We are using Loan Amount Dataset from LendingClub and Credit score from FICO

In [32]:
```
#import libraries

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from pandas import Series ,DataFrame
import seaborn as sns

%matplotlib inline
sns.set_style('whitegrid')
```

In [2]:
```
dataset = pd.read_csv('https://spark-public.s3.amazonaws.com/dataanalysis/loansDa
dataset.head()
```

Out[2]:

	Amount.Requested	Amount.Funded.By.Investors	Interest.Rate	Loan.Length	Loan.Purpose
81174	20000	20000.0	8.90%	36 months	debt_consolidatior
99592	19200	19200.0	12.12%	36 months	debt_consolidatior
80059	35000	35000.0	21.98%	60 months	debt_consolidatior
15825	10000	9975.0	9.99%	36 months	debt_consolidatior
33182	12000	12000.0	11.71%	36 months	credit_carc

In [3]:
```
dataset['Loan.Length'][0:10]
```

Out[3]:
```
81174    36 months
99592    36 months
80059    60 months
15825    36 months
33182    36 months
62403    36 months
48808    36 months
22090    60 months
76404    36 months
15867    36 months
Name: Loan.Length, dtype: object
```

In [4]:
```
dataset['Interest.Rate'][0:10]
```

Out[4]:
```
81174     8.90%
99592    12.12%
80059    21.98%
15825     9.99%
33182    11.71%
62403    15.31%
48808     7.90%
22090    17.14%
76404    14.33%
15867     6.91%
Name: Interest.Rate, dtype: object
```

In [5]:
```
dataset['FICO.Range'][0:10]
```

Out[5]:
```
81174    735-739
99592    715-719
80059    690-694
15825    695-699
33182    695-699
62403    670-674
48808    720-724
22090    705-709
76404    685-689
15867    715-719
Name: FICO.Range, dtype: object
```

In [6]:
```
#"715-719" (715,719)
```

Data Cleaning

We need to remove 'months' from Loan.Length , '%' from Interest.Rate & we need to parse the string from FICO.Range

In [7]:
```
# Now we import another file which is data cleaned

loans = pd.read_csv('C:\\Users\\bittu\\loan.csv')
```

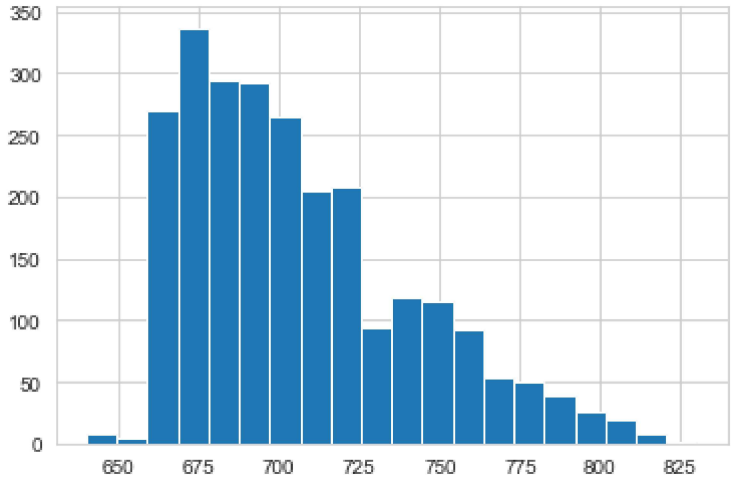
In [8]:
```
loans.head()
```

Out[8]:

	Interest.Rate	FICO.Score	Loan.Length	Monthly.Income	Loan.Amount
6	15.31	670	36	4891.67	6000
11	19.72	670	36	3575.00	2000
12	14.27	665	36	4250.00	10625
13	21.67	670	60	14166.67	28000
21	21.98	665	36	6666.67	22000

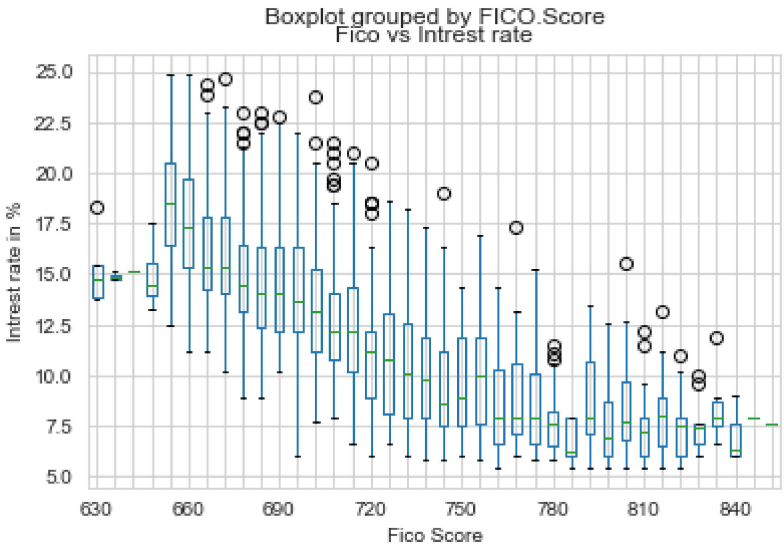
In [9]:
```
plt.figure()
loans = pd.read_csv('C:\\Users\\bittu\\loan.csv')
fico = loans['FICO.Score']
fico.hist(bins = 20)
```

