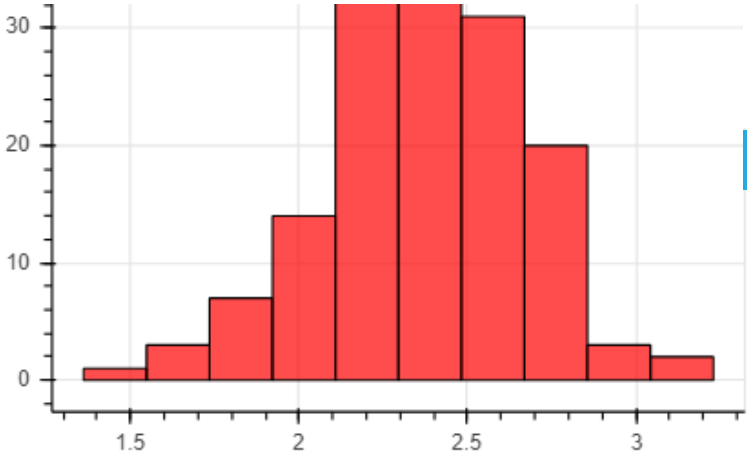
```
from matplotlib import pyplot as plt
plt.figure(figsize = [20, 100])
for i in range(1,wine_data.shape[1]):
  column=wine_data.columns[i]
  plt.subplot(13, 1, i)
  plt.title(column)
  plt.hist(wine_data[column])
```

```
import numpy as np

from bokeh.plotting import figure
from bokeh.io import output_notebook, show, output_file
from bokeh.models import ColumnDataSource, HoverTool, Panel
from bokeh.models.widgets import Tabs
from bokeh.io import output_notebook

def plot_hist(data,column):
  hist,bins = np.histogram(data[column])

  hist_info = pd.DataFrame({column: hist,
                            "left": bins[:-1],
                            "right": bins[1:]})

  hist_info['tofrom'] = [f'{left} to {right}' for left,right in zip(hist_info['left'],hist_in

  src = ColumnDataSource(hist_info) #Creating a column data source for bokeh plotting


  plot = figure(plot_height = 400, plot_width = 400,title=column)


  plot.quad(bottom = 0, top = column,left = "left", right = "right", source = src, fill_color
    line_color = "black", fill_alpha = 0.7,
    hover_fill_alpha = 1.0, hover_fill_color = 'blue')


  hover = HoverTool(tooltips = [('Interval', '@tofrom'),
                            ('Count', str("@" + column))])
  plot.add_tools(hover)

  show(plot)
  output_notebook()


for i in wine_data.columns:
  plot_hist(wine_data,i)
```

