

# Analyzing Educative Course Data

BY CHARLOTTE ESI ABRAHAM

---

# Contents

Project Description	1
Introduction	1
Project Deliverables	1
Framing the problem	2
Problem Statement	2
Problem Domain	2
Data Design	3
Data Merging	3
Data Cleaning	3
Data Summary	3
Data Visualization	3
Analysis	5
PowerBI Dashboard	5
Root Cause Analysis of Number of Subscribers by Subject	10
Root Cause Analysis of Average Rating of Courses by Subject	12
Key Performance Influencers for Pricing and Number of Subscribers	14
What influences the average price to increase?	14
What influences the number of subscribers to decrease?	16
When is the number of subscribers more likely to be low?	17
When is the number of subscribers more likely to be high?	18
Insights	19
Recommendation	20

# Project Description

---

## Introduction

The project aims to analyze the course data on the Educative platform and generate insights on how to maximize revenue and also track the performance of courses and identify key influencers of some metrics. Using Excel, the data was cleaned and pivoted tables were created to generate a summary of the insights of the data. The cleaned data was imported into PowerBI to visualize insights and key influencers. Based on the insights, recommendations were given on how the Educative management can maximize revenue.

## Project Deliverables

The scope of this project is limited to

- Analyzing the data from the company in line with Cross Industry Standard for Data Mining (CRISP-DM) and generate a report on insights and recommendations.
- Create an interactive dashboard in Excel showing your visuals and metrics.
- Create an interactive Dashboard with Power BI in addition to Excel the dashboard.

### Framing the problem

---

#### Problem Statement

The management at Educative needs to increase their course revenue by identifying different opportunities using data on their courses to increase revenue and tracking the performance of courses in the next three weeks.

#### Problem Domain

These questions were asked to further understand the data and serve as a guide when making visuals to generate insights from the course data.

- What is the number of subscribers for each subject?
- What are the top 10 subscribed courses?
- What subject has the highest and lowest rating?
- What subject generated the highest and lowest revenue?
- What factors affect the pricing of courses?
- What factors affect the number of subscribers to a subject?

# Data Design

---

## Data Merging

The folder containing the dataset in csv format was combined and loaded into excel as one csv file for easy manipulation of the dataset.

## Data Cleaning

The following steps were taken to ensure that the data was clean before analysis.

- Duplicated rows found in the dataset were removed
- Blank cells from the dataset were removed
- The various column headers were renamed using a uniform convention
- Development had to be changed to Web Development because they were inconsistent with other rows.

## Data Summary

Pivot tables and pivot tables were created to gain a summary of the insights according to the following metrics.

- The number of subscribers by subject
- The average price of courses by level
- The average number of subscribers by subject
- The average duration of courses by subject
- Average rating of courses by subjects and level

## Data Visualization

- An interactive dashboard was created with excel using the metrics described for creating the pivot chart.

- A second interactive dashboard was created using PowerBI to visualize the following domains.

1. The summary of the insights generated from the course data
2. Analyzing the root problem for the number of subscribers.
3. Analyzing the root problem for the average rating.

