

Course End Project Report
ON
ANALYSIS OF UDEMY COURSES
B.E.(IT) **IV-Sem**

BY
P. JAYADIR (160122737191)
CH. SHIVA KRISHNA (160122737173)
N. PRAVARDHAN (160122737188)

UNDER THE GUIDANCE OF

Dr. N. Sudhakar Yadav
Associate Professor,
IT Department, CBIT.



DEPARTMENT OF INFORMATION TECHNOLOGY
CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY (A)
(Affiliated to Osmania University; Accredited by NBA and NAAC, ISO 9001:2015 Certified
Institution), GANDIPET, HYDERABAD – 500 075
Website: www.cbit.ac.in

ABSTRACT

In our project "ANALYSIS OF UDEMY COURSES " we delve into the vast dataset provided by Udemy, one of the largest online learning platforms. With over thousands of courses spanning diverse subjects, Udemy's dataset offers a rich landscape for exploration and analysis. Our objective is to extract valuable insights and patterns from this dataset using Python libraries and data analysis techniques.

Through meticulous data preprocessing, exploratory data analysis, and visualization, we aim to uncover trends in course popularity, pricing dynamics, user engagement metrics, and more. By analyzing factors such as course category, price, number of subscribers, and reviews, we seek to understand the underlying patterns and behaviors of Udemy users.

The outcomes of our analysis will provide educators, course creators, and platform administrators with actionable insights to optimize course offerings, pricing strategies, and user experience. Additionally, we aim to identify potential gaps in course content and areas for improvement to enhance the overall learning experience on Udemy. Ultimately, our project aims to contribute to the enhancement of online learning platforms and empower learners worldwide with access to quality education.

OBJECTIVES AND OUTCOMES

Objectives:

- Analyze Udemy course dataset to understand course trends and user behaviour.
- Identify factors influencing course popularity, pricing, and enrolment.
- Visualize insights for easy interpretation and decision-making.

Outcomes:

- Insights into popular course categories, pricing strategies, and user engagement metrics.
- Recommendations for course creators, learners, and Udemy administrators based on analysis findings.

Dataset Collection & Preparation

The dataset for our project, "Udemy Course Dataset Analysis," was obtained from Kaggle, a renowned platform for sharing datasets and conducting data science competitions. Kaggle hosts a diverse range of datasets contributed by the community, including datasets from various domains such as finance, healthcare, education, and more.

To collect the Udemy course dataset, we utilized Kaggle's platform, where the dataset was uploaded by a user or a group of contributors. The dataset likely contains information about Udemy courses, including attributes such as course title, description, price, number of subscribers, reviews, duration, level, and more.

Upon downloading the dataset from Kaggle, we performed essential data preparation steps to ensure its suitability for analysis. This involved tasks such as data cleaning, handling missing values, removing duplicates, and converting data types as needed.

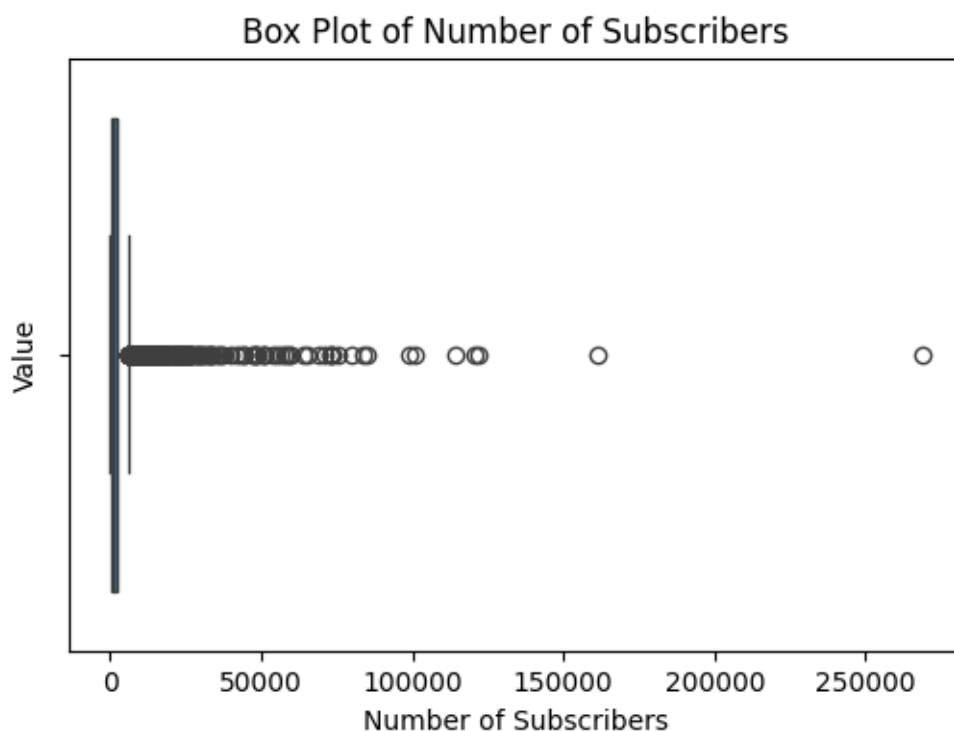
Additionally, we conducted exploratory data analysis (EDA) to gain initial insights into the dataset's structure and contents.