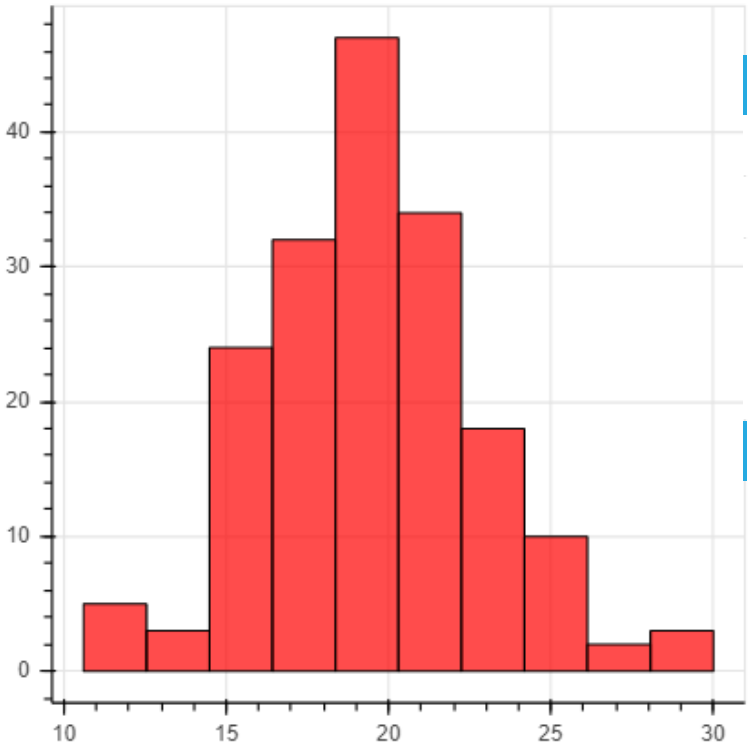

Alcohol
30
25
20
15
10
5
0
11
12
13
14
15


Malic.acid
60
50
40
30
20
10
0
1
2
3
4
5
6

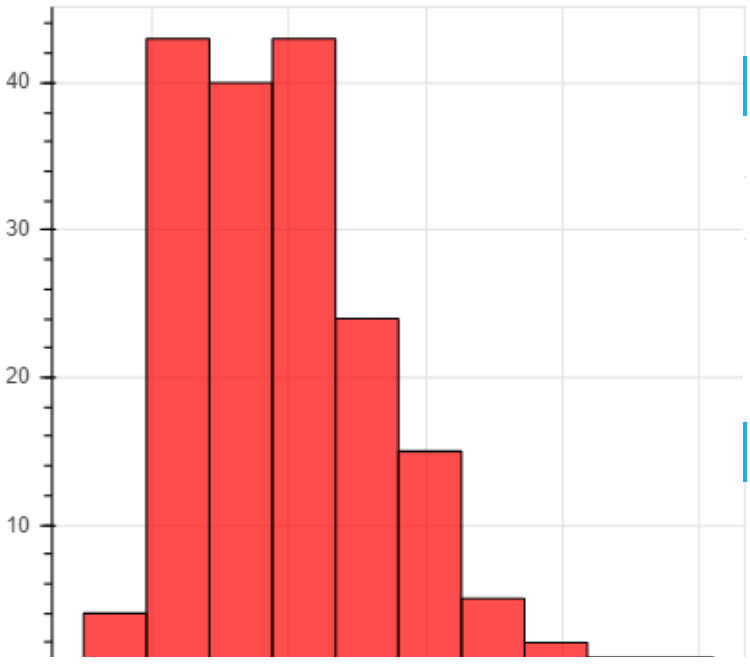

Ash
50
40

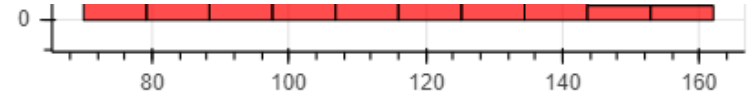


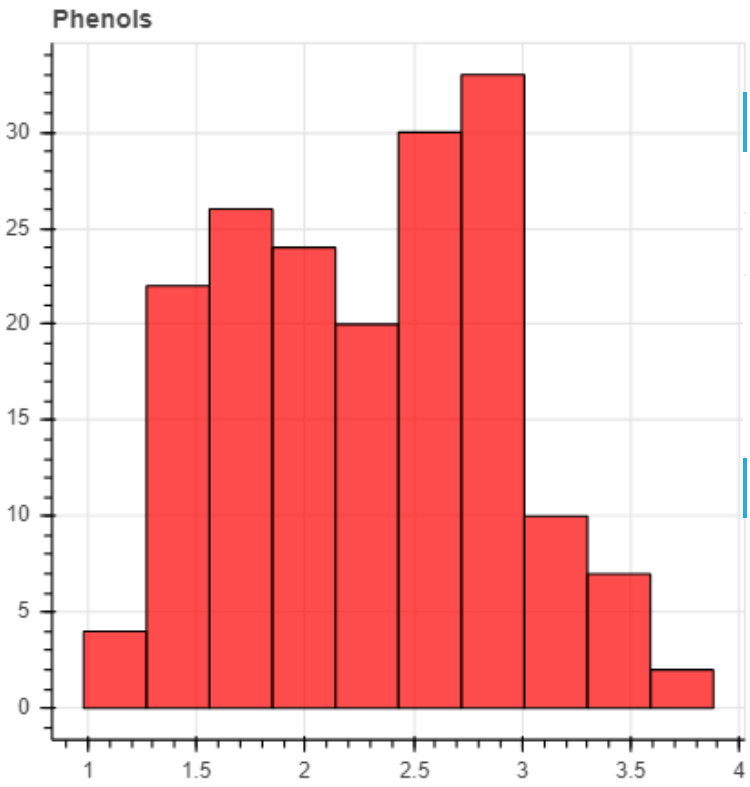
Acl

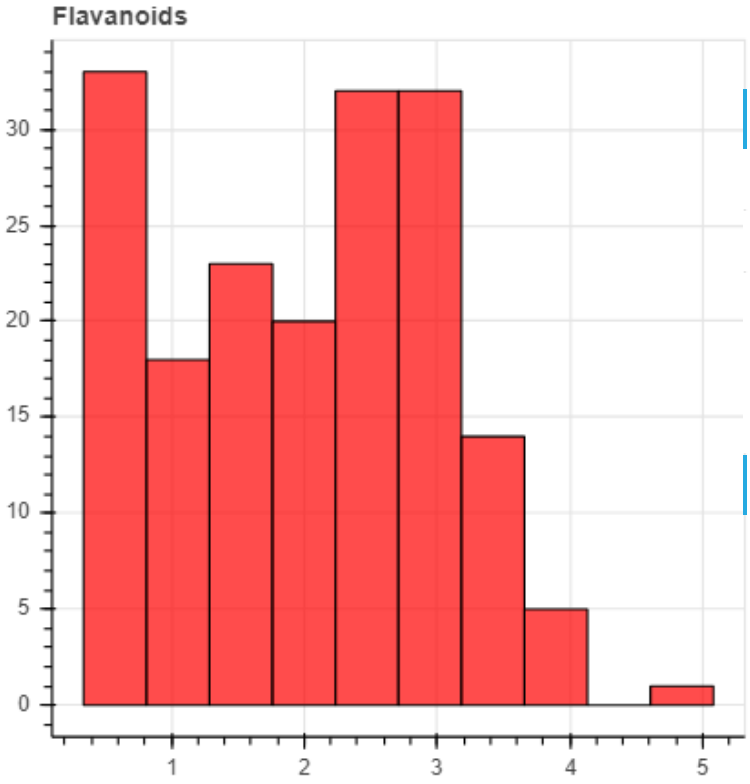


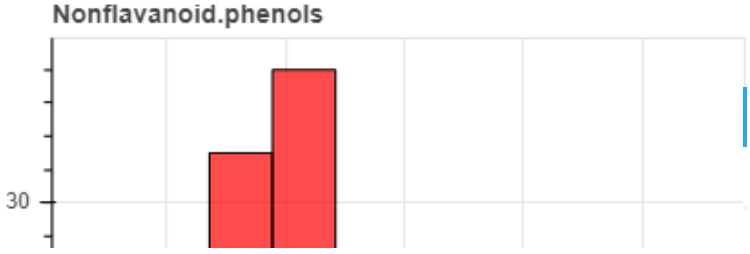
Mg

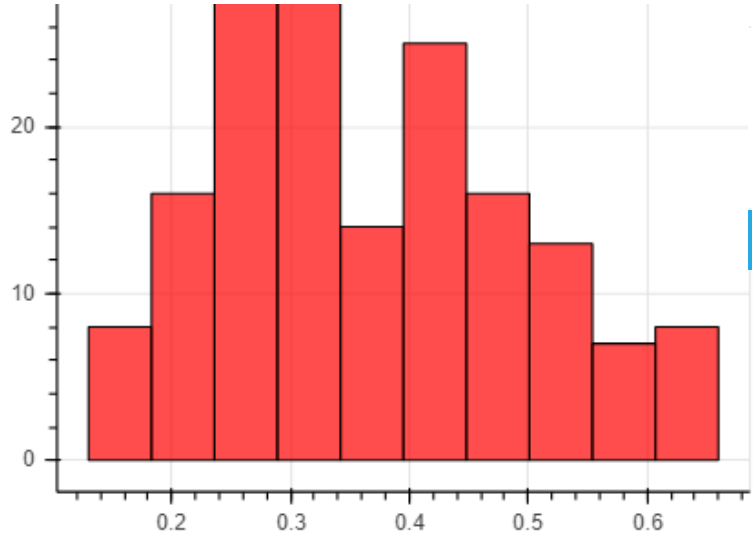



0
80
100
120
140
160

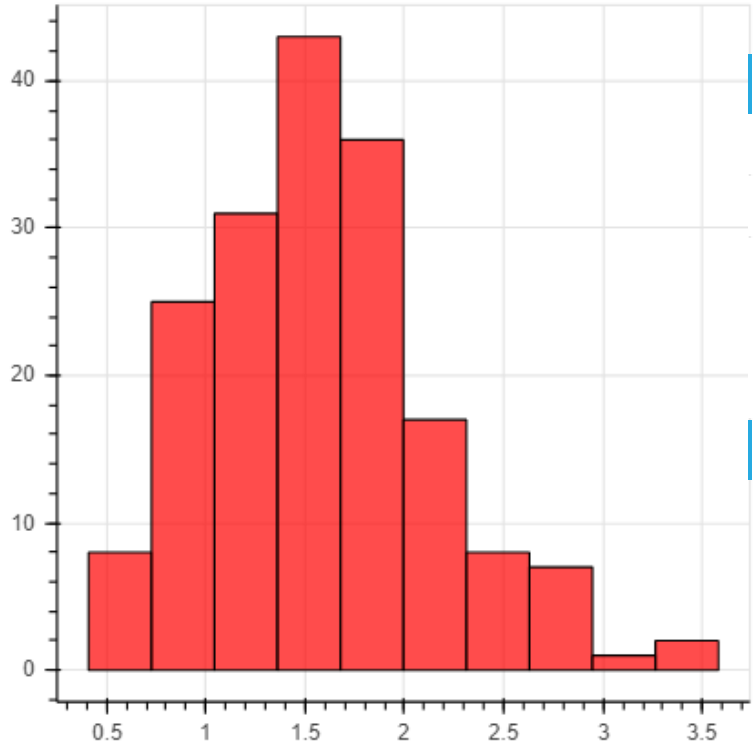

Phenols
30
25
20
15
10
5
0
1
1.5
2
2.5
3
3.5
4


Flavanoids
30
25
20
15
10
5
0
1
2
3
4
5

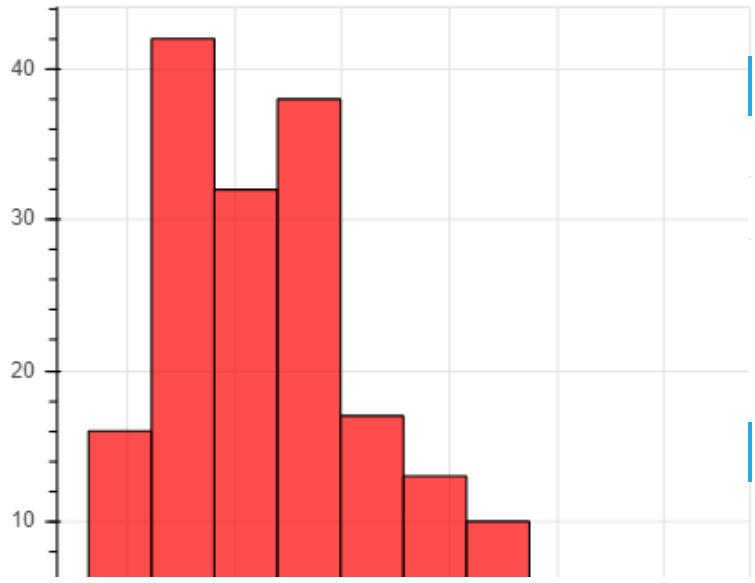

Nonflavanoid.phenols
30

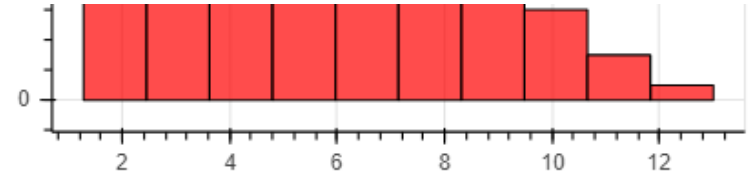


Proanth

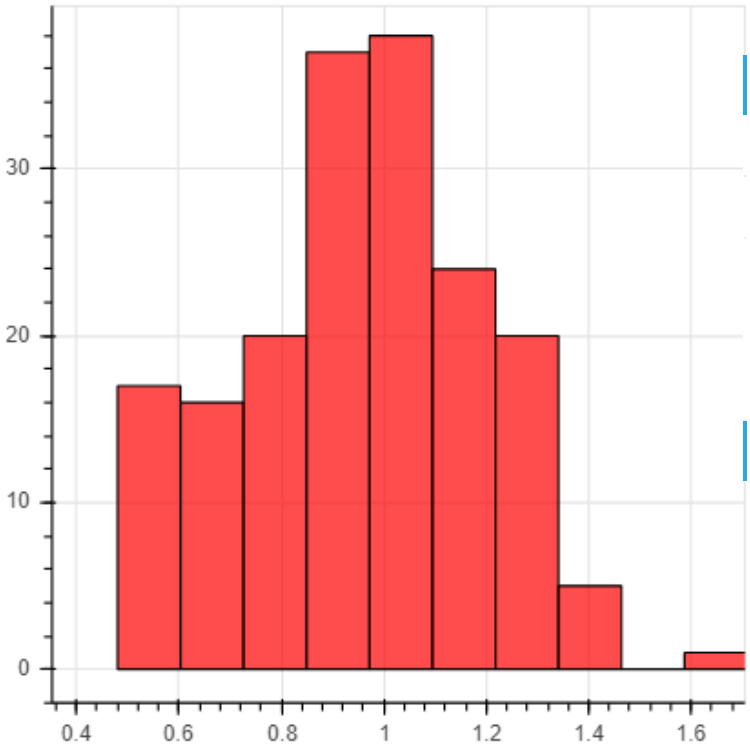


Color.int

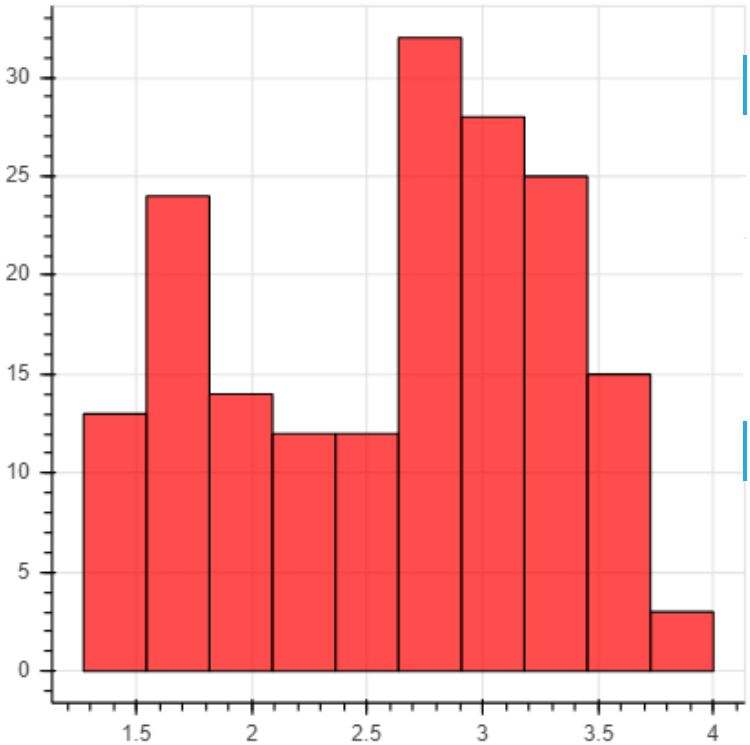