Out[9]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19b1a06b9c8>
```



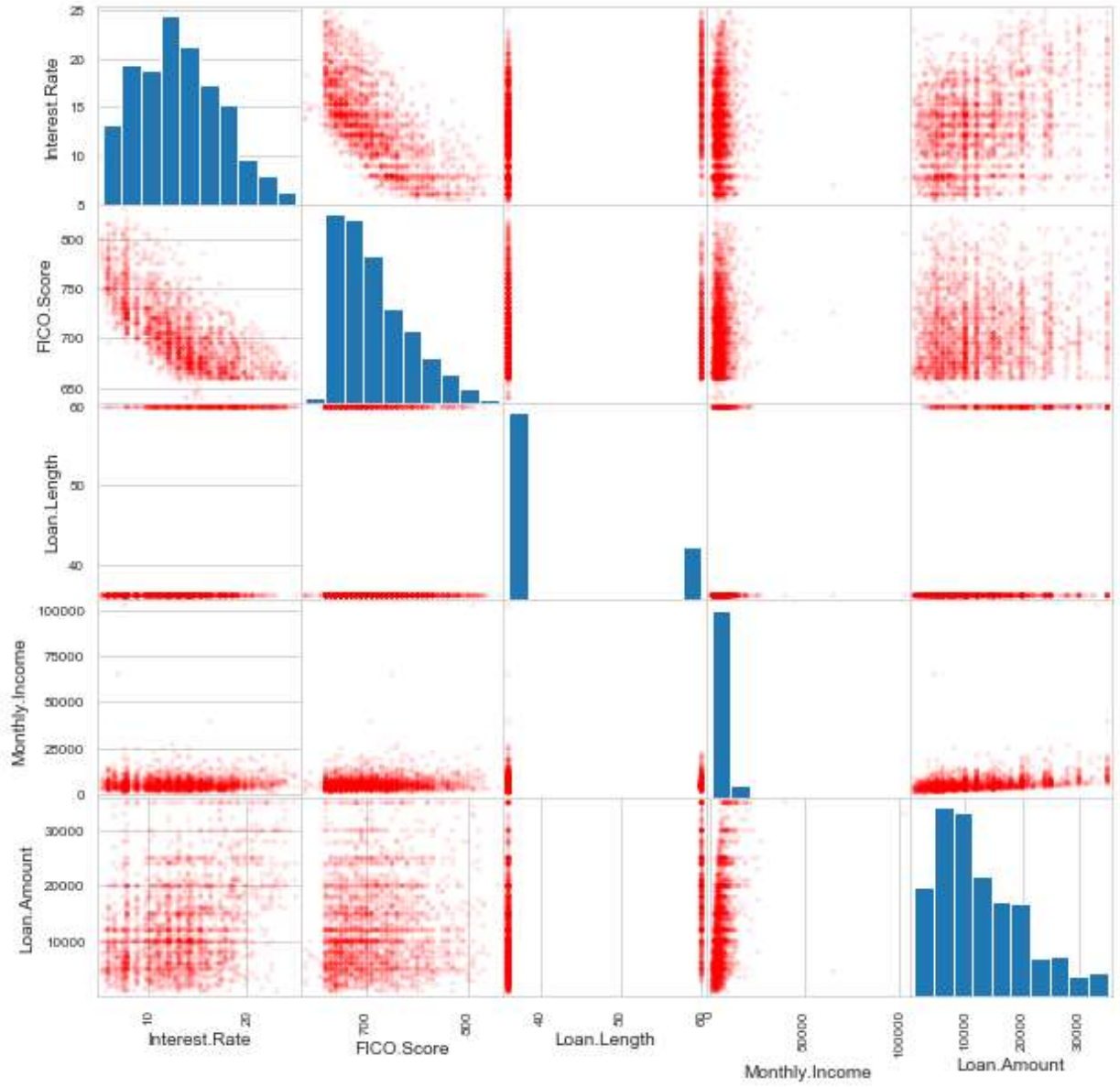
In [24]:
```
plt.figure()
loans = pd.read_csv('C:\\Users\\bittu\\loan.csv')
x = loans.boxplot('Interest.Rate','FICO.Score')
x.set_xticklabels(['630',' ' ,' ',' ',' ','660',' ' ,' ',' ',' ','690',' ',' ',' '
x.set_xlabel("Fico Score")
x.set_ylabel("Intrest rate in %")
x.set_title("Fico vs Intrest rate")
```

Out[24]:
```
Text(0.5, 1.0, 'Fico vs Intrest rate')

<Figure size 432x288 with 0 Axes>
```



In [34]:
```
#scatterplot Matrix
loans = pd.read_csv('C:\\Users\\bittu\\loan.csv')

pd.plotting.scatter_matrix(loans ,alpha = 0.1, figsize = (10,10), diagonal = 'hi
```

Out[34]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20092988>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B2009DE88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B200B0C08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FE5FE48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FE92048>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FEC7948>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FF00448>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FF3A588>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FF46188>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FF7F348>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x0000019B1FFE58C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B2001E948>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20267A48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B2029FB88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B202D8C88>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x0000019B2030FD88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20348E48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20381F48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B203BF088>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B203F81C8>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20430308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B20468388>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B204A0488>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B204D95C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x0000019B205126C8>]],
      dtype=object)
```


Interest.Rate
FICO.Score
Loan.Length
Monthly.Income
Loan.Amount