BLOSSOM ACADEMY

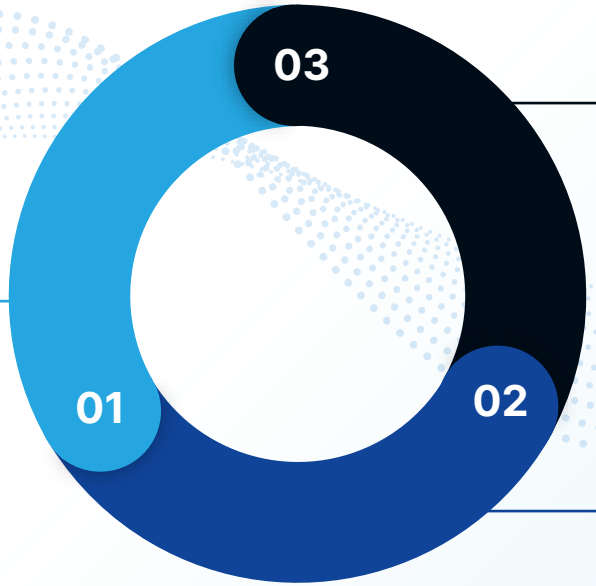
EDA on Udemy Course Datasets Using Power BI

Hello!

My name is Jammal Adeyemi

Just here to present my analysis and insights to my respective colleagues in the building.

The process taken can be divided into 3



Recommendation & Conclusion

Lastly, we made recommendation based on our insights.

Transforming the Datasets

Here, we got the datasets and transformed it so that we can make proper analysis & insights.

Analysis Questions

Answering of importants question for our analysis.

1. Udemy Courses

“

Information is the oil of the 21st century, and analytics is the combustion engine

About this Datasets

- This dataset contains 3.682 records of courses from 4 subjects (Business Finance, Graphic Design, Musical Instruments & Web Design) taken from Udemy.
- Udemy is a massive online open course (MOOC) platform that offers both free & paid courses.
- This data has the following column names; course_id, course_title, url, is_paid, price, num_subscribers, num_reviews, num_lectures, level, content_duration, published_timestamp, and subject.

2. Transforming the Data

- The dataset was provided by Blossom Academy
- Since the data is in CSV format we start by clicking on “Get Data” > “Text/CSV” > “Locate File Location and import”.
- Once done importing, I dived into the Power Query Editor to start making the necessary transformation.
- Corrected some columns with the wrong datatypes setting.
- Dropped duplicate & content_duration that was less than 1mins

- Extracted the “month”, “year”, and “year quarter” columns from the “published_timestamp” column.
- Created a new column, where I changed the “is_paid” column True/False to paid/unpaid.
- Calculated a new measure called “profit” to get the total profit Udemy made per course.
- Hide some columns that won’t be needed for visualization like url, is_paid, content_duration etc.

3. Analysis Questions

Subjects/Category –

- What is the distribution of subjects
- How many courses per subject
- Distribution of subjects per year
- How many people purchase a particular subject
- Which subject is the most popular

Published year –

- Number of courses per year
- Which year has the highest number of courses
- Which month generate the highest number of course and profits
- What is the trend of course per year

Levels –

- How many levels do we have
- What is the distribution of courses per level
- Which subject have the highest levels
- How many subscribers per level
- How many courses per level

Distribution of Courses –

- Which courses have the highest duration (paid or not)
- Which courses have a higher duration
- Duration vs number of subscribers

Subscribers –

- Which course has the highest number of subscribers
- Average number of subscribers
- Number of subscribers per subject
- Number of subscribers per year

Distribution of Courses –

- Which courses have the highest duration (paid or not)
- Which courses have a higher duration
- Duration vs number of subscribers

Correlation Questions –

- Does the number of subscribers depend on Price, and No of Review
- Does price influence Subscription and Review

