MIS500, Option #2 – Data Set Validation and Graphing

Julie Passey

Colorado State University – Global Campus

Module 1, Option #2 Assignment: Installing Python and Running a Data Set Visualization and Graphing in PyCharm Edu Program

In order to complete this assignment, PyCharm and all necessary packages were installed and launched via the Anaconda workspace. The script below was compiled and executed without error. Resultant plots, graphs and screenshots follow and are appropriately labeled. This completed document has been uploaded to <https://github.com/mycomputeris2slow/CSU_Global_Classwork> and is available for public access.

*# BEGIN  
# ##########################################################################  
# Python Statistical Data Visualization - Plotting data for presentation  
###########################################################################  
# Pandas use for data structures and data analysis  
# Import the necessary libraries***import** matplotlib.pyplot **as** plt  
**import** pandas **as** pd  
**import** seaborn **as** sns  
  
*# Create Data\_Frame from array. This could also be a data file loaded  
# Three ways to load files  
# csv read,*df = pd.DataFrame({  
 **'name'**: [**'john'**, **'mary'**, **'peter'**, **'jeff'**, **'bill'**, **'lisa'**, **'jose'**],  
 **'age'**: [23, 78, 22, 19, 45, 33, 20],  
 **'gender'**: [**'M'**, **'F'**, **'M'**, **'F'**, **'M'**, **'F'**, **'M'**],  
 **'state'**: [**'CA'**, **'DC'**, **'CA'**, **'DC'**, **'VA'**, **'NY'**, **'NY'**],  
 **'num\_children'**: [2, 0, 0, 3, 2, 1, 4],  
 **'num\_pets'**: [5, 1, 0, 5, 2, 2, 3]  
})  
*# 1 Generate a scatter plot comparing num\_children and num\_pets*df.plot(kind=**'scatter'**, x=**'num\_children'**, y=**'num\_pets'**, color=**'red'**)  
plt.show()  
  
*# 2 Generate a simple line plot*df.plot(kind=**'bar'**, x=**'name'**, y=**'age'**)  
plt.show()  
  
*# 3 Generate line plot with multiple columns*ax = plt.gca()  
df.plot(kind=**'line'**, x=**'name'**, y=**'num\_children'**, ax=ax)  
df.plot(kind=**'line'**, x=**'name'**, y=**'num\_pets'**, color=**'red'**, ax=ax)  
plt.show()  
  
*# 4 Generate stacked bar plot with two-level group by*df.groupby([**'state'**, **'gender'**])[**'name'**].size().unstack().plot(kind=**'bar'**, stacked=**True**)  
df.sample(n=3)  
plt.show()  
  
*# 5 Generate a plot with count of people by gender, splitting by state:*df.groupby([**'gender'**, **'state'**])[**'age'**].size().unstack().plot(kind=**'bar'**, stacked=**True**)  
plt.show()  
  
*# 6 Generate a violinplot*fig, ax = plt.subplots()  
ax.violinplot(df[**"age"**], vert=**False**)  
plt.show()  
  
*# 7 Generate a plot of the distribution of faculty children*num\_bins = 10  
plt.hist(df[**'num\_children'**], num\_bins, normed=1, facecolor=**'blue'**, alpha=0.5)  
plt.show()  
  
*# 8 Use Seaborn library to construct a pet plot*sns.set()  
*# Set context to "paper"*sns.set\_context(**"paper"**)  
*# Construct pets plot*sns.swarmplot(x=**"num\_pets"**, y=**"age"**, data=df)  
*# Show plot*plt.show()  
  
*# Save last plot to a file with Permanent link  
# the plot gets saved to 'graphoutput.png' image file*plt.savefig(**'graphoutput.png'**)  
*# END*

A picture containing screenshot

Description automatically generated

*Figure 1.* Scatter plot comparing number of children and number of pets.

A screenshot of a cell phone

Description automatically generated

*Figure 2.* Plot with multiple columns showing names of persons and age.

A screenshot of a social media post

Description automatically generated

*Figure 3.* Simple line graph grouped by two-levels: names and number of children; names and number of pets.

A screenshot of a cell phone

Description automatically generated

*Figure 4.* Stacked bar plot counting people by gender and splitting by state.

A picture containing screenshot

Description automatically generated

*Figure 5.* Stacked bar plot grouped by two-levels: names and number of children; names and number of pets.

A picture containing screenshot

Description automatically generated

*Figure 6.* Violin plot showing distribution of age.

A screenshot of a computer

Description automatically generated

*Figure 7.* Plot showing the distribution of faculty of children.

A screenshot of a computer

Description automatically generated

*Figure 8.* Simple line plot comparing number of pets to age.

A screenshot of a computer screen

Description automatically generated

*Figure 9:* Screenshot showing last plot saved as ‘graphoutput.png’.

A screenshot of a social media post

Description automatically generated

*Figure 10:* Screenshot showing the final output of the program: *“Process finished with exit   
code 0”.*