William Alec Akin

IN300 – Programming for Data Analysis (Python, R, Java)

Unit 8 Assignment 1

Dr. Robert Kayl

June 26th, 2022

# #1 – Java Source Code (JavaFX/JFreeChart Bar Chart)

/\*

William Alec Akin

IN300 Programming for Data Analysis

Unit 8 Assignment 1 - Java Source Code

June 26th, 2022

Description: Write a Java program using JavaFX that reads the dataset "Average\_Daily\_Traffic\_Counts.csv" file plot a bar chart.

 \*/

//Import necessary libraries

import org.jfree.chart.\*;

import org.jfree.chart.plot.PlotOrientation;

import org.jfree.data.category.CategoryDataset;

import org.jfree.data.category.DefaultCategoryDataset;

import javax.swing.\*;

import java.awt.\*;

import java.io.\*;

import java.util.\*;

public class IN300\_WilliamAlecAkin\_Unit8\_Q1 extends JFrame {

    public IN300\_WilliamAlecAkin\_Unit8\_Q1() {

        initUI();

    }

    private void initUI() {

        CategoryDataset dataset = createDataset();

        JFreeChart chart = createChart(dataset);

        ChartPanel chartPanel = new ChartPanel(chart);

        chartPanel.setBorder(BorderFactory.createEmptyBorder(15,15,15,15));

        chartPanel.setBackground(Color.white);

        add(chartPanel);

        pack();

        setTitle("Bar Chart");

        setLocationRelativeTo(null);

        setDefaultCloseOperation(JFrame.EXIT\_ON\_CLOSE);

        try {

            ChartUtils.saveChartAsPNG(new File("IN300\_Unit8\_Q1.png"), chart, 1000, 600);

        } catch (IOException e) {

            e.printStackTrace();

        }

    }

    private CategoryDataset createDataset() {

        String[][] dataArray = getMyData();

        var dataset = new DefaultCategoryDataset();

        dataset.setValue(Integer.parseInt(dataArray[1][4]), "count1", dataArray[1][2]);

        dataset.setValue(Integer.parseInt(dataArray[2][4]), "count2", dataArray[1][2]);

        dataset.setValue(Integer.parseInt(dataArray[8][4]), "count1", dataArray[8][2]);

        dataset.setValue(Integer.parseInt(dataArray[9][4]), "count2", dataArray[8][2]);

        dataset.setValue(Integer.parseInt(dataArray[40][4]), "count1", dataArray[40][2]);

        dataset.setValue(Integer.parseInt(dataArray[41][4]), "count2", dataArray[40][2]);

        return dataset;

    }

    private JFreeChart createChart(CategoryDataset dataset) {

        JFreeChart barChart = ChartFactory.createBarChart(

                "Bar Chart for Traffic Counts",

                "Street",

                "Count",

                dataset,

                PlotOrientation.VERTICAL,

                false, true, false);

        return barChart;

    }

    public String[][] getMyData() {

        String[][] csvArray = null;

        try {

            //Read in the csv file "Average\_Daily\_Traffic\_Counts.csv"

            //May need to provide absolute path to the file

            Scanner csvData = new Scanner(new File("/Users/alec/Development/PUG-IN300/Unit 8/IN300\_WilliamAlecAkin\_Unit8\_Java/IN300\_WilliamAlecAkin\_Unit8\_Java/src/Average\_Daily\_Traffic\_Counts.csv"));

            //File the array list with the data from the csv

            ArrayList<String> list = new ArrayList<String>();

            while (csvData.hasNext()) {

                list.add(csvData.nextLine());

            }

            String[] tempArray = list.toArray(new String[1]);

            csvArray = new String[tempArray.length][];

            for(int i=0; i < tempArray.length; i++) {

                csvArray[i] = tempArray[i].split(",");

            }

            csvData.close();

        }

        catch (Exception ex) {

            System.out.println("Error: " + ex.getMessage());

        }

        //Return the array with newly added data

        return csvArray;

