Quiz 1   **Name:** _____

For each of the first two exercises, identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.

1. (6 points) A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

2. (6 points) A politician is interested in the proportion of voters in his district who think he is doing a good job.

3. (3 points) Create an example of the cluster sampling technique. Be specific.

For the following exercises, identify the type of data that would be used to describe a response: numerical discrete, numerical continuous, or categorical.

4. (1 point) distance to the closest primary polling place

5. (1 point) favorite presidential candidate

6. (1 point) time in line to buy groceries

7. (1 point) 2020 Kansas City Chiefs jersey numbers

8. (5 points) Name the sampling method used in each of the following situations:

   (a) A woman in the airport is handing out questionnaires to travelers asking them to evaluate the airport's service. She does not ask travelers who are hurrying through the airport with their hands full of luggage, but instead asks all travelers who are sitting near gates and not taking naps while they wait.

   (b) A teacher wants to know if her students are doing homework, so she randomly selects rows two and five and then calls on all students in row two and all students in row five to present the solutions to homework problems to the class.

   (c) The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.

   (d) The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which she marks whether books are checked out by an adult or a child. She records this data for every fourth patron who checks out books.

   (e) A political party wants to know the reaction of voters to a debate between the candidates. The day after the debate, the party's polling staff calls 1,200 randomly selected phone numbers. If a registered voter answers the phone or is available to come to the phone, that registered voter is asked whom he or she intends to vote for and whether the debate changed his or her opinion of the candidates.

9. (3 points) The age at inauguration for each U.S. President is shown. Construct a stem and leaf plot for the data.

57 54 52 55 51 56 61 68 56 55 54 61 57 51 46 54 51 52 57 49 54 42 60
69 58 64 49 51 62 64 57 48 50 56 43 46 61 65 47 55 55 54 46 47 70

10. (3 points) Construct a relative frequency distribution using the following class intervals: 40 to < 45, 45 to < 50, etc.

11. (3 points) Construct a histogram distribution using the following class intervals: 40 to < 45, 45 to < 50, etc.

12. (4 points) In advance of the 1936 Presidential Election, a magazine titled Literary Digest released the results of an opinion poll predicting that the republican candidate Alf Landon would win by a large margin. The magazine sent post cards to approximately 10,000,000 prospective voters. These prospective voters were selected from the subscription list of the magazine, from automobile registration lists, from phone lists, and from club membership lists. Approximately 2,300,000 people returned the postcards.

   (a) Think about the state of the United States in 1936. Explain why a sample chosen from magazine subscription lists, automobile registration lists, phone books, and club membership lists was not representative of the population of the United States at that time.

   (b) What type of bias might be present in the sample.

   (c) What is the response rate of the survey?

   (d) What effect does the low response rate have on the reliability of the sample?

13. (6 points) Do people with college degrees think that their education was worth the cost? This question was posed to 2548 adults with an associate degree and to 30,151 adults with a bachelor's degree. The data, from the Gallup report "Two-Year Grads Satisfied with Cost of Degree" (www.gallup.com, April 11, 2016), are summarized in the accompanying table.
Construct a comparative bar chart for the survey responses.

| response | frequency | |
	associate degree holders	bachelor's degree holders
strongly disagree	178	1508
disagree	153	2111
neither agree or disagree	408	4221
agree	637	8743
strongly agree	1172	13568

Group Members First and Last Names:

The article "Rethinking Calcium Supplements" (U.S. Airways Magazine, October 2010) describes a study investigating whether taking calcium supplements increases the risk of heart attack. Consider the following four study descriptions. For each study, answer the following five questions:

Question 1: Is the study described an observational study or an experiment?

Question 2: Did the study use random selection from some population?

Question 3: Did the study use random assignment to experimental groups?

Question 4: Based on the study description, would it be reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack?

Question 5: Would it be reasonable to generalize conclusions from this study to some larger population? If so, what population?

Study 1: Every heart attack patient and every patient admitted for an illness other than heart attack during the month of December, 2010, at a large urban hospital was asked if he or she took calcium supplements. The proportion of heart attack patients who took calcium supplements was significantly higher than the proportion of patients admitted for other illnesses who took calcium supplements.

Study 2: Two hundred people were randomly selected from a list of all people living in Minneapolis who receive Social Security. Each person in the sample was asked whether or not they took calcium supplements. These people were followed for five years, and whether or not they had had a heart attack during the five-year period was noted. The proportion of heart attack victims in the group taking calcium supplements was significantly higher than the proportion of heart attack victims in the group not taking calcium supplements.

