By

Joel Fernandes

Daniel Tsai

Jordan Mullen

DAta Analytics Bootcamp

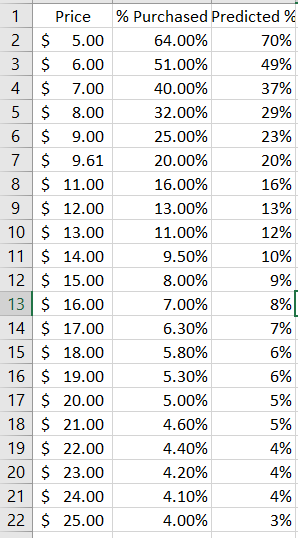
Optimized pricing strategy

***Q1. Regression Analysis***

***1a) This is a graph of percent purchased against price***

***Percent purchased- Y axis***

***Price- X axis***



b. Perform a regression using power regression to determine the predicted % column for each price

***b)***

* + - 1. ***Graph the new curve (5%)***
      2. ***Estimate the equation of the line (5%)***
      3. ***What does the R2 mean? (5%)***

**Equation:** y=14.103x-1.874

What does the R2 mean? (5%)

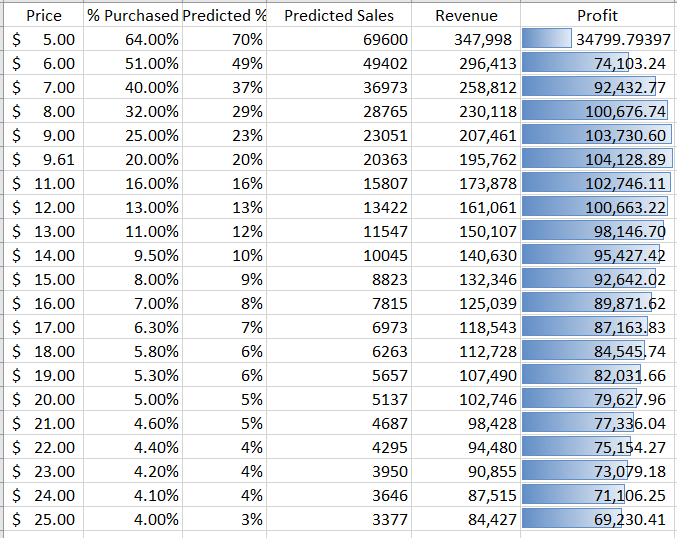
R2 is the statistical measure of how close the data is fit to the regression line. It is also called coefficient of determination. It is the percentage of response variable variation that is explained by a model.

* 0 percent indicated none of the variability of response data
* 100 percent indicates all the variability of response data

In this model R squared is 0.9918 which means the power regression curve can explain 99.18% of the predicted values. In other words, the predicted data points fits the equation at about 99.18%.

***c. Assuming there are 100,000 customers who visit your website and the publisher cost is $4.50, estimate the number of books sold (predicted sales column) for each price point (5%)***

* 1. ***Calculate the revenue column (price \* predicted sales) (5%)***
  2. ***Calculate the profit column ((price – book cost) \* predicted sales) (5%)***
  3. ***Use conditional formatting to highlight the profit values for all prices (5%)***



***2)Optimization analysis (with constraints) (35%)***

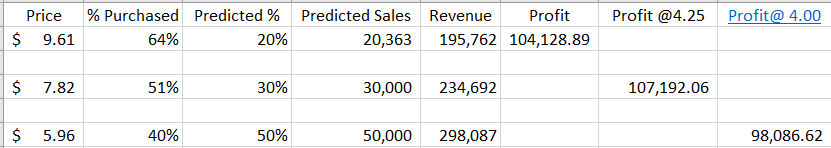
* 1. ***Calculate the price point for the highest profit possible***
     1. ***The publisher will sell the books to you at $4.50 each with no minimum order (10%)***
     2. ***The publisher has agreed to sell you the books at $4.25 each if you sell at least 30,000 (10%)***
     3. ***The publisher has agreed to sell you the books at $4.00 each if you sell at least 50,000 (10%)***

To solve these questions, we prepare a table for calculating Predicted Sales and profit by entering different values of price. The formulas for profit and sales are taken from Q.1.

So Predicted Sales = [14.344\*(Price^-1.88)]\*100000 and Profit = (Price – Book cost)\*Predicted Sales.

We then use solver to get the price point for the highest profit possible.

i. Highest profit cost point when the publisher will sell the books to you at $4.50 each with no minimum order - $10.00. By setting profit as the objective and price as the variable in solver without any constraint, we get the price point as $9.61 and the highest profit possible as $1,04,128.9.



***b. Which cost point (from the publisher) and price (to your customer) maximizes your profit?***

Taking the lowest publishers book cost of $4, using solver without any constraint we get the highest profit of $115500.92 at a price of $8.54.

Using solver with constraint that price is an integer, we get the highest profit of 115256.22 at a price of $9. We can see the same from the table above.

From the table, we can see can see that the highest profit of $115,256.22 is earned at a price of $9. We can see this in the last column of the above table which shows different profit points for different prices when the publisher charges a book cost of $4.

***3. Discussion***

***a. What are the risks of using Harry Potter 7 data in predicting your new demand curve for the Harry Potter sequel?***

The risks of using Harry Potter 7 data to predict the demand curve for Harry Potter 8 is that the model will omit the capture of essential variables. The main risk of using Harry Potter 7 data is the time in which it was sold. Since Harry Potter 7 was sold in 2006, and Harry Potter 8 was sold in 2009, the model will fail to account for the financial crisis of 2008. Hence the demand curve will predict upward(overpredict).

The next factor is the accuracy of regression analysis. Since we only have the sales figures of 2007 data, relying entirely on regression analysis is not a good idea because regression analysis becomes more accurate with time. Since we only have 2007 data it would be risky to predict the new demand curve. A better regression curve could be obtained by including historical data from previous books(Harry Potter 1-6) to obtain a more accurate demand curve.

The risk of using the sales data of a previous book is the problem of establishing a cause and effect relationship. The sales of Harry Potter 7 does not necessarily indicate that Harry potter 8 could be as well received. Neither does it indicate a positive or negative relationship which could disprove existing theories. We are also unable to capture time variant conditions that are relevant to the demand of the book.

***b. What other data would you like to have to perform your analysis?***

Relevant data that we’d like to have while performing analysis is- Information of the prevailing market conditions. The specific financial conditions of relevant areas of the book sold and the demand for each of these geographic locations would help understand the consumer demand of each of these locations. A list of customers pre orders would provide an increased insight into the possible demand for the book. Marketing efforts for the book, could enable us to gauge the consumer behavior better in estimating demand. The most relevant data that we feel would help us perform a better analysis is the demand data for all the Harry Potter books and data about prevailing market conditions during launch would enable us to make an accurate estimation of the demand for the 8th book.