    }

    public static void main(String[] args) {

        EventQueue.invokeLater(() -> {

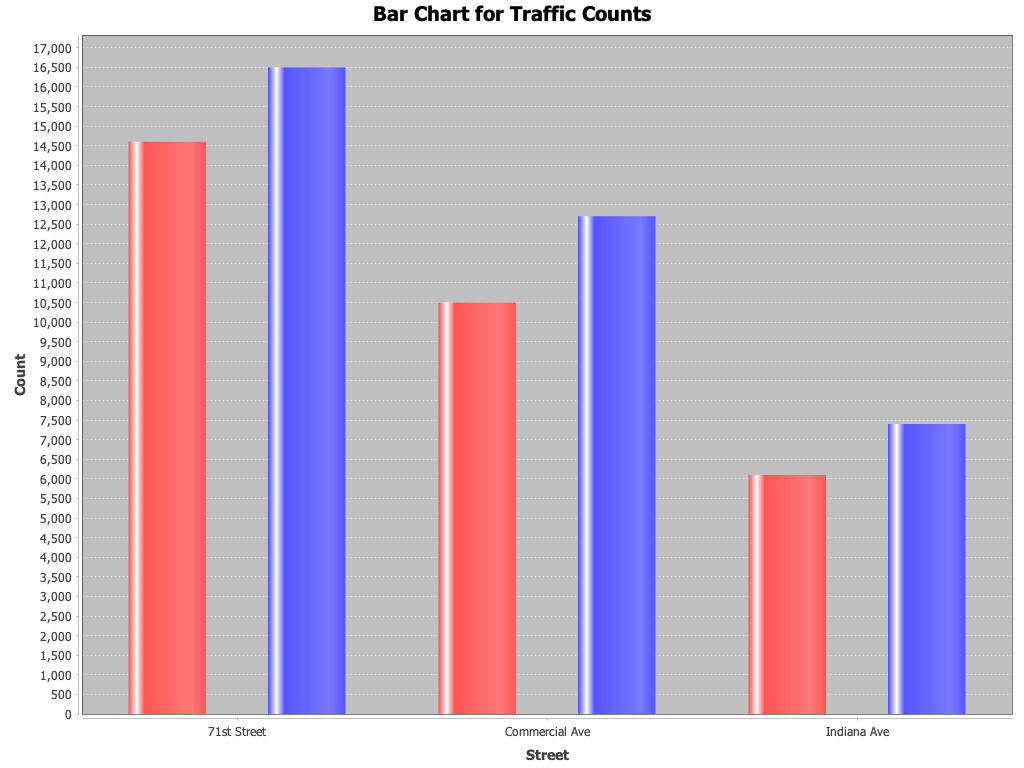
            var ex = new IN300\_WilliamAlecAkin\_Unit8\_Q1();

            ex.setVisible(true);

        });

        }

    }



# #2 – Java Source Code (JavaFX/JFreeChart, Scatter Plot)

/\*

William Alec Akin

IN300 Programming for Data Analysis

Unit 8 Assignment 1 - Java Source Code

June 26th, 2022

Description: Write a Java program using JavaFX that reads the dataset "Demographic\_Statistics\_By\_Zip\_Code.csv" data

file to create a scatter plot of count\_female and count\_male only.

 \*/

//import necessary libraries

import org.jfree.chart.ChartFactory;

import org.jfree.chart.ChartUtils;

import org.jfree.chart.JFreeChart;

import org.jfree.data.xy.XYSeries;

import org.jfree.data.xy.XYSeriesCollection;

import java.io.File;

import java.io.IOException;

import java.util.Scanner;

public class IN300\_WilliamAlecAkin\_Unit8\_Q2 {

    public static void main(String[] args) throws IOException {

        XYSeriesCollection dataset = new XYSeriesCollection();

        XYSeries series1 = new XYSeries("Scatter Chart Sample");

        //Read the csv file in

        //NOTE: The file is in the same directory as the .java file, but still may need absolute path?

        try(Scanner dataReader = new Scanner(

            new File("/Users/alec/Development/PUG-IN300/Unit 8/IN300\_WilliamAlecAkin\_Unit8\_Java/IN300\_WilliamAlecAkin\_Unit8\_Java/src/Demographic\_Statistics\_By\_Zip\_Code.csv")))

        {

            dataReader.nextLine();

            String line = null;

            String[] row = null;

            //Fill the scanner chart with data from the csv file via a loop

            while(dataReader.hasNext()) {

                line = dataReader.nextLine();

                row = line.split(",");

                int female = Integer.parseInt(row[2]);

                int male = Integer.parseInt(row[4]);

                series1.add(female, male);

                System.out.println("Female Count: " + female + " Male Count: " + male);

            }

        }

        catch(Exception e) {

            System.out.println(e);

        }

        dataset.addSeries(series1);

        JFreeChart scatterPlot = ChartFactory.createScatterPlot(

            "Demographic Scatter Plot Graph", //This is the chart title

            "Count\_Female", //X-Axis Label

            "Count\_Male", //Y-Axis Label

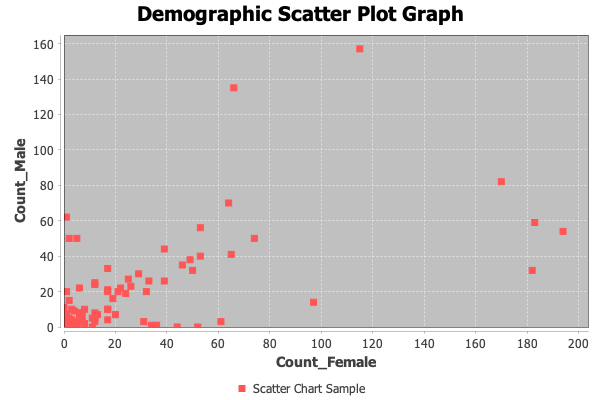
            dataset //Dataset used to construct the chart