Study 3: Two hundred people were randomly selected from a list of all people living in Minneapolis who receive Social Security. Each person was asked to participate in a statistical study, and all agreed to participate. Those who had no previous history of heart problems

were instructed to take calcium supplements. Those with a previous history of heart problems were instructed not to take calcium supplements. The participants were followed for five years, and whether or not they had had a heart attack during the five-year period was noted. The proportion of heart attack victims in the calcium supplement group was significantly higher than the proportion of heart attack victims in the no calcium supplement group.

Study 4: Four hundred people volunteered to participate in a study. Each volunteer was assigned at random to either group 1 or group 2. Those in group 1 took a daily calcium supplement. Those in group 2 did not take a calcium supplement. The proportion who suffered a heart attack during the study period was noted for each group. The proportion of heart attack victims in group 1 was significantly higher than the proportion of heart attack victims in group 2.

Quiz 1   Name: _Key_

For each of the first two exercises, identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.

1. (6 points) A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

a) All of the cardiologist's patients who have had a heart attack.

b) a group of patients of this cardiologist — their recovery times.

c) the mean recovery time (in days or weeks) of all the cardiologist's patients who have had a heart attack

d) the mean recovery time of the sample

e) X = the recovery time (in days or weeks)

f) values for X, such as 5 weeks, 6 weeks, 12 weeks, and so on.

2. (6 points) A politician is interested in the proportion of voters in his district who think he is doing a good job.

a) All voters in the politician's district

b) a group of voters from the politician's district

c) the percentage of voters in the district who think the politician is doing a good job.

d) the percentage from the sample

e) X = the number of voters who "yes" to the question: "Do you think the politician is doing a good job?"

f) A list of "yes" or "no" responses

3. (3 points) Create an example of the cluster sampling technique. Be specific. A list of Italian restaurants are gathered for a list of cities in the county. 3 or 5 cities are randomly selected, and the list of italian restaurants from the 3 or 5 cities makes up the sample.

For the following exercises, identify the type of data that would be used to describe a response: numerical discrete, numerical continuous, or categorical.

N C 4. (1 point) distance to the closest primary polling place

C 5. (1 point) favorite presidential candidate

N C 6. (1 point) time in line to buy groceries

C 7. (1 point) 2020 Kansas City Chiefs jersey numbers

8. (5 points) Name the sampling method used in each of the following situations:

Voluntary Response Sampling

(a) A woman in the airport is handing out questionnaires to travelers asking them to evaluate the airport's service. She does not ask travelers who are hurrying through the airport with their hands full of luggage, but instead asks all travelers who are sitting near gates and not taking naps while they wait.

Cluster Sampling

(b) A teacher wants to know if her students are doing homework, so she randomly selects rows two and five and then calls on all students in row two and all students in row five to present the solutions to homework problems to the class.

Stratified Sampling

(c) The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.
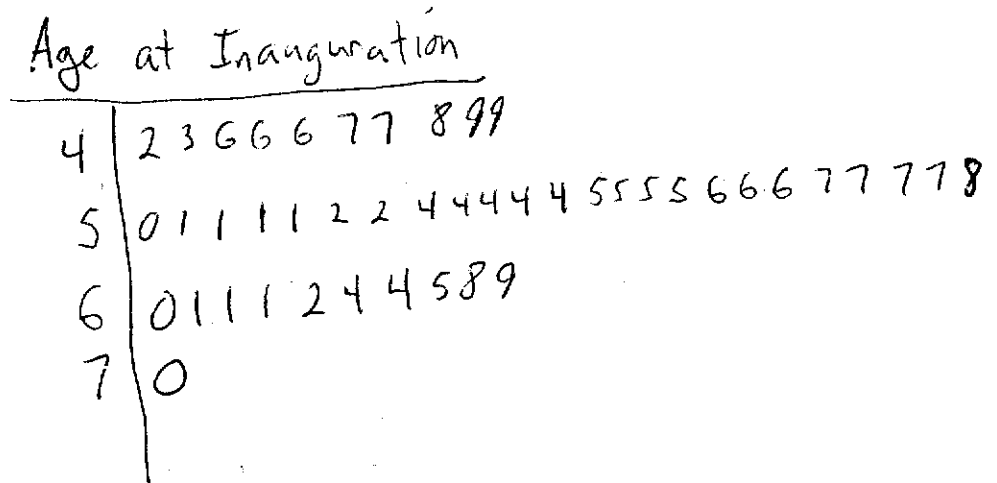
Systematic Sampling

(d) The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which she marks whether books are checked out by an adult or a child. She records this data for every fourth patron who checks out books.