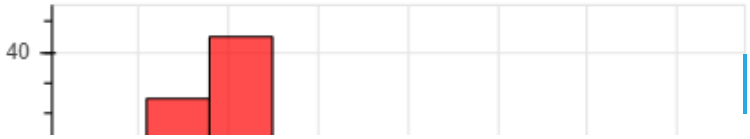
Hue

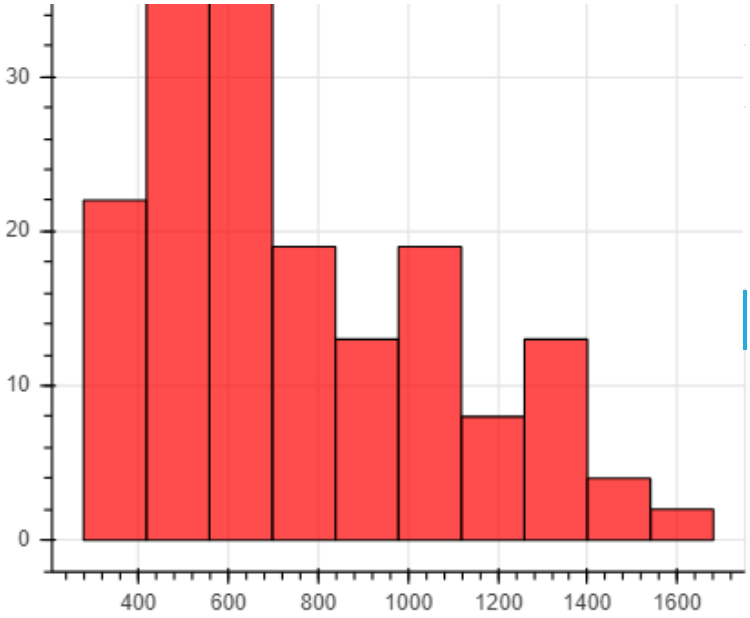


OD



Proline






```
swine_data=wine_data
from sklearn import preprocessing
#for i in range(1,wine_data.shape[1]):
#  column=wine_data.columns[i]
#  print(preprocessing.scale(wine_data[column]))
for i in range(1,swine_data.shape[1]):
  column=wine_data.columns[i]
  swine_data[column]=preprocessing.scale(wine_data[column])
swine_data
```

	Wine	Alcohol	Malic.acid	Ash	Acl	Mg	Phenols	Flavanoids	No
0	1	1.518613	-0.562250	0.232053	-1.169593	1.913905	0.808997	1.034819	
1	1	0.246290	-0.499413	-0.827996	-2.490847	0.018145	0.568648	0.733629	
2	1	0.196879	0.021231	1.109334	-0.268738	0.088358	0.808997	1.215533	
3	1	1.691550	-0.346811	0.487926	-0.809251	0.930918	2.491446	1.466525	