            );

            ChartUtils.saveChartAsPNG(new File("IN300\_WilliamAlecAkin\_Unit8\_Q2.png"), scatterPlot, 600, 400);

    }

}



# #3 – Python Source Code (MatPlotLib, Bar Chart)

'''William Alec Akin

Purdue University Global

IN300 - Programming for Data Analysis (Python, R, Java)

Dr. Robert Kayl

Unit 8 Assignment - Python Source Code

June 26th, 2022

Description: Wrte a Python program to plot a bar chart using MatPlotLib and the average for the

"Average\_Daily\_Traffic\_Counts.csv" dataset file. Be sure to get the Mean of the dataset. For example,

data.mean(axis = 0) would return the mean of the dataset for each column. (Hint: data is in the Pandas dataframe)

'''

# Import required modules and libraries

import pandas as pd

import matplotlib.pyplot as plt

# Read in the file Average\_Daily\_Traffic\_Counts.csv with pandas

data = pd.read\_csv('Average\_Daily\_Traffic\_Counts.csv')

# Get the mean of the data

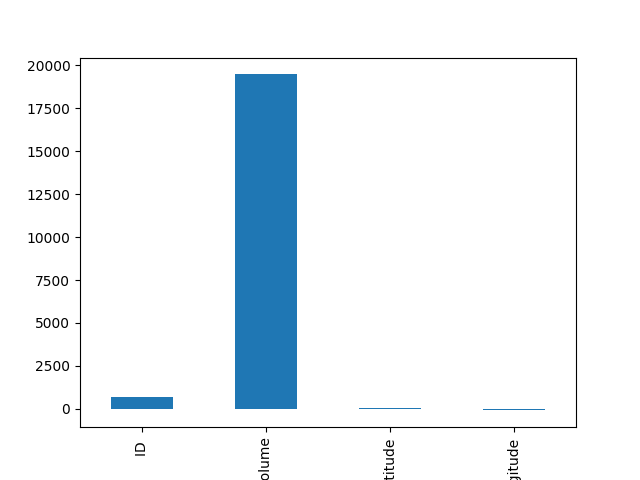
result = data.mean(axis = 0)

# Plot the mean values of the data set

result.plot(kind = 'bar')

# Display the bar chart

plt.savefig("IN300\_WilliamAlecAkin\_Unit8\_Q3\_Python.png")



# #4 – Python Source Code (MatPlotLib, Scatter Plot)

'''William Alec Akin

Purdue University Global

IN300 - Programming for Data Analysis (Python, R, Java)

Dr. Robert Kayl

Unit 8 Assignment - Python Source Code

June 26th, 2022

Description: Write a Python program to plot a scatter chart using MatPlotLib from the "Demographic\_Statistics\_By\_Zip\_Code.csv" dataset file.

You will plot the count\_female and count\_male columns.

'''

# Import necessary libraries/modules

import pandas as pd

import matplotlib.pyplot as plt

# Read the data in with pandas

# Absolute file path may be necessary depending on circumstances

data = pd.read\_csv('Demographic\_Statistics\_By\_Zip\_Code.csv')

# Remove sapces in column names

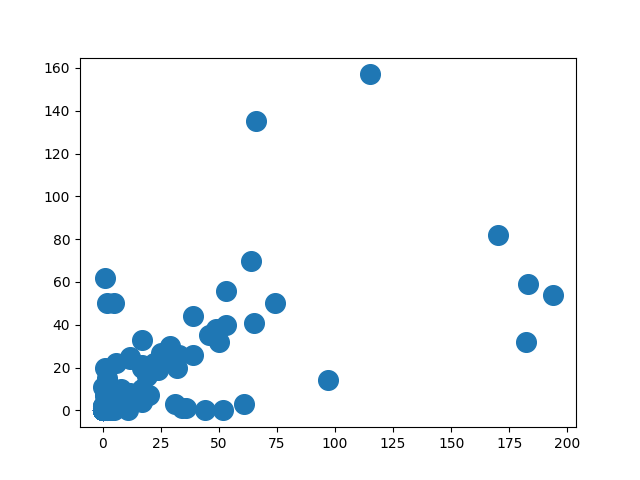
data.columns = data.columns.str.replace(" ", "\_")

# Get the counts from the file to plot as a scatter plot

plt.scatter(data.COUNT\_FEMALE, data.COUNT\_MALE, s=200)

# Save the file to local disk

plt.savefig("IN300\_WilliamAlecAkin\_Unit8\_Q4\_Python.png")