4. Data Visualization

I. Subscribers Analysis

Published_Year

- [ ] 2011
- [ ] 2012
- [ ] 2013
- [ ] 2014
- [ ] 2015
- [ ] 2016
- [ ] 2017

Published_Month

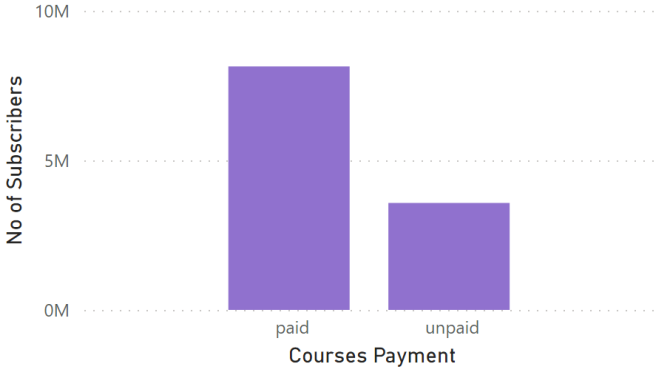
- [ ] Apr
- [ ] Aug
- [ ] Dec
- [ ] Feb
- [ ] Jan
- [ ] Jul
- [ ] Jun
- [ ] Mar
- [ ] May
- [ ] Nov
- [ ] Oct
- [ ] Sep

12M
Total Subscribers

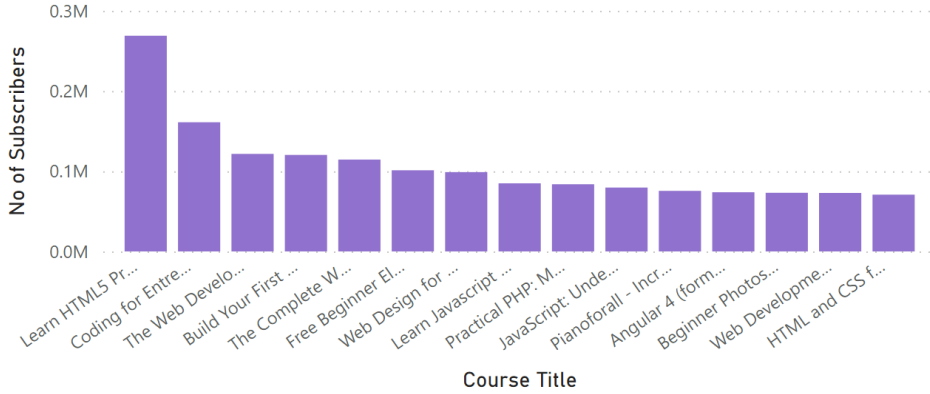
3K
Avg. No of Subscrib...

269K
Max No of Subscribers

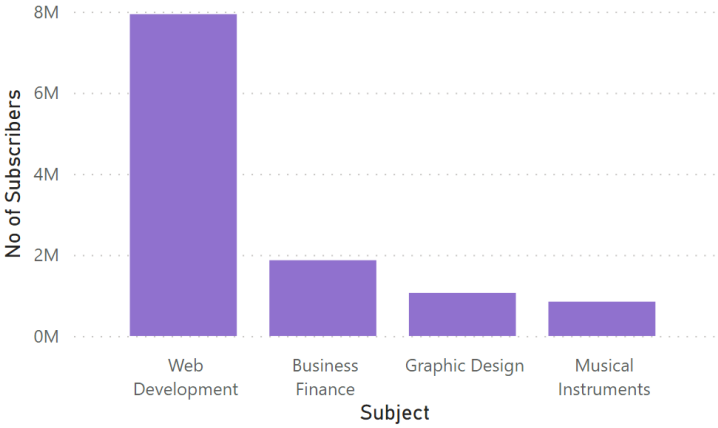
Total Number of Paid & Unpaid Subscribers



