|  |
| --- |
| **ANALYSIS OF EDUCATIVE COURSE DATA** |
|  |
| This documents contains the analytic methods used and the insights derived from analyzing an Educative Course Dataset to understand where opportunities to increase revenue may lie and track the performance of courses |
|  |

|  |  |  |
| --- | --- | --- |
| Contents  [INTRODUCTION 2](#_Toc88359102)  [ PROJECT DESCRIPTION 2](#_Toc88359103)  [ STAKEHOLDERS 2](#_Toc88359104)  [ DATA COLLECTION 2](#_Toc88359105)  [ FUTHER QUESTIONS 2](#_Toc88359106)  [DESIGN 2](#_Toc88359107)  [ SOFTWARES 2](#_Toc88359108)  [ DATA CLEANING 2](#_Toc88359109)  [DATA ANALYSIS RESULTS 3](#_Toc88359110)  [ Top 20 MOST SUBSCRIBER PAID OR FREE COURSES 3](#_Toc88359111)  [ GENERAL STATSTITICS 4](#_Toc88359112)  [ SUBSCRIBERS COUNT BY LEVELS 6](#_Toc88359113)  [ AVERAGE PRICE (PAID COURSES ONLY) 7](#_Toc88359114)  [ RATINGS AND REVIEWS 8](#_Toc88359115)  [ PUBLISHING DATE ANALYSIS 10](#_Toc88359116)  [CONCLUSIONS 12](#_Toc88359117)  [APPENDIX 13](#_Toc88359118) |  |  |

# INTRODUCTION

## PROJECT DESCRIPTION

As a Data Analyst working with Educative, an online interactive course platform, have been tasked by the Head of Curriculum, to present an analysis on courses where opportunities to increase revenue may exist, as well as track the performance of courses. This analysis is to focus mainly on the four main subject groupings of Educative (Web Development, Business and Finance, Graphic Design and Musical Instruments).

The Head of Curriculum has suggested encouraging Web Development courses to charge more, because she believes that these are the most popular courses. Furthermore, the Head of Curriculum has requested for the list of the top 20 most subscribed courses. This analysis is to be completed within the next two weeks

## STAKEHOLDERS

1. The Head of Curriculum
2. The CEO of Educative
3. Course Content Creators

## DATA COLLECTION

Educative provided 4 CSV files, each containing data on the 4 subject groupings (Web Development, Business and Finance, Graphic Design and Musical Instruments).

## FUTHER QUESTIONS

The questions below were asked in other to acquire more insights for the assigned task.

1. The specific criteria used to determine an increase in the cost of a subject.
2. The existence of any previous analysis related to the task.
3. Enquire if there’s data available to aid in grouping reviews into negative and positive reviews

# DESIGN

## SOFTWARES

Due to the relatively small size of the dataset, Microsoft Excel was used for the analysis. Microsoft Power BI was also used to create interactive charts for visualizing the results of the analysis.

## DATA CLEANING

The dataset provided was contained in four different CSV files with each file having data on only a specific subject. Before any data cleaning was carried out, these four CSV files were consolidated into a single Excel Worksheet to aid in a more comparative analysis of all the four subjects grouping offered by Educative. The following data cleaning procedures below were carried out to ensure the quality of the dataset before analysis.

1. **Removal of duplicates:** rows with duplicate values were removed .leaving only one instance of occurrence.
2. **Removal of blank cells:** All blank cells within the subject column were removed from the merged dataset.
3. **Formatting of headers:** All headers were formattedin other to ensure that they were clear and concise. Also underscores were use in between words to enable easy parsing later**.**
4. **Find and Replace:** The subject ***“Subject: Web Development***” was replaced with “***Web Development”*** to ensure uniformity within the subject column.
5. **Formatting of Data types:** the data types of each of the columns were cross checked to ensure that every data value was well formatted.

