udemy


# PREDICTING ONLINE CORSES RATING

By: Ruqaiah Abdullah and Alanoud Alzhrani

# Outline

- Introduction
- Dataset
- Target
- EDA

# Outline

- Data Split
- Regression Models
- Best Model
- Conclusion


# Introduction

udemy

- American educational platform

- Founded in May of 2010

- Contains many categorical fields
(Development-design-business-marketing-technology)


# In 2021,

- 44 million students
- 183,000 courses
- 65 instructors teaching courses in 75 languages
- Over 594 million courses enrollments

# Dataset

Course ID

Course Title

Enrollment

Subscribers

Reviews

Stars

Rating

Published Lectures

Published Time

Category

Course Price

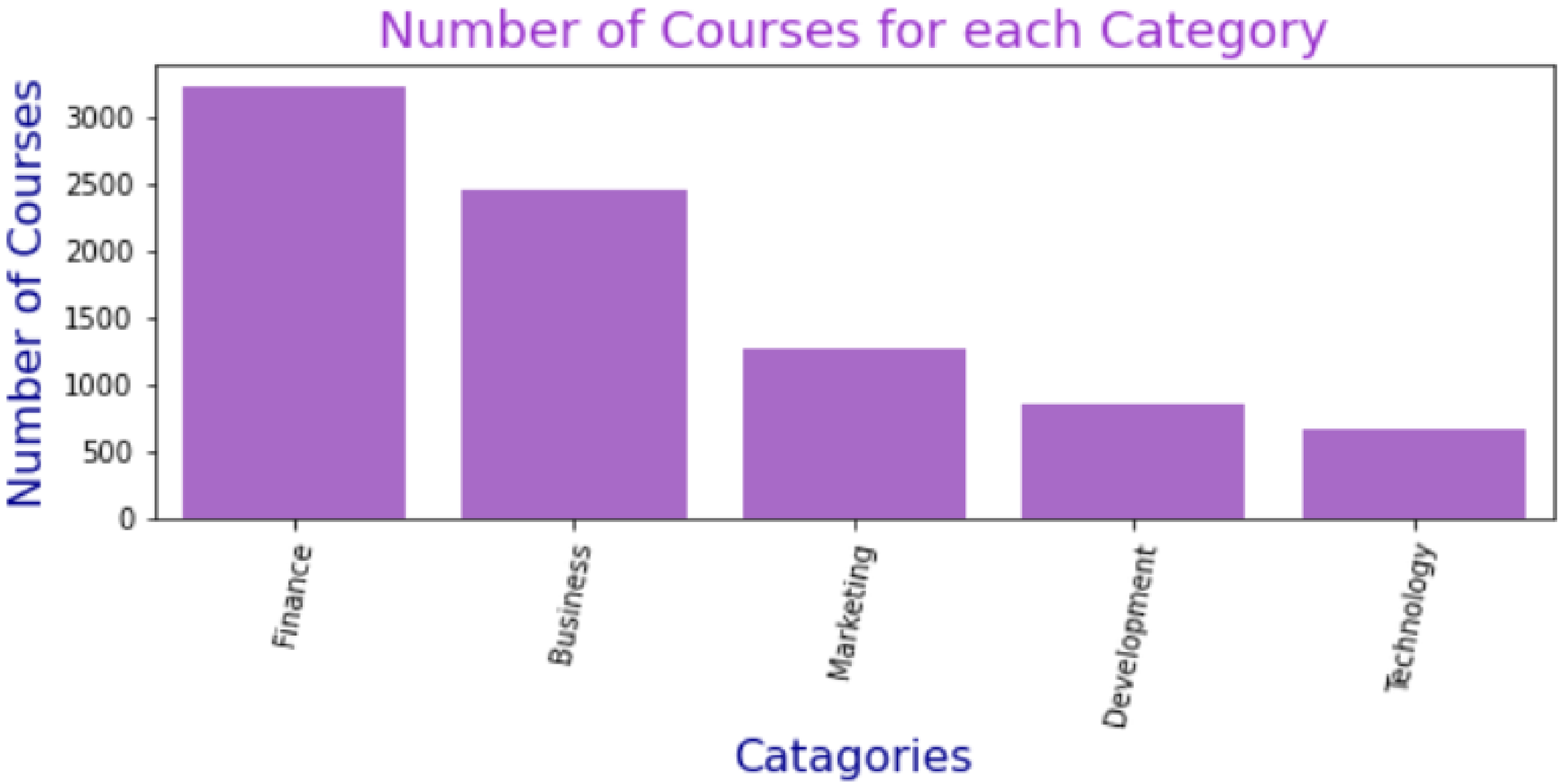
Course Discount

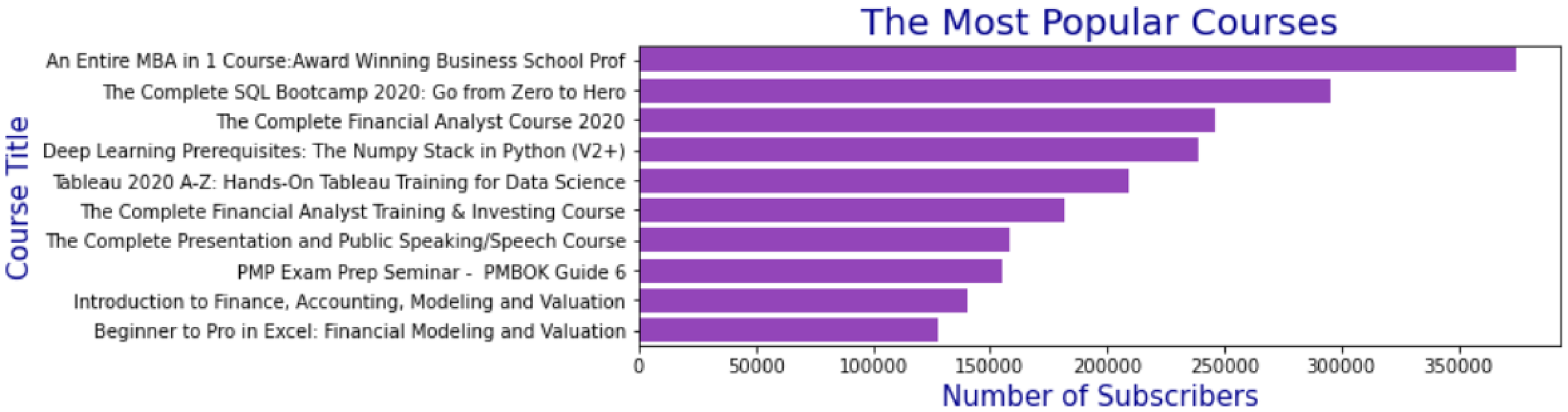
Price Curenncy

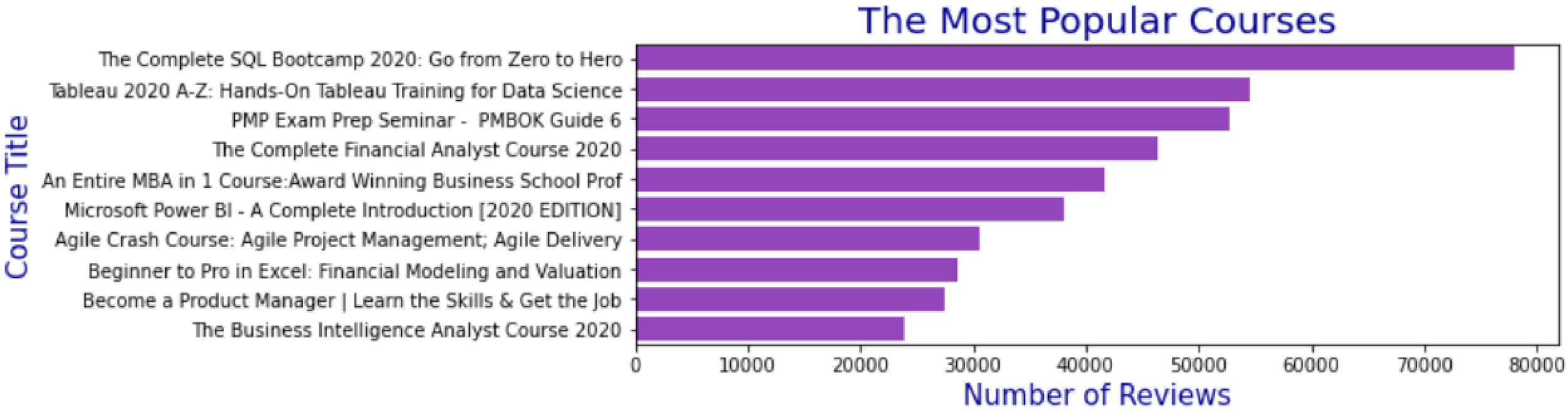
# Target

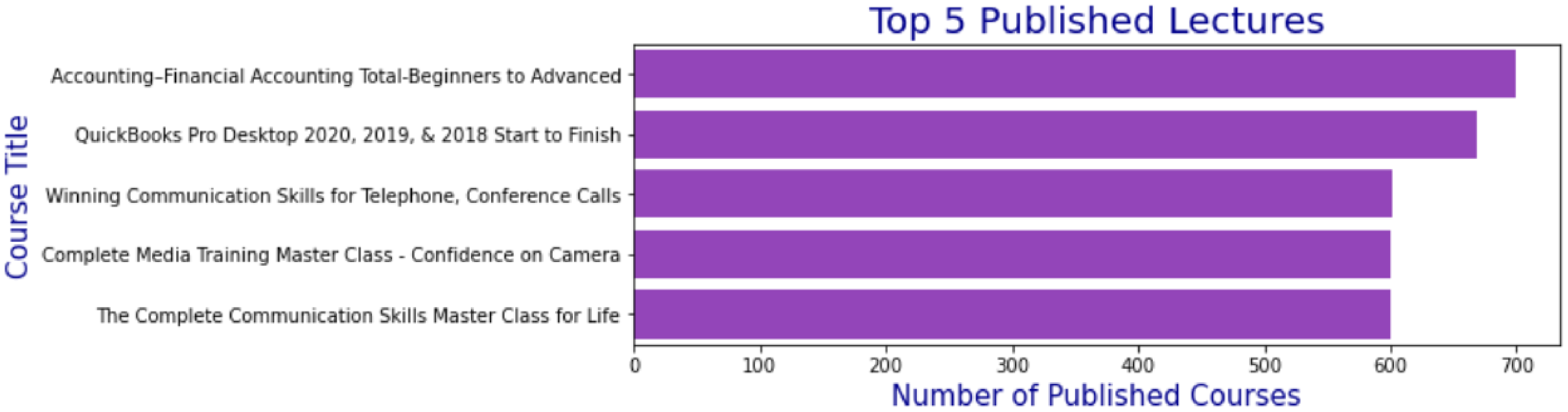
## Predicting the number of people who raited each online courses

