**BAN 5733**

**Exercise 2**

**Data Set Description**

A data set with 308 random observations and 6 variables will be used for most of the questions in this exercise. The data are in an excel file called **DIAMONDS**. It contains the prices of cut diamonds, along with data on color, clarity, and ratings agency.

|  |  |
| --- | --- |
| **IDNO** | an identification number for each data point |
| **WEIGHT** | weight of the diamond, in carats |
| **COLOR** | degree of color purity in the diamond |
| **CLARITY** | diamond clarity (presence or absence of minute flaws) |
| **RATER** | the diamond was evaluated by one of three independent rating agencies |
| **PRICE** | in dollars |

1. Open JMP Pro 14 and import the Diamonds worksheet from the DIAMONDS.XLS file. **(2 points)**
   1. Using Analyze > Distribution > Generate bar charts and frequency distribution for the variables **COLOR** (horizontal bar chart) and **CLARITY** (vertical bar chart).
      1. Change the Layout to order the bars by Ascending Count.
   2. Using the Graph Builder> Generate a vertical bar chart **RATER** (X) by **PRICE** (Y)
      1. Ensure the **PRICE** is the mean price.
2. Mosaic Plot and Contingency Table (**2 points)**
3. Using Analyze > Fit Y by X generate Contingency Tab for **CLARITY** by **RATER**. The table should have the Count, Row% and Col%.
4. Generate a Mosaic plot for the variables **CLARITY** & **RATER**. **Display Percent** in the Mosaic plot.
5. The average price in the market for a diamond is $4,000.00. A jewelry store manager wants to determine if his sales are within the market range. Run the Distribution of **PRICE** and answer the following questions using this information. **(2 points)**
   1. Write out the manager’s problem in testable null and alternative hypothesis terms.
   2. Test your hypothesis with a two-sided t-test to determine if the price in the dataset is over or under the market average price.
   3. What suggestion would you give to the jewelry store manager regarding his prices according to this data?
6. The jewelry store manager is concerned that one of the three rating companies is undervaluing his diamonds. Run the appropriate test for the following comparisons to see if the manager’s concerns are valid regarding the **RATER** and **PRICE** variables: **(2 points)**
   * + GIA compared to HRD
     + GIA compared to IGI
   1. Summarize your findings for the manager and provide a suggestion supported by your results.
7. There are assumptions that must be met for most statistical tests to be valid. The one-sample t-test has the following assumptions: **(2 points)**

* The dependent variable must be continuous (interval/ratio).
* The observations are independent of one another.
* The dependent variable should be approximately normally distributed.
* The dependent variable should not contain any outliers.
  1. Does the above analysis meet these assumptions? Talk to each of the assumptions in terms of your analysis.
  2. Would your response to the store manager change based on this information?

Deliverables (please follow these instructions):

* As you complete the exercise, create a report in Microsoft Word. In this report, answer the questions in the exercise description.
* Make sure you comment or explain and not just provide snapshots of data.
* Limit your report to no more than 5 pages including tables and diagrams.
* Copy and paste or screen shot supporting tables/diagrams as needed to justify any of your answer. You may need to shrink your table/ diagrams but please ensure they are readable.
* Make sure you print your name, student ID#, student email on the cover page of the report and turn-in the report as communicated by your instructor.
* Please also put a running header/footer with your name, on each page of your exercise solution report.

Failure to follow these instructions will result in deduction of points