# DATA ANALYSIS RESULTS

## Top 20 MOST SUBSCRIBER PAID OR FREE COURSES



## GENERAL STATSTITICS

**Paid Courses**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Subjects | Total Number of Subscribers | Average Number of Subscribers | Number of Courses | Average of Content Duration(hours) | Number of Reviews | Average of Rating | Average Price |
| Web Development | 5,588,168 | 5,227 | 1,069 | 5.97 | 339,168 | 0.65 | $ 86.69 |
| Business Finance | 1,265,483 | 1,156 | 1,095 | 3.68 | 51,772 | 0.69 | $ 74.72 |
| Graphic Design | 778,327 | 1,373 | 567 | 3.69 | 26,494 | 0.73 | $ 61.46 |
| Musical Instruments | 541,954 | 855 | 634 | 2.95 | 24,895 | 0.31 | $ 53.15 |
| **Grand Total** | **8,173,932** | **2,429** | **3,365** | **4.27** | **442,329** | **0.61** | **$ 72.23** |

**Free Courses**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Subjects | Total Number of Subscribers | Average Number of Subscribers | Number of Courses | Average of Content Duration (hours) | Number of Reviews | Average of Rating | Average Price |
| Web Development | 2,393,767 | 17,864 | 134 | 2.55 | 90,731 | 0.60 | 0 |
| Business Finance | 603,228 | 6,284 | 96 | 2.15 | 24,130 | 0.66 | 0 |
| Musical Instruments | 304,735 | 6,625 | 46 | 1.55 | 6,829 | 0.31 | 0 |
| Graphic Design | 284,821 | 8,138 | 35 | 1.92 | 10,576 | 0.78 | 0 |
| **Grand Total** | **3,586,551** | **11,532** | **311** | **2.21** | **132,266** | **0.60** | **0** |

**Observations**

1. There exist both paid courses and free courses within all the four subject groupings
2. Generally Subscribers participated in paid courses more than free courses.
3. Despite the fact that paid Web Development courses have the second highest number of courses ***(1,069),*** it is the subject which ranked highest in terms of the total number of subscribers **(7,981,935).** These paid courses within Web Development were also ***averagely quite expensive($86.69)*** and required ***longer amount of hours*** compared to the average cost and content duration for the other subject groupings
4. The average price for all the paid courses within the 4 major subject groupings ranges between ***$86 and $53***.
5. The average rating for Musical Instruments is ***relatively low***(for both paid and free courses)compared to the remaining subjects which have average ratings ranking approximately between ***0.6 and 0.8***
6. Surprisingly, despite ***the low numbers of Graphic Design Subscribers*** especially for free courses, it has the ***overall highest rating of over 0.7*** for both ***paid and free courses***.
7. Reviews were also received from only ***4.9%*** of the total number of Subscribers.

## SUBSCRIBERS COUNT BY LEVELS

**Paid Courses**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Subscriber Counts** | Subjects |  |  |  |  |
| Levels | Web Development | Business Finance | Graphic Design | Musical Instruments | Grand Total |
| All Levels | 2,830,259 | 692,611 | 451,180 | 254,418 | 4,228,468 |
| Beginner Level | 2,153,160 | 397,835 | 263,988 | 223,486 | 3,038,469 |
| Intermediate Level | 553,722 | 155,493 | 56,828 | 47,719 | 813,762 |
| Expert Level | 51,027 | 19,544 | 6,331 | 16,331 | 93,233 |
| **Grand Total** | **5,588,168** | **1,265,483** | **778,327** | **541,954** | **8,173,932** |

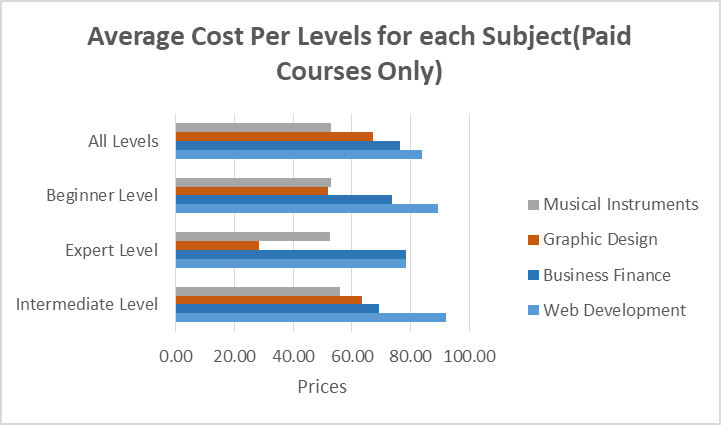
**Free Courses**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Subscriber Counts** | Subjects |  |  |  |  |
| Levels | Web Development | Business Finance | Musical Instruments | Graphic Design | Grand Total |
| All Levels | 1,251,319 | 390,849 | 210,626 | 179,107 | 2,031,901 |
| Beginner Level | 806,244 | 140,186 | 62,045 | 95,600 | 1,104,075 |
| Expert Level | 170,862 | 15,955 | 61 |  | 186,878 |
| Intermediate Level | 165,342 | 56,238 | 32,003 | 10,114 | 263,697 |
| **Grand Total** | **2,393,767** | **603,228** | **304,735** | **284,821** | **3,586,551** |

