**SECTION MEETING RECAP: 9/17/14**

**Instructor**: Jess Karanian

**Homework**: Posted on Canvas under Assignments within our section meeting:

“Assignment 2: Stem and Leaf” and “Assignment 3: Nominal Data

**Due**: Friday, 9/19 at 11:59pm.

Key Concepts

**Part 1. Summation Notation**

Distribution of Scores: 2, 3, 3, 4

∑X = (2) + (3) + (3) + (4) = 12

∑X2 = (22) + (32) + (32) + (42) = 38

(∑X)2 = (12)2 = 144

MEAN: $\frac{∑X}{N}$ = $\frac{12}{4}$ = 3

SUM OF SQUARES: ∑ (X - MEAN)2 = [(2-3)2 + (3-3) 2 + (3-3) 2 + (4-3) 2]= [(-1) 2 + (0) 2 + (0) 2 + (1) 2] = 2

(∑ (X - MEAN))2 = [(2-3)+ (3-3) + (3-3) + (4-3)]= [-1 + 0 + 0 + 1] 2  = [0]2 = 0

**Part 2. Stem & Leaf Plots**



-Pick a good stem. Look at the data, find the lowest and highest score, and decide the best way to present it.

-Each score must take up the same amount of space/area to assure proper representation

-Observe the general shape of the distribution.

-Estimate the mean. Where do most of the scores lie?

-Are the distributions skewed? Look for asymmetry.

-What does the variability of scores look like? Look at how spread out or tightly clumped the scores are.

**Part 3. Nominal Data**

-Assure that categories are well defined for the respondents.

-A way to improve the scale: provide a fixed set of alternatives for respondents.

-How many categories are ideal? Offering many categories retains a lot of detail, but adds clutter.

-“Other” category is useful so that each hand every respondent has some answer to provide

-Always consider the population being measured. Adjust scales based on your sample.

-Ethical concern: Entire data set should not be published without permission to assure anonymity.

-A way to assure anonymity: save data file in parts to remove linkage of identifying information (ex. height)

**Part 4. Mean vs. Median / Skew**

Distribution of Scores: 100, 120, 100,000

Median = 120

Mean = 33,407

-The mean is heavily influenced by extreme values as it is sensitive to outliers 