4	1	0.295700	0.227694	1.840403	0.451946	1.281985	0.808997	0.663351	
...	...	...	...	...	...	...	...	...	
173	3	0.876275	2.974543	0.305159	0.301803	-0.332922	-0.985614	-1.424900	
174	3	0.493343	1.412609	0.414820	1.052516	0.158572	-0.793334	-1.284344	
175	3	0.332758	1.744744	-0.389355	0.151661	1.422412	-1.129824	-1.344582	
176	3	0.209232	0.227694	0.012732	0.151661	1.422412	-1.033684	-1.354622	
177	3	1.395086	1.583165	1.365208	1.502943	-0.262708	-0.392751	-1.274305	

178 rows × 14 columns

```
SMean_result=list()
from sklearn import preprocessing
```

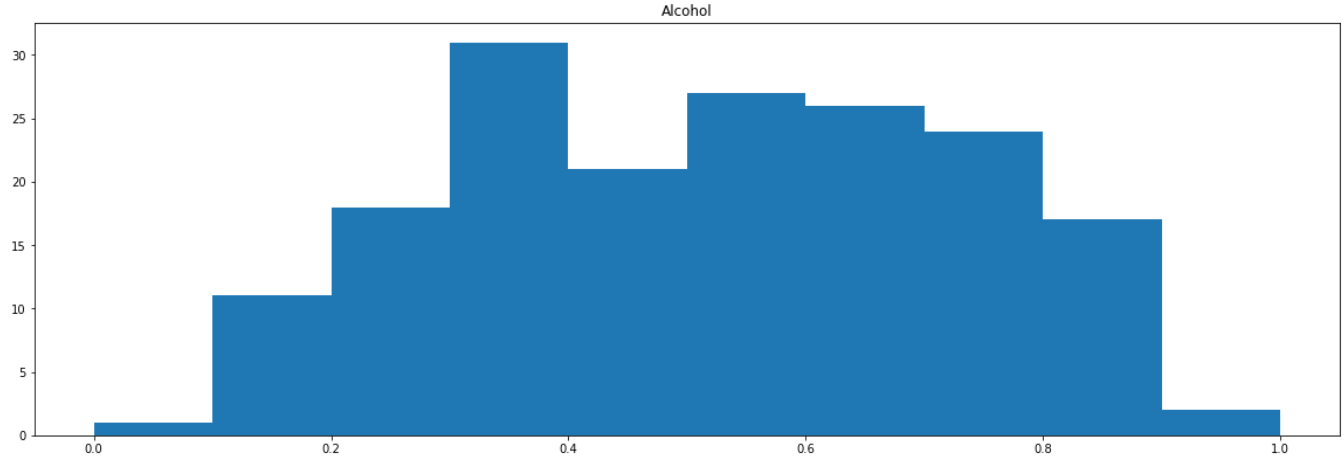
```
for i in range(1,swine_data.shape[1]):
  column=swine_data.columns[i]
  SMean_result.append(round(swine_data[column].mean()))
print(SMean_result)
```