**Observations**

1. In all four Subjects, subscribers always preferred to complete ***all levels***of the ***paid or free course*** they undertook. This level is followed in rank by Beginner level courses and Intermediate Level Courses
2. Generally the ***least*** patronized level for ***all*** ***paid courses*** was ***the Expert level***
3. There were ***no subscribers*** for ***Expert Level free courses in Graphic Design*** and also, Subscribers for ***Musical Instruments Expert free courses*** were also ***relatively low***.

## AVERAGE PRICE (PAID COURSES ONLY)

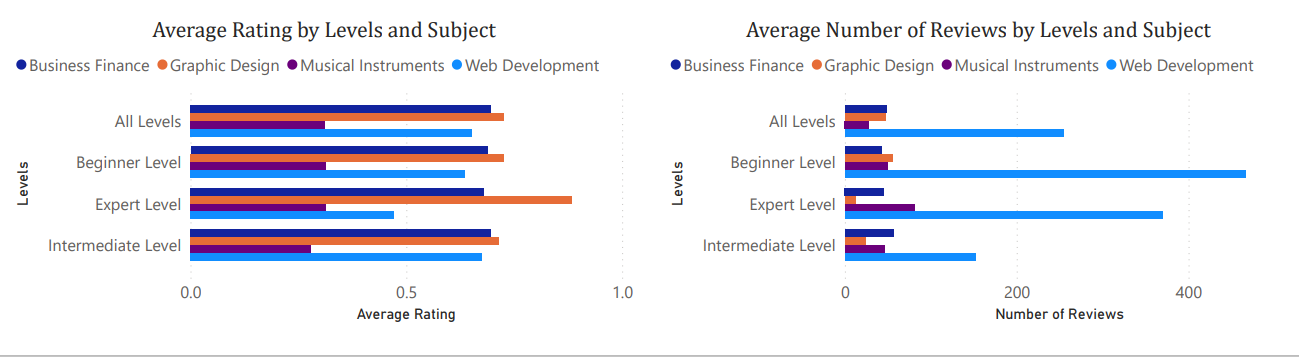


**Observations**

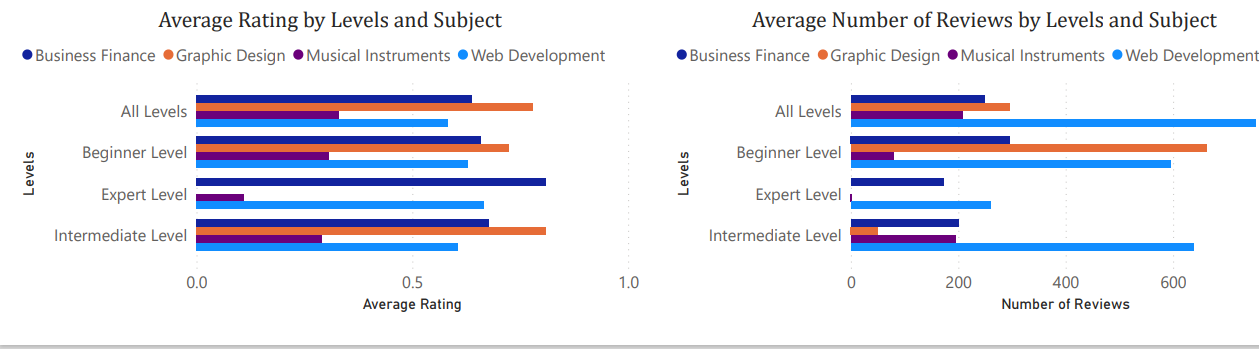
1. For almost all Levels ***Web Development courses*** have the ***highest average price***. Except in the case of ***Expert Courses*** were is averagely priced the ***same*** as ***Business Finance Courses***.
2. The other subjects are ranked from ***higher to lowest average price*** in the order ,***Business Finance, Graphic Design and Musical Instruments***
3. In most cases, courses have ***almost the same average price*** for ***each level.***
4. The ***least priced*** courses are ***Graphic Design Expert Level Courses.***

