

- Exam 2 Review
- Institutional Review Board
 - Reviews studies to protect subjects from possible harm
 - Does NOT decide whether study produces valuable information
 - At least five members
 - Scientist, non-scientist, one unaffiliated
- Informed consent before data collected
 - Informed in advance of any risk of harm
 - usually in writing
- Confidentiality (no individual data, only summaries)
- Anonymity (subjects' names not known - makes follow-up difficult)
- Clinical Trials
 - Phase 1 - small group (20-80) evaluate safety, determine safe dosage
 - Phase 2- larger group (100-300), evaluate effectiveness, further evaluate safety
 - Phase 3, large groups (1,000-3,000) confirm effectiveness
 - Phase 4 - post marketing status
- Valid measure of a property - how a rate can be more valid measure than a count
- Predictive validity - measure can be used to predict success on tasks related to property measured
- Bias - systematically overstates or understates true value
 - Reduce bias by getting a better instrument
- Reliability - result is repeatable (random error is small)
 - Increase reliability by averaging several measurements
- Percent Change = $(\text{current} - \text{previous}) / \text{previous} \times 100$
 - positive result -> percent increase
 - negative result -> percent decrease (can't have a decrease of >100%)
- Types of variables:
 - Categorical -

- Nominal - categories, not ordered
- Ordinal - categories, ordered
 - Ex: months
- Quantitative
 - Discrete - countable
 - Ex: number of friends, number of classes
 - Continuous - can be measured
 - Ex: time, weight, height
- Interquartile range = $Q3 - Q1$
- Outliers: $<Q1 - 1.5(IQR)$ or $>Q3 + 1.5(IQR)$
- 5 number summary (box plots)
 - Minimum
 - Q1 (median of data below the distribution median)
 - Median (Q2)
 - Q3 (median of data above the distribution median)
 - Maximum
- Mean is strongly affected by extreme values
- Standard deviation - average deviation of observations from the mean
 - if the data is more spread out, there will be a larger standard deviation
 - $s = 0$ only if there is no spread or variability in the data (all of the values in a set of data are exactly the same)
- Median is NOT strongly affected by extreme values