Geraldo Braho  
Homework #1

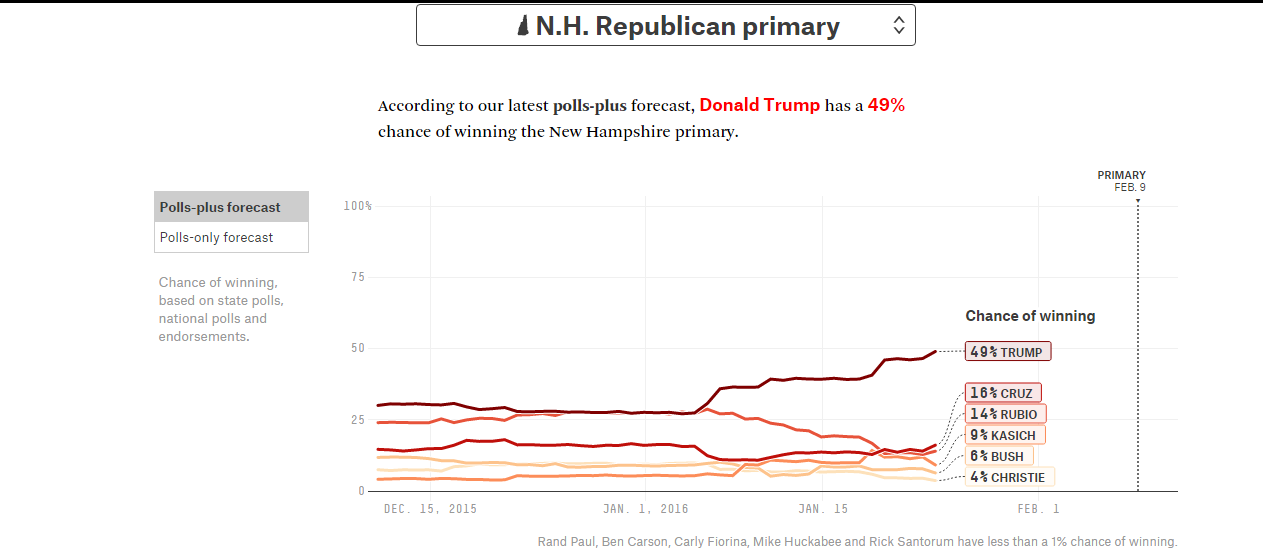
**Chapter #1 Review Problems**

1-Find an example of

a) Descriptive study

In the first three quarters since launching the iPhone 6, Apple has sold a total number of 183.17 million iPhones (iPhone 6, iPhone 6 Plus, iPhone 5s and iPhone 5c)

Reference: Apple.com

b) Inferential study   
Trump will win the election with 49 % of the votes

Reference: CNN Politics

2. Almost any inferential study involves aspects of descriptive statistics. Explain why.

In order to give out inferential study conclusions, we still need some solid data (processed firstly with descriptive statistics) about a sample chosen to represent the population

3. College Football Scores. On October 20, 2008, we obtained the following scores for week 8 of the college football season from the Sports Illustrated Web site, SI.com. Is this study descriptive or inferential? Explain your answer.   
**Big Ten Scoreboard**

Wisconsin 16, Iowa 38

Purdue 26, Northwestern 48

Ohio State 45, Michigan State 7

Michigan 17, Penn State 46

Indiana11Illinois55

The study is descriptive, because it gives us the exact data containing the teams and their scores

6. Teen Drug Abuse. In an article dated April 24, 2005, USA TODAY reported on the 17th annual study on teen drug abuse conducted by the Partnership for a Drug-Free America. According to the survey of 7300 teens, the most popular prescription drug abused by teens was Vicodin, with 18%—or about 4.3 million youths—reporting that they had used it to get high. OxyContin and drugs for attention-deficit disorder, such as Ritalin/ Adderall, followed, with 1 in 10 teens reporting that they had tried them. Answer the following questions and explain your answers.

a) Is the statement about 18% of youths abusing Vicodin inferential or descriptive?   
It is descriptive because it tell us the percent of youths

b) Is the statement about 4.3 million youths abusing Vicodin inferential or descriptive?   
It is inferential because it tell us that 4.3 millions youths abusing vicodin and conclusion made for whole teen population

7. Regarding observational studies and designed experiments:

b) With respect to possible conclusions, what important difference exists between these two types of statistical studies?  
The way observation and conclusions are made, in observational study, conclusions are made on previous data, while in designed experiment, and various experiments are designed to understand the cause, effects and to make conclusions.

