Data Analysis Portfolio

I am a Data Scientist Aspirant. I'm familiar with basics of Pandas, Numpy, Data Visualization, Basics of Tableau and Excel.

Professional Background

I am pursuing 6th semester of Computer Science Engineering.

I am familiar with Java, C++, Python, C, HTML5, CSS3, JavaScript, SQL. I completed my first internship as a Data Science & Business Analytics Intern at The Sparks Foundation (GRIP) and completed a task on Supervised Machine Learning. I had learned the basics of Pandas, Numpy, Basic Data Visualization & few concepts of Machine Learning in the internship.

My Mini project is a Web Application titled as "FITNESS GUIDE".

Fitness Guide helps to take care of health based on BMI.
This application takes the height and weight of a person and gives BMI (Body Mass Index). Based on BMI(Low, Medium, High), one is given a proper diet and appropriate workouts to maintain a normal BMI. This application performs information collection and analysis.
Technology : HTML5, CSS3, JavaScript, Bootstrap, PHP, SQL

Table of Contents

Udemy Project Description

Given Problem statement

“ You’re a Data Analyst working for the education tech company Udemy. You have been asked by your manager, Head of Curriculum at Udemy, to present the data on course revenue, and you have been provided with data on courses from different topics to understand where opportunities to increase revenue may lie, and track the performance of courses.

Your manager has suggested encouraging Web Development courses to charge more because she believes that these are the most popular courses. She needs to send a report to the CEO in the next three weeks on how they will increase their next quarterly earnings ”

Situation: I was given a task to present data on course revenue and was provided with data on courses from different topics to understand where opportunities to increase revenue may lie, and track the performance of courses.

Task: The purpose of the project is to present revenue details of the courses. I had to find insights from the data given by using methods and narrate it.

Action:

Outline Actions

- Understand the type of data given.
- Perform basic Data Cleaning and Use basics Visualization techniques to understand data.
- Methods used are Data consolidation, Data Cleaning, Data Manipulation & Wrangling, Data Visualization
- Narrate a story to convey my insights.

Result: Web Development & Business Finanace Courses give good revenue where Web development stands first . Business Finance and Graphic Design stood second and third place respectively

The Problem

Business Problem: To present data on course revenue

Timeline: I was given 3 three weeks to work on this project and submit the report

Questions that I asked to understand the business problem

- What exactly do you want to find out?
- Which data sources are available to work with and of what type?
- Will I be provided with previous trends in your data? Like previous visualizations or insights you had.
- How will the results be used?
- Do you currently analyze any of your data? If yes, what tools do you do and how often do you look at them?

Design

The given data is in structured format(.xlsx, .csv). The data was spread upto 15000 rows and atmost 15 columns.

Steps that were performed for Data Cleaning

1. Remove Duplicates
   I have found many duplicate rows so I selected the entire data sheet data to remove duplicates.
2. Removing blank cells
   Removed Blank cells using Create filter and filter by condition option
3. Headers
   I ensured to have clear and concise names for headers and used dashes or underscores in between words to make it easier to parse later on.
4. Used Find and replace function to make subject data consistent
5. Used Pivot Tables and charts in Excel for Visualizations
6. I used Tableau to create additional data visualizations that will support my findings in the previous steps.
7. At last created Dashboards and stories to convey my insights.

Findings

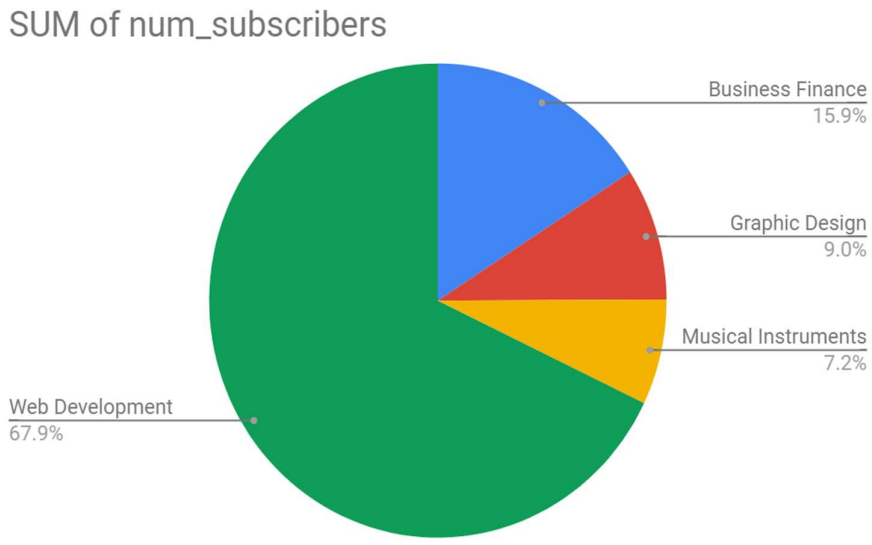
Finding 1

Figure 1: Total Subscriptions by Category

I created a Pie Chart for Total Subscritptions by Category to find out category with highest subscriptions. I found that Web Development Category stood first.