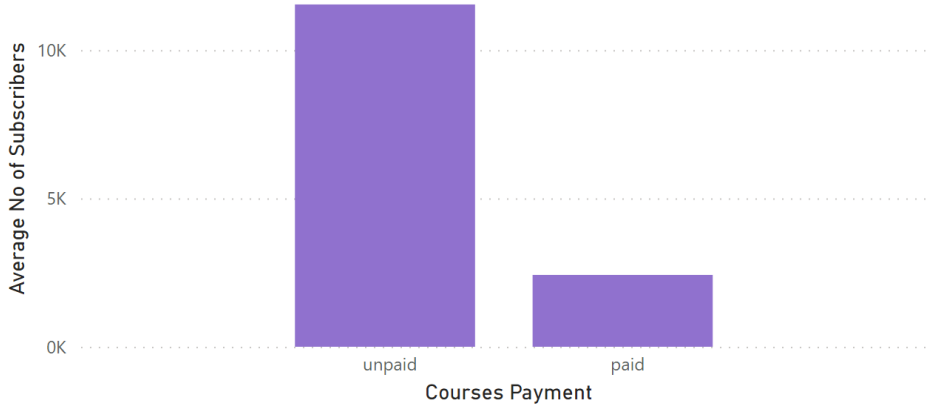
No of Subscribers by Course Title



Popularity of Each Course Type



Average No of Subscribers per Course Type



- The most popular subject amongst the subscribers is Web Development, followed by Business Analysis
- Udemy has 12M total subscribers between the duration of 2011-2017
- Average number of subscribers: “3193”
- The maximum number of subscribers for a course is “269k”, which is “Learn HTML5 Programming from scratch” which falls under “Web development” Subject category.
- Filtering further, I found that the year it occurs was 2013 which has a high number of unpaid subscribers.

II. Level Analysis

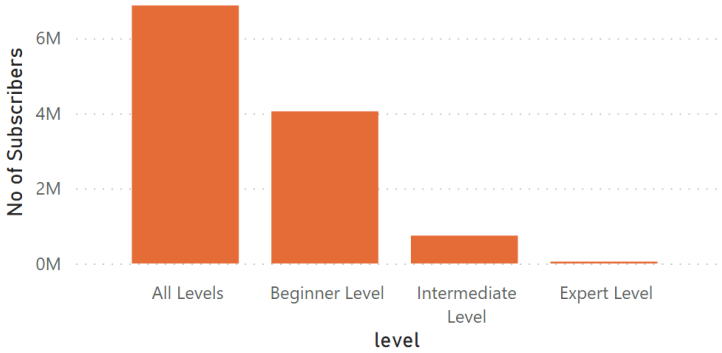
Published Year

- [ ] 2011
- [ ] 2012
- [ ] 2013
- [ ] 2014
- [ ] 2015
- [ ] 2016
- [ ] 2017

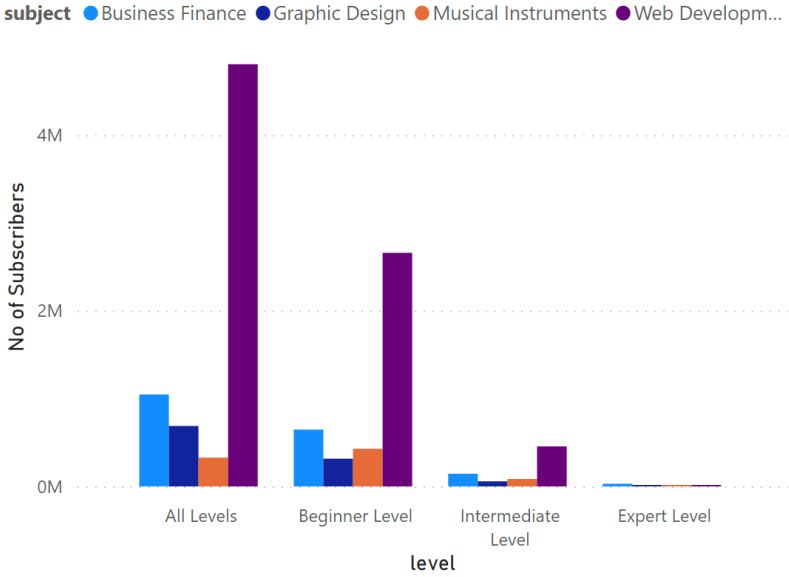
Published Month

- [ ] Apr
- [ ] Aug
- [ ] Dec
- [ ] Feb
- [ ] Jan
- [ ] Jul
- [ ] Jun
- [ ] Mar
- [ ] May
- [ ] Nov
- [ ] Oct
- [ ] Sep

