Demand modeling and Price Optimization for Book Sales

Team 2:

BB, NK, RS, Jennifer Mead

**Table of Contents**

[Introduction 2](#_30j0zll)

[1. Regression analysis 3](#_1fob9te)

[a. Graph the percent purchased against price 3](#_3znysh7)

[b. Perform a regression using power regression to determine the predicted % column. 3](#_tyjcwt)

[c. Assuming there are 100,000 customers who visit your website and the publisher cost is $5.00, estimate the number of books sold 6](#_2s8eyo1)

[d. Calculate the revenue column 6](#_17dp8vu)

[e. Calculate the profit column 6](#_3rdcrjn)

[f. Use conditional formatting to highlight the profit values for all prices 6](#_26in1rg)

[2. Optimization analysis (with constraints) 8](#_lnxbz9)

[a. Calculate the price point for the highest profit possible 8](#_35nkun2)

[b. Run a constrained optimization for each of the above situations to determine which cost point (from the publisher) and price (to your customer) maximizes your profit. Which cost point should you accept from the publisher? 9](#_1ksv4uv)

[3. Discussion 9](#_44sinio)

[a. What are the risks of using Harry Potter 7 data in predicting your new demand curve for the Harry Potter sequel? 9](#_2jxsxqh)

[b. What other data would you like to have to perform your analysis? 10](#_z337ya)

# Introduction

Background

The Book Emporium wants to price books to optimize profits. The spreadsheet for this homework has sales data on Harry Potter book 7. For each week, the Book Emporium varied prices on Harry Potter 7 to determine a demand curve. The percent of customers who visited BookEmporium.com and purchased Harry Potter book 7 is shown in the spreadsheet. J.K. Rowling has announced a sequel to the Harry Potter series. Determine the price for the sequel.

Assumptions

1. Assume that the demand for the book sequel will be similar to Harry Potter 7.

2. Assume that 100,000 customers will consider purchasing a book from you.

3. The data is not an entirely accurate prediction of the demand, but a regression on the data using a power model will give a reasonable prediction.

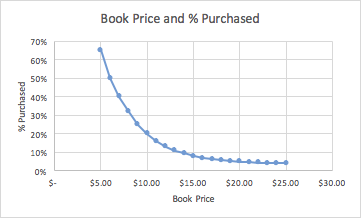
4. Assume that you pay the publisher $5.00 for each book.

Definitions

* **Price** what you will charge each customer who purchases the new book
* **Book Cost** what you must pay the publisher for each book
* **Purchased %** in your pricing test, the percent of people who bought at that price
* **Predicted %** your regression model estimate of the percent sold based on price
* **Predicted sales** estimate of number of customers who buy the book from you
* **Revenue** total revenue generated (price \* predicted sales)
* **Profit** (price – book cost) \* predicted sales

# 1. Regression analysis

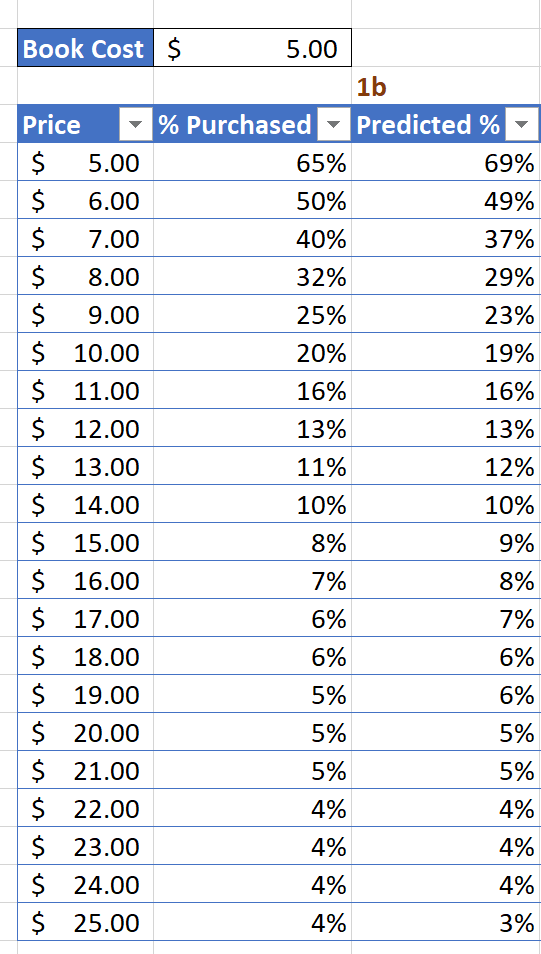
## a. Graph the percent purchased against price



The percent purchased versus price forms an approximate power curve. This is consistent with our expectations of supply and demand. When the price is low, the demand is high. When the price goes up, the demand goes down.

