**Udemy Courses Analysis and Visualization**

**Problem Statement:**

The objective of this project is to analyze a dataset from an online learning platform similar to Udemy. The analysis aims to provide insights into various aspects of the dataset, including course subjects, levels, pricing, popularity among subscribers, and other relevant metrics. The goal is to understand trends and patterns within the dataset to inform strategic decisions related to course offerings, marketing strategies, and platform optimization.

**About Dataset:**

The dataset consists of information related to online courses available on the platform. It includes details such as course titles, subjects, levels (e.g., beginner, intermediate, advanced), pricing (paid or free), number of lectures, number of subscribers, number of reviews, and other relevant attributes. The dataset offers a comprehensive view of the courses offered on the platform and their respective metrics.

**Conclusion:**

**Data Exploration:** The dataset was initially explored by displaying the top 10 rows and checking the last 5 rows to understand its structure and contents.

**Data Shape:** The shape of the dataset, including the number of rows and columns, was determined to provide an overview of its size and dimensions.

**Data Information:** Information about the dataset, such as total number of rows, total number of columns, data types of each column, and memory requirement, was gathered to understand its composition and properties.

**Data Quality:** Null values in the dataset were checked to assess data completeness, and duplicate data was identified and dropped to ensure data integrity.

**Course Analysis:** The number of courses per subject and per level was determined to understand the distribution of courses across different subjects and levels.

**Pricing Analysis:** The count of paid and free courses was displayed to analyze the pricing strategy and distribution of courses based on pricing.

**Lecture Analysis:** The course with more lectures, categorized by free or paid, was identified to understand the depth and breadth of course content.

**Subscriber Analysis:** Courses with higher numbers of subscribers, categorized by free or paid, were determined to assess course popularity and subscriber engagement.

**Level Analysis:** The level with the highest number of subscribers was identified to understand the popularity of courses at different proficiency levels.

**Popularity Analysis:** The most popular course title and the 10 most popular courses based on number of subscribers were identified to highlight top-performing courses.

**Review Analysis:** The course with the highest number of reviews was identified, and the impact of price on the number of reviews was analyzed to understand subscriber feedback and engagement.

**Python Courses:** The total number of courses related to Python and the 10 most popular Python courses based on number of subscribers were determined to assess the popularity and demand for Python-related content.

**Yearly Analysis:** The year with the highest number of courses posted was identified, and category-wise count of posted subjects for each year was displayed to understand temporal trends in course offerings.