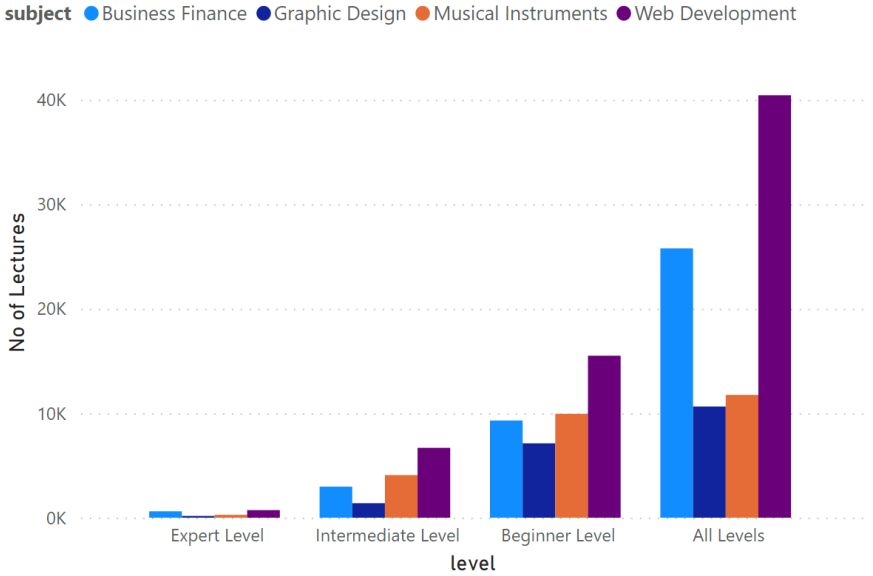
No of Subscribers per Level



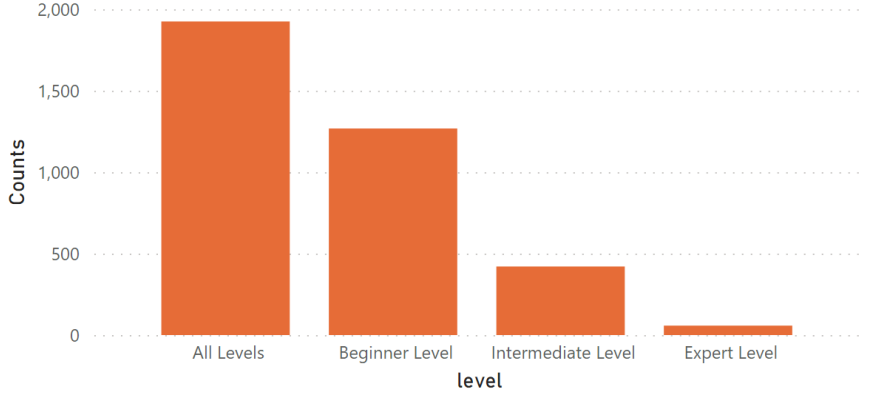
Levels & Subscribers per Subject Category



Levels & Lectures per Subject Category



Distribution of courses per levels



- We have 4 levels as we can see from the visualization.
- “All Levels” has the highest distribution based on subscribers and courses.
- “Web development” subject generated the highest per lectures & subscribers across all levels.
- “2015” had the highest number of subscribers per level & courses.

III. Price & Lecture Analysis

Published Year

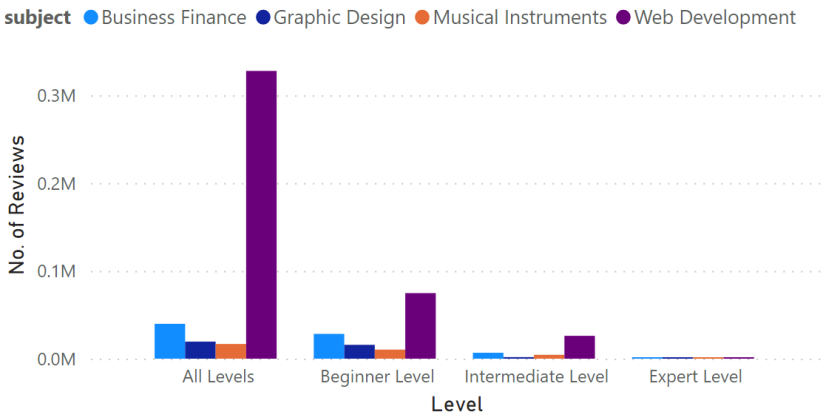
- [ ] 2011
- [ ] 2012
- [ ] 2013
- [ ] 2014
- [ ] 2015
- [ ] 2016
- [ ] 2017

Published ...

