Airbnb Descriptive Statistics Lab Worksheet (Chapter 1): Winter 2019

Names: _____   _____

Objectives: To analyze a statistical study, display data, and draw conclusions from the data.
To explain how the data appear to support the conclusions.

Part A: Type of Study, Variables & Data and Sampling Method

Researchers at Harvard Business School conducted a study to determine if there was racial discrimination by hosts (people offering rentals) on Airbnb.com. They investigated whether renters (people requesting to rent a property) were accepted or rejected more often by hosts, depending on the perceived race of the renter.

The researchers created profiles on Airbnb that were identical in every way except for name. Based on hospital birth records, 10 "distinctively white" names and 10 "distinctively African-American (black)" names were chosen. Each group had 5 male and 5 female names. Thus there were a total of 20 created profiles separated into 4 groups (black, female), (black, male), (white, female), (white male). No photos were used.

From 5 randomly selected and geographically disperse metropolitan areas, the researchers used these 20 profiles to attempt to rent 6269 similar properties listed on Airbnb. Each rental was for dates about two months in the future. They recorded whether the host accepted or rejected the guest's attempt to rent the property.

1. **Type of Study:** Is this an experiment or observational study?

2. **Variable and Data:** The variables are: "Type of Name" and "Whether the rental attempt was accepted or rejected". Another variable that the Harvard study examined but that we are not examining here is "Gender".

   Are the variables and data quantitative continuous, quantitative discrete, or qualitative? Explain why.

3. **Sampling:** In summary, the sampling method was to randomly choose 5 diverse metropolitan areas and sampling all available listings in those selected metropolitan areas that met their criteria.

   Of the sampling methods we learned about (simple random, stratified, cluster, systematic, convenience), which sampling method (or combination of methods) best fits the sampling procedure used in this study. Explain why you selected this. (If you don't think any of the listed methods fit what was done, explain why.)

Part B. Data: The data table below is called a "contingency table". The table summarizes the frequencies (counts of the number of people) for each category of request and response.

	Renter Name sounds "Black"	Renter Name sounds "White"	Total
Renter Accepted	1210	1498	2708
Renter Rejected	1900	1661	3561
Total	3110	3159	6269

We can summarize qualitative data in a table by counting the number of responses in each data category; then we can display the results as counts and/or as percents.

We can use pie or bar charts to display qualitative data graphically to make it visually easier to understand.

4. Display percent accepted or rejected for each category of name. Each column should add to 100%

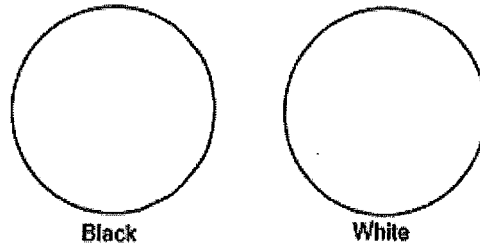|	Renter Name sounds "Black"	Renter Name sounds "White"
Renter Accepted		
Renter Rejected		
Total	100%	100%

5. Display percent for each type of name, separately for each outcome. (accepted vs rejected). Each row should add to 100%

	Renter Name sounds "Black"	Renter Name sounds "White"	Total
Renter Accepted			100%
Renter Rejected			100%

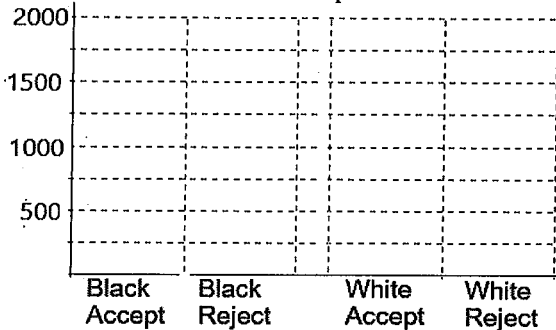*All graphs should be reasonably accurate and should be neat. Use rulers to draw straight lines.*

6. **Pie Chart:**
   Create pie charts for each type of name category. Each circle (or pie) gets divided into two sectors, one sector for accepted, one sector for rejected. The size of each sector should be related to the percent for the category.

Black   White

7. **Frequency Bar Chart:**
   Create a bar graph for each category below, using the counts in the original table. The counts are called frequencies.

8. **Relative Frequency Bar Chart:**
   Create a bar graph for each category below, using the decimal values corresponding to percents in the pie charts. The proportions are called relative frequencies.

Part C. Drawing a conclusion from the data: Type (or write neatly) your answer on a separate sheet of paper and staple to this page to hand in. Answer each question using complete sentences.

9. Do you think there was discrimination against black renters based on this study? Using the results in some of the data tables and graphs, explain specifically what you see in the data that leads you to your conclusion.

10. Which table do you think best illustrates the results of this study? Which graph(s) do you think best illustrates the results of this study? Why do you think the table and graph you chose are the best representation of the data to support your conclusion (rather than the other tables that you did not choose as best)?

Additional Info: The researchers wanted to extend the study to more metropolitan areas but Airbnb became aware of the profiles and the researchers could no longer use them. The researchers' analysis of data by gender showed another interesting result: female renters were accepted by hosts at a higher rate than male renters.

Sources: StatProcessWS, 12/1/2018, Mo Geraghty, De Anza College

Edelman, Benjamin, Michael Luca, and Daniel Svirsky. "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment" American Economic Journal: Applied Economics 9, no. 2 (Apr. 2017):1–22 https://www.hbs.edu/faculty/Pages/item.aspx?num=51647

Suggested Additional Reading: To see what actions were taken by Airbnb as a result of the Harvard University study analyzing this data, read New York Times article 9/9/2016 Airbnb Adopts Rules to Fight Discrimination by Its Hosts https://www.nytimes.com/2016/09/09/technology/airbnb-anti-discrimination-rules.html