



M2-06: Basic Data Visualization in Python

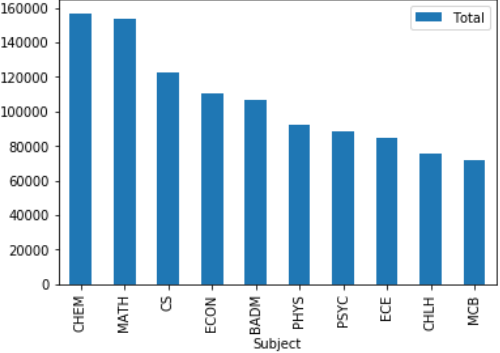
Part of the "Exploratory Data Analysis" Learning Badge

Video Walkthrough: <https://discovery.cs.illinois.edu/m2-06/>

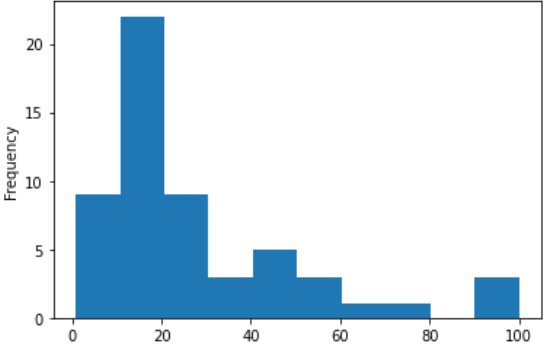
Basic Data Visualization in Python

Using the GPA dataset, both a histogram and a bar chart can be used to describe the data.

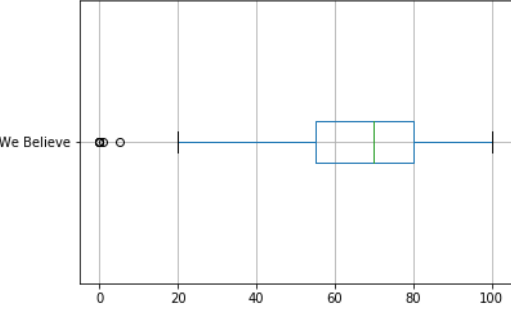
Creating a Bar Graph

Python:	General Syntax: <code>df.plot.bar()</code>	
Example Shown:	<pre># df contains "GPA dataset" g = df.groupby("Subject") df_subject = g.agg("sum") df_subject["Total"] = df_subject["A+"] + df_subject["A"] + ... df_subject = df_subject[["Total"]] df_subject = df_subject.nlargest(10, "Total") df_subject.plot.bar()</pre>	

Creating a Histogram

Python:	General Syntax: <code>df.plot.hist()</code> <code>df.hist()</code> Options: <code>bins = 15</code> (# of bins) <code>density = True</code> (y-axis %s)	
Example Shown:	<pre># df contains "Party dataset" g = df["Many"].hist()</pre>	

Creating a Boxplot

Python:	General Syntax: <code>df['Column Name'].plot.box()</code> Options: <code>grid = False</code> # gridlines <code>vert = False</code> # create horizontal	
Example Shown:	<pre>df = pd.read_csv(...) df["We Believe"].plot.box(vert = False)</pre>	