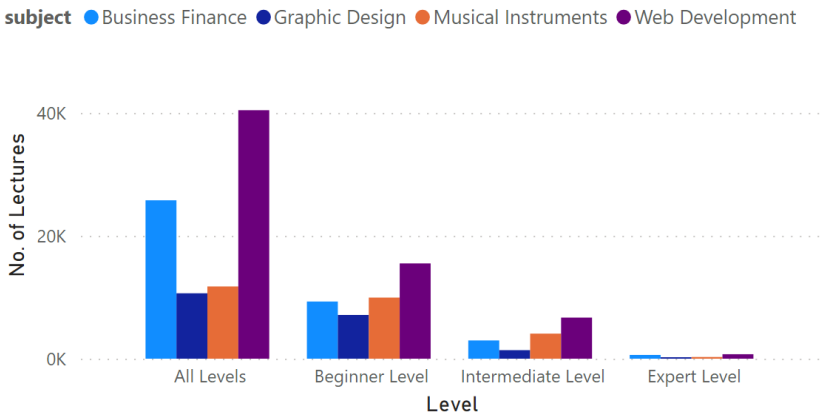
- [ ] Apr
- [ ] Aug
- [ ] Dec
- [ ] Feb
- [ ] Jan
- [ ] Jul
- [ ] Jun
- [ ] Mar
- [ ] May
- [ ] Nov
- [ ] Oct
- [ ] Sep

$243K	$66	$200	$24M	12M	574K	147K
Total Price	Avg. Price	Max. Price	Most Profitable	Total Subscrib...	Total Reviews	Total Lectures

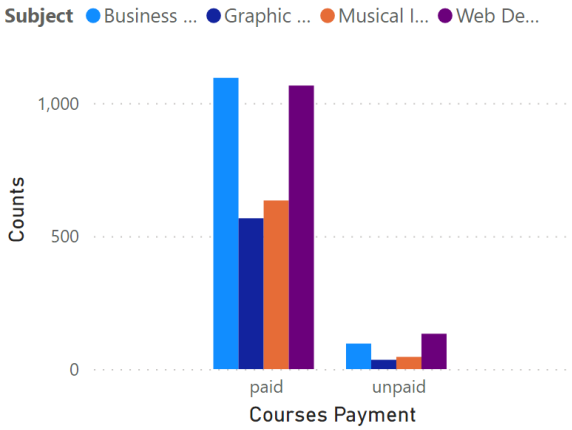
No. of Reviews by Level and subject



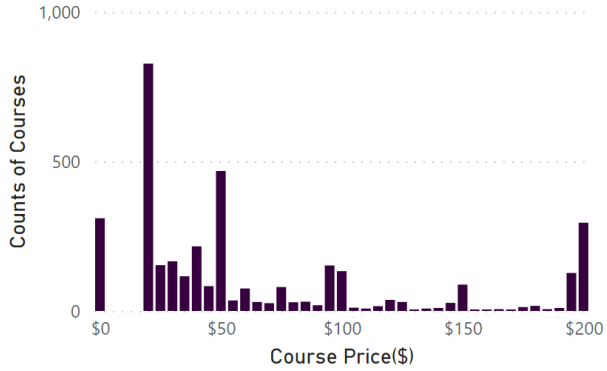
No. of Lectures by Level and subject



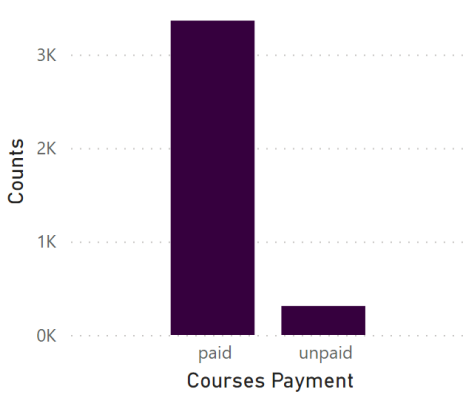
Subject Category paid/unpaid



Distribution of course cost



Courses on Udemy that are paid/unpaid



- The max price is “$200”, filtering further to get the course title and subject which are “Ultimate Investment Banking course” and “Business Finance” respectively.

