Chapter 1 Statistics Project   (100 pts)   Name:

1. **Collect quantitative data from two groups (25 or more from each group)**
    -Collect the data from the Random Sampler at http://www.amstat.org/censusatschool
    -You must <u>print</u> or copy this data and cite its source (only print what you need; delete the columns you are not researching)
    -Write a research question relevant to this data. (Eg: Are 10th grade Hawaiian students taller than 10th grade Pennsylvanian students?)
    -<u>Write</u> a couple sentences describing your data and what you are comparing. Sketch what you *hypothesize* the 2 dotplots would look like (don't actually do the data analysis yet)

2. **Collect categorical data from the same two groups (25 or more from each group)**
    -Collect the data from the Random Sampler at http://www.amstat.org/censusatschool
    -You must <u>print</u> or copy this data and cite its source (only print what you need; delete the columns you are not researching)
    -Write a research question relevant to this data. (Eg: Is left-handedness independent of state of origin, or do left-handed people tend to live in warmer states?)
    -<u>Write</u> a couple sentences describing your data and what you are comparing. Sketch what you *hypothesize* a side-by-side bar graph would look like.

The report must be typed and should include the following sections:

*   Introduction (give background information; what is your research question and why is it of interest to you?)
*   Data Collection (describe where your data came from and how you got it)
*   Data Analysis
    1. **Quantitative Data:**
        o   Mean and Standard Deviation
        o   Five number summary
        o   IQR and Range
        o   Describe what these numbers tell you about your data, in context.
        o   Show your calculations of an **outlier test** for each set of data. Are there any outliers? If there are, how do they affect your results?
        o   Display your data using 2 different graphics (dotplots, boxplots, stemplots, histograms)
        o   Describe pros and cons of each graphic
        o   Compare the 2 distributions (CSSCO)
    .2. **Categorical Data: Compare the distributions for 2 groups.**
        o   Two-Way Table or frequency table
        o   Side-by-Side or Segmented Bar Graphs
        o   Compare the distributions
        o   Describe what these graphs tell you about independence of your response variable from your explanatory variable.
*   Conclusions
*   Reflections/Error analysis; ideas for further study

Report Checklist: (think about these items as you put together your report, and make sure they are all included)

- o **Title / Research Question**
- o **Demonstrate how and why the particular topic was chosen**
- o **Background Information (why is this question of interest? What do you expect to find? Where did the data come from?)**
- o **Include the collected data and its analysis**
- o **Appropriate and well-scaled, labeled graphical displays (side-by-side boxplots, histograms, stemplots, dotplots) – do they tell a story? Which is the best display/ what are the advantages/disadvantages of each?**
- o **Have you considered and discussed Center, Shape, Spread, Outliers, Clusters? What do you notice about the data? What interesting details did your comparative graphs and statistics reveal about the two groups?**
- o **Is it neat? Does it make sense? Is it easy to read and follow?**
- o **Discuss the strengths and weaknesses of the selected statistical methods**