Table 1: Subject vs Sum of num_subscribers

subject	SUM of num_subscribers
Business Finance	1868711
Graphic Design	1063148
Musical Instruments	846689
Web Development	7981935
Grand Total	**11760483**



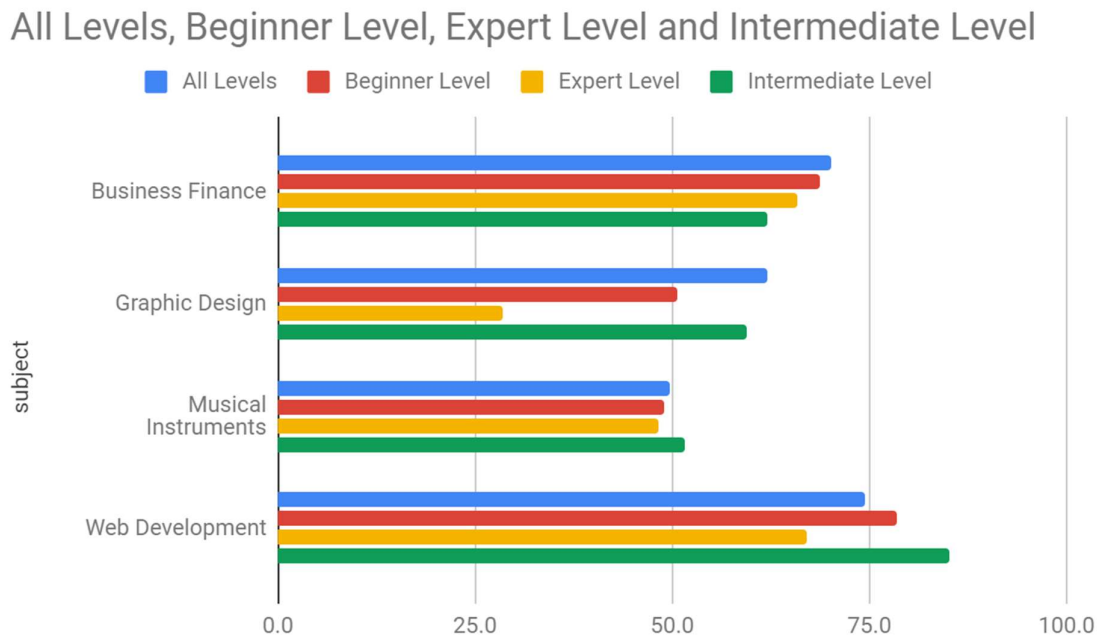
Finding 2

To find Average price per level I plotted a bar chart
I found that Web development stood first in all levels with highest average price.
Business Finance stood second.

AVERAGE of price	*level*				
subject	All Levels	Beginner Level	Expert Level	Intermediate Level	Grand Total
Business Finance	70.2	68.7	65.8	62.0	68.7
Graphic Design	62.1	50.7	28.6	59.4	57.9
Musical Instruments	49.6	49.0	48.3	51.6	49.6
Web Development	74.5	78.5	67.1	85.1	77.0
Grand Total	**66.8**	**65.2**	**58.0**	**66.9**	**66.1**