- The most profitable course made over $24M, which is the “Web Developer Bootcamp”

- $20 has the highest distribution of course cost, followed by $50 then $0.

- “Web Development” has the highest number of unpaid subject.

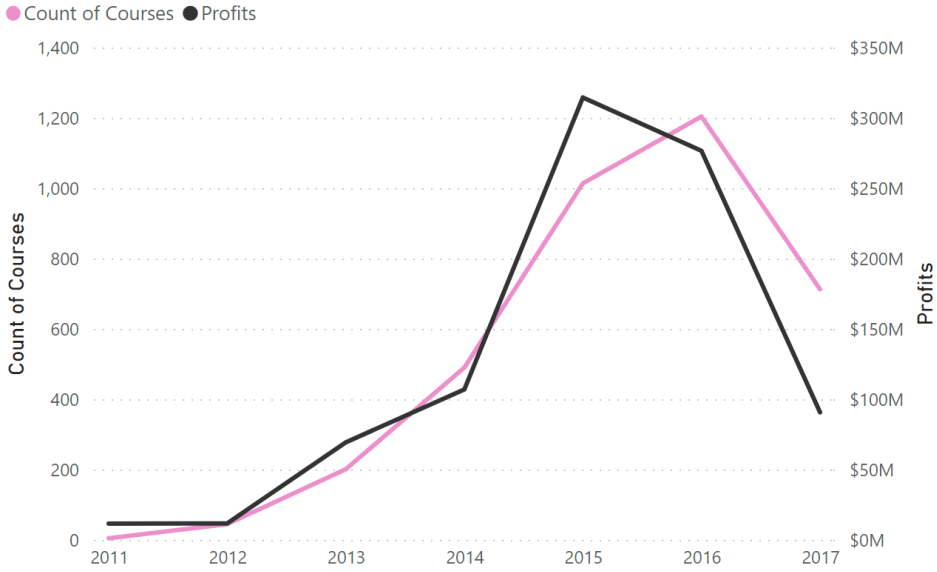
- “Business Finance” has the highest number of paid subject.

- “Web Development” has the highest number of subject per review across all level, followed by “Business Finance”.

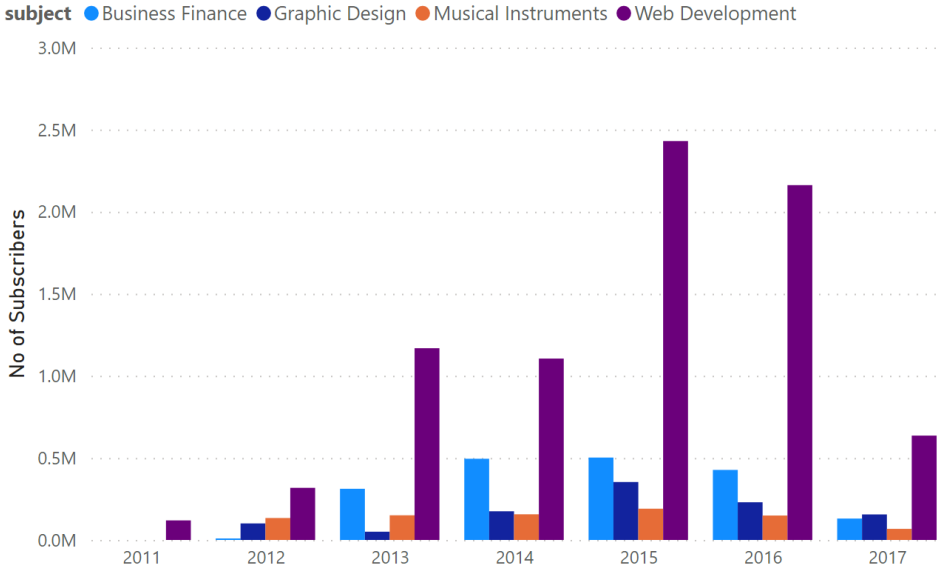
- “Web Development” has the highest number of subject per number of lectures across all level, followed by “Business Finance”.