## Analysis

### PowerBI Dashboard

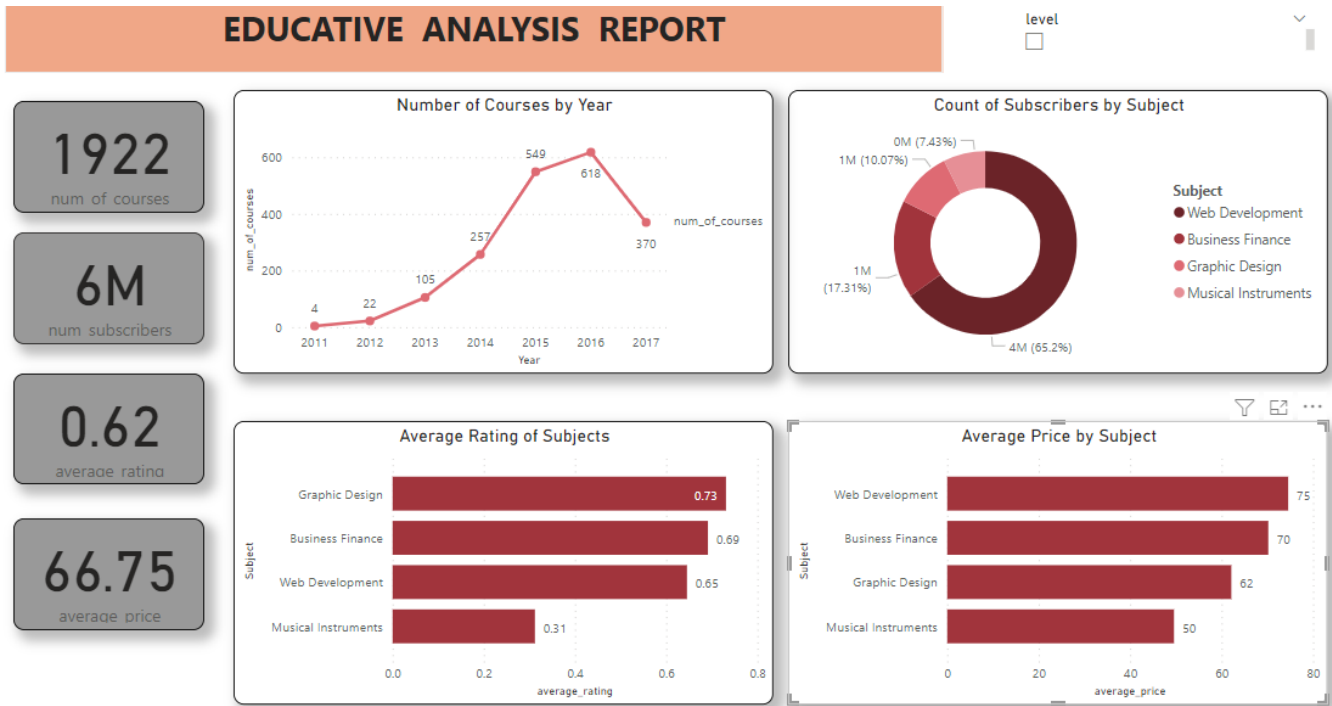


Figure 1: A dashboard showing the number of courses, the number of subscribers, average rating and price of subjects

## 3.1 Number of Courses by Year

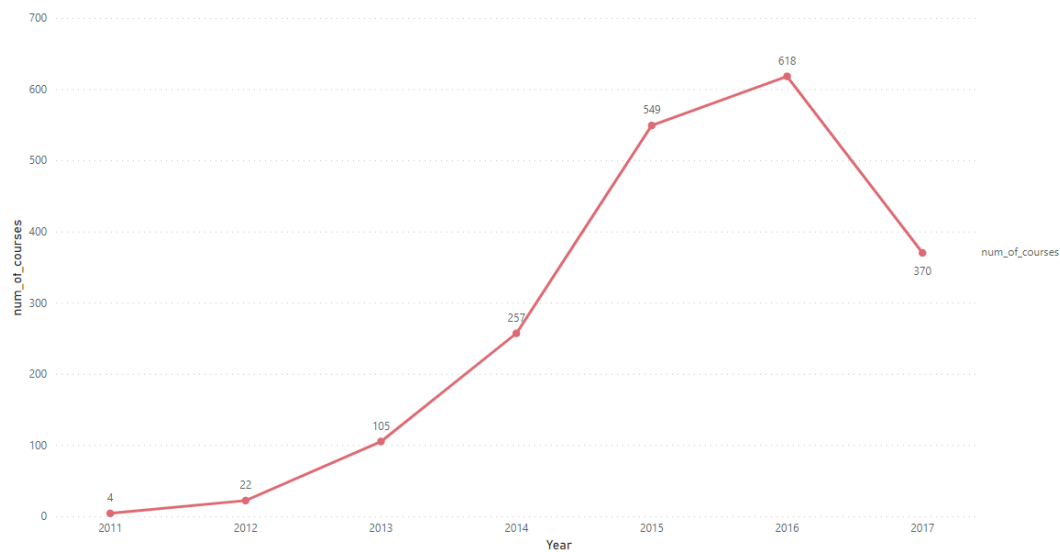


Figure 2: Line plot of the number of courses by year

From the figure above, the number of courses started trending up on 2013. The number of courses increased from 105 to 618 during its steepest incline between 2013 and 2016 and observed a decrease in. The number of courses published experienced a peak value in the year 2016.



## 3.2 Count of Subscribers by Subjects

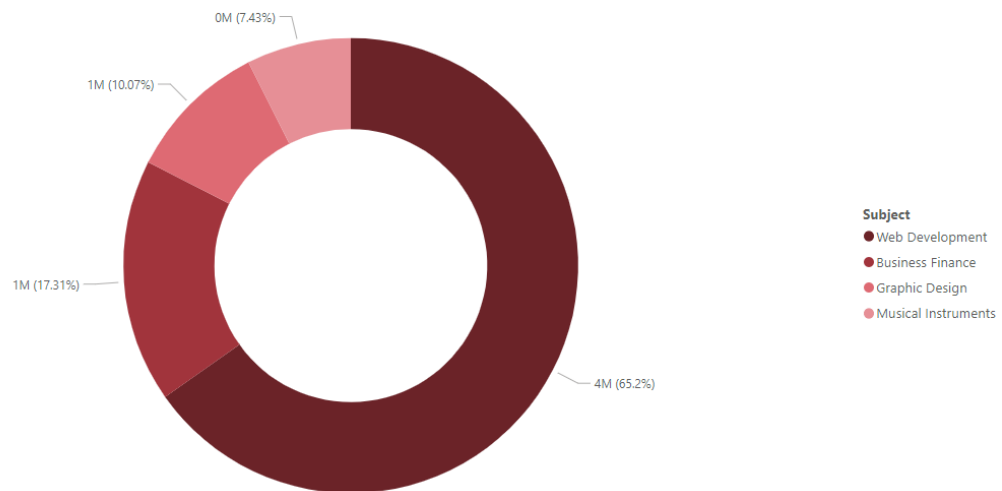


Figure 3: A donut chart showing the count of subscribers by subjects

From the figure above, Web Development had the highest number of subscribers at 4081578, followed by Business Finance, Graphic Design, and Musical Instruments. Web Development accounted for 65.20% of the number of subscribers while Musical Instruments accounted for 7.4% of the subscribers.

