

Udemy Courses Descriptive Analysis

Fateme Zahra Alinejad Harikandei

810100420

Project Abstract

This project is about analyzing the features of Udemy courses in 4 categories Business, Design, Music and web development. We want to see what features in a course affects its number of reviews and ratings. We also can see what courses to watch in each category based on different features. We want to do an Exploratory data analysis on this dataset and understand it.

Dataset Description

The dataset is Udemy courses and contains information about courses in Udemy in four categories Business, Design, Music and web development. It has 11 columns:

- 1- course_title: The title of the Udemy course. (string)
- 2- url : The URL of the Udemy course. (string)
- 3- price: The price of the Udemy course. (float)
- 4- num_subscribers: The number of people who subscribed in the Udemy course. (integer)
- 5- num_reviews: The number of reviews for the Udemy course. (integer)
- 6- num_lectures: The number of lectures in the Udemy course. (integer)
- 7- level: The level of hardness of the Udemy course. (string)
- 8- Rating: The rating of the Udemy course. (float)
- 9- content_duration: The content duration of the course. (Float)
- 10- published_timestamp: The date when the course was published. (datetime)
- 11- subject: The subject of the course which is the four categories Business, Design, Music and web development. (string)

Why This Subject & Dataset

This Subject is important because it helps content creators to know what features affect the reviews and ratings of their content. Also it helps them to assign a reasonable price for their content or make a decision whether to charge for their course or not. It also helps people who want to buy a course or watch a course not to waist their time and money and shows the

important features which they should focus on. Udemy is one of the most popular websites to learn so exploring this dataset can be very helpful for our purposes.

The Questions

- 1- Is setting a price for a course affects its number of review and ratings? What happens if the price gets higher?
- 2- Which category has the highest mean of reviews and ratings?
- 3- How many courses are in each level of each category? Which level has the highest number of courses?
- 4- How does the level of course affect the price and number of reviews and ratings in each category?
- 5- How does the number of lectures and content duration affect the number of reviews and ratings?
- 6- How does the published date affect the number of views?
- 7- What is the relationship between number of reviews and rating?
- 8- Which Courses have the highest and lowest review and rating in each category?

Dataset URL

<https://www.kaggle.com/datasets/thedevastator/udemy-courses-revenue-generation-and-course-anal>