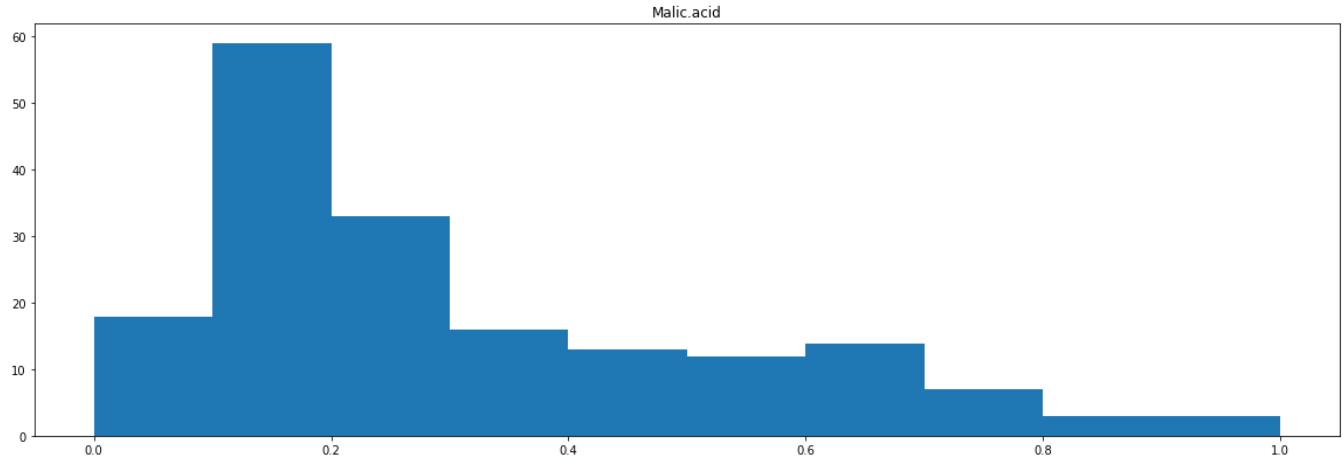
```
    [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

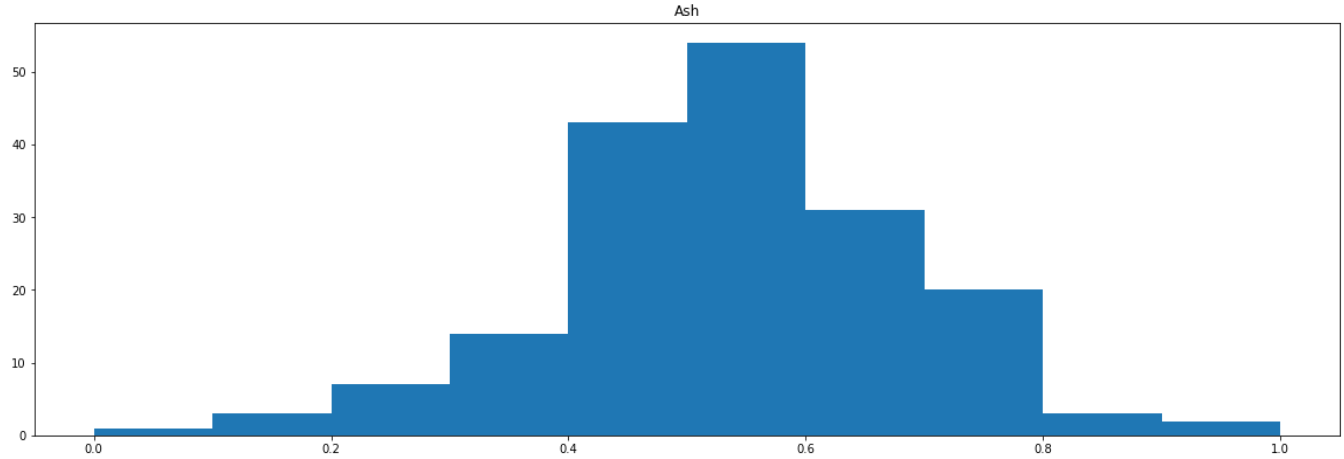
```
SSD_result=list()
from sklearn import preprocessing
for i in range(1,swine_data.shape[1]):
  column=swine_data.columns[i]
  SSD_result.append(round(swine_data[column].std()))
print(SSD_result)
```

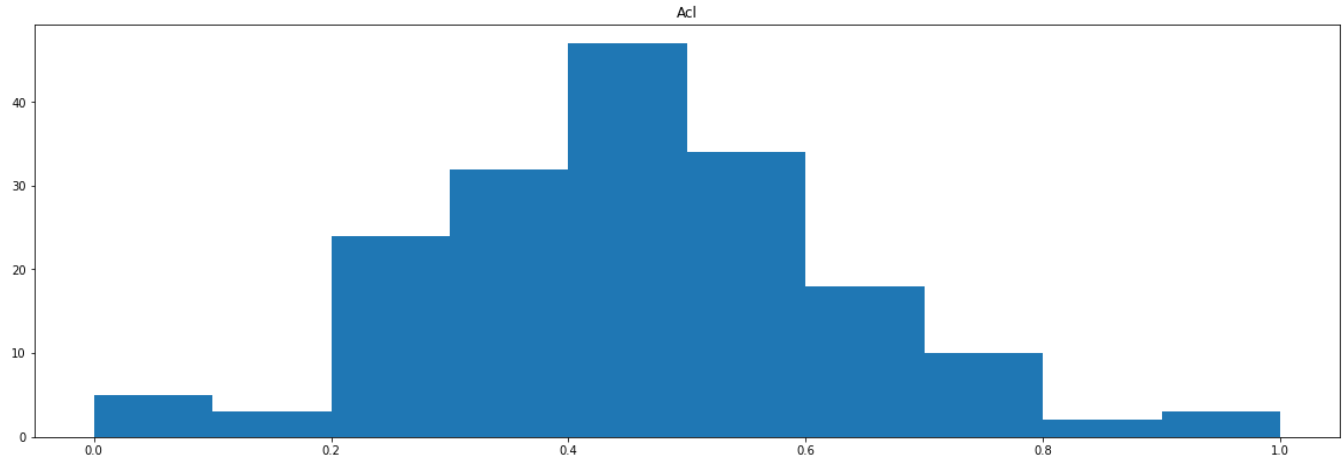
```
    [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