Describe each type of statistical study   
Observational study- observing and getting conclusion based one what we see or collect during observation   
Designed experiment - after the flow of experiments, treatments and controls, observations and conclusions are made.

8. Persistent Poverty and IQ. An article appearing in an is- sues of the Arizona Republic reported on a study conducted by G. Duncan of the University of Michigan. According to the re- port, “Persistent poverty during the first 5 years of life leaves children with IQs 9.1 points lower at age 5 than children who suffer no poverty during that period. . . .” Is this statistical study an observational study or is it a designed experiment? Explain your answer.   
This is an observational study, because conclusion is based on the already existing data.

10. Before planning and conducting a study to obtain information, what should be done?   
 Research about what we are going to do and check if three are already existing data and works about chosen topic.

11. Explain the meaning of a:

a) Representative samples -a subset of a statistical population that accurately reflects the members of the entire population. A representative sample should be an unbiased indication of what the population is like  
b) probability sampling -- Probability sampling is a sampling technique wherein the samples are gathered in a process that gives all the individuals in the population equal chances of being selected.  
c) Simple random sampling - Simple Random Sample' A subset of a statistical population in which each member of the subset has an equal probability of being chosen. A simple random sample is meant to be an unbiased representation of a group.

12. Incomes of College Students’ Parents. A researcher wants to estimate the average income of parents of college students. To accomplish that, he surveys a sample of 250 students at Yale. Is this a representative sample? Explain your answer.   
This is not a representative sample because the researcher is taking sample just from Yale.In order to be a representative sample he should select student from other schools as well.

13. Which of the following sampling procedures involve the use of probability sampling?

a) A college student is hired to interview a sample of voters in her town. She stays on campus and interviews 100 studentsinthecafeteria.   
No, because not all people of the city had the same chance to be interviewed.

b) A pollster wants to interview 20 gas station managers in Baltimore. He posts a list of all such managers on his wall, closes his eyes, and tosses a dart at the list 20 times. He interviews the people whose names the dart hits.   
Yes, because all managers on the list had the same chance to be chosen.

14. On-Time Airlines. From USA TODAY’s Today in the Sky with Ben Mutzabaugh, we found information on the on-time performance of passenger flights arriving in the United States during June 2008. The five airlines with the highest percentage of on- time arrivals were Hawaiian Airlines (H), Pinnacle Airlines (P), SkyWest Airlines (S), Alaska Airlines (A), and Atlantic Southeast Airlines (E).

a) List the 10 possible samples (without replacement) of size 3 that can be obtained from the population of five airlines. Use the parenthetical abbreviations in your list.  
HPS, HPA, HPE, PSA, PSE, SAE, HAS, HSE, HAE, PAE

b) If a simple random sampling procedure is used to obtain a sample of three of these five airlines, what are the chances that it is the first sample on your list in part? The second sample? The tenth sample?  
All samples have the same chance, it will be 1/10 for each of them.

c) Describe three methods for obtaining a simple random sample of three of these five airlines.  
Replacement, without replacement and simple random sampling from sample and or population.

d) Use one of the methods that you described in part c to obtain a simple random sample of three of these five airlines.

Method – simple random sampling without replacement:  
 from H, P, S, A and E we choose H  
put back again H so we again have H, P, S, A ,E from here we choose P ( but we still might choose H) we put back P so again we have H, P, S, A ,E from these we choose A ( but still is possibility we choose H or P or any other letter)

Selected ones: H, P and A

**Chapter #2 Understanding the Concept and Skills**

Section1

1. Give an example, other than those presented in this section, of a  
    a) qualitative variable - Height  
    b) discrete, quantitative variable - temperature  
    c) continuous, quantitative variable - time needed to go from Houston to Dallas
2. Doctor Disciplinary Actions. The Public Citizen Health Research Group (the “group”) calculated the rate of serious disciplinary actions per 1000 doctors in each state. Using state-by- state data from the Federation of State Medical Boards (FSMB) on the number of disciplinary actions taken against doctors in 2007, combined with data from earlier FSMB reports covering 2005 and 2006, the group compiled a national report ranking state boards by the rate of serious disciplinary actions per 1000 doctors for the years 2005–2007. Following are data for the 10 states with the highest rates. Note: According to the group, “Absent any evidence that the prevalence of physicians deserving of discipline varies substantially from state to state, this variability must be considered the result of the boards’ practices.”

