Introductory comments
Describe how familiarity with statistical methods can
- be associated with “power” in the workplace
- lead to advancement in the workplace
- provide a type of literacy that crosses many domains
- help to control for cognitive biases and illusions
- be useful in improving communication, computer skills, information management, technical literacy, career advancement and quality assessment
- be applied to several domains within business

Statistics is the “science of collecting, organizing, presenting, analyzing, and interpreting data to assist in making effective decisions” Lind, page 4

Biases and Illusions
- Describe how the rat/man demonstration illustrates how expectations affect our perceptions of the world
- Describe how the ambiguous (blob) figures illustrate how knowledge affects our perceptions of the world
- Describe how the attention blindness demonstration (gorilla) illustrates our vulnerability to bias
- Describe how awareness of these illusions allows us to be better critical thinkers
- Please note: Chapter 1 on Plous also addresses these phenomena

Contrast the measurement of observable actions (and/or characteristics) with the theoretical constructs associated with those measurements Lind, page 565 – 566
- Be able to identify examples of measurements versus constructs
- Describe what is meant by operational definitions
- Describe what is meant by validity versus reliability in measuring a construct
  - Provide an example of a measure that has high reliability and low validity?
  - Provide an example of a measure that has low reliability and high validity?
  - Describe the “target” example of reliability and validity
- Describe what is meant by a “meta-analysis” Lind, page 567

Contrast a dependent variable (the data) versus independent variable (how the groups differ)? Lind, page 572
- be able to identify them in an example
A study with a single independent variable would be associated with a univariate data set
- while a study with two variables (as in a correlation) would be associated with a bivariate data set
- a study with more than two variables would be associated with a multivariate data set
Contrast control groups versus experimental groups Lind, page 573
- be able to identify them in an example.
What is a within versus between participant design? Lind, page 573
- also known as within and between subject designs?
What is random assignment - what’s it good for?
What is random sampling - what’s it good for?
Populations versus samples - be able to define and identify examples of each
- Define and contrast a statistic with a parameter Lind, page 59, 60
- Sample data
- Population data
- Sample mean ( \( \bar{x} \))
- Population mean (\( \mu \))
Describe what is meant by a placebo \textit{Lind, page 575}

Describe what is meant by a placebo effect \textit{Lind, page 575}

Contrast the double-blind procedure with the single-blind procedure \textit{Lind, page 572 – 575}

Describe how this is related to experimenter bias \textit{Lind, page 573}

Contrast true experimental designs with quasi-experimental designs \textit{Lind, page 572}
- remember “random assignment” is key
- subject variables like gender or political affiliation are associated with quasi-experimental designs
- with no random assignment we are especially vulnerable to selection bias and confounds \textit{Lind, 572}
- be able to identify each type of design from a description of a study

Describe what is meant by “cause-and-effect” \textit{Lind, page 572}

Describe why we can talk about cause-and-effect in ‘true’ experiments, but we can only describe relationships in quasi-experimental designs (as in correlational studies)

Contrast inferential statistics (\textit{Lind, page 583}) with descriptive statistics (\textit{Lind, page 579}) \textit{See also Lind, page 6}

Be able to identify each from a description of a study

Define continuous versus discrete variables \textit{Lind, page 9}

Define categorical versus numerical data
- These are also known as qualitative versus quantitative data \textit{Lind, page 8}
- Be able to identify each from examples
- Contrast verbal and coded labels of categorical data
  Remember, that just because someone might code male = 1 and female = 2, this does not make “gender” a quantitative variable
- Describe what is meant by a binary variable

Levels of measurement: be able to identify each from an example \textit{Lind, page 9 - 13}
- Nominal
- Ordinal
- Interval
- Ratio

Define what is meant by naturalistic observation and field observation

Describe time series (also called longitudinal) versus cross-sectional comparisons \textit{Lind, page 570}

Describe the “Likert Scales” \textit{an example is presented in Lind, page 569}

Describe surveys and what the goal of a survey would be

Contrast a census with a sample

Contrast a population with a sample \textit{Lind, page 7, & 577}
- Review situations where sample or census might be preferred
- Describe how a sample can be representative versus biased
- Define random selection and random assignment relative to sampling for administering survey

What is meant by a random sample \textit{Lind, page 577}

Contrast random sampling with random assignment

Define a sampling frame
- How might a “sampling frame” be different from a “target population”