```

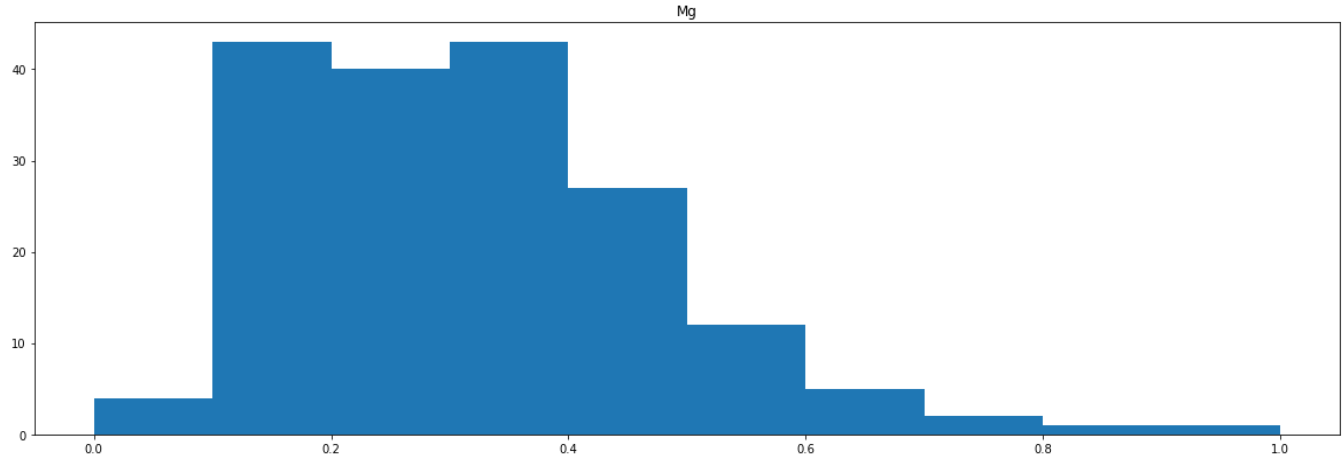
```
plt.figure(figsize = [20, 100])
for i in range(1,swine_data.shape[1]):
  column=swine_data.columns[i]
  plt.subplot(13, 1, i)
  plt.title(column)
  plt.hist(swine_data[column])
```

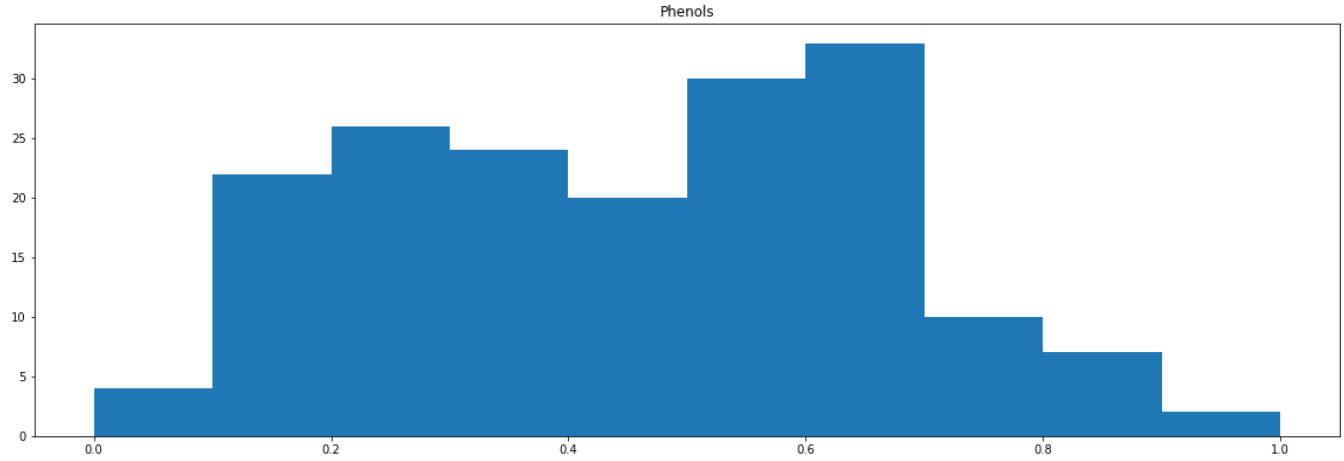

Alcohol
0
5
10
15
20
25
30
0.0
0.2
0.4
0.6
0.8
1.0

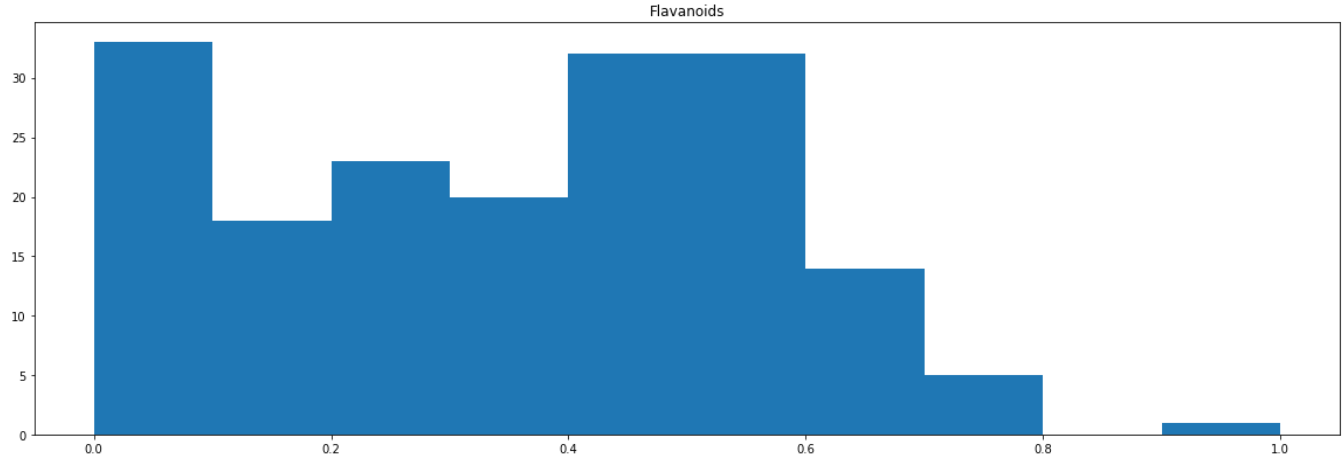

Malic.acid
0
10
20
30
40
50
60
0.0
0.2
0.4
0.6
0.8
1.0

