

## This document is subject to change

- February 15
- worth 250 pts (25% of the final grade)
- Chapter 2
  1. type of variable. (i.e. categorical, quantitative (discrete or continuous))
  2. histograms
    - construction
    - shape (modality and skewed vs symmetric)
    - difference from time plot
  3. time plots (recognize trends)
  4. mean
  5. median
  6. mode
  7. resistance to outliers
  8. how to tell if the distribution is skewed using the mean and median. understand reasons for the skew.
  9. standard deviation
  10. range (usefulness of?)
  11. empirical rule
  12. position / 5 number summary
  13. boxplots
  14. meaning of percentiles.
  15. interquartile range
  16. using the IQR to identify outliers
  17.  $z$  scores
- Chapter 3
  1. explanatory variables
  2. response variables
  3. association
  4. contingency table
  5. conditional proportions
  6. determining association using a contingency table
  7. scatterplot
  8. positive/negative association

9. correlation
  10. linear relationships
  11. prediction using regression
  12. prediction error / residuals
  13. calculating the regression coefficients
  14.  $r^2$  interpretation of (?)
  15. dangers of extrapolation
  16. regression outliers (what to do with them, how they might influence the regression prediction)
  17. correlation  $\neq$  causation
  18. lurking variables
- Chapter 4
    1. types of studies (experiment vs observation)
    2. benefits of each type
    3. causation can be determined by experiments
    4. sample surveys
    5. types of bias
    6. convenience samples
  - Chapter 5
    1. probability long run behavior
    2. independent trials
    3. sample space
    4. outcomes
    5. events
    6. how to calculate a simple probability
    7. complement
    8. intersection (and)
    9. union (or/both)
    10. disjoint events
    11. probability for union and intersection
    12. probability for intersection of two independent events
    13. determining independence of events using conditional probabilities