### 3.3 Average Rating of Subjects

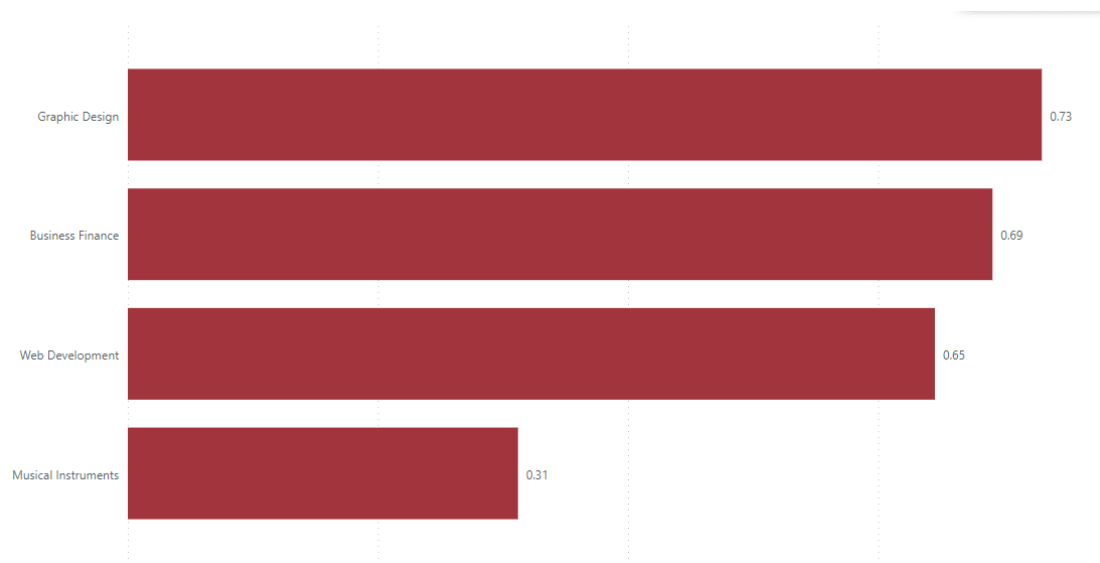


Figure 4: Bar chart showing the average rating of courses by subject

At 0.73, Graphic Design had the highest average rating and whiles Musical Instruments had the lowest average rating at 0.31. Graphic Design had the highest average rating at 0.73, followed by Business Finance, Web Development, and Musical Instruments. Across all four subjects, the average rating of subjects ranged from 0.31 to 0.73.