## b. Perform a regression using power regression to determine the predicted % column.

*i. Graph the new curve*



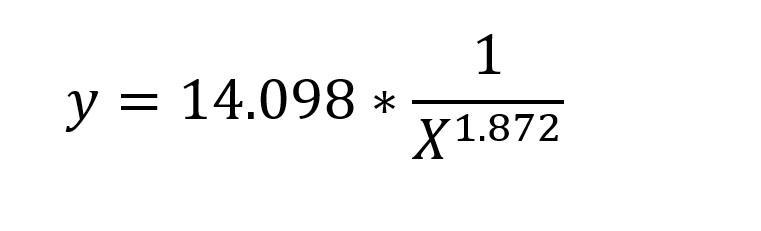
The new curve shows that our R squared is 0.99075. It also gives us our equation to predict sales: y = 14.098x ^ (-1.872).



Note that the data for the % purchased and the % predicted are very, very similar.

*ii. Estimate the equation of the line*

The equation for the power regression is as shown below:



where “y” is the projected “percent sold based on price” and “x” is the selling book price.

*iii. What does the R2 mean?*

In linear models, R-squared is the amount of variability in the percentage of predicted purchases that are explained by changes in the price; in this case 99%.

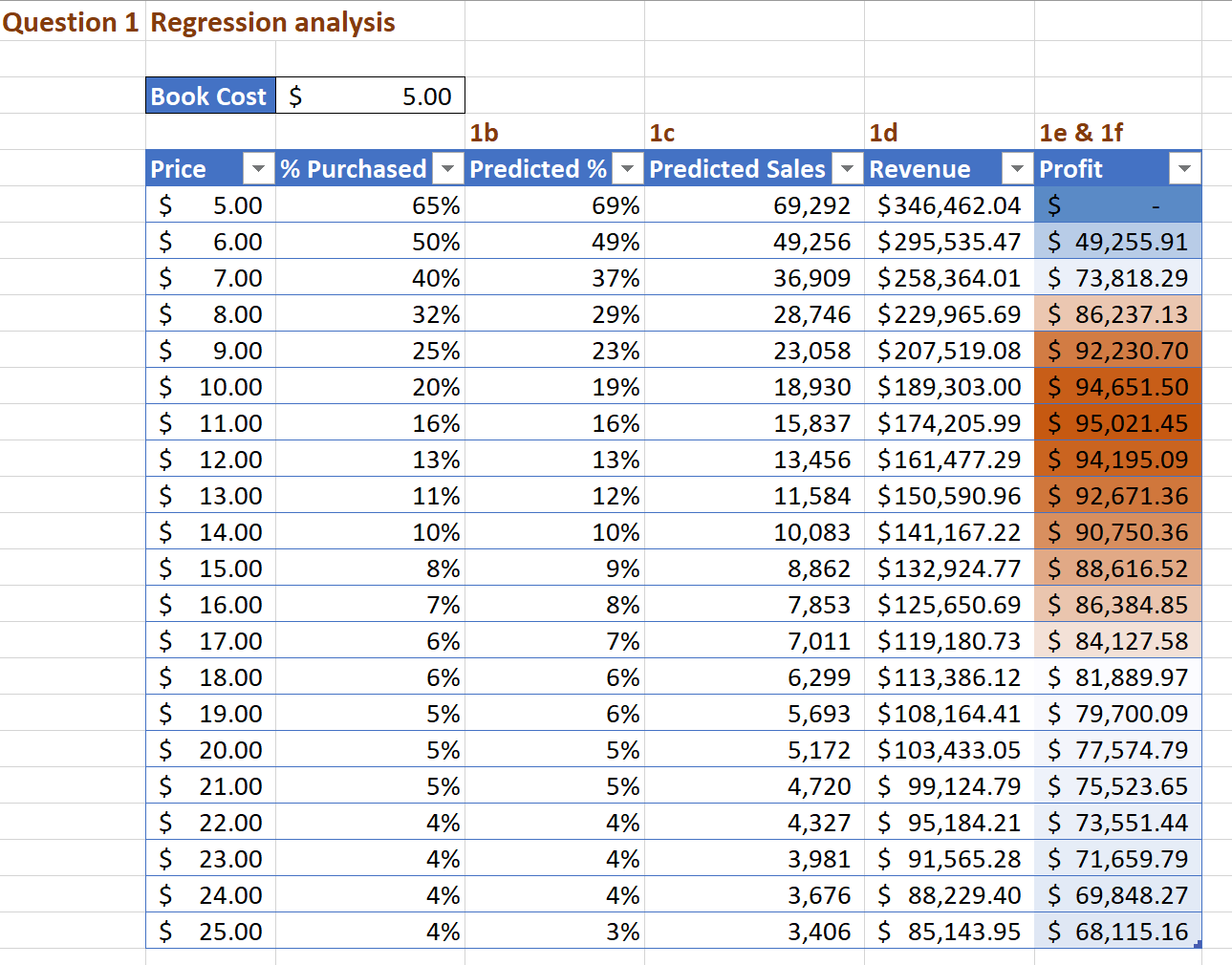
The translation is the following: as the selling book price increases the demand for higher priced books decreases significantly. In other words, bigger price changes are proportionately more important than small changes. If we decide to make the selling book price equal to $25, it would translate to much smaller demand than the price set at $5. We would need to think twice before drastically changing our pricing strategies. (A small change would be less noticeable in the total amount of revenue, while tripling book price would become a deal-breaker for the bottom line).

