KEY

Statistics/Univariate Data Unit

Vocabulary: Below is a list of terms you will need to be familiar with for this unit. Please write the definitions and examples for each.

1. Variable
   a characteristic that can be measured/counted
2. Individuals (elements)
   people/objects in a study
3. Observational Study
   understand cause + effect by observation
4. Sample Survey
   small set of people surveyed to get results
5. Population
   entire pool from which a sample is drawn
6. Sample
   subset of population
7. Census
   acquiring info about the population
8. Experiment
   procedure to make a discovery based on a hypothesis
9. Bias
   systematic error that has prejudice in favor of something
10. Convenience Sample
    sample of most available subjects
11. Voluntary Sample
    people self-select into survey
12. Simple Random Sample
    people chosen randomly
13. Mean
    average
14. Median
    middle # of list
15. Mode
    appears most
16. Midrange
    midpoint of range (avg of max + min)
17. Five-Number Summary
    - min
    - 1st quartile
    - median
    - 3rd quartile
    - max
18. Outliers
    # detached from main data group
19. Quartile
    one of 4 equal groups into which a population can be divided
20. Variance / Standard Deviation
    extent of deviation from mean (spread of data)

KEY

Identifying Populations and Samples

Name:

Use the scenario to identifying populations and samplings.

1) A beverage company wanted to see if people in the United States liked their new logo.

Which choice **best** represents a population?

A. A selection of logo artists.
B. Every person in the United States.
C. A selection of shoppers from different states.
D. 3,800 children age 5 - 15

2) A musician wanted to see what people who bought his last album thought about the songs.

Which choice **best** represents a sample?

A. Every person who bought the album.
B. A selection of people who didn't want to buy the album.
C. 250 girls who bought the album.
D. A selection of 3,294 people who bought the album.

3) A gaming website wanted to find out which console its visitors owned.

Which choice **best** represents a population?

A. Visitors to the 3DS section.
B. All of the website visitors.
C. Visitors to the PS4 section.
D. Visitors who are on the website for more than 5 minutes.

4) Before a nation wide election, a polling place was trying to see who would win.

Which choice **best** represents a sample?

A. A selection of voters over age 50.
B. A selection of male voters.
C. A selection of voters of different ages.
D. All voters.

5) A toy store owner tracking how much kids spend each month on toys.

Which choice **best** represents a population?

A. All of the kids who buy toys.
B. 227 rich kids.
C. 228 boys age 7 - 15
D. 235 kids from age 10 to 15.

6) A mayor wanted to see if the people in his town thought he was doing a good job.

Which choice **best** represents a sample?

A. 1,000 unemployed voters.
B. The mayor's family.
C. The residents of the town.
D. 242 voters.

Answers

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____

AFM
Univariate Data
Categorical vs. Quantitative Data

Name _____
Date _____

Determine whether the following variables are **categorical (C)** or **quantitative (Q)**

1. Brand of vehicle purchased by a customer C
2. Price of a CD Q
3. Type of M&Ms preferred by students (peanut, plain) C
4. Phone number of each student C
5. Height of a 1-year old child Q
6. Term paper status (turned in on time or turned in late) C
7. Gender of the next baby born at a particular hospital. C
8. Amount of fluid (oz) dispensed by a machine used to fill bottles with soda Q
9. Thickness of the gelatin coating on a Vitamin C capsule Q
10. Brand of computer purchased by a customer C
11. State that a person is born in C
12. Price of a textbook Q
13. Zip code of each student in this class C
14. Actual weight of coffee in a one pound can Q
15. Length of a rattlesnake Q

Types of Data:

d. Univariate Data: 1 variable data (This is what we're studying THIS unit)
e. Quantitative data: numbers a quantity (ex. height, weight, age, etc)
f. Categorical data: group names (categories)
(makes sense to take an average)
- Each individual falls into a category (Ex: Freshmen, sophomore, junior, senior; male, female)

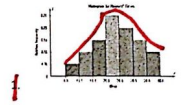
Concerned with how many or what percent fall into each category

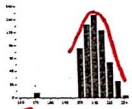
g. Continuous Data: every value within the same range is possible

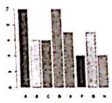
The following vocabulary words can be used to describe graphical displays

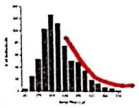
Uniform Each bin has approximately the same height	**Gaps** Spaces between data points 3	**Uni-Modal** One bin has the highest value 1, 2, 3, 4, 5	**Bi-Modal** Two bins tie for the highest value 6
Multi-Modal There are more than two ties for the highest bin	**Outliers** Extreme values that don't appear to belong with the rest of the data 3	**Symmetric** The two halves look like approximate mirror images 4	**Normal** Looks like a hill with the highest peak near the middle 4
Long Tails The edges slowly drop off 2, 5	**Short Tails** The edges drop off quickly 1, 3	**Skewed Left** The longer tail reaches to the left 5	**Skewed Right** The longer tail reaches to the right 2

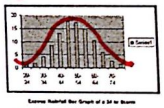
Use as many of these vocabulary words to describe the following displays