Table 2: Average of Price vs level

Figure 2: Average Price Vs Course and Level

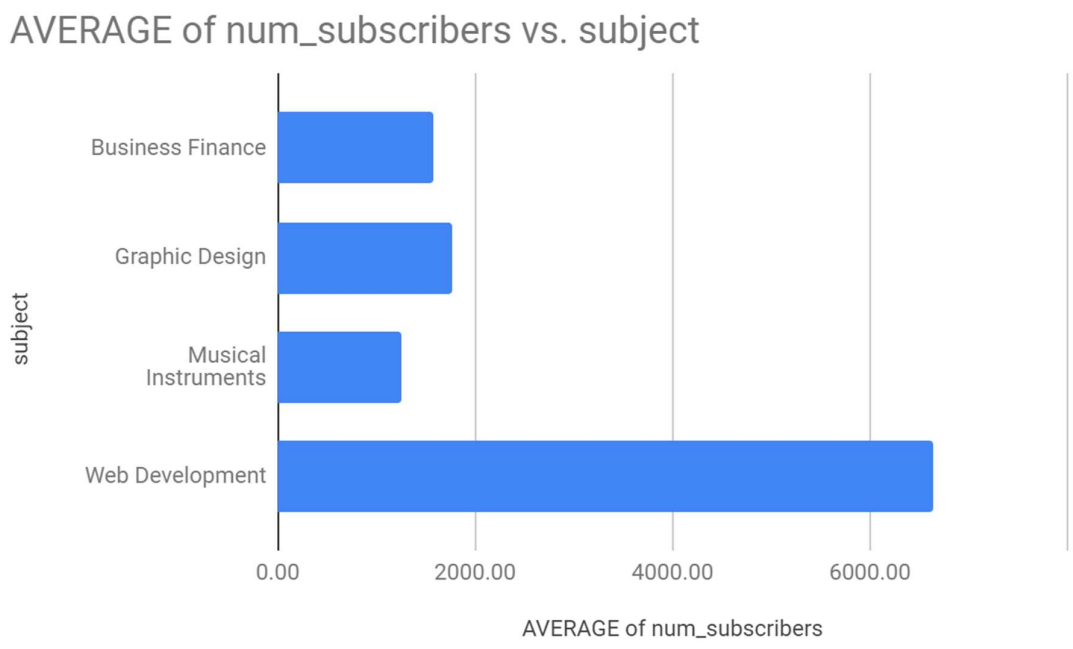


Finding 3

Table 3: Average Subscribers by Subject

subject	AVERAGE of num_subscribers
Business Finance	1569.03
Graphic Design	1766.03
Musical Instruments	1245.13
Web Development	6635.02
Grand Total	**3199.26**

Figure 3: Average Subscribers by Course

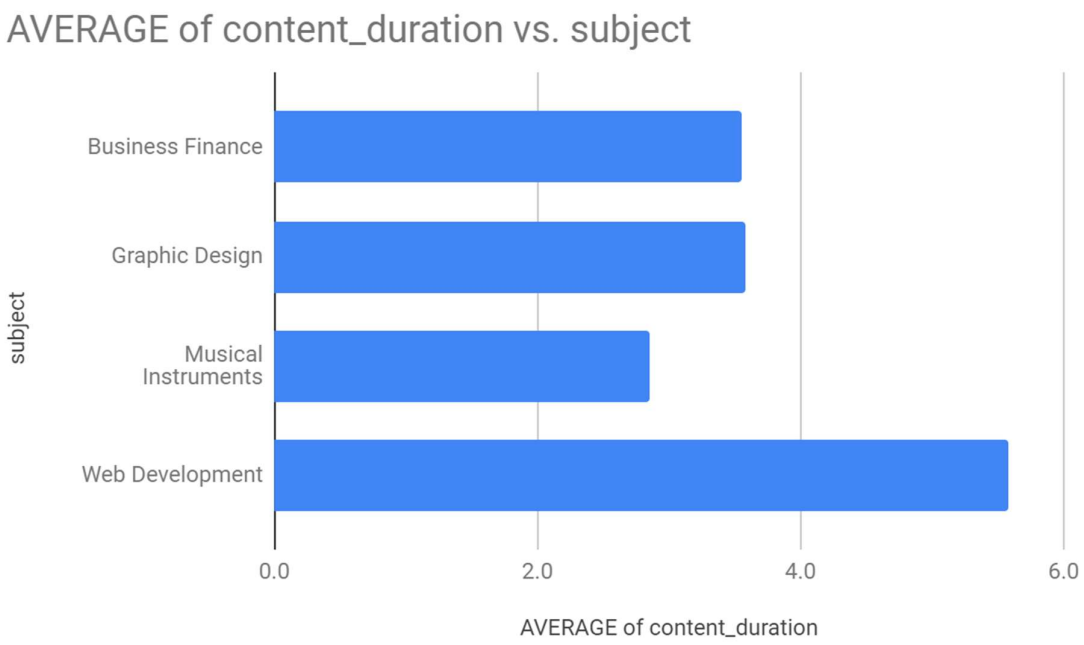


Finding 4

Table 4: Average Content Duration by Subject

subject	AVERAGE of content_duration
Business Finance	3.6
Graphic Design	3.6
Musical Instruments	2.9
Web Development	5.6
Grand Total	**4.1**