Describe why larger samples can be preferable to smaller samples

Describe why a higher response rate percentage is better than a smaller response rate
- Describe why response rate is so important for a sample to remain unbiased \textit{(Lind, page 617)}
Define and contrast probability sampling with non-probability sampling (Lind, page 613)

Describe these sampling techniques (Summarized in a table, Lind, page 616)
- simple random sampling (this is a probability sampling technique) (Lind, page 613)
  - what is a random numbers table good for?
  - how would you find a random number using excel
- systematic random sampling (this is a probability sampling technique)
- stratified random sampling (this is a probability sampling technique) (Lind, page 613)
  - proportional and disproportional stratified random sampling
- cluster sampling (this is a probability sampling technique) (Lind, page 614)
- convenience sampling (aka haphazard sampling and is a non-probability sampling technique) (Lind, page 615)
- snowball sampling (this is a non-probability sampling technique)
- judgment sampling (this is a non-probability sampling technique)

Describe what is meant by an interview
- contrast structured with unstructured interviews
- contrast gathering quantitative versus qualitative interview data
- contrast informal conversational interview vs highly structured closed-ended interview

Describe a focus group
- information gathered can be qualitative or quantitative
- one can use probability or non-probability sampling technique

Describe what is meant by a questionnaire
As “consumers” of questionnaire data what should we be evaluating?
- number of responders versus percentage of responders
- methodologies used in sampling
- operational definitions of constructs
- wording and ordering of items
- See also chapter 6 in Plous text

Describe the 5 principles of questionnaire construction
1. Make sure items match research objectives and identify what constructs you are trying to understand
   (Be explicit in identifying your constructs)
2. Responders have the answers to our questions
   • We are tapping into their attitudes/beliefs/ knowledge
   • Understand your research participants
     • “think like” the responders / consider their sensibilities
     • use appropriate, natural and familiar language (for them)
   • Be simple, be clear, precise and concise (short questions) (Lind, page 600)
     • don’t make them work any harder than they need to
3. Assessment should feel easy and clear, unthreatening
   • Start with most friendly (least threatening) questions first
     • then at the end “now a couple questions about you” (foot in the door phenomenon)
   • Avoid double negatives (Lind, page 601)
     • For example: Agree or disagree? “Teachers shouldn’t have less contact with parents”
4. Avoid ambiguity and bias in your items
   • Avoid “double-barreled” questions - Difficult to interpret answers(Lind, page 601)
   • Avoid leading or loaded questions - Can introduce bias(Lind, page 601)
   • Consider problem of acquiescence – Ask question in different ways (careful with coding)
     - also “Yea-saying” or “Nay-saying” (Lind, page 601)
5. Consider lots of different formats for responses
   • Consider open-ended vs closed-ended questions (Lind, page 602)
   • Consider complementing your questionnaire with other forms of data collection

Pilot – feedback – fix - pilot – analyze – fix - pilot – etc (Exploit process of empirical approach)
Describe and give examples of rating scales (Lind, page 603)
- describe what is meant by anchor points versus fully anchored rating scales

Describe what is meant by ranking scales

Describe and give an example of a semantic differential technique

Describe how summated rating scales work - give an example of how the responses can be summed up

Describe how pilot testing a questionnaire can be useful

Describe how peer review of questionnaire can be useful

Administering surveys - identify pros and cons of each (Lind, page 607 - 608)
- personal administering to groups or individuals - face-to-face
- mail surveys
- internet surveys
- telephone surveys
- focus groups

Review Ethics Guidelines for gathering data (Lind, page 586-587)

Describing Data Visually

Describe how raw data can be organized and presented in a dot plot Lind, page 96 - 98
Describe a frequency distribution for grouped data
Describe a frequency distribution for ungrouped data
Review the guidelines for constructing frequency distributions
- classes should be mutually exclusive
- set of classes should be exhaustive
- all classes should have equal intervals
- selection of number of classes and interval width is multi-faceted
- class width should be round (easy numbers)
- avoid open-ended classes
Describe what is meant by cumulative frequencies (be able to calculate them)
Describe what is meant by relative frequencies and cumulative relative frequencies (be able to calculate)
Describe a frequency histogram and a cumulative frequency histogram
Describe a frequency polygon and a cumulative frequency ogive (how do you pronounce "ogive")
Describe a contingency table Lind, page 113, 143
Describe a tree diagrams Lind, page 145