## 3.4 Average Price by Subject

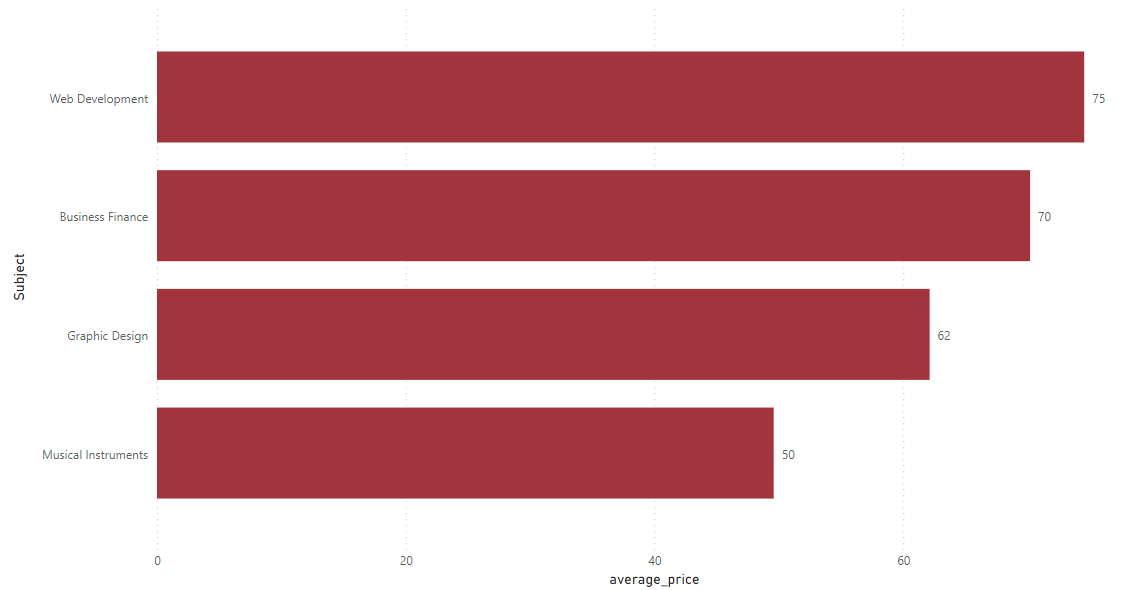


Figure 5: A bar chart of the average rating by subject

At 74.55 dollars Web Development had the highest average price Musical Instruments had the lowest average price at 49.58. Web Development had the highest average price at 74.55, followed by Business Finance, Graphic Design, and Musical Instruments. Across all four subjects, average price ranged from 49.58 to 74.55.

## Root Cause Analysis of Number of Subscribers by Subject

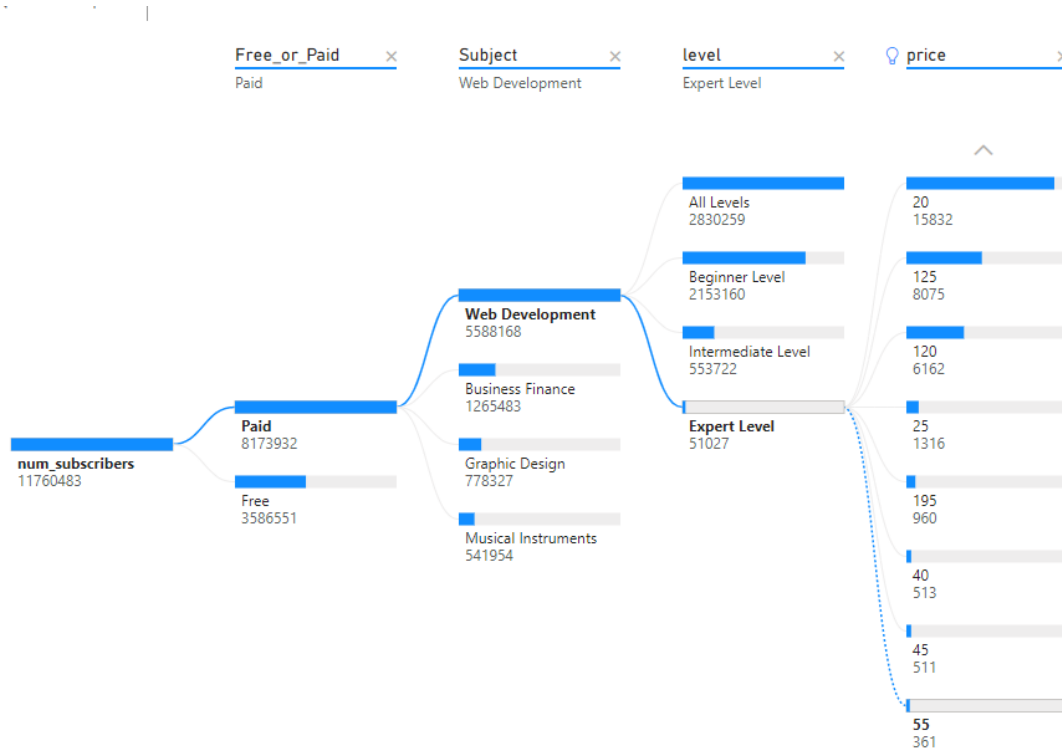


Figure 6: A decomposition tree investigating the breakdown of the number of subscribers

The number of subscribers for paid courses is higher than that of free courses. Web development had the highest number of subscribers compared to the other subjects and it generates the most revenue from the beginner level courses

