**PSYCH 613 Homework 1, Elizabeth Trinh**

1. This study is confirmatory because we are testing a hypothesis. We have an idea about the relationship between the variables (i.e., when the weather is gloomy – compared to when it is sunny – consumers are more likely to be drawn to or click on negative ads). We are conducting this study because we are trying to see if our theory and hypotheses are supported by empirical data of the Psychology Subject Pool. We have hypotheses, and we know what we are trying to measure in advance. (Note – another hint was replication. In my experience, many replication studies are more confirmatory rather than exploratory.)

2. The independent or predictor variable is weather, which is operationalized as sunny (1) or gloomy (0). This is based on ratings from a meteorologist. The level of measurement is nominal. The dependent variable is purpose points for each advertised item. Participants could rate how many points they are willing to pay for each item. The level of measurement is ratio because there is an absolutely zero that is meaningful. By choosing 0, the participant is saying that the participant would not spend any money on the advertised item.

3. This study uses a within-subjects design, which means that the same participant tests all the conditions. In this case, there are 2 conditions: sunny weather and gloomy weather. I know that this is a within-subjects design because in the weather.ads.csv, the same participant undergoes to both conditions. There are several potential threats to internal validity for within-subjects designs. For example, there could be confounding factors from environmental variables. Maybe during gloomy days, people increase the brightness of their rooms by turning on more lamps than during sunny days. Or maybe they increase the room temperature during gloomy days compared to sunny days. This may affect the dependent variable of participants allocating purchasing points. There are also time-related factors, like history (perhaps some event happened between time 1 and time 2 that affects the DV), or maturation (perhaps the participants aged between time 1 and time 2). There is also the issue of participant attribution and ordering effects that pose as potential threats to internal validity.

4. Null hypothesis: The weather does not affect consumer purchasing responses to marketing advertisements.

Alternative hypothesis. The weather does affect consumer purchasing responses to marketing advertisements.

Null hypothesis: Consumers are not more likely to buy advertised items when it is gloomy outside.

Alternative hypothesis: Consumers are more likely to buy advertised items when it is gloomy outside.

5. If we repeated this study 100 times with N=40, the distribution of the sample means will be approximately be normally distributed. We know this because according to the Central Limit Theorem, the sample variation creates a sampling distribution or a frequency distribution of the sample means, and as the sample increases, the distribution becomes normal. Further, when we do not have information about population parameters, we can use sampling to estimate. We use inferential statistics and statistical techniques to estimate with some desired degree of accuracy with a desired likelihood of being correct. In particular, there are formulas that allow us to relate the mean and standard deviation of the sample mean to the population mean and standard deviation, as the sample is drawn from the population. This has to do with the bell curve and probabilities from a normal distribution. 68% of data is within one standard deviation from the mean, 95% of the data is within two standard deviations from the mean, and 99% of the data is within three standard deviations from the mean. I would not change my answers to the previous three questions if I learned that the distribution was bimodal (assuming the sample size was more than, for example, 2). With a large enough sample size, the sampling distribution starts to look like a normal distribution.

6.

A screenshot of a computer screen

Description automatically generated

7. Overall mean (i.e., across weather): 486.0727

Overall median (i.e., across weather): 439

Overall mode (i.e., across weather): 220, 244, 369, 662

Missing cases: 11

Skewness: Based on the histogram, the dataset seems to have positive skewness (1.526921).

Kurtosis: 5.050642

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

8. Median: The median of study points for gloomy weather is higher than that of sunny weather.

Interquartile range: The minimum value of study points for gloomy weather is lower than that of sunny weather. The maximum value of study points for gloomy weather is higher than that of sunny weather. The first quartile (Q1) of study points for gloomy weather is higher than that of sunny weather. The third quartile (Q3) of study points for gloomy weather is higher than that of sunny weather. In fact, Q1 for gloomy weather is around the same as Q3 for sunny weather. The interquartile range of study points for gloomy weather is larger than that of study points for sunny weather. The interquartile range of study points for sunny weather is more compressed than that of gloomy weather.

Outliers: Whereas there do not seem to be outliers for study points for gloomy weather, there are two outliers for study points for sunny weather. One of these outliers is the maximum value for sunny weather.

A screenshot of a cell phone

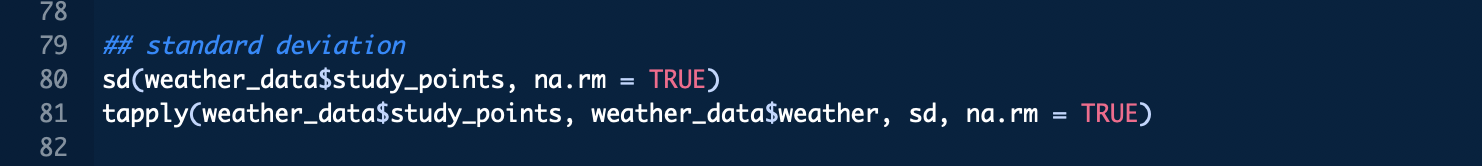
Description automatically generatedA picture containing drawing

Description automatically generated

9. Standard deviation on gloomy days: 245.5645

Standard deviation on sunny days: 167.8764

Based on only these statistics, gloomy days has a distribution of study\_points that is more spread out cross the x-axis. This is because the standard deviation on gloomy days is larger.



10. Mean: -1.094399e-16

Median: -0.1893744

Mode: -1.07041647373409, -0.97386391944812, -0.470986032542028, 0.707759734365854

Skewness: 1.526921

Kurtosis: 5.050642

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A close up of a logo

Description automatically generated

Comparison for question 10:

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated