For the project, there was a provision of a "Course-wise" dataset with aggregation on the grounds of number of subscribers, duration, level, price, reviews, lectures, rating and the date of publish of the course.

The main background of the given project was to find out why and where a product is lacking and where work could be done which would help in increasing revenues and tracking the performance of courses.

Design

The dataset being scattered and unsorted following steps were considered to clean the data in order to gain some insights-

->All the empty cells were deleted

->The date on which they were published was extracted from published_timestamp column using LEFT FUNCTION.

->Using IF FUNCTION and modified IF FUNCTION the category of each course was found i.e. whether the course is free/paid and depending on the output of the whether the course is free or paid a column returning answer in yes/no for the course being free was created.

->Next, the courses with top 20 subscribers were found using sorting the subscribers column and the top 20 courses were compiled into a separate sheet.

->Next, using the VLOOKUP FUNCTION the level, category, duration, date & the subject of each of the Top 20 courses was found.

->Pivot tables were then created for feature represent all of the important features that could act as an aid to potential insights.

->To represent subject wise subscribers a pie-chart was used as it represents part-whole relationships the neatest and the best.

-> To represent average number of subscribers per subject a bar chart was used as it efficiently represents comparative metrics and are super easy to read.
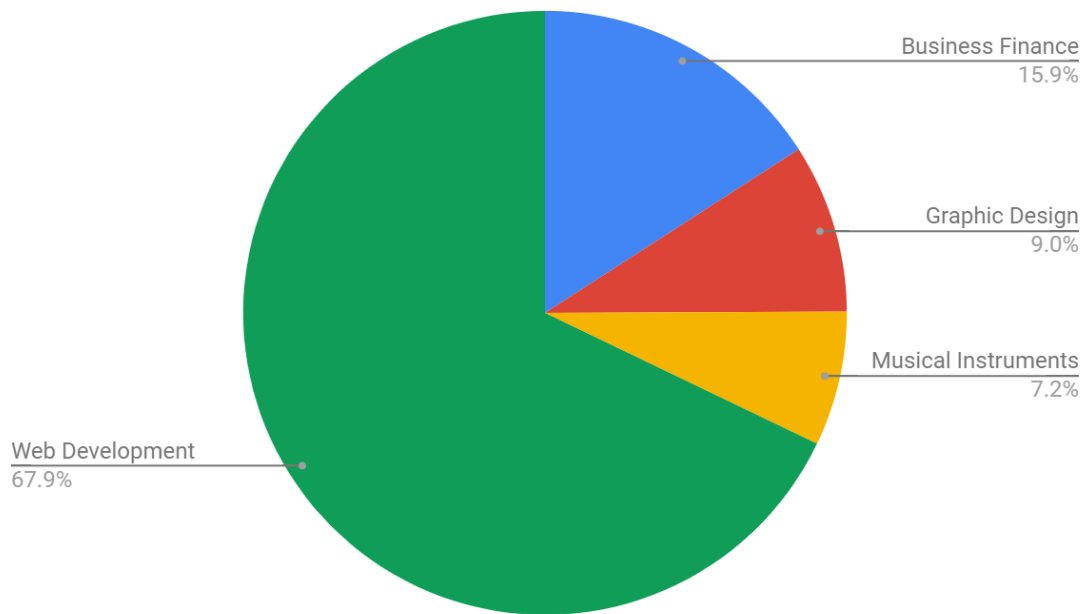
-> To find average cost per subject at each level a bar chart gave a pellucid outlook as to how each course corresponds with respect to each other.

->Finding average content duration for each subject was made simple using a bar chart.