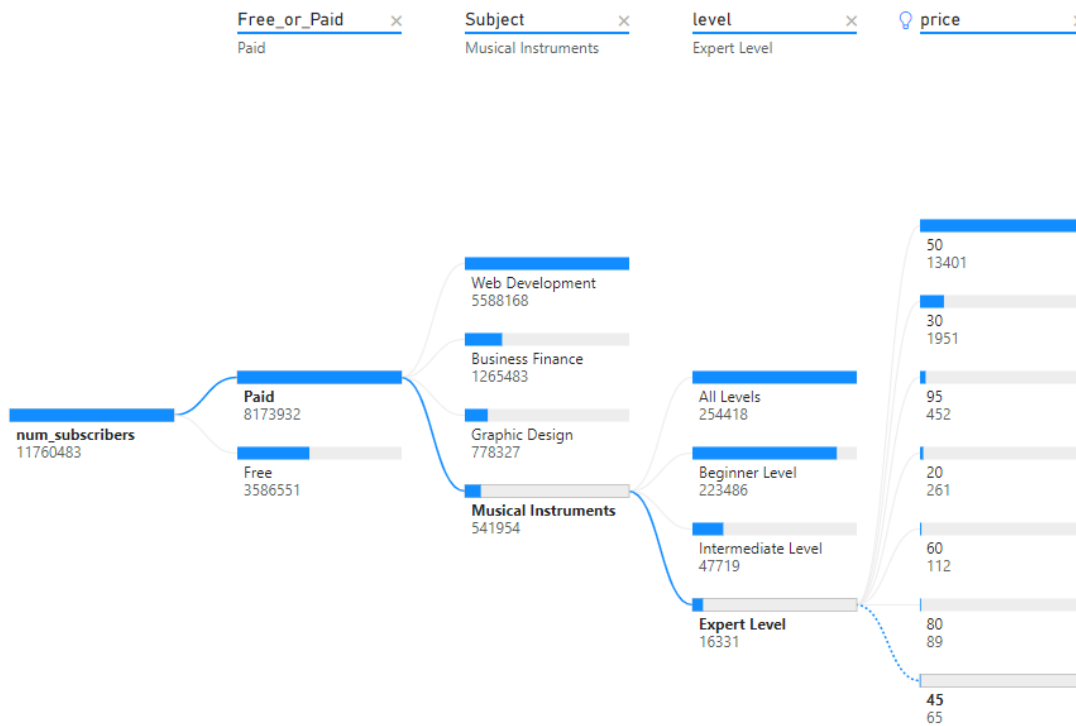


Figure 7: A decomposition tree investigating the breakdown of the number of subscribers for Musical Instruments

From the figure above, Musical Instruments has the least number of subscribers for all courses. Most of its subscribers originate from the beginner level

## Root Cause Analysis of Average Rating of Courses by Subject

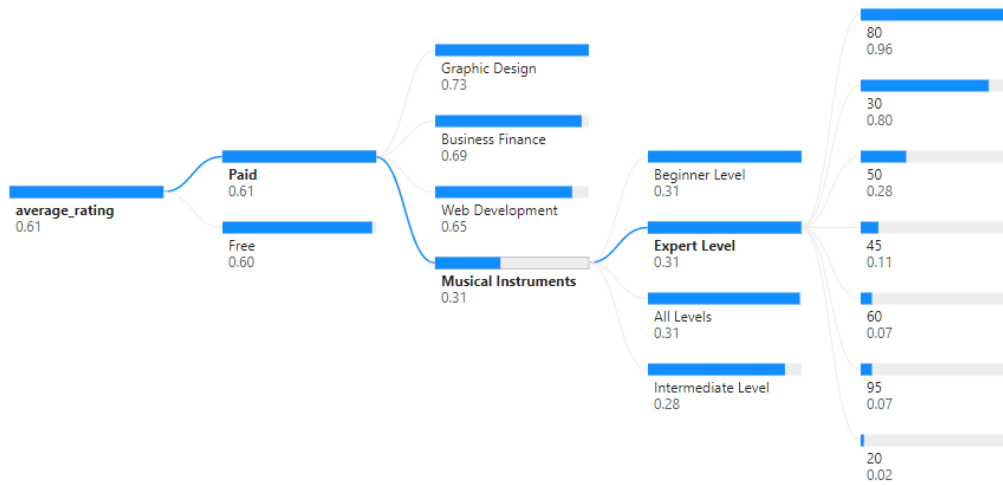


Figure 8: A decomposition tree investigating the breakdown of the average rating for Musical Instruments

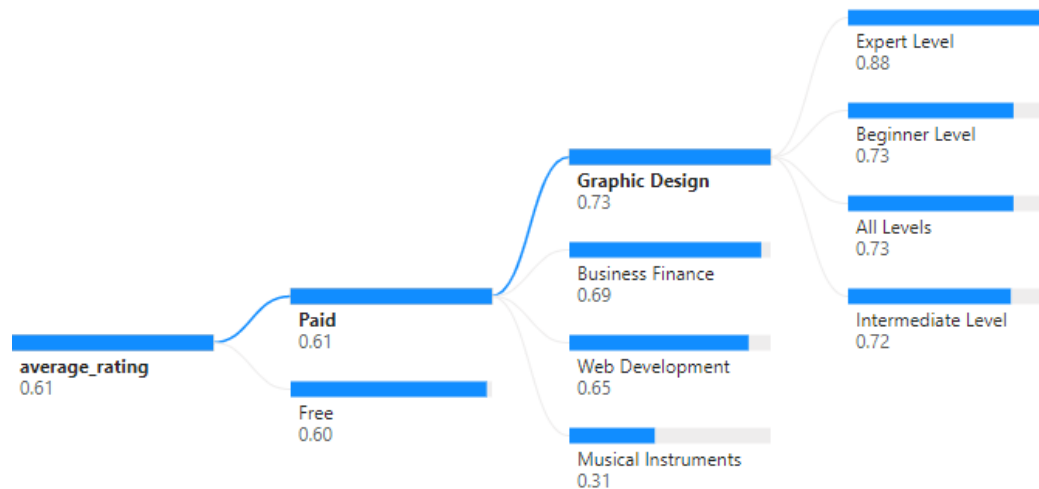


Figure 9: A decomposition tree investigating the breakdown of the average rating for Graphic Design