# #5 – R Source Code (Corrgram, Scatter Plot)

# W. Alec Akin

# IN300 Programmging for Data Analysis (R, Python, Java)

# Dr. Robert Kayl

# Unit 8 Assignment 1 - R Source Code

# Description: Wrte an R program using the corrgram library to plot the

# "Average\_Daily\_Traffic\_Counts.csv" file. Use the lower.panel=panel.conf,

# uppoer.panel=panel.pts settings for the plot

# Import required libraries

library(corrgram)

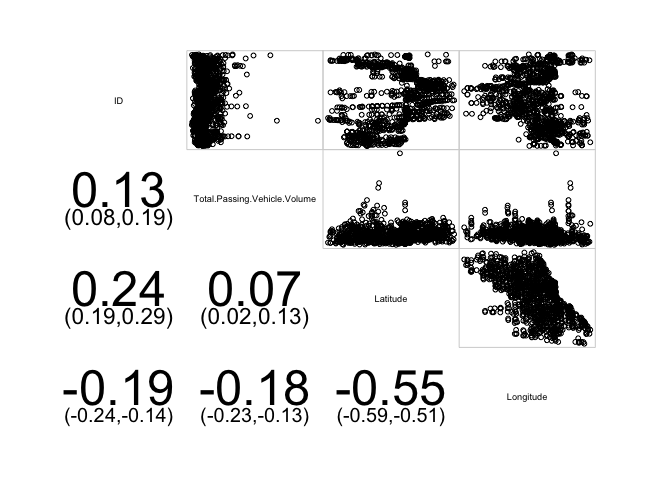
# Read in the CSV file "Average\_Daily\_Traffic\_Counts.csv"

# Absolute path may be required

traffic <- read.csv("/Users/alec/Development/PUG-IN300/Unit 8/Average\_Daily\_Traffic\_Counts.csv")

# Print the corrgram

print(corrgram(traffic, lower.panel=panel.conf, upper.panel=panel.pts))



# #6 – R Source Code (ggplot2 library, scatter chart)

# W. Alec Akin

# IN300 Programmging for Data Analysis (R, Python, Java)

# Dr. Robert Kayl

# Unit 8 Assignment 1 - R Source Code

# Description: Write an R program using the ggplot2 library to print a scatter

# plot of count\_female and count\_male columns in the following file:

# "Demographic\_Statistics\_By\_Zip\_Code.csv" dataset.

# Import required libraries

library(ggplot2)

# Read in the "Demographic\_Statistics\_By\_Zip\_Code.csv" file

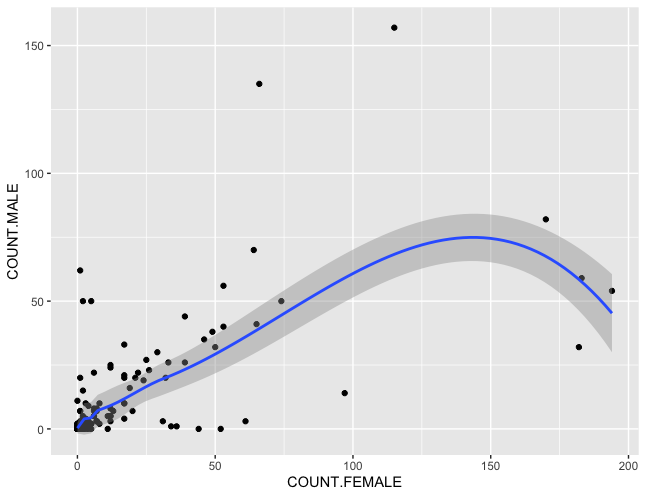
# Absolute path seems to be required

data <- read.csv("/Users/alec/Development/PUG-IN300/Unit 8/Demographic\_Statistics\_By\_Zip\_Code.csv")

# Print the scatter plot

# Note that output can vary depending on parameters used

print(qplot(COUNT.FEMALE, COUNT.MALE, data=data, geom=c("point", "smooth")))



# #7 – Compare and Contrast Languages

For every HW assignment we’ve had; Python has always taken the cake, albeit barely, against R regarding ease of use, the structure of syntax, and often lines of code required to produce the intended result. However, this time, R is far and wide the most effective at making the results with simple syntax and the fewest lines of code. This is quite rare from what I’ve experienced in programming, but R knocked it out of the park. It accomplished the objectives in approximately three lines of code each, and in my opinion, it produced the best-looking and most detailed graphs of all of the programs we wrote. The Scatter chart in question number six looks excellent! Java was also the most difficult of all languages – something it seems to repeat. I’m still unsure why someone would choose Java over R and Python for many of these assignments. It’s more complicated and typically slower than the other two, but perhaps I’m missing something?