Math 106 Spring 2010, 01/20/10
Chapter 1 The Role of Statistics and Data Analysis Process
Chapter 3 Graphical Methods for Describing Data

Review of vocabulary:
Population
Sample
Descriptive statistics
Inferential statistics
Categorical and numerical variable
Univariate/Bivariate/Multivariate data sets
Discrete and continuous data

Example: For the following numerical attributes, state whether each is discrete or continuous.

1. The number of suitcases lost by an airline
2. The time it takes for a car battery to die
3. The length of a 1-year old rattlesnake
4. The number of insufficient-funds checks received by a grocery store during a given month
5. Number of heads obtained after flipping a coin 5 times
6. Number of cars arrive at a McDonald's between 12-1

Graphical Displays
An appropriate graphical or tabular display of data can be an effective way to summarize and communicate information.

1. We can examine the overall shape of the distribution.
2. We can detect gaps or unusual observations.
3. We can locate the center of distribution, etc.

1. **Frequency distribution for categorical data**
   Frequency
   Relative frequency

2. **Bar Charts for categorical data**
   A bar chart is a graph of the frequency distribution of categorical data.
   - Use Graph → Bar Chart.
   - Under "Bars represent:" choose Values from a table
   - Click on Simple
   - For "Graph variables" choose the column with frequencies
   - For "Categorical variable" choose the column with the categorical variables and click OK.

   If you want your bars to be labelled go to Labels.

   Look at Example 1.9 Page 17.

3. **Dot Plots for numerical data**
   - Use Graph → Dotplot.
   - Choose Simple and enter OK.
   - For "Graph variables" choose the column with frequencies.

Use it when you have small numerical data sets.

Example 1.10: The Chronicle of Higher Education reported graduation rates for NCAA Division I schools. The rates reported are the percentages of full time students in fall 93 who had earned a BA by August 99. Data from the two largest states are as follows:
California: 64 41 44 31 37 73 72 68 35 37 81 90 82 74 79 67 66 66 70 63
Texas: .....

4. **Pie Charts**
   Mostly used when we have categorical data with a relatively small number of possible categories.

   - Use Graph → Pie Chart. Choose the appropriate categorical variables and their frequencies.

   Look at Example 3.5 on page 78.

5. **Stem and Leaf Plots**
   Mostly used when we have numerical data sets with a small to moderate number of observations.

   Stem is the first part of the number and consists of the beginning digit(s).
   Leaf is the last part of the number and consists of the final digits(s)

   - Separate each observation into a stem and a leaf.
   - List the stems vertically in increasing order from top to bottom.
   - Add the leaves (typically the last digit) to the right of the stems.

   Stem and Leaf (Graph → Stem-and-Leaf)

   Look at Example 3.7 Page 83.

6. **Histograms**

   - Divide the range of the data into classes of equal width.
   - Count the number of observations in each class.
   - Compute the frequency or relative frequency for each class.
   - Erect bars over each class interval.

   Histograms (Graph → Histogram)

7. **Time Plots**

Important!!! To open a data from the P-Drive:

1. Start Minitab
2. Go to File→ Open Worksheet
3. Go to P→Data→Math→STATS and choose the required file.

Every data set that is in the book including the exercises and the examples from the text are in the P-drive. To open them

1. Start Minitab
2. Go to File→ Open Worksheet
3. Go to P→Data→Math→STATS→Peck-Devore-Data→Minitab format and choose the relevant chapter and exercises/examples. required file.