

## Homework 5

**Due:** Thursday September 15th

**Instructions:** For the General Knowledge Questions, answer as succinctly as possible. For the Programming Assignment, show your code as well as a partial screenshot of your output.

### 1. Producing Vertical Bar Charts and Pie Charts

Use the **personl** data set and a WHERE statement to produce the three charts requested below for the ticket agents only (i.e., where **JobCode** values are TA1, TA2, and TA3.)

- Produce a vertical bar chart that displays the number of male and female ticket agents. (**Gender** values are M and F.)
- Save the chart to an HTML document file. Add an appropriate title.
- Create a pie chart to compare salaries of each ticket agent job level. Each pie slice should represent the average salary for one of the three **Jobcode** values. Send the pie chart as an ActiveX control and output to an HTML document. Use the ODS style **banker**, and add an appropriate title. Explore the interactivity of the graph in the completed HTML document.
- Enhance the pie chart by exploding the slice that represents the TA3 value of **JobCode**. Change the font color for the title to red.

### 2. Producing a Horizontal Bar Chart

Use the **chicago** data set to produce a horizontal bar chart that displays the total number of passengers boarded (**Boarded**) each day of the week. First, write a Data Step to create a new variable, **Day**, which contains the day of the week, where 1 represents Sunday, 2 represents Monday, and so on (hint: you can use a SAS function to return the Day of the Week).

- Specify the GIF file format for the chart.
- Place an appropriate title on the chart.
- Use the label Day of the Week for the variable **Day** and the label Passengers for the variable **Boarded**.

If the chart did not generate seven bars, add the DISCRETE option to the HBAR statement and generate the chart again.

1. You will need to write a Data Step to create a new variable, called **Day**. Your data step should create a new data set, called (for example) temp\_chicago. In the data step, use a SAS function to assign the day of the week (hint: Week Day) to the new variable **Day**, using the existing variable **Date** as an argument to the function.

2. Your bar chart should graph the total number of passengers who boarded (hint: the sum of **Boarded**, not the frequency).

3. You should assign a label to each of the graphed variables. If you graph the Sum of Boarded, the label you assign to the **Day** variable will appear on the chart. If you only graph the frequency, the label will not appear on the chart.
4. The bar chart should be a GIF image. You change the image type by using the Options Device= statement. The ODS file type will still be HTML.

### 3. Producing a Line Graph

The data set **delay** contains dates and delays in minutes for International Airlines flights. Use the data set and an appropriate WHERE statement to select flights to Copenhagen (**Dest='CPH'**) and produce the plot described below:

- Use ODS to save the graph to a PDF format document (e.g., ODS PDF file='linegraph.pdf';).
- Use the ODS style named Journal.
- Plot **Delay** on the vertical axis and **Date** along the horizontal axis.
- Adjust the scale on the vertical axis to start at -15 and end at 30 with a tick mark every 15 minutes.
- Display the title Flights to Copenhagen in red.
- Display the points as red squares.
- Use the NEEDLE interpolation technique to connect the points to the horizontal axis.

### 4. Producing Scatterplots

The data set **diabetes** contains physical measurements (age, height, weight, fasting glucose and post-meal glucose) for patients with diabetes. Create the plots below using this data set (NOTE: you can list more than one plot statement within a single PROC GPLOT step).

- Use ODS to save the plots in a HTML format document.
- Use the ODS style named Torn.
- Clear all previous symbols and titles.
- Set a title of 'Diabetes Patients'.
- Create a plot with **Weight** on the vertical axis and **Age** along the horizontal axis.
  - Assign the label 'Patient Age' to the **Age** axis, and 'Patient Weight (lbs)' to the **Weight** axis.
  - Adjust the scale on the **Weight** axis to range from 75 lbs to 250 lbs with a label every 25 lbs.
  - Display the points as blue dots. Add a regression line, but not a confidence interval around the line (hint: use the RL interpolation).
- Create a second plot by adding another Plot statement to the same GPLOT step. Create a plot with **Postgluc** on the vertical axis and **Fastgluc** along the horizontal axis.

- Assign the label 'Fasting Glucose' to the **Fastgluc** axis, and 'Post Meal Glucose' to the **Postgluc** axis.
  - Adjust the scales on the both axes to range from 100 to 750 lbs with a label every 50.
- Point your mouse at one of dots on the plot to explore the interactive functionality of the Activex image.
  - What does the Activex image allow you to see?