By leveraging Python libraries such as Pandas, NumPy, and Matplotlib, we facilitated dataset manipulation, visualization, and analysis. These tools provided a robust framework for preprocessing the dataset, conducting exploratory analysis, and preparing the data for further investigation.

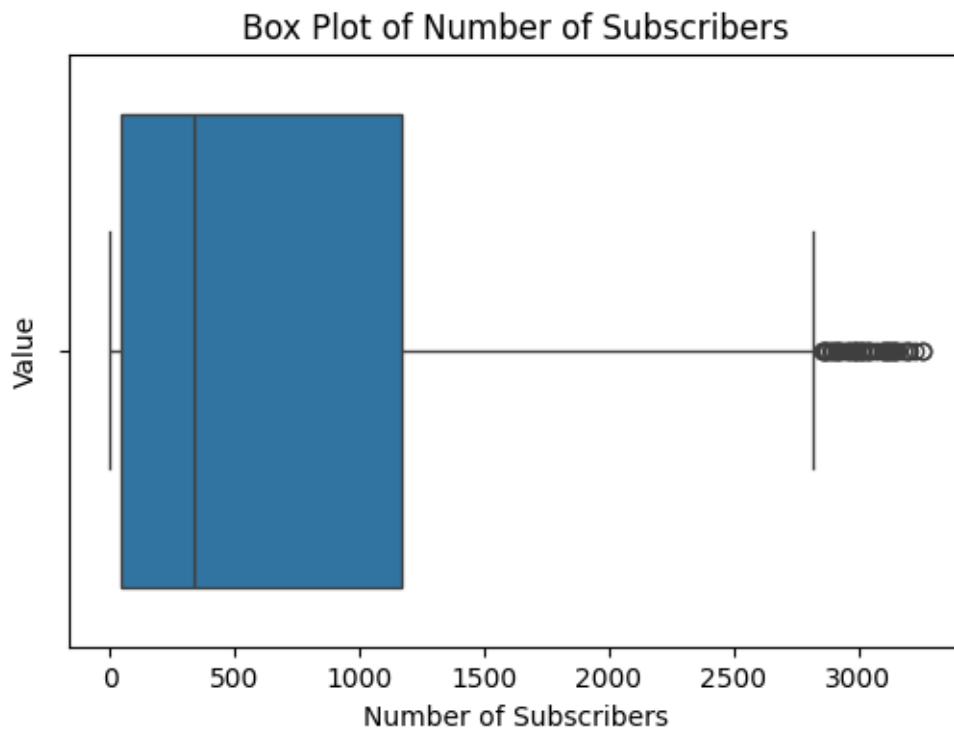
Overall, the dataset collection and preparation phase laid the foundation for our subsequent analysis, enabling us to extract meaningful insights and draw conclusions about Udemy courses' characteristics and trends.

HANDLING OF OUTLIERS:

With outliers:



Without outliers:



Design of Architecture / Model / Solution / Procedure / Algorithm / Mathematical Model etc

- Exploratory Data Analysis (EDA): Understanding dataset structure, summary statistics, and initial insights.
- Feature selection: Identifying relevant features for analysis based on project objectives.
- Data visualization: Creating visualizations such as histograms, scatter plots, and heatmaps for insights.

Results Visualization, Analysis & Conclusion