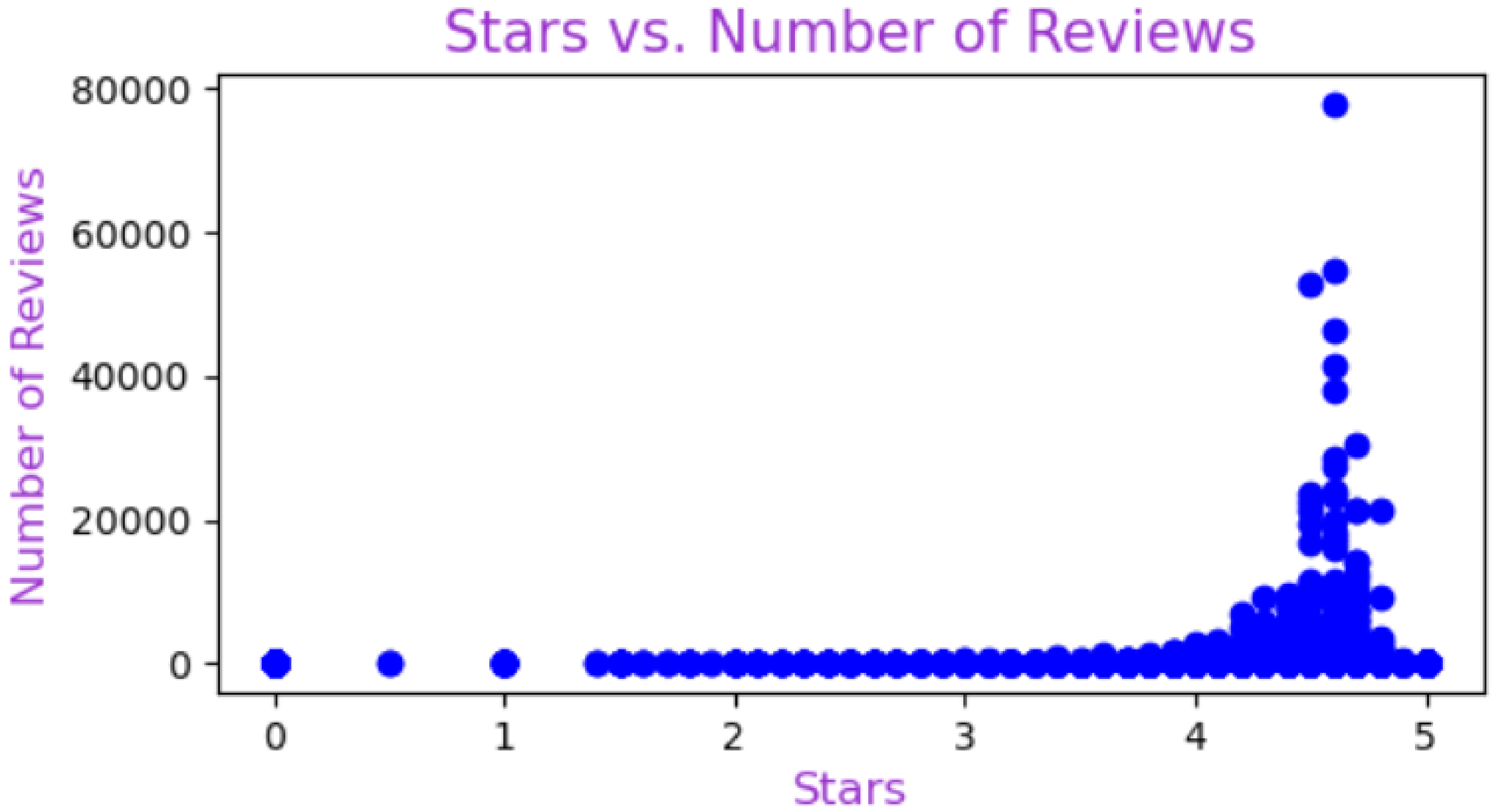
# EDA

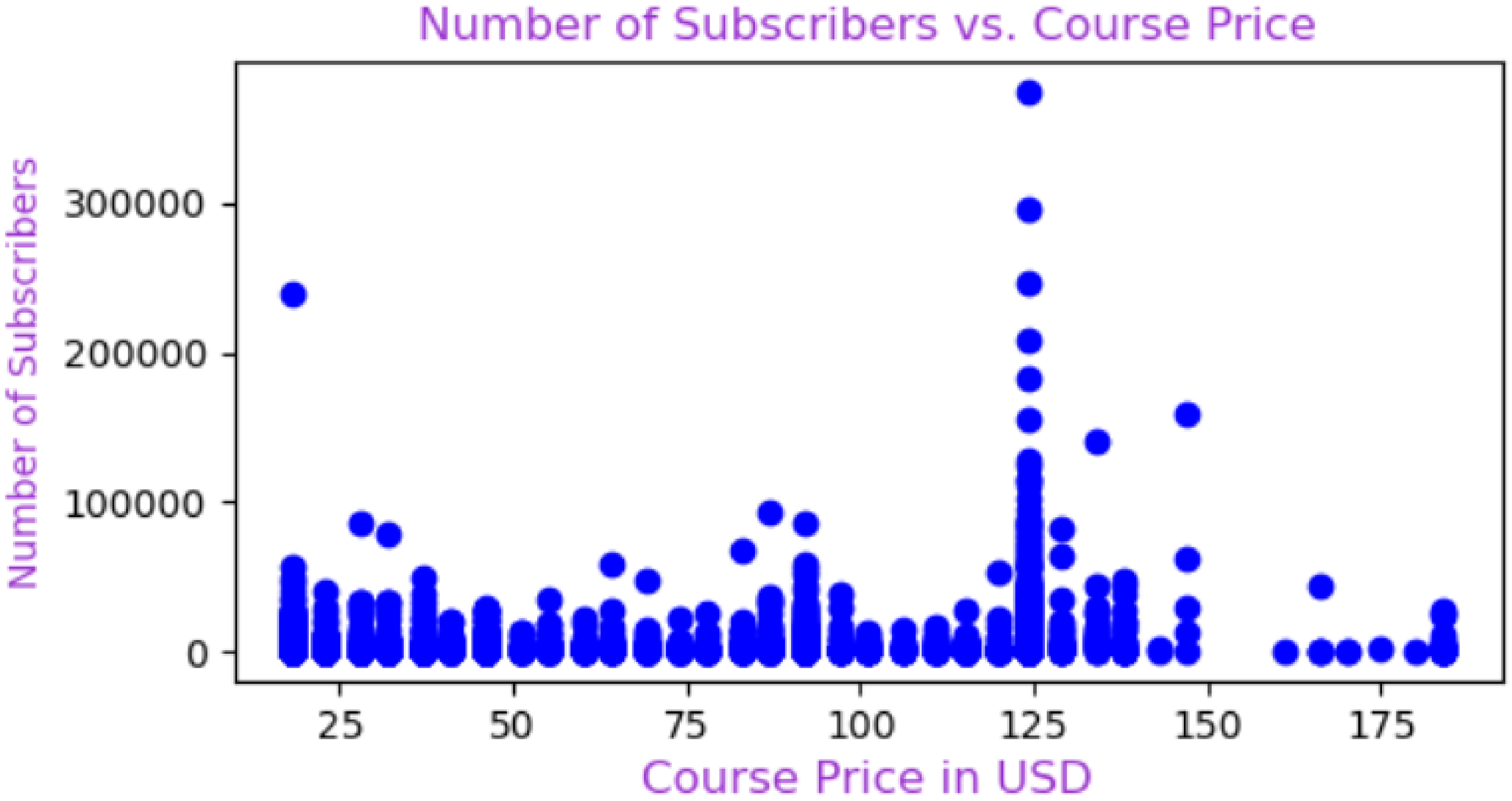

Number of Courses for each Category
Number of Courses
3000
2500
2000
1500
1000
500
0
Finance
Business
Marketing
Development
Technology
Catagories


The Most Popular Courses
An Entire MBA in 1 Course:Award Winning Business School Prof
The Complete SQL Bootcamp 2020: Go from Zero to Hero
The Complete Financial Analyst Course 2020
Deep Learning Prerequisites: The Numpy Stack in Python (V2+)
Tableau 2020 A-Z: Hands-On Tableau Training for Data Science
The Complete Financial Analyst Training & Investing Course
The Complete Presentation and Public Speaking/Speech Course
PMP Exam Prep Seminar - PMBOK Guide 6
Introduction to Finance, Accounting, Modeling and Valuation
Beginner to Pro in Excel: Financial Modeling and Valuation
Course Title
0
50000
100000
150000
200000
250000
300000
350000
Number of Subscribers