Describe the different shapes of frequency histograms and be able to interpret examples
- skewed left, skewed right (negatively or positively skewed) Lind, page 69, 106, 107
- bimodal skewed left, bimodal skewed right
- skewed left with outliers, skewed right with outliers
- symmetric, multimodal symmetric, symmetric with outliers
- how does the skew of the distribution affect the order of the mean, median and mode Lind, page 107
Describe a Bar chart Lind, page 24
Describe a Pareto chart and be able to interpret one and when it should be used
Describe a Stacked Bar chart and be able to interpret one and when it should be used
Describe a Simple line chart and be able to interpret one and when it should be used Lind, page 25

Define a correlation
Describe and be able to interpret scatterplots
- strong positive pattern, strong negative, weak positive, weak negative, zero pattern, curvilinear pattern
- See also Lind 111, 112
Describe the relationship between the strength and direction of the correlation coefficient
Direction of correlation (positive, negative)
- Be able to identify a positive correlation from a correlation coefficient, a verbal description of the data, a scatterplot, or raw data
- Be able to generate examples of both negative and positive correlations
Strength of correlation (0 - 1.0 or 0 - 1.0)
- Be able to identify (or estimate) the strength of a correlation from a correlation coefficient, a scatterplot, or raw data
Correlation vs causation (when does a correlation imply, or provide evidence for causation?)
Describe linear vs curvilinear relationship -
- Be able to identify a linear or curvilinear relationship from raw data, a verbal description of data, or a scatterplot

Review the guidelines for constructing frequency distributions
- classes should be mutually exclusive
- set of classes should be exhaustive
- all classes should have equal intervals
- selection of number of classes and interval width is multi-faceted
- class width should be round (easy numbers)
- avoid open-ended classes
- review how to find class midpoint and class interval. *Lind, page 33*
Describe cumulative frequencies (be able to calculate them)
Describe relative frequencies and cumulative relative frequencies (be able to calculate) *Lind, page 34*
Describe a frequency histogram and a cumulative frequency histogram *Lind, page 36*
Describe a frequency polygon and a cumulative frequency ogive  *Lind, page 39, 43*

**Distributions**

Describe the dot plot display and how it represents central tendency and dispersion and shape
Define and contrast these three characteristics of distributions: central tendency (measure of location), dispersion (measure of variability), and shape

Define the characteristics of the normal curve
- Measured on continuous scale
  *Note: when range is large we often treat a discrete variable as continuous (exam scores for example)*
- Possess clear central tendency
- Have only one peak (unimodal)
- Exhibit tapering tails
- Be symmetric around the mean (equal tails)
- Be bell-shaped

What are the three measures of “location” or measures of “central tendency?” *Lind, page 58*
- Mode: most commonly occurring score – also tallest point on normal distribution *Lind, page 66*
- Median: middle score also the midpoint score: (remember also the 50th percentile)  *Lind, page 63*
- Describe procedure for finding quartiles  *Lind, page 99, 100*
- Describe how box plots use medians, quartiles, maximum and minimum scores  *(Lind, page 103)*
  *Note: just plots the minimum score, maximum score and first three quartiles*

- Mean: average score also balance point of distribution  *Lind, page 59*
  - define what is meant by a “trimmed mean”
  - define what is meant by a weighted mean  *Lind, page 63*
In a normal distribution mean = median = mode
In a positively skewed distribution mean > median > mode
In a negatively skewed distribution mean < median < mode
*Note: the mean is most influenced by extremely large or extremely small scores*

Each measure of central tendency is useful for which type of data (nominal, ordinal, interval and ratio)
Describe the empirical rule for the normal curve *Lind, page 85*

Normal distribution
Describe what is meant by a z-score and how area under the curve relates to a particular z-score
What is the probability that a score will fall above a z of 0 (50%)
What is the probability that a score will fall between -1 and +1 standard deviation of the mean? - 68%
What is the probability that a score will fall between -1 or +1 standard deviation of the mean? - 34%
- notice, z = 1
What is the probability that a score will fall between -2 and +2 standard deviation of the mean? - 95%
What is the probability that a score will fall between -1 or +1 standard deviation of the mean? - 47.5%
- notice, z = 2
What is the probability that a score will fall between -3 and +3 standard deviation of the mean? - 99.7%
What is the probability that a score will fall between -1 or +1 standard deviation of the mean? - 49.85% - notice, z = 3

