MKT 317 Module 3 Lab

Name(s): \_Amit Grewal\_

# Instructions

You may complete this lab independently or work with classmates. There will be time to work on this lab with classmates in breakout rooms in your section’s Zoom meeting.

* If you work with classmates on the lab who are registered in the same section as you, then you may choose to either each upload your own copy of the lab or to upload one completed copy of the lab for the entire group. If you submit your lab as a group, it is important that every group member’s full name appears on the lab.
* If you work with classmates in different sections, then please each upload your own completed copy of the lab.
* If you primarily work independently on this lab, please upload your own completed copy of the lab.

You must upload your responses to the appropriate Assignment dropbox in D2L in either **MS Word format** or **.pdf format**. We can not accept .pages files or links to Google documents.

**Due date:** Sunday, February 7, at 11:50 PM Eastern Time.

# Data

In this lab, please use the **SuperstoreExtract4** data set that appears in Top Hat. These data were generated from a large data set that is originally provided by Tableau, and records purchasing information of customers in a very large store.

# Lab Questions

**We can analyze this data set using either Tableau, RStudio, or both.**

These data set contains two variables about customers at a large store: UnitPrice and Sales. The UnitPrice is the average unit price of items purchased by each customer, and the Sales is the sum of sales for each customer.

* Note that if a customer has a low value of UnitPrice but a very high value for Sales, it indicates that this customer purchases large amounts of relatively cheap items.
* Customers who have high values for UnitPrice and low Sales purchased a very small number of high-price items.

Suppose we would like to create a model that predicts a Sales based on UnitPrice (i.e. **Y = Sales**, and **X = UnitPrice**)

Sales is the Y-variable (dependent variable, graphed on y-axis) and

UnitPrice is the X-variable (independent variable, graphed on x-axis)

1. We have learned about 4 types of models: linear, exponential, logarithmic, and power.

Create four scatterplots (using combinations of logs and variables without taking a log) to determine which of these four models is the most appropriate for the data.

Include screenshots of your four scatter plots in the appropriate places below. You may use either R or Tableau to create these plots.

* The scatter plot that we use to determine if a linear model is appropriate a model the relationship between UnitPrice and Sales:

[paste your screenshot here]

The linear plot of the linear model shows a **poor fit** between the model and the data

Graphical user interface, chart

Description automatically generated

* The scatter plot that we use to determine if an exponential model is appropriate a model the relationship between UnitPrice and Sales

[paste your screenshot here]

The linear plot of the exponential model shows a **poor fit** between the model and the data

Graphical user interface, chart

Description automatically generated

* The scatter plot that we use to determine if a logarithmic model is appropriate a model the relationship between UnitPrice and Sales

[paste your screenshot here]

The linear plot of the logarithmic model shows a **poor fit** between the model and the data

Graphical user interface

Description automatically generated

* The scatter plot that we use to determine if a power model is appropriate a model the relationship between UnitPrice and Sales

[paste your screenshot here]

The linear plot of the power law model shows a **good fit** between the model and the data

Chart, scatter chart

Description automatically generated

1. Which of the four model types is the most appropriate to model the relationship between UnitPrice and Sales? (*select either linear, logarithmic, exponential, or* **power**)

*Note that in some data sets, there can be more than one correct answer. For the variables in this data set, only one of the four options is reasonably accurate.*

The power law model is the most appropriate to model the UnitPrice and Sales relationship. The linear plot of the power law model [linear plot with log(y) and log(x) scales] is the only one with a clear fit between the model’s equation and the data, as proven by the red lines.

1. In question 2, you determined which was the best type of model: linear, logarithmic, exponential, or power.

Create that model and do any necessary algebra to solve for the Y-variable, Sales. *Hint: the correct answer to question 2 was not linear. You need to have a reasonable answer to question 2 in order to receive credit for question 3.*

Your answer should be an equation:

16.108 \* X1.196115

Power Law model: = a \* Xb, or = eb0 \* Xb1

b0 = 2.779313, so eb0 = 16.108; b1 = 1.196115

1. For this question, you will either complete part (a), or part (b), but not both.

Depending on your model type, the change in UnitPrice (i.e. the “change in X”) might mean a dollar amount change or it might mean a percentage change.

**For the “change in Y” that you calculate, include units** (i.e. either include a dollar sign if it’s a dollar amount increase, or a % sign if it is a percentage increase).

1. ~~If the model you created in question 3 has an interpretation where the “change in X” is a dollar amount change in unit price, then complete the interpretation below.~~

*~~If the model you created in question 3 has an interpretation where the “change in X” is a percentage change in unit price, then omit or delete part a.~~*

~~For every additional $500 in UnitPrice the predicted average sales increases \_\_\_\_\_\_ .~~

1. If the model you created in question 3 has an interpretation where the “change in X” is a percentage change in unit price, then complete the interpretation below.

*If the model you created in question 3 has an interpretation where the “change in X” is a dollar amount change in unit price, then omit or delete part b.*

For every 50% increase in UnitPrice, the predicted sales increases **62.4%**

(1 + X)b1 = (1 + ), so = (1 + 0.5)1.196115 1 = 1.624 1 = 62.4% increase