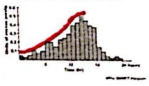

1


3


6


2


4

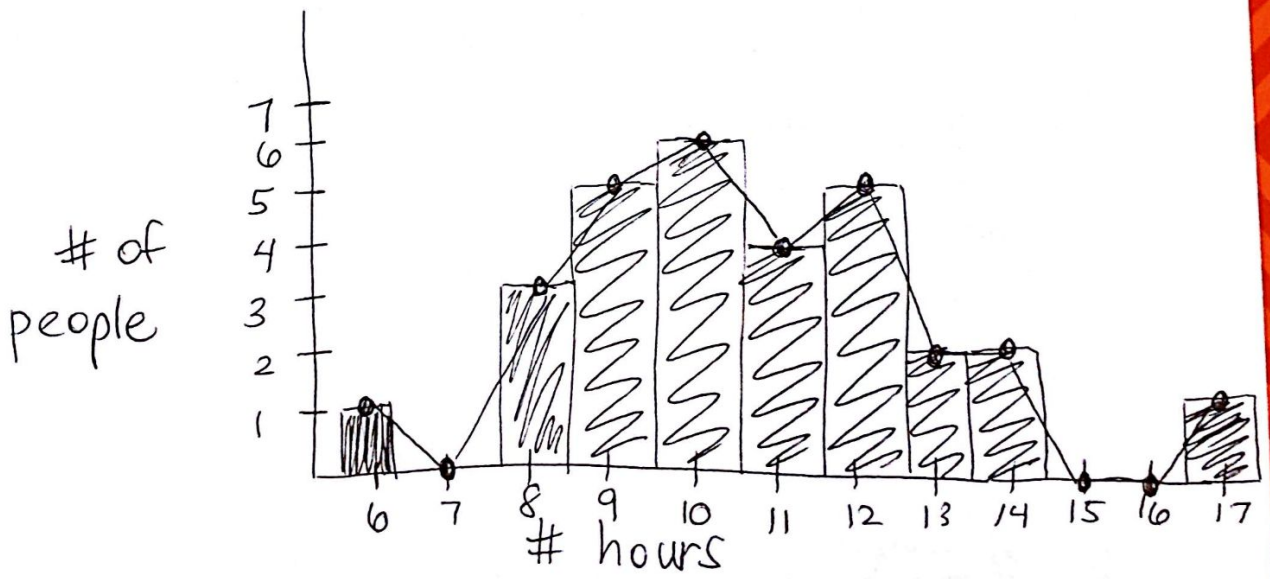

5

Day 1 HW KEY:

1) 100,000 residents
2) households of city A
3) C
4) C

5)

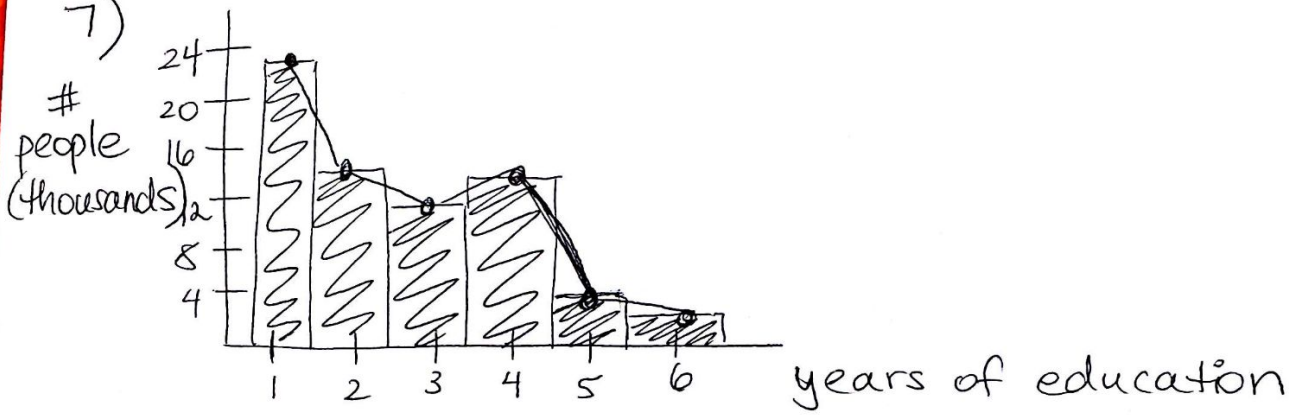
Hours	Tally	Frequency
6	1	1
7	0	
8	3	III
9	5	~~IIII~~
10	6	~~IIII~~ I
11	4	IIII
12	5	~~IIII~~
13	2	II
14	2	II
15	0	
16	0	
17	2	II



6)

Age Group	Tally	Frequency
17-26	𝍸 𝍸 I	11
27-36	𝍸 IIII	9
37-46	𝍸 I	6
47-56	IIII	4
57-66	II	2

7)



8)

Stem	Leaf
3	9 9
4	0 5 1 2 9 8 10
5	2 5 1 3 4 4 8
6	8 5 2 1 3 4 3
7	3 1 0 7
8	2 1