Measures of variability (dispersion) *Lind, page 73, 74*
- range: smallest score subtracted from the largest score *Lind, page 75*
  - note range makes no reference to scores between the largest and smallest scores
- variance: standard deviation squared *Lind, page 78, 79*
- standard deviation: typical amount observations deviate on either side of their mean *Lind, page 78, 79*
  - \( \Sigma(x - x) = 0 \) also \( \Sigma(x - \mu) = 0 \)
- Describe the definitional formula for the standard deviation?
- Describe a “deviation score”
- memorize the standard deviation and variance (definitional) formula for samples and populations
  *Lind, page 79 - 82*
- how is definitional formula different from the calculation formula for the standard deviation
  (be able to calculate standard deviation using calculation formula)
- how are the formula for standard deviations for samples different than for population
- how are standard deviation formula different from variance formulas
- describe what is meant by a deviation score
- what would happen if we took the average of the deviation scores (without taking the square or absolute values of each deviation)?
- describe the “mean deviation” - it uses the absolute values of the deviation scores *Lind, page 75, 76*
- what does it mean to say that the standard deviation is calculated relative to the mean
- be able to calculate standard deviation and variance from a set of scores
- be able to estimate it from a normal distribution when given an example deviation
- Note: standard deviation can also be estimated by range / 6

Describe what it means to say that the "Mean is a measure of 'position', (it lives on one location of the curve) and than the standard deviation is a distance score of the spread of the distribution

Describe what is meant by
- "not unusual scores" (z of less than 2), “unusual scores” (z between 2 up to 3),
- “outliers” (z between 3 up to 4) and “extreme outliers” (z of 4 and above)
Selective Perception (Chapter 1)

Describe how expectations can:
- affect simple perception, consider the example of the playing cards (*Plous, page 15*)
- cause a “placebo effect” in alcohol-related behaviors (*Plous, page 17*)
- affect perception of a complex event (football game) (*Plous, page 18*)
- affect how we *perceive* the neutrality of the media in an election year (*Plous, page 20*)

Plasticity (Chapter 5)

Describe the term “plasticity” relative to malleability of responses to survey items
(*Plous, page 53, also page 58*)
Describe the idea of “order effects” and consistency demonstrated by Schuman and Presser (1981)
- Especially example of responses to accessibility of Communist versus US reporters
(*Plous, page 52*)
Describe what is meant by pseudo-opinions
- generally pseudo-opinions are offered by what percentage of respondents (*Plous, page 55*)
When researchers are “filtering out pseudo-opinions”, describe what they are attempting to do?
Describe what is meant by “consistency” and how it differs from “plasticity”
Provide an example of “Attitudes about abstract propositions are often unrelated to attitudes about specific applications of the same propositions” (*Plous, page 59*)
- Note: this is attitude-attitude inconsistency
Provide two examples of “attitude-behavior” inconsistency (e.g. cheating, discrimination, and helping)
(*Plous, page 59 – 61*)
- These examples will help explain what Wicker (1969) meant when he said “it is considerably more likely that attitudes will be unrelated, or only slightly related to overt behaviors than that attitudes will be closely related to actions”
What 4 conditions can help improve the problem of attitude-behavior inconsistency (*Plous, page 62*)

Wording and Framing (Chapter 6)

Describe the shift in responding (described by Schuman & Scott, 1987) when they compared responses to an open-ended question to a closed-ended question
- Note especially how responses to otherwise uncommon options were exaggerated
(*Plous, page 65-66*)
Describe what is meant by (*Plous, page 67*)
- “Response scales are not simply ‘measurement devices’ that respondents use to report their behaviors. Rather…respondents may use the range of behaviors described in the response alternative as a frame of reference in estimating and evaluation of their own behavior”
Describe the findings comparing the word “forbid” versus “allow” in question construction (page 69)
Describe what is meant by a decision frame (*Plous, page 69*)
Describe what is meant by psychological accounting, especially the “lost ticket” example (page 74)

Anchoring and Adjustment (Chapter 13)

Describe what is meant by anchoring and adjustment, especially the “wheel of fortune”, “nuclear war” and real estate examples (*Plous, page 145, 147 & 148*)
Describe how even outrageously extreme anchors can affect responses

Note: Lecture notes can be found at
http://courses.eller.arizona.edu/mgmt/delaney/