IV. Year Analysis

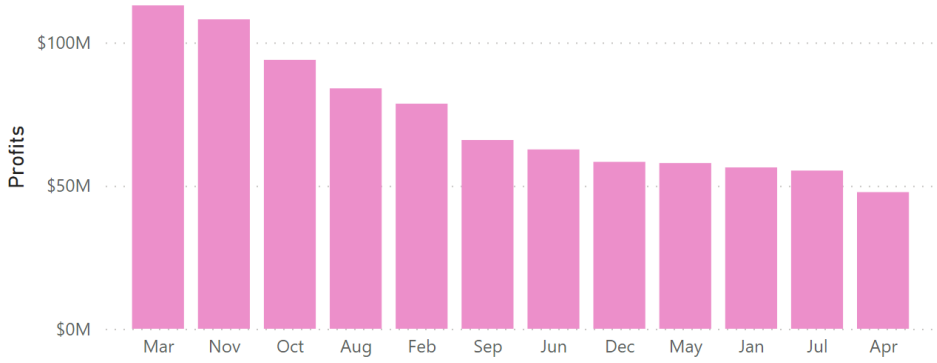
Number of Courses per Year and Profits



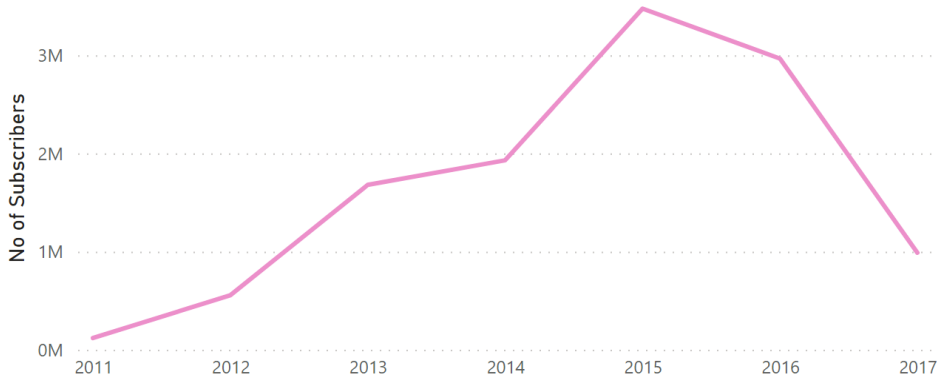
Subscribers per course across the years



Profits by Month



Total Udemy Subscription across the years



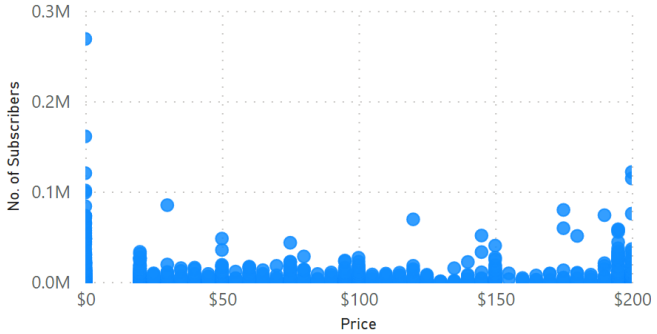
- 2015 has highest number of subscribers across the years
- 2015 generated the highest profit across the years
- 2016 has the highest number of courses published on Udemy
- Web Development in 2015 topped the chart when it comes total subscribers across the years
- March generated the highest profit in the month section

V. Correlation Analysis

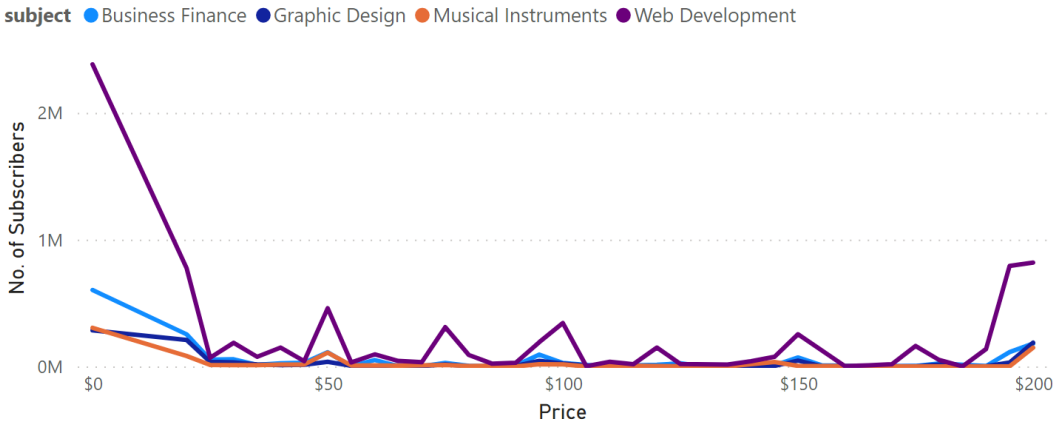
Published ...