Identify the type of data provided by the information in the:

A) first column of the table = qualitative  
B) second column of the table = discrete, quantitative  
B) third column of the table = continuous, qualitative

13. Smartphones. Several companies conduct reviews and perform rankings of products of special interest to consumers. One such company is TopTenReviews, Inc. As of October 2008, the top 10 smartphones, according to TopTenReviews, Inc., are as shown in the second column of the following table. Identify the type of data provided by the information in each column of the table.

First column = discrete, quantitative Second column = qualitative -Third column = discrete, quantitative Fourth column = qualitative  
Fifth column = continuous, quantitative

**Section 2**

For each data set in Exercises 2.18–2.23,   
a. determine a frequency distribution.  
 b. obtain a relative-frequency distribution.   
c. draw a pie chart.   
d. construct a bar chart.

19. NCAA Wrestling Champs. From NCAA.com—the official Web site for NCAA sports—we obtained the National Collegiate Athletic Association wrestling champions for the years 1984–2008. They are displayed in the following table.

a) Freq. b) Rel.Freq. c) pie chart - angle (excel) d) bar chart (excel)

Iowa = 13 Iowa = 13/25 x 100 x 3.6 = 187.2  
Iowa St = 1 Iowa St = 1/25 x 100 x 3.6 = 14.4  
Arizona St = 1 Arizona St = 1/25 x 100 x 3.6 = 14.4  
Oklahoma St = 7 Oklahoma St = 7/25 x 100 x 3.6 = 100.8  
Minnesota = 3 Minnesota = 3/25 x 100 x 3.6 = 43.2  
 Total =13+1+1+7+3=25

23. Road Rage. The report Controlling Road Rage: A Literature Review and Pilot Study was prepared for the AAA Foundation for Traffic Safety by D. Rathbone and J. Huckabee. The authors discuss the results of a literature review and pilot study on how to prevent aggressive driving and road rage. As described in the study, road rage is criminal behavior by motorists characterized by uncontrolled anger that results in violence or threatened violence on the road. One of the goals of the study was to determine when road rage occurs most often. The days on which 69 road rage incidents occurred are presented in the following table.

a) Freq. b) Rel.Freq. c) pie chart - angle (excel) b) bar chart (excel)

M = 5 M = 5/69 x 100 x 3.6 = 26.09 Tu = 11 Tu = 11/69 x 100 x 3.6 = 57.39  
 W = 11 W = 11/69 x 100 x 3.6 = 57.39 Th = 11 Th = 11/69 x 100 x 3.6 = 57.39  
 F = 19 F = 19/69 x 100 x 3.6 = 99.13 Sa = 7 Sa = 7/69 x 100 x 3.6 = 36.52  
 Su = 5 Su = 5/69 x 100 x 3.6 = 26.09  
  
 Total 5+11+19+5+11+11+7= 69

In each of Exercises 2.24–2.29, we have presented a frequency distribution of qualitative data. For each exercise,  
 a. obtain a relative-frequency distribution.  
 b. draw a pie chart.   
c. construct a bar chart.

25. M&M Colors. Observing that the proportion of blue M&Ms in his bowl of candy appeared to be less than that of the other colors, R. Fricker, Jr., decided to compare the color distribution in randomly chosen bags of M&Ms to the theoretical distribution reported by M&M/MARS consumer affairs. Fricker published his findings in the article “The Mysterious Case of the Blue M&Ms” (Chance, Vol. 9(4), pp. 19–22). For his study, Fricker bought three bags of M&Ms from local stores and counted the number of each color. The average number of each color in the three bags was distributed as shown in the following table.