From the two diagrams above, average rating of the overall subjects was greatly affected by the low average rating of Musical Instruments. Musical Instruments recorded a very low rating for all levels which ranged from 0.28 to 0.3

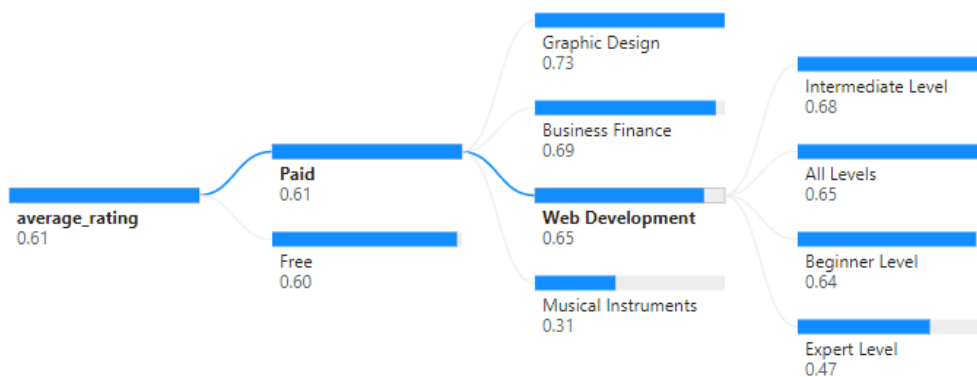


Figure 10 : A decomposition tree investigating the breakdown of the average rating for Web development

Web development recorded a fairly good average rating. However on further analysis, the expert level courses recorded very low rating compared to the average rating of beginner level and intermediate level.

## Key Performance Influencers for Pricing and Number of Subscribers

### What influences the average price to increase?

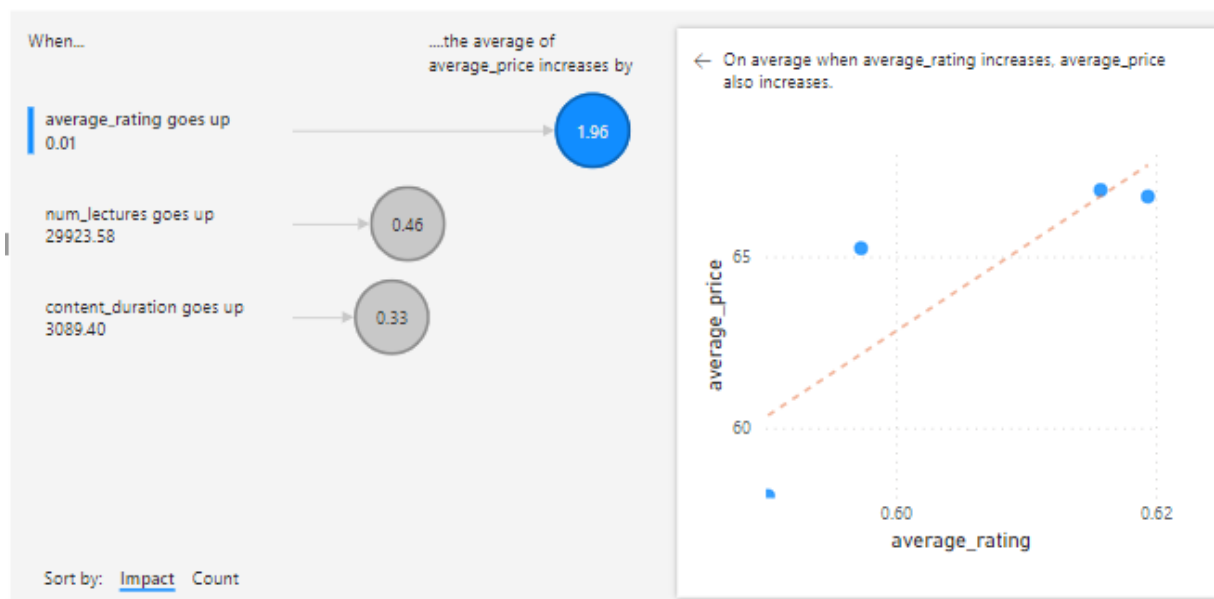


Figure 11: Key performance influencers for increasing the price by average rating

The average price of subject course is likely to increase when the average of rating increases. The price of a course is likely to increase by 1.96 when the average rating increase by 0.01



