Descriptive Statistics

1. Syllabus, logistics, etc.

2. Discussion of descriptive statistics

Statistics is the making of inferences about a population based on information contained in a sample from that population and to provide an associated measure of goodness for the inference.

We study statistics using *random variables* typically denoted X, Y, Z, etc. Variables can be **numerical**, both *continuous* and *discrete*, or **categorical**.

Descriptive statistics are used to characterize a set of numerical measurements.

MEASURES OF....	Population (Greek letters)	Sample (Roman letters)
CENTRAL TENDENCY:		
mean	μ	$\bar{x} =$
median		
SPREAD OR DISPERSION:		
variance	σ^2	$s^2 = \sum_{i=1}^{n} \frac{1}{n-1}(y_i - \bar{y})^2$
standard deviation	$\sigma = \sqrt{\sigma^2}$	$s = \sqrt{s^2}$

Empirical Rule: For a distribution of measurements that are approximately normally distributed,

$\mu \pm \sigma$ contains approximately _____ % of the measurements.

$\mu \pm 2\sigma$ contains approximately _____ % of the measurements.

$\mu \pm 3\sigma$ contains approximately _____ % of the measurements.

Boxplots, frequency and relative frequency histograms, quantiles, deciles, etc.