a) Rel.Freq b) Pie Chart (angle) excel c) Bar chart

152/509 x 100 x 3.6 = 107.5  
114/509 x 100 x 3.6 = 80.63  
106/509 x 100 x 3.6 = 74.97  
51/509 x 100 x 3.6 = 36.07  
43/509 x 100 x 3.6 = 30.41  
43/509 x 100 x 3.6 = 30.41

28. Hospitalization Payments. From the Florida State Center for Health Statistics report Women and Cardiovascular Dis- ease Hospitalizations, we obtained the following frequency distribution showing who paid for the hospitalization of female cardiovascular patients under 65 years of age in Florida during one year.

a) Rel.Freq b) Pie Chart (angles) in excel c) Bar Chart in excel

9983/52389 x 100 x 3.6 = 68.9 **To get the pie chart we will need to multiply 360/100=3.6**8142/52389 x 100 x 3.6 = 55.95  
26825/52389 x 100 x 3.6 = 184.33  
1777/52389 x 100 x 3.6 = 12.21  
5512/52389 x 100 x 3.6 = 37.88  
150/52389 x 100 x 3.6 = 1.03

**Section 3**

For each data set in Exercises 2.52–2.63, use the specified grouping method to

1. determine a frequency distribution.
2. obtain a relative-frequency distribution.
3. construct a frequency histogram based on your result from part (a).
4. construct a relative-frequency histogram based on your result from part (b)

53. Household Size. The U.S. Census Bureau conducts nation- wide surveys on characteristics of U.S. households and publishes the results in Current Population Reports. Following are data on the number of people per household for a sample of 40 households. Use single-value grouping.

a) Freq. b) Rel.Freq c) Histogram on part (a) excel d) Histogram on part (b) excel

**How many times is each number repeated   
(Freq) Rel.Freq  
1 - 7 1 - 7/40  
2 - 13 2 - 13/40  
3 - 9 3 - 9/40  
4 - 5 4 - 5/40  
5 --4 5 - 4/40  
6 - 1 6 - 1/40  
7 - 1 7 - 1/40**

(7+13+9+5+4+1+1=40)

54. Cottonmouth Litter Size. In the paper “The Eastern Cottonmouth (Agkistrodon piscivorus) at the Northern Edge of Its Range” (Journal of Herpetology, Vol. 29, No. 3, pp. 391–398), C. Blem and L. Blem examined the reproductive characteristics of the eastern cottonmouth, a once widely distributed snake whose numbers have decreased recently due to encroachment by humans. A simple random sample of 24 female cottonmouths in Florida yielded the following data on number of young per litter. Use single-value grouping.

a) Freq. b) Rel. Freq. c) Histogram on part (a) excel d) Histogram on part (b) excel

**Freq Rel.Freq  
 1 -1 1 - 1/24  
3 -1 3 -1/24  
4 - 3 4 - 3/24  
5 - 7 5 - 7/24  
6 - 6 6 - 6/24  
7 - 4 7 - 4/24  
8 -2 8 - 2/24**

57. Early-Onset Dementia. Dementia is a person’s loss of intellectual and social abilities that is severe enough to interfere with judgment, behavior, and daily functioning. Alzheimer’s dis- ease is the most common type of dementia. In the article “Living with Early Onset Dementia: Exploring the Experience and Developing Evidence-Based Guidelines for Practice” (Alzheimer’ s Care Quarterly, Vol. 5, Issue 2, pp. 111–122), P. Harris and J. Keady explored the experience and struggles of people diagnosed with dementia and their families. A simple random sample of 21 people with early-onset dementia gave the following data on age, in years, at diagnosis. Use limit grouping with a first class of 40–44 and a class width of 5.

a) Freq. b) Rel. Freq. c) excel d) excel

Freq Rel.Freq  
40-44 - 4 4/21  
45-49 - 3 3/21  
50-54 - 4 4/21  
55-59 - 8 8/21  
60-64 - 2 2/21

58. **C**heese Consumption. The U.S. Department of Agriculture reports in Food Consumption, Prices, and Expenditures that the average American consumed about 32 lb of cheese in 2007. Cheese consumption has increased steadily since 1960, when the average American ate only 8.3 lb of cheese annually. The following table provides last year’s cheese consumption, in pounds, for 35 randomly selected Americans. Use limit grouping with a first class of 20–22 and a class width of 3.