## RATINGS AND REVIEWS

**Paid Courses**

****

**Free Courses**

****

**Observations**

1. ***Reviews***

***Web Development paid courses*** have the ***highest average reviews*** for all ***4 levels***.The other subjects lag behind greatly for paid courses***.*** However, this situation is different in the case of **free courses** where it can be observed that ***Web Development free courses,*** have the ***highest average reviews*** for only ***3 levels except*** the Beginner Level.This Level rather has ***Graphic Design free Courses,*** having ***the highest number of reviews***

There are ***no review*** for free ***Expert Level Business Finance Courses*** and ***free Graphic Design Courses***

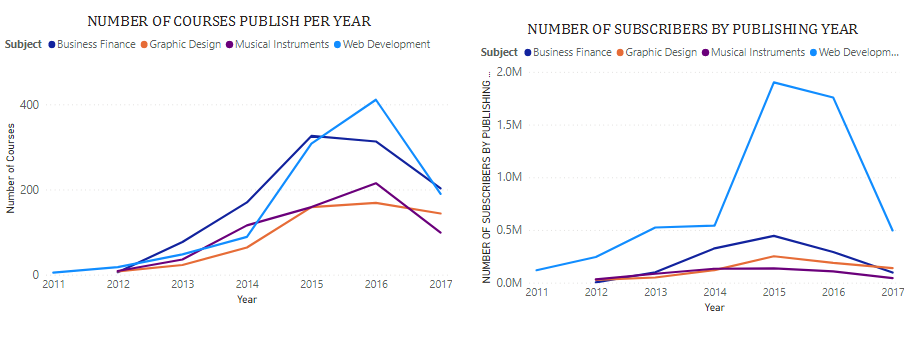
1. ***Ratings***

For ratings, ***Graphic Design paid or free Courses*** have the ***highest rating for all 4 levels***, except in the case of ***Expert free courses*** were there’s ***no rating*** for Graphic Design Courses. As such**, *Business Finance free courses*** has the ***highest rating*** for the Expert Free Course level

At all 4 Levels, ***Web Development Courses*** have the ***third highest course rating*** for both free or paidcourses, and ***Musical Instruments courses*** have the ***least course rating*** for both free and paid courses

## PUBLISHING DATE ANALYSIS

**Paid Courses**

****

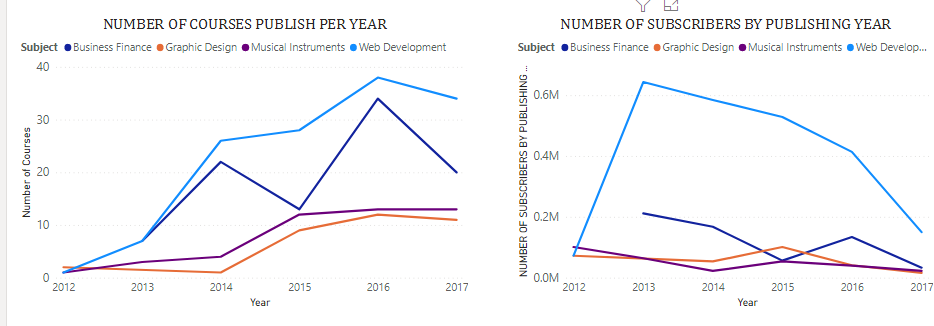
**Observations**

1. **Web Development Paid Courses:** Courses related to this subject were ***included*** from the year ***2011*.** The number of courses ***relatively increased*** until ***2016*** when it ***decreased sharply*** from ***400 courses*** to ***200 courses***. This decreases in courses may have contributed to the ***dip in subscriber numbers*** from the year ***2016-2017***

Despite there being an ***increase*** in the number of courses between ***2015 and 2016.*** The number of subscribers decreased steadily from 2015 to 2016