- [ ] 2011
- [ ] 2012
- [ ] 2013
- [ ] 2014
- [ ] 2015
- [ ] 2016
- [ ] 2017

Does Price Influence Subscription



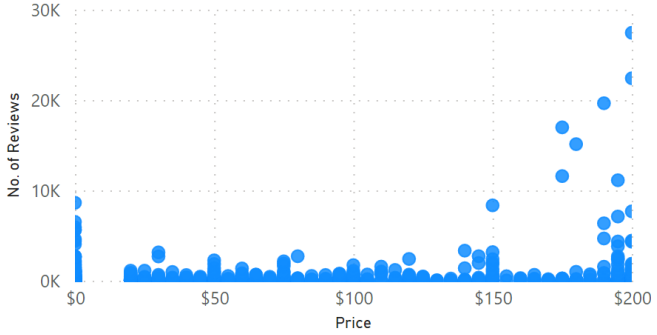
Does Price Influence Subscription Per Subject Category



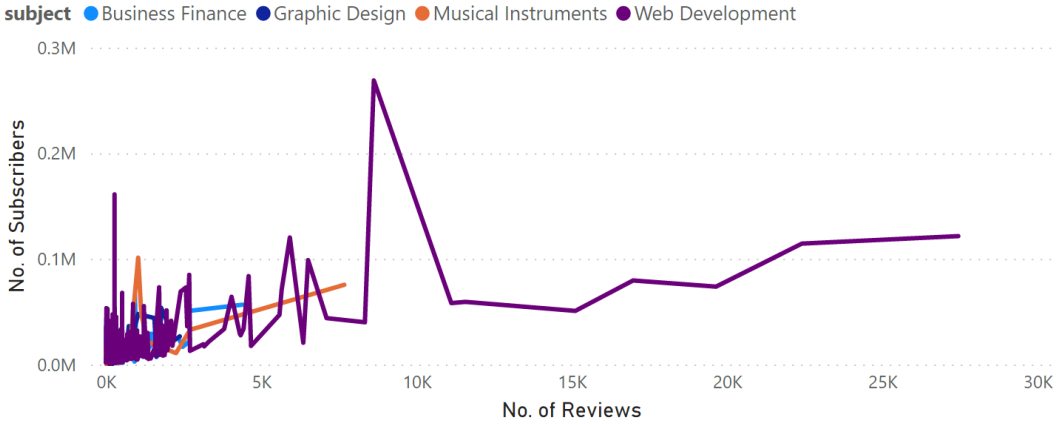
Published ...

- [ ] Apr
- [ ] Aug
- [ ] Dec
- [ ] Feb
- [ ] Jan
- [ ] Jul
- [ ] Jun
- [ ] Mar
- [ ] May
- [ ] Nov
- [ ] Oct
- [ ] Sep

Does Price Influence Review



Does Reviews Influence Subscription Per Subject Category



- The Price actually influence users subscription as there was high intake on free courses on the platform.
- The price influences the reviews as users that subscribe for a course has a high reviews compare to the free courses.
- The price influences what kind of subscription the user made, as there were more subscribers taking “Web Development” free subject.
- Lastly, users review had a some influence to user subscription but not as much has prices.

RECOMMENDATION & CONCLUSION

From my analysis,

- Udemy has had its fair share of glory. From the data available, the peak was in 2015 after which it had a decline.
- This could be due to a surge of other learning platforms.
- The most popular course across all time, price and level is web development.
- So, if I am to put a course on Udemy it would be under web development.
- Price range of 20-50 USD tagged “All Levels” and between 10-12hrs long.

Thanks!

Any questions?

You can send me a mail

- oabass7@gmail.com