Simple Random Sample

(e) A political party wants to know the reaction of voters to a debate between the candidates. The day after the debate, the party's polling staff calls 1,200 randomly selected phone numbers. If a registered voter answers the phone or is available to come to the phone, that registered voter is asked whom he or she intends to vote for and whether the debate changed his or her opinion of the candidates.

9. (3 points) The age at inauguration for each U.S. President is shown. Construct a stem and leaf plot for the data.
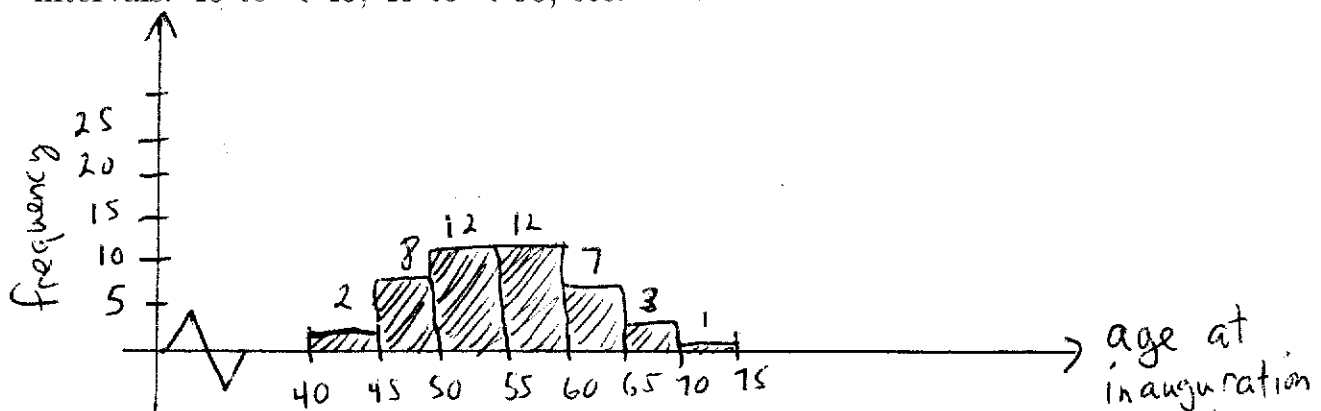57 54 52 55 51 56 61 68 56 55 54 61 57 51 46 54 51 52 57 49 54 42 60
69 58 64 49 51 62 64 57 48 50 56 43 46 61 65 47 55 55 54 46 47 70

Age at Inauguration

4	2 3 6 6 6 7 7 8 9 9
5	0 1 1 1 1 2 2 4 4 4 4 4 5 5 5 5 6 6 6 7 7 7 7 8
6	0 1 1 1 2 4 4 5 8 9
7	0

10. (3 points) Construct a relative frequency distribution using the following class intervals: 40 to < 45, 45 to < 50, etc.

Class	frequency	relative frequency
40 to < 45	2	4.44%
45 to < 50	8	17.78%
50 to < 55	12	26.67%
55 to < 60	12	26.61%
60 to < 65	7	15.56%
65 to < 70	3	6.67%
70 to < 75	1	2.22%

11. (3 points) Construct a histogram distribution using the following class intervals: 40 to < 45, 45 to < 50, etc.

12. (4 points) In advance of the 1936 Presidential Election, a magazine titled Literary Digest released the results of an opinion poll predicting that the republican candidate Alf Landon would win by a large margin. The magazine sent post cards to approximately 10,000,000 prospective voters. These prospective voters were selected from the subscription list of the magazine, from automobile registration lists, from phone lists, and from club membership lists. Approximately 2,300,000 people returned the postcards.

(a) Think about the state of the United States in 1936. Explain why a sample chosen from magazine subscription lists, automobile registration lists, phone books, and club membership lists was not representative of the population of the United States at that time.

(b) What type of bias might be present in the sample.

(c) What is the response rate of the survey?

(d) What effect does the low response rate have on the reliability of the sample?

a) The "great depression" was during the 1930s. U.S. unemployment was 25%, or 1 out of every 4 workers did not have a job. It is likely that the prospective voters that were selected had jobs and/or money/wealth. Therefore, it is likely that at least 25% of the working class was not surveyed.

b) [Selection bias] since a part of the population was excluded; AND [non-response bias] since responses were not gathered by over 7 million of those selected to receive a post card.

c) $\frac{2,300,000}{10,000,000} = \frac{23}{100} = 0.23 = \boxed{23\%}$

d) It can distort results if those who respond differ in important ways from those who do not respond.

1. (4 points) Create an example of stratified sampling.

