**Math 10 MPS - Lab 3 – Cross-tabulation and Two Way Tables**

|  |  |
| --- | --- |
|  |  |
|  |  |

In 2016, researchers from Harvard Business School conducted an experiment on the online rental company Airbnb to explore if there is racial discrimination. Airbnb allows owners (hosts) to rent out all or part of their properties to guests. Renters (guests) create a profile on Airbnb and hosts are allowed to accept or reject potential renters.

The researchers created 20 profiles on Airbnb that were identical in every way except for name. Based on hospital birth records from 1974 to 1979, 10 "distinctively white" names were chosen and 10 "distinctively African-American (black)" names were chosen. Each of these groups had 5 male and 5 female names. No photos were added to these 20 profiles. From 5 randomly selected and geographically disperse metropolitan areas, the researchers using these 20 profiles attempted to rent approximately 6400 similar properties listed on Airbnb. Each rental was for dates about two months in the future.

The Minitab file contains 6269 records recreated from this study: data for name, race, gender, and whether or not the applicant was accepted by the host.

1. For the variables **Name**, **Race**, **Gender** and **Accepted**, identify the explanatory variables and the response variables.
2. What lurking variables did the researchers control for?
3. You can make a **stacked bar** 100% graph: **MINITAB>graph>bar graph** and then select the **stack** option.



Choose the three variables Race, Gender, Accepted, then click **Chart Options** and choose **Show Y as a Percent** and **Within categories at level 2**. Create the graph and paste it here. Analyze and explain what this graph is showing.

In questions 4 to 10 you will analyze the claim that the perceived **race** of the guest affects the **acceptance** rate.

1. Cross Tabulation is a method of taking pairs of categorical variables and creating a two-way table. The command can be found on the menu bar **STAT>TABLES>CROSSTABULATION**. Choose two data items and check that you want **count**, **row percents** and **column percents.** Cross tabulate the variables **Race** and **Accepted** and paste the table here.
2. You can also make a clustered bar graph **GRAPHS>BAR GRAPH>CLUSTERED.** Created a clustered bar graph of the variables **Race** and **Accepted.** Paste the graph here.
3. What is the probability that a randomly selected guest is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
4. What is the probability that a randomly selected guest is White **and** is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
5. What is the probability that a randomly selected African-American guest is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
6. What is the probability that a randomly selected White guest is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
7. Are **Race** and **Acceptance** rates independent events? Explain your answer by using both probabilities from the table and the graph.

In questions 11 to 17 you will analyze the claim that the perceived **gender** of the guest affects the **acceptance** rate.

1. Cross tabulate the variables **Gender** and **Accepted** and paste the table here.
2. You can also make a clustered bar graph **GRAPHS>BAR GRAPH>CLUSTERED.** Created a clustered bar graph of the variables **Gender** and **Accepted.** Paste the graph here.
3. What is the probability that a randomly selected guest is a woman? What type of probability is this (Marginal, Joint, or Conditional)?
4. What is the probability that a randomly selected guest is a woman **and** is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
5. What is the probability that a randomly selected woman is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
6. What is the probability that a randomly selected Man guest is accepted by the host? What type of probability is this (Marginal, Joint, or Conditional)?
7. Are **Gender** and **Acceptance** rates independent events? Explain your answer by using both probabilities form the table the table and the graph.
8. Based on this study, would you support the claim that there is discrimination by Airbnb hosts based on race and gender? Explain your answer. What changes (if any) do you think Airbnb should make to their platform?
9. Read the linked New York Times article about changes that Airbnb made due to backlash they received from the Harvard study. Do you think Airbnb has adequately addressed this issue?