1. **Business and Finance Paid Courses:** Courses related to this subject were ***only included*** from the year ***2012*.** The number of courses ***relatively increased*** until ***2015*** when it began to experience a decrease in numbers until 2017 This decreases in courses may have contributed to the ***dip in subscriber numbers*** from the year ***2015-2017***
2. **Musical Instrument Paid Courses:** Courses related to this subject were ***only included*** from the year ***2012*.** The number of courses ***relatively increased*** until ***2016*** when it decreased sharply in numbers until 2017. The number of subscribers seems not to have been affected by the trends in the number of ***courses .it rather increased sharply*** until ***2014,***then remained ***relatively constant*** from ***2014-2015*** and then ***decreased*** from ***2015 to 2017*** .
3. **Graphic Design Paid Courses:** Courses related to this subject were ***only included*** from the year ***2012*.** The number of courses ***relatively increased*** from ***2012 to 2015***, remained ***constant*** from ***2015-2016*** and ***dip slightly*** from ***2016-2017***.The ***constant and gradual*** ***dip*** from ***2015- 2016***, may have caused the ***fall*** in subscriber numbers for these years.

**Free Courses**

****

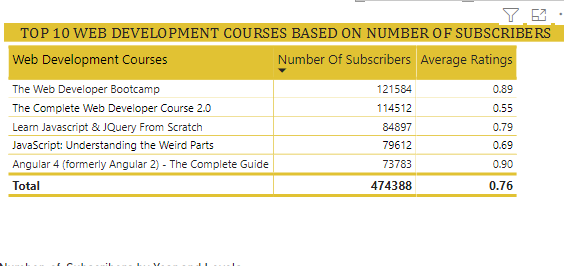
**Observations**

1. **Web Development Free Courses:** Courses related to this subject were ***included*** from the year ***2012*.** The number of courses ***relatively increased*** until ***2016*** when it ***decreased sharply*** until 2017. Despite the relative increase in courses from 2012-2016, the number of subscribers did not follow the same trend, it ***decrease sharply from 2013 to 2015*** and continued dipping ***until 2017***.
2. **Business and Finance Free Courses:** Courses related to this subject were ***only included*** from the year ***2013*.** The number of courses ***increased sharply and decreased sharply alternatively*** between the years ***2013-201***.The number of subscribers didn’t follow this trend. Surprisingly, 2013 which had the ***least*** number of courses actually had ***the highest number of subscribers*** for this subject.
3. **Musical Instrument Free Courses:** Courses related to this subject were ***included*** from the year ***2012*.** The number of courses ***relatively increased*** throughout the period ***(2012-2015)*** and remained ***constant*** from ***2015 to 2017***. The number of subscribers did not follow the same trend.
4. **Graphic Design Free Courses:** Courses related to this subject were ***included*** from the year ***2012*.** The number of courses ***decreased*** from ***2012 to 2014***, increased from 2014***-2016*** and ***dip slightly*** from ***2016-2017*** .The subscriber numbers from 2012 to 2014 were affected by the decrease in courses for these years. This however ***increased sharply*** from ***2014-2015*** and ***decreased sharply*** again from ***2015-2016***. And it continued to ***dip*** until ***2017***

# CONCLUSIONS

* The analysis confirmed the suspicion that Web Development courses were the most popular courses. Whether in terms of Paid or Free Courses. Despite being the most expensive subject averagely, it had the highest number of subscriptions, it also had the highest number of courses and the highest number of reviews. However its rating was a bit lower compared to the Graphic Design and Business Finance.

A deeper dive into Web Development Paid Courses revealed that the following 5 paid courses were amongst the top 10 list of Web Development Paid Courses.



This could be an indication that a slight increase in Web Development Paid Courses may not have an effect on Subscriber Numbers

* The few Graphic Design Paid Courses ***(567)*** and Free Courses ***(35)*** had the highest average rating of ***0.73*** and ***0.78*** for Paid and Free Courses respectively. Even though the subscriptions numbers are just ***9%*** of the total. It could potentially become a subject to explore once subscribers begin to see the benefit of the Graphic Design Courses as it’s the highest rated subject
* More than Half (**53%)** of Subscribers are willing to complete all levels for each of the Subjects. This is a positive indication that subscribers are indeed interested in the courses available on the Educative interactive Platform.

# APPENDIX

Excel Data Set for Educative Project

<https://docs.google.com/spreadsheets/d/1okKr_aCwKBvig3b_GJr1pJMKshkez4NW/edit?usp=sharing&ouid=112561061263128095673&rtpof=true&sd=true>

Power BI Data Visualization for Educative Project

<https://drive.google.com/file/d/1HzXtnVC_owaMz2xZu-ndIbEoJ2sSmFly/view?usp=sharing>

For More Information on Educative

<https://www.educative.io/>