A 50 acre farm is divided into 2 acre subplots. To determine the amount of crops that are being eaten by wildlife, a farmer randomly selects samples or crops from each subplot and analyzes the damage

2. (6 points) Do people with college degrees think that their education was worth the cost? This question was posed to 2548 adults with an associate degree and to 30,151 adults with a bachelor's degree. The data, from the Gallup report "Two-Year Grads Satisfied with Cost of Degree" (www.gallup.com, April 11, 2016), are summarized in the accompanying table. Construct a comparative bar chart for the data.

response	frequency		rel freq A.D.H	rel freq B.D.H
	associate degree holders	bachelor's degree holders		
strongly disagree (sd)	178	1508	0.06986	0.05001
disagree (d)	153	2111	0.06005	0.07001
neither agree or disagree (n)	408	4221	0.1601	0.1400
agree (a)	637	8743	0.25	0.2900
strongly agree (sa)	1172	13568	0.4600	0.4500
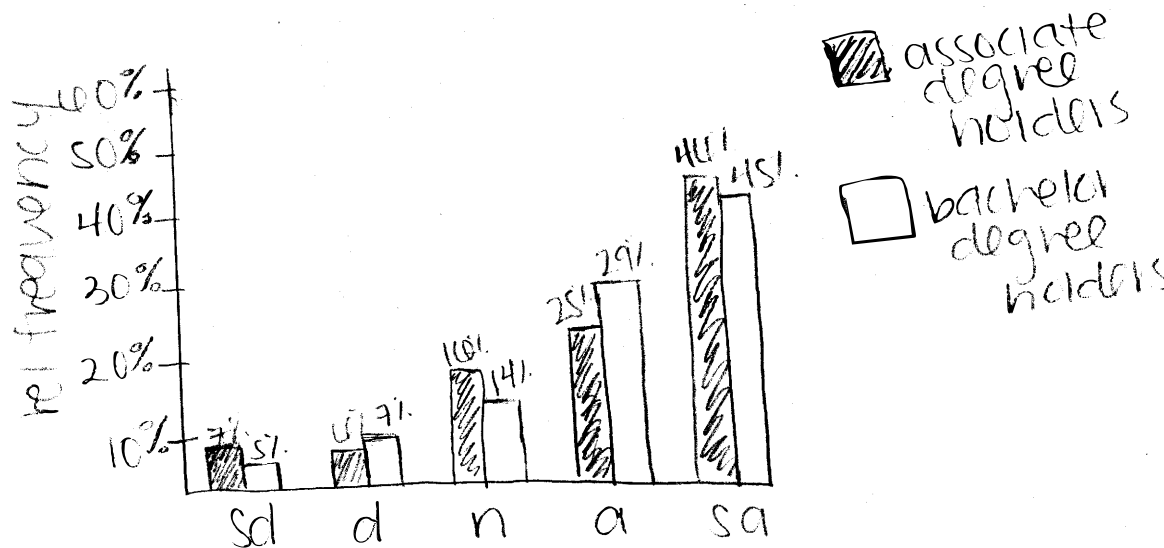
Study 1:

Question 1: The study described is an observational study.

Question 2: No, there was no random selection from a population.

Question 3: No, there was no random assignment to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk.

Question 5: No, it is not reasonable to generalize conclusions from this study to a larger population because there was no random selection from a larger population.

Study 2:

Question 1: The study described is an observational study.

Question 2: Yes, there was random selection from the population of people living in Minneapolis who receive Social Security.

Question 3: No, there was no random assignment of subjects to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk.

Question 5: Yes, it is reasonable to generalize the results of this study to the population of people living in Minneapolis who receive Social Security.

Study 3:

Question 1: The study described is an experiment.

Question 2: Yes, there was random selection from the population of people living in Minneapolis who receive Social Security.

Question 3: No, there was no random assignment of subjects to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack due to confounding and the lack of random assignment of subjects to experimental conditions. The participants in this study who did not have a previous history of heart problems were given the calcium supplement, and those with a history of heart problems were not given the supplement. It is not possible to determine the role of the calcium supplement because only those study participants who did not have a history of heart problems were given the supplement.

Question 5: It is possible to generalize the results from this study to the population of all people living in Minneapolis who receive Social Security. However, it is unclear (due to the confounding described in Question (4) what the conclusion would be.

Study 4:

Question 1: The study described is an experiment because there was random assignment of subjects to experimental conditions.

Question 2: No, there was no random selection from some larger population.

Question 3: Yes, there was random assignment of study participants to experimental groups.

Question 4: Yes, it is reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack.

Question 5: No, it is not reasonable to generalize conclusions from this study to some larger population because of the lack of random selection of study participants from a population.