Excel Capstone Project

On

Analyzing Udemy Course Data

By

Temitope Fafure

Project Description

This Project aims to help an education tech company to track the performance of courses and help to understand where opportunities to increase revenue lies.

Datasets of courses covering four subjects (web-development, Business Finances, Musical instruments, and Graphic design) were given.

Each dataset contains twelve columns: course id, course title, URL of course, price of course, number of course, number of reviews, level, rating, content duration, published timestamp and subject.

Questions & hypothesis

1. Will Increasing the prices of web development courses increases the earning of the company?

Hypothesis: Yes, because web development courses is believed to be the most popular courses.

2. Does Web development courses have highest rating?

Hypothesis: Yes, because they probably have the highest subscriber

3. Based on level, which categories of courses are the most expensive?

Hypothesis: All-Level and expert category are the most expensive because they are more technical than other levels.

4. Which category of courses based on level will bring in more revenues?

Hypothesis: Beginner and Intermediate levels should bring in revenue across all subjects

5. Does the length of courses have an influence on the price?

Hypothesis: Yes, the longer the length of courses, the higher the price of the course and vice versa

Approach and Analysis

The following approaches and analysis were carried out to achieve the aim of the project:

1. The four datasets were formatted, compiled and cleaned into one Excel sheet file
2. Duplicates and blanks cells were removed.
3. Proper naming of headers and subject title "Web development"
4. Two columns were dropped. (Course id and URL of courses)
5. New columns were created
6. New column "Total revenue" was generated by multiplying prices of course with number of subscribers.
7. Statistical inference and graphical visualization will be employed using Pivot table and chart on Microsoft excel

Technical challenges

There was no challenge as the dataset were clean with about 1 percent missing data.