Stat 350 Lab
GSI: Sahar Zangeneh

Section 19: Monday 5:30-7:00 pm, 444-B MH
Section 49: Monday 8:00-9:30 pm, 444-B MH

Agenda
• Introductions
• Administrative stuff
• Module 1 – Introduction to SPSS
• Module 2 – Activity 1
• Example Lab Work

Highlights from the Syllabus
• Lab Structure
• Homework
• Breakdown of grades
  – 15% Homework
  – 5% Lab Participation
    (Mini-Projects, attention, discussion, etc.)
• General Policies

Websites
Old Stat 350 Website:
http://www.stat.lsa.umich.edu/~bkg/stat350/
Check C-Tools and course website frequently
Slides will be posted on:
www.umich.edu/~saharzz

Categorical Variables
• A variable that places an individual in to one of several categories
• Examples: ______________________
• Can be numerically summarized with a count, relative frequency, or percent
• Can be visually summarized with ________
• If there is an order in the categories, this is often called an _____ variable

Now let’s get started!
Quantitative Variables

- Takes numerical values for which mathematical operations make sense
- Also called numerical or measurement variables
- The distribution of the data is important to show the usefulness of meaningful numerical calculations
- The histogram is a useful graphical tool to determine the distribution

Quantitative Variables

- The distribution is often described by its shape. Some words we use include:
  - Bell-shaped
  - Uniform
  - Skewed
  - Bimodal

Quantitative Variables: Location and Spread

- Numerical summaries (when appropriate) of quantitative data can be used to describe the location or the spread of the data
- Location is described by:
  - Mean: The arithmetic average of the data
  - Median: Middle values of the data
- Spread of the data is shown by:
  - Range: the highest value minus the lowest value
  - Interquartile range: Q3-Q1
    - Q3: the third quartile, or 75th percentile.
    - Q1: the first quartile, or 25th percentile
  - Standard deviation

Standard Deviation

- Measure of spread of the observations from the mean
- The square of the standard deviation is called the variance.
- Standard deviation is given by
  \[ s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \]

Interpretation of Standard Deviation:
- In general (for any distribution), we say the points are roughly (s) away from the sample mean of (\(\bar{x}\)), on average

Summarize Quantitative Variable

- Graphical
- Numerical

Module 1: Activity 1

- Data: employee.sav
- What type of variable is gender?
- What type of graphs would be good to make for this variable?
- What type of variable is current salary?
- What type of graphs would be good to make for this variable?
Module 1: Activity 1

• Make histogram for Current Salary
• Shape of histogram?
• Customize your graphs: Put title, color...
• Always remember to put your names on all your graphs!!
• Read page 7 for how to save output

• Note: the mean and standard deviation are given with the histogram by default

Module 1: Activity 1

• Basic summary measures of current salary
  Analyze ➔ Descriptive Statistics ➔ Descriptives and click the OPTIONS button.

• What does the Mean in the output being ***** mean?

Module 1: Activity 1

• Five number summary:
  – Min, 1st quartile, median, 3rd quartile, max
  Analyze ➔ Descriptive Statistics ➔ Frequencies and click the Statistics button.
• Can also un-check the 'Display frequency tables'.
• Is the median smaller than the mean?

Module 1: Activity 2

The Mean and Median

• Open the descriptives applet from the applet link on the Stat 350 Lab Info website.

• Follow through the steps of this activity.
• Focus on the numbers on the top left hand side

Module 1: Activity 2

The Mean and Median

• How do the mean and median compare?
• Play around and explore. Fill the table below with what you observe:

<table>
<thead>
<tr>
<th>Shape of Distribution</th>
<th>Relationship btw Mean and Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric (like bell-shape)</td>
<td></td>
</tr>
<tr>
<td>Positive skewed (longer tails to the right)</td>
<td></td>
</tr>
<tr>
<td>Negative skewed (longer tails to the left)</td>
<td></td>
</tr>
</tbody>
</table>

• Play around and explore!
Module 1: Activity 2
The Mean and Median

• The mean and median are both measures of the center of a distribution

• When there are extreme observations, such as in skewed distributions, the mean will be pulled in the direction of the skew and may not be as good measure of center

• The median is not be affected by extreme observations.

Module 2: Activity 2

• Data: SSHA.sav

• Problem: How do the genders compare on SSHA scores?

• Task: Produce appropriate graphs and numerical summaries to examine the distribution of the scores (score) by gender.

Statistics 350 Qwizdom Steps

Step 1: Turn on your remotes

Hold down the MENU key for a few seconds to turn the remote on – you should see Q4 v1.16

Step 2: Check Screen – what needs to be entered?

If SESSION ID, type in the 3-digit LAB section number = _______
And press the SEND button

If Find Net or No Net … you may have a different session ID entered from a previous class …

✓ Press MENU
✓ Press Right arrow one time
✓ Session ID should be on the screen
✓ Press SEND
✓ Press YES when asked Change?
✓ Enter correct Session ID and press SEND

When screen says QWIZDOM, you are ready!
Next week: roster used and you enter 8-digit UMID