a) Freq. b) Rel. Freq. c) excel d) excel

**Freq Rel.Freq  
20-22 - 2 2/35  
23-25 - 3 3/25  
26-28 - 4 4/25  
29-31 - 7 7/35  
32-34 - 6 6/35  
35-37 - 5 5/35  
38-40 - 3 3/35  
41-43 - 3 3/35  
44-46 - 2 2/35**

61. Clocking the Cheetah. The cheetah (Acinonyx jubatus) is the fastest land mammal and is highly specialized to run down prey. The cheetah often exceeds speeds of 60 mph and, according to the online document “Cheetah Conservation in Southern Africa” (Trade & Environment Database (TED) Case Studies, Vol. 8, No. 2) by J. Urbaniak, the cheetah is capable of speeds up to 72 mph. The following table gives the speeds, in miles per hour, over 1/4 mile for 35 cheetahs. Use cut point grouping with 52 as the first cut point and classes of equal width 2.

a) Freq. b) Rel.Freq. c) excel d) excel

52–under 54 --> 2 2/35  
54–under 56 --> 5 5/35  
56–under 58 --> 6 6/35  
58–under 60 --> 8 8/35  
60–under 62 --> 7 7/35  
62–under 64 --> 3 3/35  
64–under 66 --> 2 2/35  
66–under 68 --> 1 1/35  
68–under 70 --> 0 -  
70–under 72 --> 0 -  
72–under 74 --> 0 -  
74–under 76 -->1 1/35

62. Fuel Tank Capacity. Consumer Reports provides information on new automobile models, including price, mileage ratings, engine size, body size, and indicators of features. A simple random sample of 35 new models yielded the following data on fuel tank capacity, in gallons. Use cut point grouping with 12 as the first cut point and classes of equal width 2.

a) Freq. b) Rel. Freq. c) Histogram on part (a) excel d) Histogram on part (b) excel

**12-under 14 - 2 2/35  
14-under 16 - 6 6/35  
16-under 18 - 7 7/35  
18-under 20 - 6 6/35  
20-under 22 - 6 6/35  
22-under 24 - 3 3/35  
24-under 26 - 3 3/35  
26-under 28 - 2 2/35**

75. Cholesterol Levels. According to the National Health and Nutrition Examination Survey, published by the Centers for Disease Control and Prevention, the average cholesterol level for children between 4 and 19 years of age is 165 mg/dL. A pediatrician who tested the cholesterol levels of several young patients was alarmed to find that many had levels higher than 200 mg/dL. The following relative-frequency histogram shows the readings for some patients who had high cholesterol levels.

a) What percentage of the patients have cholesterol levels between 205 and 209, inclusive?   
Rel.Freq. = 0.2   
 **%= Rel.Freq. x 100 = 0.2 x 100 = 20 %**

b) What percentage of the patients have levels of 215 or higher?   
Rel.Freq = 0.2 (216-219) + 0.05 ( 220-225) = 0.25   
**--> % = Rel.Freq. x 100 = 0.25 x 100 = 25%**

c) If the number of patients is 20, how many have levels between 210 and 214, inclusive?   
A/20 = 0.35 --> A = 0.35 x 20 = 7

**Section 4**

In each of Exercises 2.98–2.107, we have provided a graphical display of a data set. For each exercise,  
 a. identify the overall shape of the distribution by referring to Fig. 2.11 on page 72.  
 b. state whether the distribution is (roughly) symmetric, right skewed, or left skewed.

**104.** Adjusted Gross Incomes. The Internal Revenue Service (IRS) publishes data on adjusted gross incomes in the document Statistics of Income, Individual Income Tax Returns. The preceding relative-frequency histogram shows one year’s individual in-come tax returns for adjusted gross incomes of less than $50,000.

**a) Reverse J-shape b) Right skewed**

105. Cholesterol Levels. According to the National Health and Nutrition Examination Survey, published by the Centers for Disease Control and Prevention, the average cholesterol level for children between 4 and 19 years of age is 165 mg/dL. A pediatrician who tested the cholesterol levels of several young patients was alarmed to find that many had levels higher than 200 mg/dL. The following relative-frequency histogram shows the readings for some patients who had high cholesterol levels.

a) Left Skewed b) Left Skewed