**REPORT**

Exploratory Data Analysis on Udemy Courses

Udemy is a massive online open course (MOOC) platform that offers both free and paid courses. Anybody can create a course, a business model by which allowed Udemy to have hundreds of thousands of courses.

**The purpose of this Exploratory Data Analysis is to determine these:**

1.How many courses on Udemy are paid and how are free?

2.What is the total number of subscribers, reviews and lectures on Udemy?

3.Which subject category has the highest subscribers, lectures, reviews and courses?

4.Which level has the highest subscribers and courses?

5.How many subscribers are paying and how many are not?

6.How many courses are paid or free in each subject?

7.Total sales earned by Udemy?

8.How much sales was made per subject?

9.Which level brought in the highest sales??

10.Is there a relation between price and the number of subscribers?

11.Is there a relation between price and the number of reviews?

12.Is there a relation between number of lectures and the number of subscribers?

13.Does the content duration have an effect on the number of subscribers?

**Data Preprocessing**

The analysis began with the cleaning of the data. After cleaning the data, we discovered that:

* There were no missing values.
* There were no duplicates.
* The dataset contained integers, objects, boolean and float datatypes.

**Data Examination**

Courses offered by Udemy can be grouped under 4 subject categories .These subject categories are Business Finance, Web Development, Musical Instruments and Graphic Design . These courses are offered in all levels, beginner level, intermediate level or expert level. Udemy has both paid and free courses. The price ranges of paid courses are between 20 dollars and 200 dollars.

**VISUALIZATIONS AND INSIGHTS**

On the subjects, we wanted to find the number of courses in each subject category. Based on this visualization we discovered that:

Web Development is the subject category with the highest number of courses with the number being **1200**, followed closely by Business Finance with **1195** then Musical Instruments with **680** and Graphic Design with **603** making it the subject category with the least number of courses.

Icon

Description automatically generated

Some courses on Udemy are offered free and some are paid.On further analysis , it came to light that **3368** courses on Udmey have to be paid before access whereas **310** courses are offered for free.

A picture containing bar chart

Description automatically generated

Breaking down the courses based on level, we discovered that 1929 courses are offered in all levels representing 52.4%,1270 (34.5%) courses in beginner level, 421(11.4%) in intermediate level and 58 (1.6%) in expert level. Observing this insight, we realized that more courses are offered in all levels whereas the expert level has the least courses.

Chart, pie chart

Description automatically generated

**Analysis on the number of subscribers in relation to other categories.**

Insights from the chart above shows that, Web Development has 7,980,572 subscribers making it the subject category with the highest number of subscribers .This might be due to many factors but primarily may be due to the surge of interest in website activities such as writing of blogs, creating the handling of organisational websites etc.

Bar chart

Description automatically generated with medium confidence

Chart, bar chart

Description automatically generated

Based on the chart above, it is observed that there are more subscribers subscribed to courses offered in all levels than in other levels. This may be because subscribers find it to be cost effective and easier to follow than to undertake the course in separate levels. There is also a notable number of subscribers in the beginner levels.

Upon analyzing the subscribers further, it is observed that 69.6% of the subscribers on Udemy are undertaking paid courses and 30.4% of the subscribers are undertaking courses that are for free. There are more subscribers undertaking paid courses because paid courses may be seen to be of more quality and credibility.

Chart, pie chart

Description automatically generated

Udemy earned approximately 885 million dollars from the paid courses. Breaking down this figure, we discovered that 71.3% of the sales came from Web Development making it the highest earning subject category on Udemy,14% of the sales from Business Finance ,8.7& from Graphic Design and 6% of the sales from Musical Instruments.

Chart, pie chart

Description automatically generated

Based on the visualization ,68.9% of sales earned came from courses offered in all levels making it the highest earning level followed by the beginner level with 24.8% .

Chart, pie chart

Description automatically generated

We analyzed to check if the content duration influenced the number of subscribers, it was discovered that content duration has significant influence on the subscriptions since there is a concentration of subscribers in range of 0-30hrs.

Chart, scatter chart

Description automatically generated

**Recommendations**

**Based on this analysis, we recommend that**:

* Udemy creates on develops more courses in Web Development since it has more subscribers and brings in the highest sales.
* In order to increase subscription with other subject category ,Udemy has to focus on offering most of the courses in all level since most subscribers prefer that.
* Although more subscribers are paying for the courses, a significant number are also subscribed to free courses therefore they can create some more courses for free and subsequently advertise the benefits of undertaken paid courses.
* To increase subscription to the other subject category apart from Web development, they can reduces the price slightly .
* The duration of content has an influence on the number of subscribers therefore Udemy can try to reduce the duration to 40 and below. This will be beneficial since their most profitable course has most of it subscribers signing up to contents with duration below 20 hrs.

**List of participants**

* Josephine Annan
* Emmanuel Mawuli Atitso