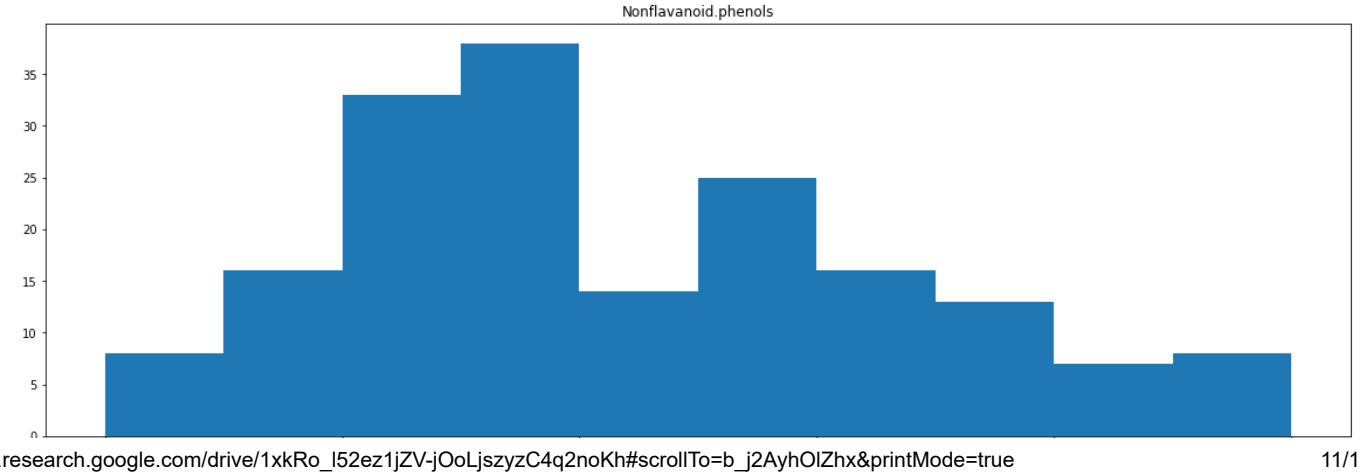

Ash
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0


Acl
0
10
20
30
40
0.0
0.2
0.4
0.6
0.8
1.0

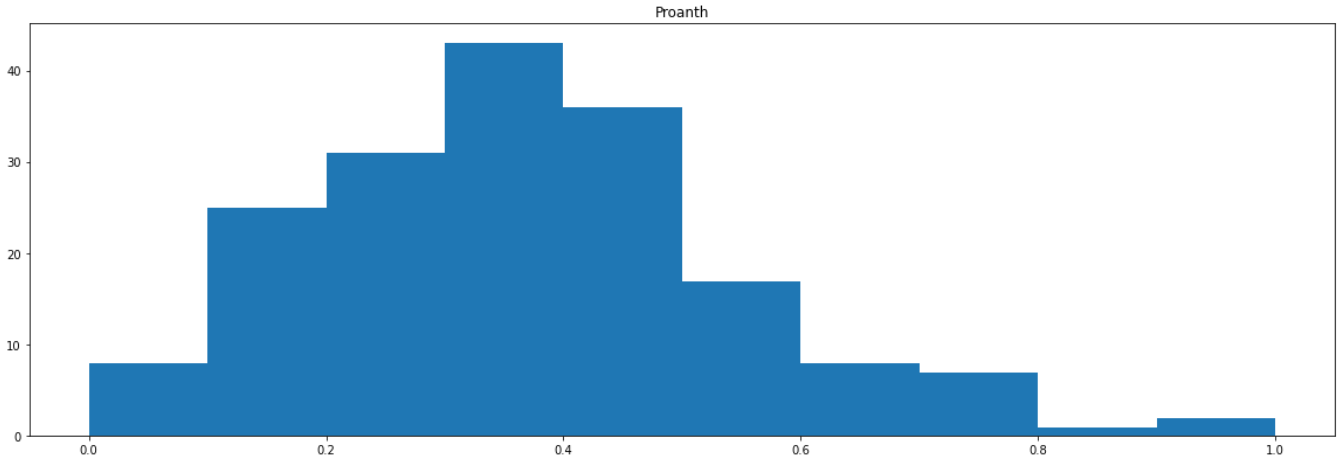

Mg
0
10
20
30
40
0.0
0.2
0.4
0.6
0.8
1.0

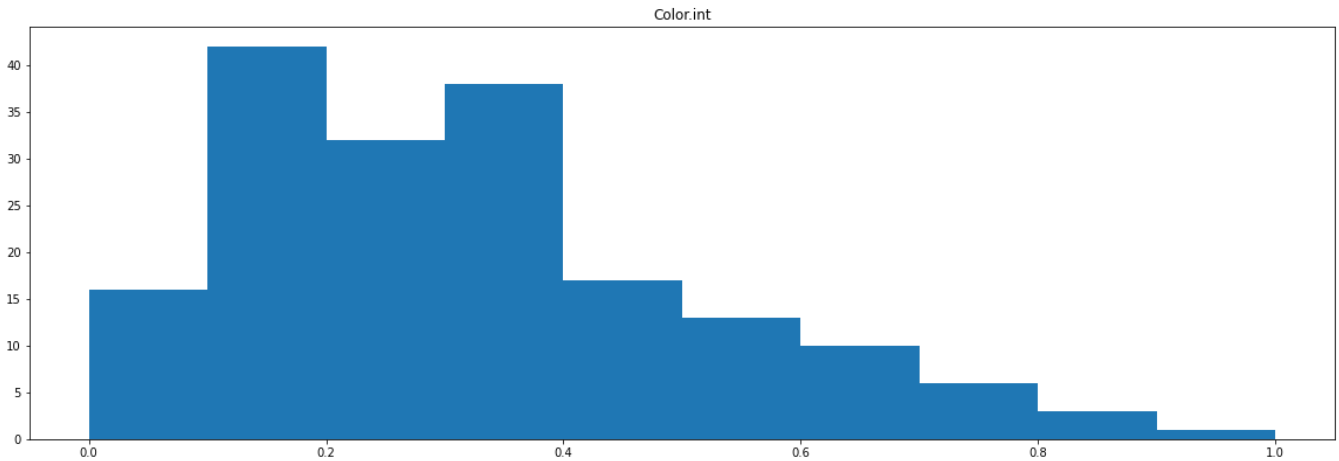

Phenols
0
5
10
15
20
25
30
0.0
0.2
0.4
0.6
0.8
1.0

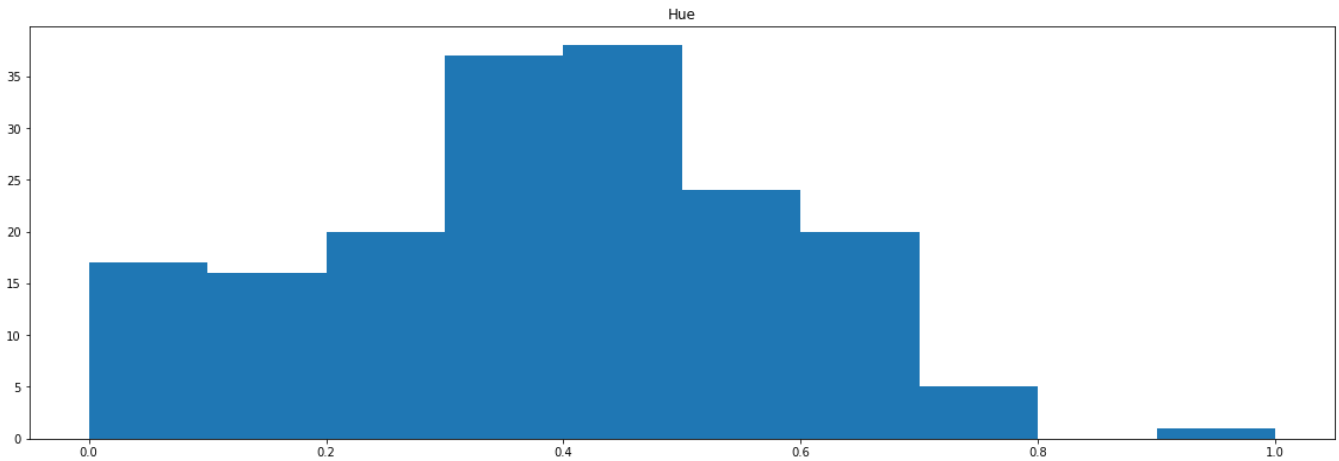

Flavanoids
0
5
10
15
20
25
30
0.0
0.2
0.4
0.6
0.8
1.0

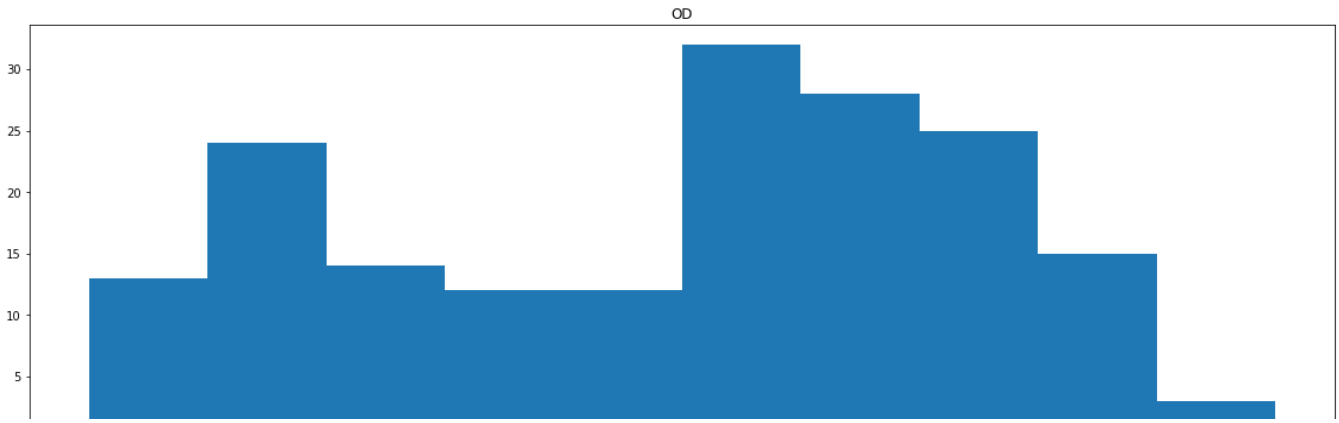

