MGT3028 Coursework

Submission date: 10th April 2022 at 7:00pm, via Turnitin only. Late submission is not permitted.

Your entire submission should be no more than 1,500 words, excluding references, tables, and figures.

The attached data set is an Excel file with three spreadsheets (Sales data.xlsx). The Excel file contains sales, pricing and distribution figures of the different variants of a particular product for a full year.

The spreadsheets each contain different data, as described as below:

  a. Weekly Sales (Spreadsheet 1). This contains weekly sales figures for 13 different *variants* of the same product from a particular supermarket chain. There are total 52 weeks of data, that is, one full year of sales. The sales figures are in number of units sold.
  b. Unit Price (Spreadsheet 2). This contains the average unit price charged per variant per week for all 13 variants. The figures are in average prices in pence. So a 100 implies £1.
  c. Distribution (Spreadsheet 3). This contains the *percentage of stores* of the supermarket chain that stored each product variant in its shelves per week. A 100 for a particular variant for a given week meant that this supermarket listed this product on its shelves in all its stores that week.

Use this data set to answer the following questions.

You are free to use any statistical software of your choice. Whichever software you use, its name and version should be clearly indicated at the beginning of your report. All figures and tables need to be clearly labelled. Please note that some marks are allocated for visual clarity and ease of interpretation of the tables and figures.

Question 1 (20)

  a. Provide a visual representation of the volume of sales for all variants across all weeks. Also provide the summary statistic of the sales volume of each of the variants. The summary statistics should contain a measures of representative sales and measures of spread. (10)
     Hint: Line charts with sales trajectories of all product variants should be presented separately. The summary statistics should provide the mean, median, standard deviation, min and max of the sales values for each of the 13 variants.
  b. Identify the top 4 selling variants among the 13 in the data. Explain your answer and illustrate your answer using a pie chart. (10)
     .

Question 2 (40)

  a. Provide a correlation table indicating overall relationships between the various prices. (10)
  b. Can you identify those variants, whose prices match each other relatively closely. Explain using the correlation table. Please propose methods for detecting and solving multicollinearity (10)
  c. Conduct an exploratory factor analysis of distribution variants and generate an aggregated index. Please present results in tables. (20)

Question 3 (40)

a. Using the multivariate regression methodology, can you identify which prices directly affect the sales of Variant 2? (20)

b. Interpret the regression results and discuss the model explanation power (20)