However, we are using a power model here, not a linear one and some sites (such as <http://blog.minitab.com>) recommend against using R-squared in this situation. We may want to take this in consideration if we do not get clear results in the analysis.

|  |  |
| --- | --- |
| Question | Answer |
| c. Assuming there are 100,000 customers who visit your website and the publisher cost is $5.00, estimate the number of books sold (predicted sales column) | Depending on the sale price, our results show that books sold could be between approximately 3,406 and 69,292. Please see chart below. |
| d. Calculate the revenue column (price \* predicted sales) | Depending on the sale price, our results show that revenue could be between approximately $85,143 and $346,462. Please see chart below. |
| e. Calculate the profit column ((price – book cost) \* predicted sales) | Depending on the sales price, our results show that profit could be between $0 and $95,021.45. Please see chart below. |

## f. Use conditional formatting to highlight the profit values for all prices

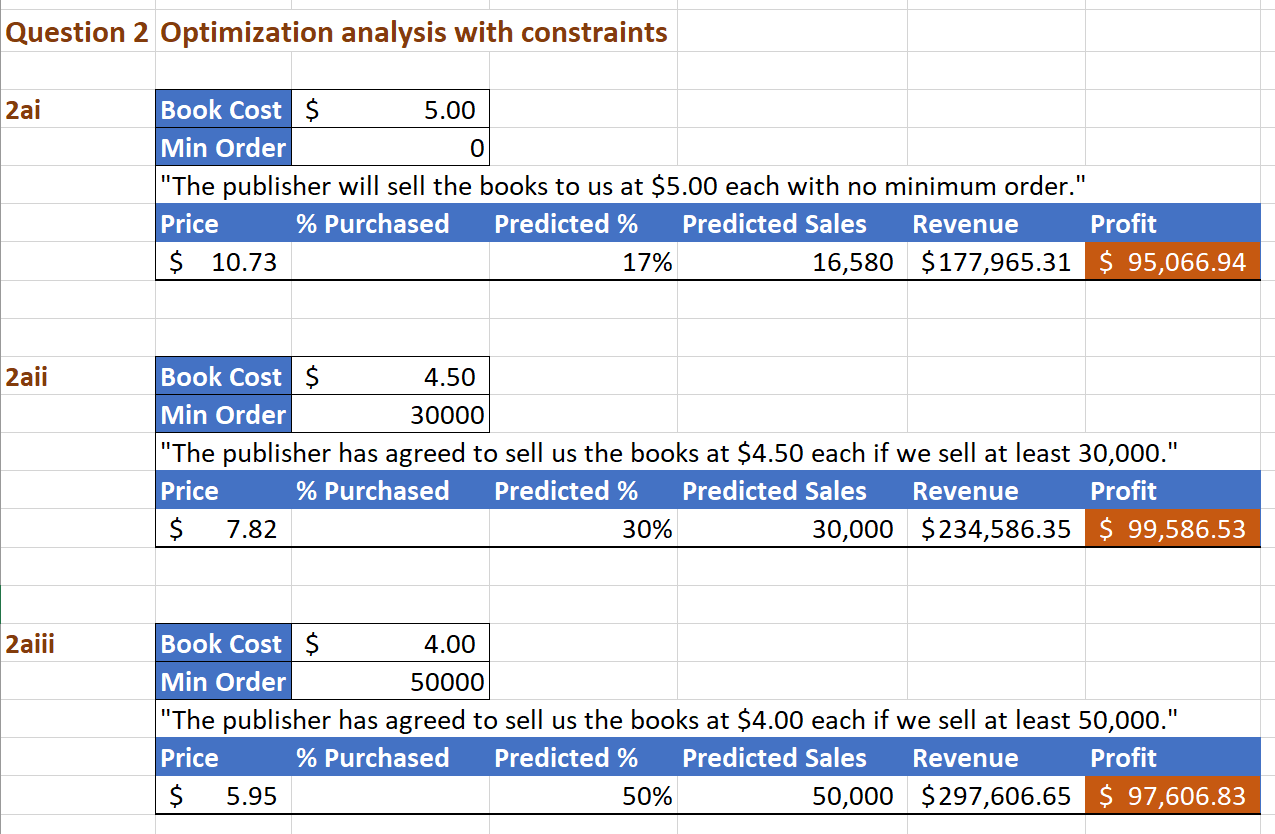
The highest profit is $95K with the selling book price of $11 based on the predicted sales by the power model. Other profits, selling price and predicted sales are shown below.



