**Quantitative methods 5643**

Database Analysis 1

Due: 10/7/21

(25 points)

These questions require using Stata software. Please upload your solution on D2L under the assignment section following the instructions below:

* List the homework questions, answers, any appropriate calculations, formulas plus any appropriate computer output from Stata (numbers, tables or graphs). Please embed the computer output from each program with the question and answer they belong to. Please use a fixed width font for your output so that the columns align. The Courier New font works well Stata output.
* Hand in relevant Stata commands **(your .do file)**. Include enough comments in your command files so that I can follow your work.
* Hand in relevant Stata results **(your log file)**.

1. Twenty-five college students were asked four questions about their TV and movie viewing habits:
2. Which TV service do you use most often to watch television shows and movies?
3. How many hours a week do you spend watching television or movies?
4. How often do you binge watch television (watching more than three episodes of the same show in a row)? They could choose from (a) not at all, (b) sometimes—one to three times per week, and (c) frequently—more than three times per week.
5. Gender

Based on their responses that are in the Table 1, enter these data into Stata and then complete the following tasks that will allow you to practice labeling variables, developing value labels, saving your work, and generating frequency tables.

Table 1. Student Reponses to Survey about TV and Movie Viewing Habits

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Student | TV Source | Hours per week | Binge Frequency | Gender |
| 1 | Hulu plus | 14 | Not at all | Male |
| 2 | Amazon Prime | 18 | Sometimes | Female |
| 3 | Hulu plus | 20 | Frequently | Female |
| 4 | Netflix | 5 | Frequently | Male |
| 5 | Netflix | 12 | Frequently | Male |
| 6 | iTunes | 10 | Not at all | Female |
| 7 | iTunes | 8 | Frequently | Female |
| 8 | iTunes | 7 | Sometimes | Male |
| 9 | Cable | 24 | Frequently | Male |
| 10 | Hulu plus | 30 | Sometimes | Female |
| 11 | Netflix | 10 | Sometimes | Male |
| 12 | Netflix | 15 | Frequently | Female |
| 13 | iTunes | 12 | Sometimes | Male |
| 14 | iTunes | 14 | Sometimes | Female |
| 15 | Amazon Prime | 2 | Sometimes | Female |
| 16 | Netflix | 8 | Frequently | Male |
| 17 | iTunes | 10 | Frequently | Male |
| 18 | Cable | 19 | Sometimes | Female |
| 19 | Netflix | 18 | Sometimes | Female |
| 20 | iTunes | 21 | Frequently | Male |
| 21 | iTunes | 22 | Sometimes | Male |
| 22 | iTunes | 23 | Frequently | Female |
| 23 | Netflix | 17 | Frequently | Female |
| 24 | Netflix | 14 | Sometimes | Male |
| 25 | Amazon Prime | 13 | Sometimes | Female |

1. (1 point) Open a do file and name it DataAnalysis-1.
2. (1 point) Open a log file and name it DA-1.
3. (1 point) Enter the data for each of the four variables. For the three categorical variables, create a numeric code for each response. For example, for TV source, 1 = Amazon Prime, 2 = Cable, 3 = Hulu Plus, 4 = iTunes, and 5 = Netflix.

(Instruction: open “Data Editor” from the menu on top of the results pane, and directly enter the data into the cells. You do not need to include any commands in the do file for part a).

1. (1 point) Once you have entered the data, rename each variable.
2. (1 point) Give each variable a variable label.
3. (1 point) Give each numeric code a value label.
4. (1 point) Save your data file and name it DA1-TV.
5. (1 point) What percentages of the sample identify as males and females?
6. (1 point) What type of TV service is used most frequently?
7. (1 point) What percentage of students binge watch TV frequently?
8. (1 point) Identify both the variable type and scale of measurement for each of the four variables (TV Source, Hours per week, Binge frequency, and Sex) in your dataset.
9. (1 point) What is the mod for “Binge Frequency”?
10. (1 point) What is the 25th percentile value for “Hours per week”? Explain what this means.
11. (1 point) What is the variance for “Hours per week”?
12. (1 point) Make a table that shows “Sex” in the rows and the mean and median of “Hours per week” in the columns.
13. (1 point) Generate a cross tabulation of “Sex” and “Binge frequency”. Be sure to think about whether the rows or columns should add up to 100%. Based on your table, what percentage of women binge watch frequently and what percentage of men binge watch frequently?
14. (3 point) Generate a bar chart that shows the average “Hours per week” that men and women binge watch TV. Explain what this means.
15. (3 point) Generate a histogram of “Hours per week”. Explain what this means.
16. (3 point) Generate a pie chart of “Binge frequency”. Label each slice of the chart with the percentage value. Explain what this means.