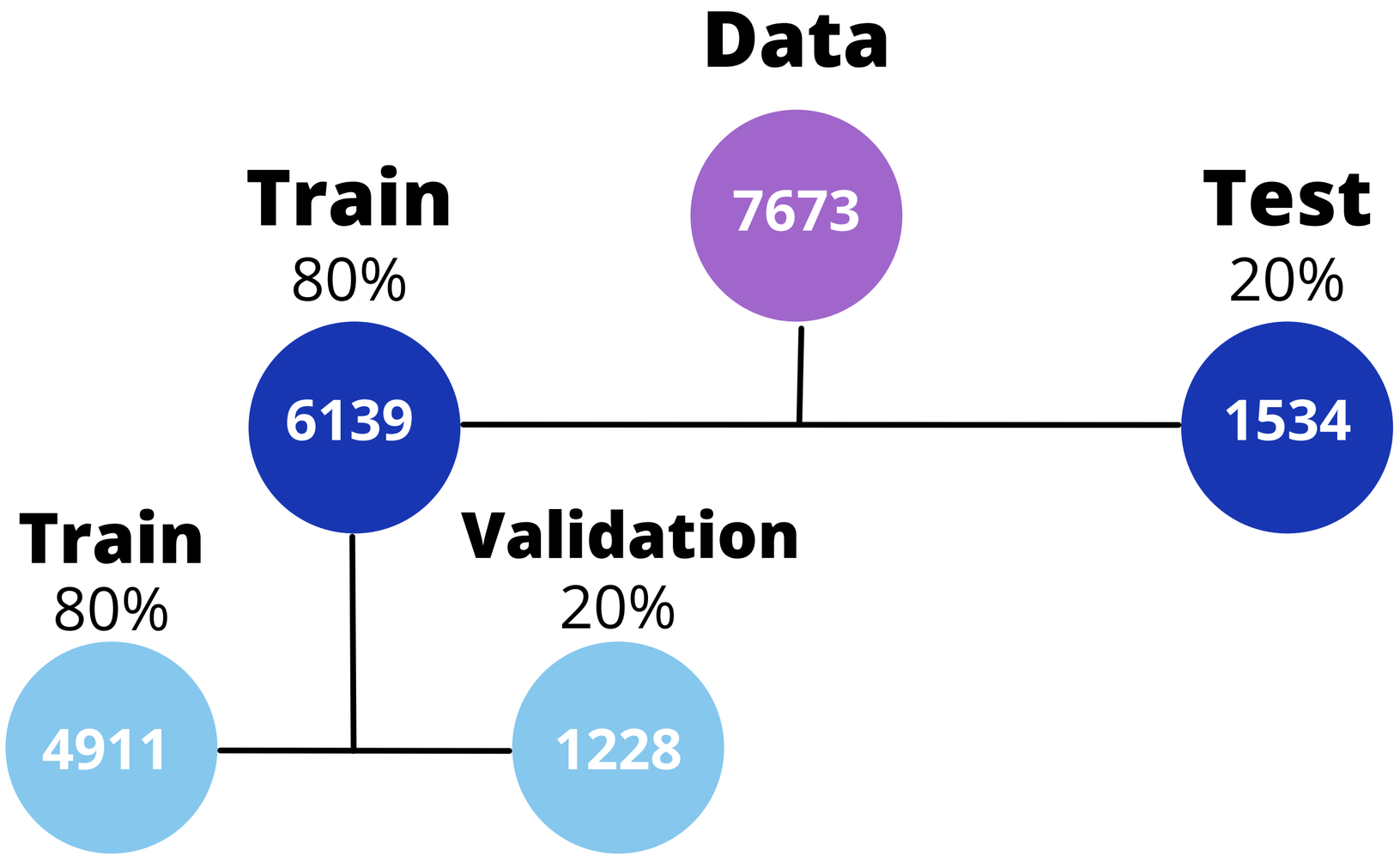

The Most Popular Courses
Course Title
The Complete SQL Bootcamp 2020: Go from Zero to Hero
Tableau 2020 A-Z: Hands-On Tableau Training for Data Science
PMP Exam Prep Seminar - PMBOK Guide 6
The Complete Financial Analyst Course 2020
An Entire MBA in 1 Course:Award Winning Business School Prof
Microsoft Power BI - A Complete Introduction [2020 EDITION]
Agile Crash Course: Agile Project Management; Agile Delivery
Beginner to Pro in Excel: Financial Modeling and Valuation
Become a Product Manager | Learn the Skills & Get the Job
The Business Intelligence Analyst Course 2020
0
10000
20000
30000
40000
50000
60000
70000
80000
Number of Reviews


Top 5 Published Lectures
Course Title
Accounting–Financial Accounting Total-Beginners to Advanced
QuickBooks Pro Desktop 2020, 2019, & 2018 Start to Finish
Winning Communication Skills for Telephone, Conference Calls
Complete Media Training Master Class - Confidence on Camera
The Complete Communication Skills Master Class for Life
0
100
200
300
400
500
600
700
Number of Published Courses


Stars vs. Number of Reviews
Number of Reviews
80000
60000
40000
20000
0
0
1
2
3
4
5
Stars


Number of Subscribers vs. Course Price
Number of Subscribers
300000
200000
100000
0
25
50
75
100
125
150
175
Course Price in USD

# Data Split


Data
7673
Train
80%
6139
Test
20%
1534
Train
80%
4911
Validation
20%
1228

# Regression

| Regression Model | Train Score | Validation Score |
|---|---|---|
| Baseline Model | 0.827 | 0.833 |
| Basic Feature Engineering (Polynomial terms) | 0.847 | 0.948 |
| Basic Feature Engineering (Interaction terms) | 0.915 | 0.936 |
| Intermediate Feature engineering (dummy variables method) | 0.934 | 0.953 |

# Best Model

| Regression Model | Train Score | Validation Score |
|---|---|---|
| Baseline Model | 0.827 | 0.833 |
| Basic Feature Engineering (Polynomial terms) | 0.847 | 0.948 |
| Basic Feature Engineering (Interaction terms) | 0.915 | 0.936 |
| Intermediate Feature engineering (dummy variables method) | 0.934 | 0.953 |

# Conclusion

We tried several regression models to predct the best model for our target which is Predicting the number of people who ratied each course

**Best Model:** feature engineering dummy variables method

**Best Score**

**0.953**

# Thanks!