Nonflavanoid.phenols
0
5
10
15
20
25
30
35


0.0
0.2
0.4
0.6
0.8
1.0

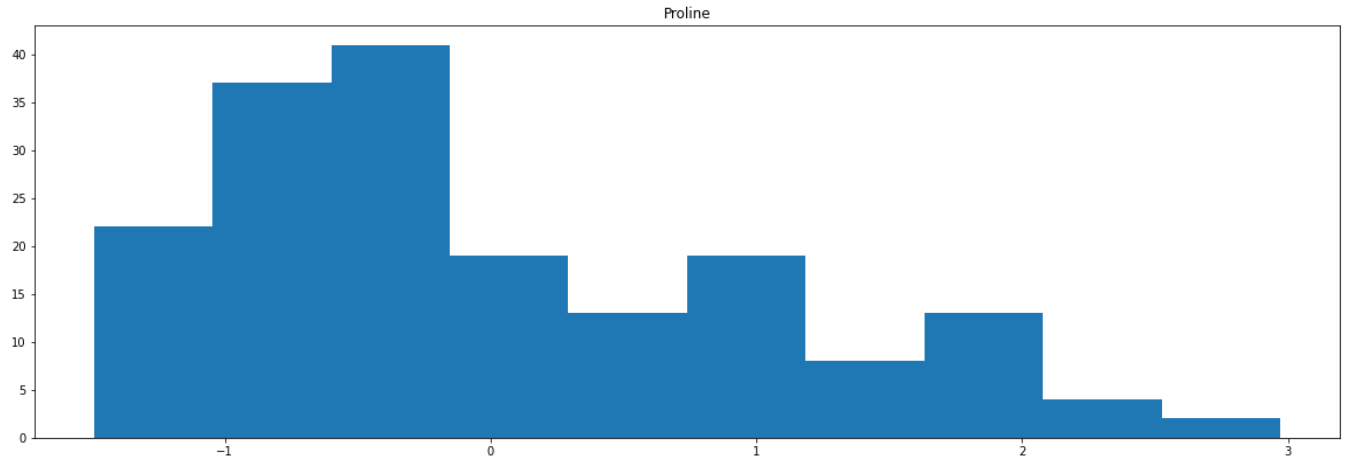

Proanth
0
10
20
30
40
0.0
0.2
0.4
0.6
0.8
1.0


Color.int
0
5
10
15
20
25
30
35
40
0.0
0.2
0.4
0.6
0.8
1.0


Hue
0
5
10
15
20
25
30
35
0.0
0.2
0.4
0.6
0.8
1.0


OD
5
10
15
20
25
30


0
0.0
0.2
0.4
0.6
0.8
1.0


Proline
40
35
30
25
20
15
10
5
0
−1
0
1
2
3

```
for i in swine_data.columns:
  plot_hist(swine_data,i)
```