# 2. Optimization analysis (with constraints)

## a. Calculate the price point for the highest profit possible

|  |  |
| --- | --- |
| Question | Answer |
| i. The publisher will sell the books to you at $5.00 each with no minimum order (10%) | Highest possible profit is $95,066.94. See below |
| ii. The publisher has agreed to sell you the books at $4.50 each if you sell at least 30,000 (10%) | Highest possible profit is $99,586.53. See below |
| iii. The publisher has agreed to sell you the books at $4.00 each if you sell at least 50,000 (10%) | Highest possible profit is $97,606.83. See below |



## b. Run a constrained optimization for each of the above situations to determine which cost point (from the publisher) and price (to your customer) maximizes your profit. Which cost point should you accept from the publisher?

The best results for profit are $99,586.53, which we get in the second option (Option 2) where books cost $4.50 and we make a minimum order of 30,000. If we are 99.9% confident that we can sell this number of books, based on the data we have, then we would select this option.

The selling book price in our second choice is also $7.82 which is in between no minimum order (Option 1) selling book price of $10.73 and a minimum order quantity of 50,000 (Option 3) with the selling book price of $5.95. If we are uncertain about our ability to sell 30,000 we would eliminate Option 3 from our selection process as it requires us to sell even more books (50,000). We potentially could select Option 1 as an alternative because the profit difference is immaterial ($4,500) while we are not constrained by a minimum order.

# 3. Discussion

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

There are numerous risks in using the Harry Potter 7 data to predict the sequel sales as we have limited data. The future demand might not be predicted accurately using the past data and a small sample size of 21 observations. We don’t know the time frame when the sample was collected. During a holiday season (Christmas time) we would expect to see a higher demand for books. Hence, our predictions of the demand and profit might be overstated if we plan to sell during off season.

The Harry Potter sequel release data is also not given and does not specify if another rival book of the same genre may be releasing before or at the same time which may reduce sales. The sequel might be something other than a novel (like a script).

The demographics and customer base might have changed since the data have been collected.

Another risk is not knowing the effect of the adoption of electronic book-readers since the original books were published. Are paper books less popular now? More children have been reached reading age since the original novels were published; does that mean that the fan base has increased or decreased since some readers may have outgrown the genre of Harry Potter?

As a result, our predictions of the demand and the profit might be overstated. Hence, we might end up with lots of books which we were not able to sell but for which we paid to the publisher. If we blindly take the data and assume the future will repeat its past and every customer from the same location will make a purchase, we might not be optimizing properly.

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

*Additional data would include:*

* The data (surveys) on overall interest in reading (increasing or decreasing)
* The data on interest in fiction (increasing or decreasing)
* Previous book reviews/ratings which would help to predict the reception of the sequel
* Information from Bloomsbury (UK publisher) and Scholastic (US publisher) on other planned Harry Potter materials
* The information on when the data was collected (right after the Harry Potter book 7 release or it was a lagged data by a few months since the book release)
* Sales figures based on demographics (age) and regional sales. What was the average age of readers and how old they are now (do they still have interest in the book series)
* Sales numbers for young-adult fantasy in the last twenty years
* The projected time frame when we are planning to begin selling books
* Was there any advertising campaign prior to the book 7 release and when the data was collected?
* How much entertainment and advertisement influence future readers (amusement parks and Harry Potter merchandise)?
* The country/region the data was collected – UK or US or another region
* The language the book was published – English or translated version
* Number of sales of books 1-7 in the years since book 7 was published. Run a regression and predict demand similar to Book 7 to see how accurate the model is to predict future sales.
* Sales figures for Harry Potter movies as this can help gauge the interest in the book series so far
* Information about the adoption of e-readers and the impact on paper book sales.