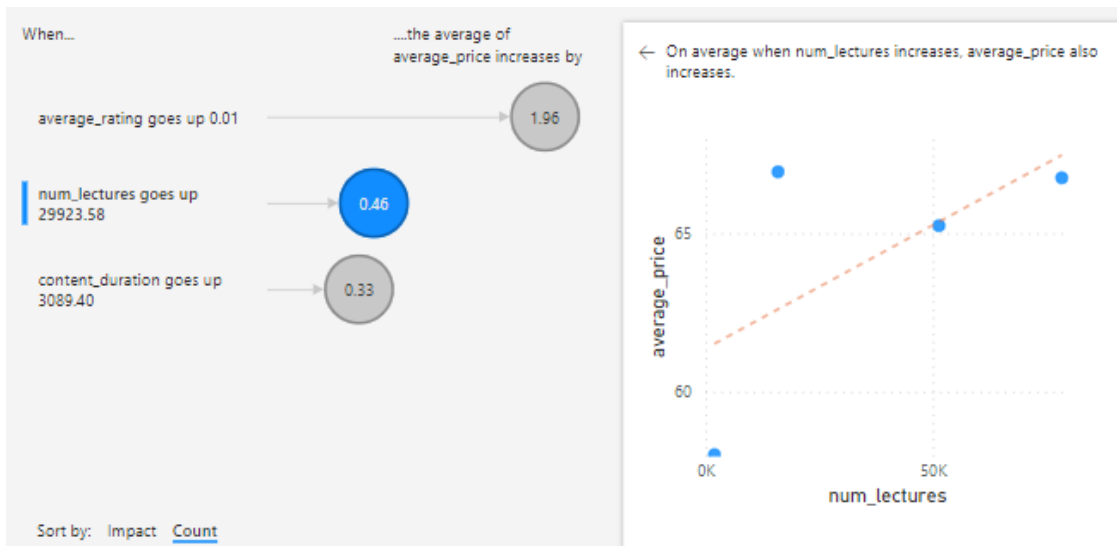


Figure 12: Key performance influencers for increasing the price by number of lecturers

The average price of a subject course is likely to increase by 0.46 when the number of lecturers increases to 29923.58.

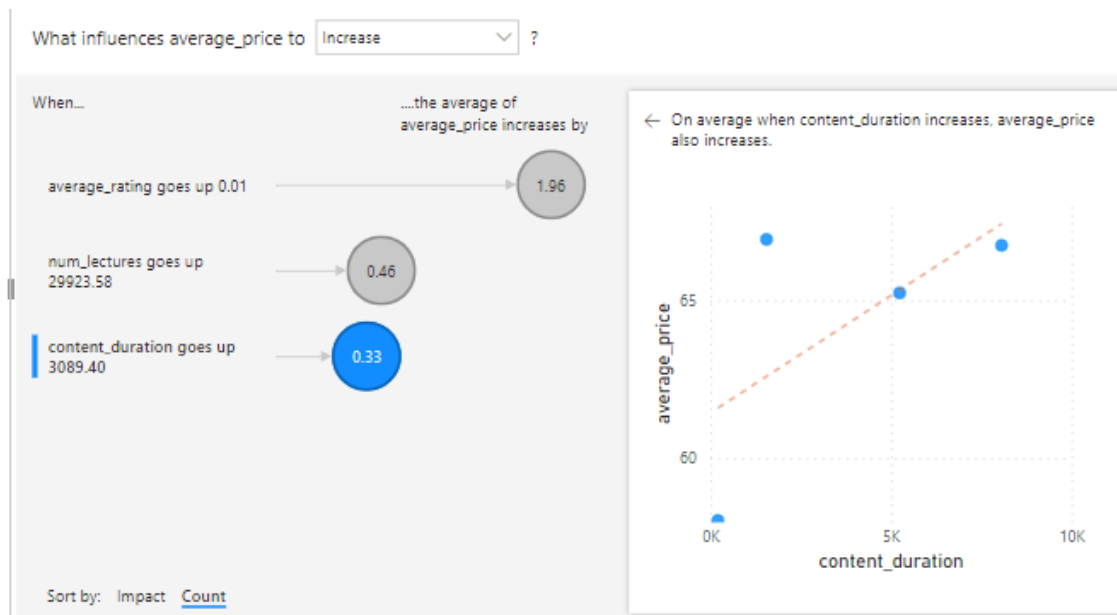


Figure 13: Key performance influencers for increasing the price by content duration

The average price of a subject course is likely to increase by 0.33 when content duration increases to 3089.40

