Notes:

1. Needs to convert .csv to .csv(utf-8) in case of UnicodeError
2. Data source: http://www.stateair.net/web/historical/1/1.html
3. Air Quality Index Scale and Color Legend: http://aqicn.org/scale/cn/

Improvements:

1. Bokeh to generate dynamic graphs, even on the same graph
2. More types of graphs to show
3. Bokeh server connection to display on the Internet
4. Display multiple years on one chart using Widgets

Exploratory Topics:

1. Whats the extreme days (PM2.5>300/200/150/100) percentage/distribution for each year?
2. Whats the good days (PM2.5<50) percentage/distribution for each year?

[81]
```
import numpy as np
import pandas as pd
from bokeh.charts import Histogram
from bokeh.plotting import show, output_notebook

#2008-2017
result = {}
for m in range(0,10):
    d = pd.read_csv("./Beijing_"+str(int(2008)+m)+".csv")
    s = d['column9'].loc[4:].astype(str).astype(float)
    #p = Histogram(s, title='Beijing Air Pollution Histogram')
    #output_notebook()
    #show(p)

    count_extreme = 0
    count_good = 0

    for i in range(0, len(s)):
        individualvalue = s.iloc[i]
        if individualvalue > 200:
            count_extreme += 1
        elif individualvalue < 50:
            count_good += 1

    percent_extreme = count_extreme/len(s)
    percent_good = count_good/len(s)
    #print ("The average PM2.5 value in",str(int(2008)+i),"is",round(np.
    mean = round(np.mean(s))

    ratio = pd.Series(["{:.2%}".format(percent_extreme), "{:.2%}".format
```

```
            result[int(2008)+m] = ratio
```

[75]
```
from bokeh.io import output_notebook, show
from bokeh.layouts import widgetbox
from bokeh.models.widgets import Select

# create some widgets
select = Select(title="Option:", value="2008", options=["foo", "bar", "ba

# put the results in a row
show(widgetbox(select, width=300))
```

[82]
```
df = pd.DataFrame(result)
df
```

	2008	2009	2010	2011	2012	2013
percent_extreme	5.37%	11.27%	13.88%	13.94%	9.58%	13.30%
percent_good	35.99%	30.33%	34.27%	39.54%	39.76%	36.66%
mean	85	102	104	99	91	102

[83]
```
# Create a Pandas Excel writer using XlsxWriter as the engine.
writer = pd.ExcelWriter('AirPollutionResult_Beijing.xlsx', engine='xlsxwr

# Convert the dataframe to an XlsxWriter Excel object.
df.to_excel(writer, sheet_name='Status Ratio')

# Close the Pandas Excel writer and output the Excel file.
writer.save()
```