Figure 4: Average Content Duration Vs subject

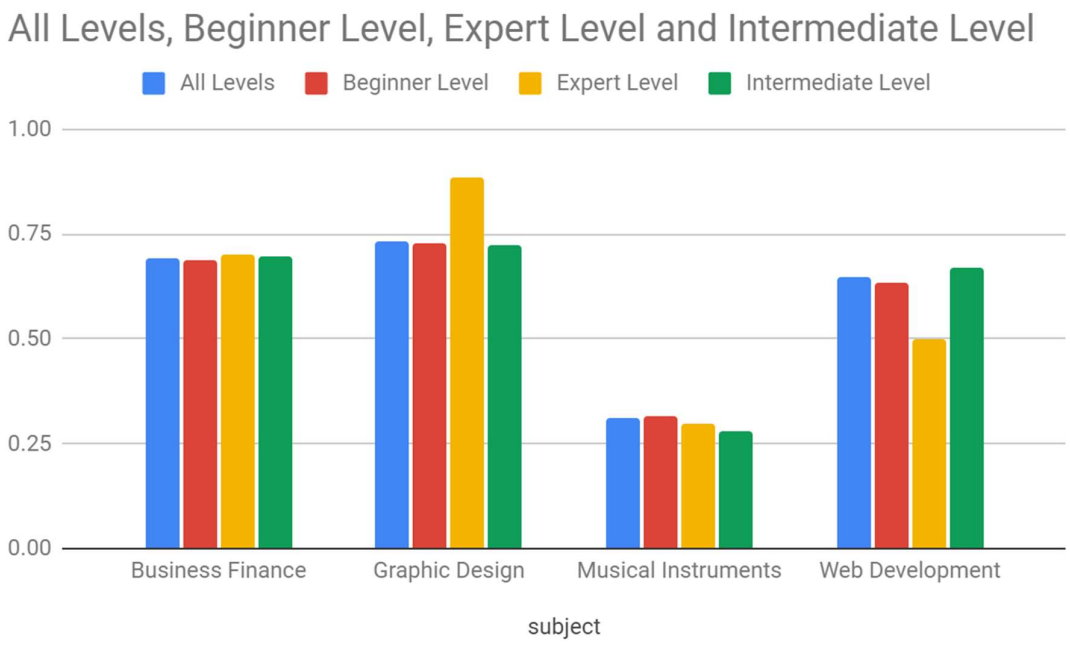


The Average content duration of Web Development is the highest and Graphic Design took second place

Table 5: Average Rating by Subject and Level

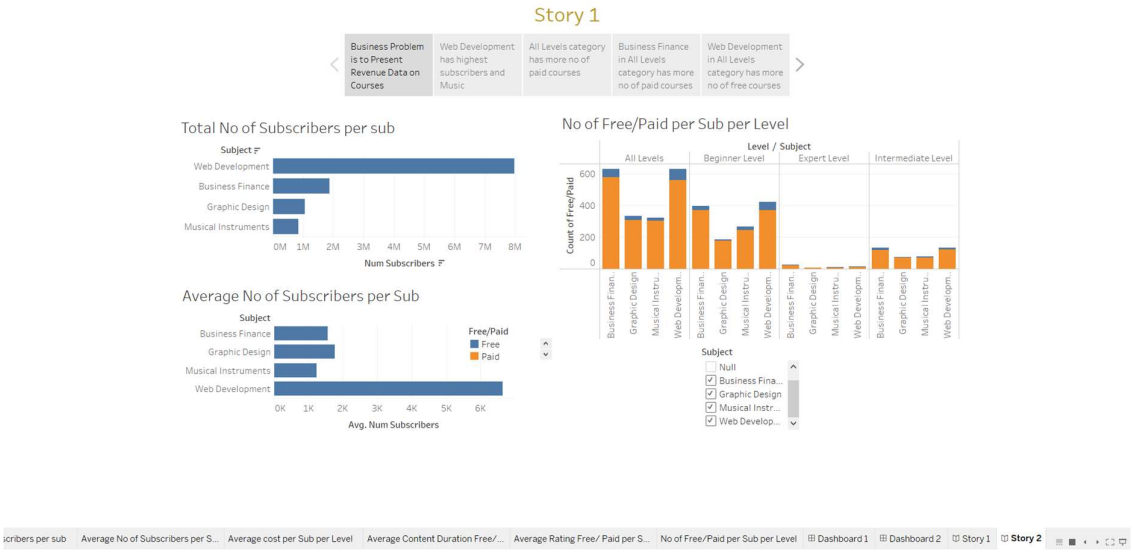
AVERAGE of Rating	*level*				
subject	All Levels	Beginner Level	Expert Level	Intermediate Level	Grand Total
Business Finance	0.69	0.69	0.70	0.70	0.69
Graphic Design	0.73	0.73	0.88	0.72	0.73
Musical Instruments	0.31	0.31	0.30	0.28	0.31
Web Development	0.65	0.64	0.50	0.67	0.64
Grand Total	**0.62**	**0.60**	**0.59**	**0.62**	**0.61**

Figure 5: Average Rating by Subject and Level



Graphic Design has the highest Average Ratings in all given levels

Analysis
Tableau Story(Narrative)



Conclusions

Web Development stood first and has major share in revenue.
Business Finance would give good revenue and stood second.
Graphic Design has Good ratings and its price is between Business Finance and Web Development. It stood third in having major share of revenue.
Graphic Design courses can be increased to get more revenue. This is because it has the best ratings than others, price less than Web Development. The reason for Graphic Design being third in contribution to revenue is, its has less no. of courses(paid).