## What influences the number of subscribers to decrease?

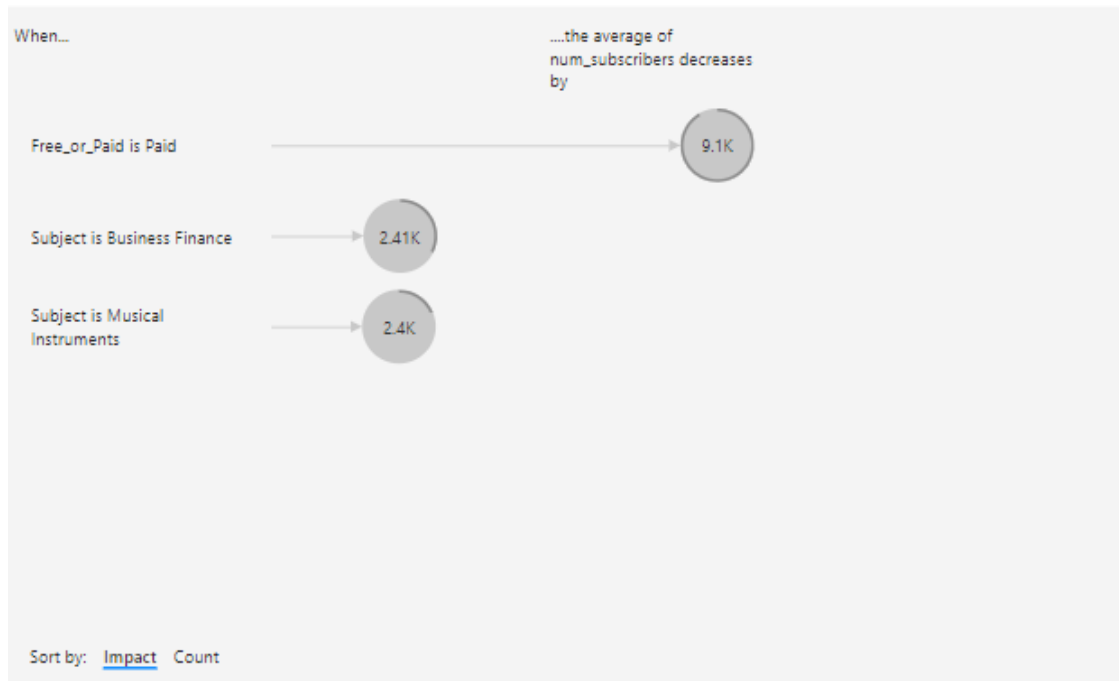


Figure 14: Key performance influencers for increasing the price by content duration

The number of subscribers is likely to reduce when the course is a paid course and the subject is either Business Finance or Musical Instruments.

## When is the number of subscribers more likely to be low?

### Segment 1

Free\_or\_Paid is Paid

Subject is Musical Instruments

In segment 1, the average num\_subscribers is 854.8. This is 2.34K units lower than the overall average, 3.2K.



Segment 1 contains 634 data points (17.2% of the data).



Figure 15: Segment 1

The number of subscribers is likely to reduce when the subject is Musical Instrument and is Paid

### Segment 2

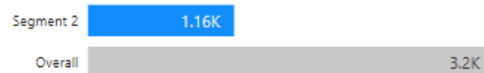
Free\_or\_Paid is Paid

Subject is not Web Development

Subject is not Musical Instruments

Subject is not Graphic Design

In segment 2, the average num\_subscribers is 1.16K. This is 2.04K units lower than the overall average, 3.2K.



Segment 2 contains 1,095 data points (29.8% of the data).

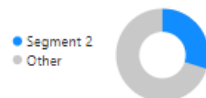


Figure 16: Segment 2

The number of subscribers is likely to reduce when the subject is Business Finance and is Paid

## When is the number of subscribers more likely to be high?

### Segment 1

Free\_or\_Paid is Free

In segment 1, the average num\_subscribers is 11.53K. This is 8.33K units higher than the overall average, 3.2K.



Segment 1 contains 312 data points (8.5% of the data).

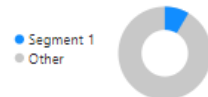


Figure 17: Segment for subscribers

The number of subscribers is likely to increase when the course is Free.

### Segment 2

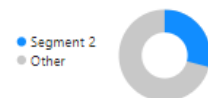
Free\_or\_Paid is Paid

Subject is Web Development

In segment 2, the average num\_subscribers is 5.23K. This is 2.03K units higher than the overall average, 3.2K.



Segment 2 contains 1,069 data points (29.1% of the data).



The number of subscribers is likely to increase when the course is Paid and the subject is Web development

## Insights

---

- Graphic Design had the highest average rating.
- The expert level had least rating for all subjects
- Musical instruments had the lowest rating for all levels.
- Web development recorded a mediocre average rating for all levels
- The average price of a subject is likely to increase when the rating is high, number of lectures increase or the content duration increase. However, the leading influencer is the rating of the subject.
- Web development has the highest content duration, and number of subscribers as well as price.
- The number of subjects is likely to decrease when the course is paid and either a Business or Musical instruments course
- Most of the subscribed courses were beginner level courses
- Increase the price of web development courses since it is popular and has the highest demand

## Recommendation

---

- Management should increase the revenue generated by other levels by introducing learning packages that will combine a beginner course and an expert level course or a beginner course and an expert level course
- Increase the prices of web development courses especially for the beginner level

### Post Three Weeks Recommendation

- Set a target to increase the rating of web development courses which in effect may increase referrals to the courses and cause an increase in the demand for the courses
- Run promotional sales for less subscribed subjects such as business and musical instruments at a discount
- Work on increasing the content duration for Graphics since it has a high rating

### Post Three Weeks Recommendation

- Management should revisit the quality of the expert level courses and re-evaluate performance influencer metrics such as rating, content of duration and the number of lectures of courses.