-> Average rating per subject for each level was visualized using a column chart as it is known to be the best way to plot several items in a specific range of values. Moreover, it is super easy to compare column lengths to draw conclusions.
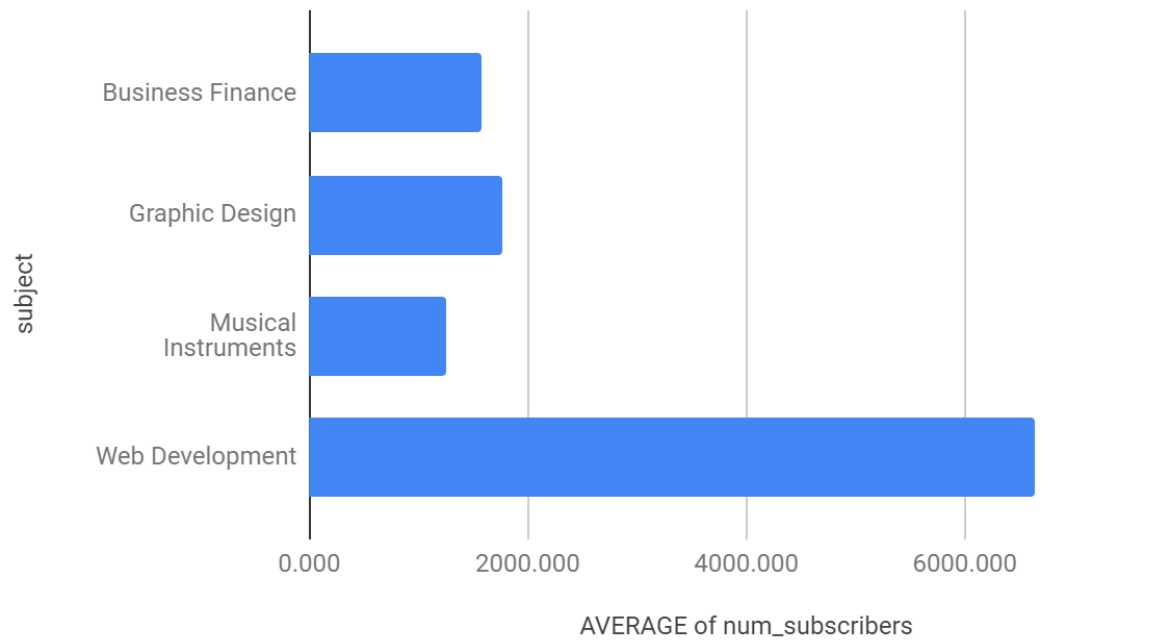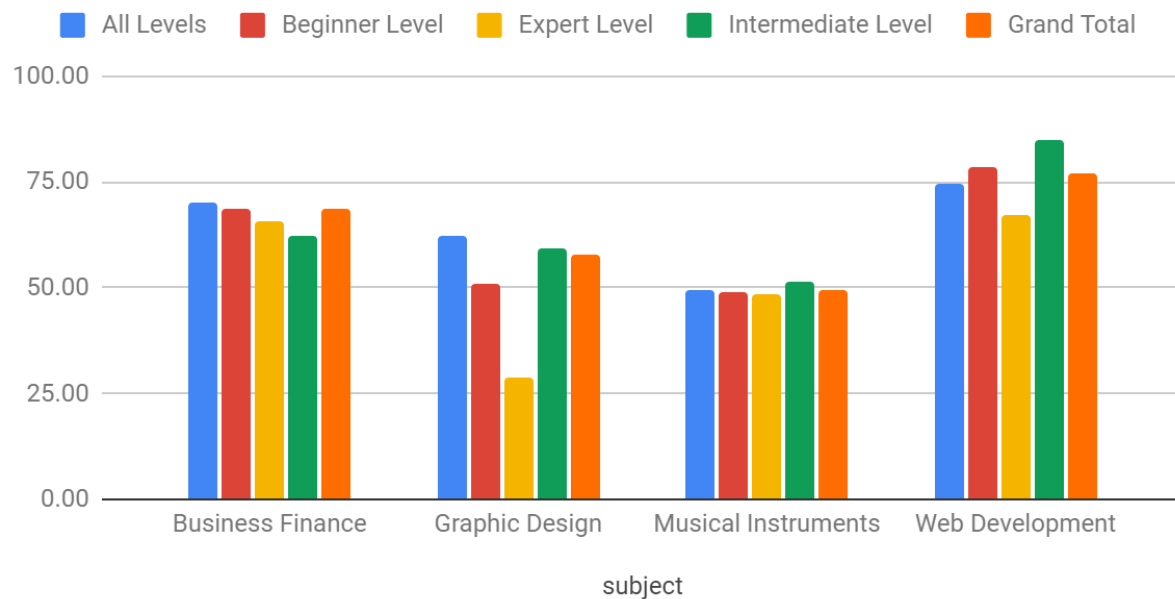
FINDINGS

SUM of num_subscribers

Business Finance
15.9%

Graphic Design
9.0%

Musical Instruments
7.2%

Web Development
67.9%

The Web Development subject had the maximum number of subscriber following Business Finance, Graphic Design and Musical Instruments respectively.

AVERAGE of num_subscribers

Here again, Web Development leaded chart by having the maximum average number of subscribers per subject.
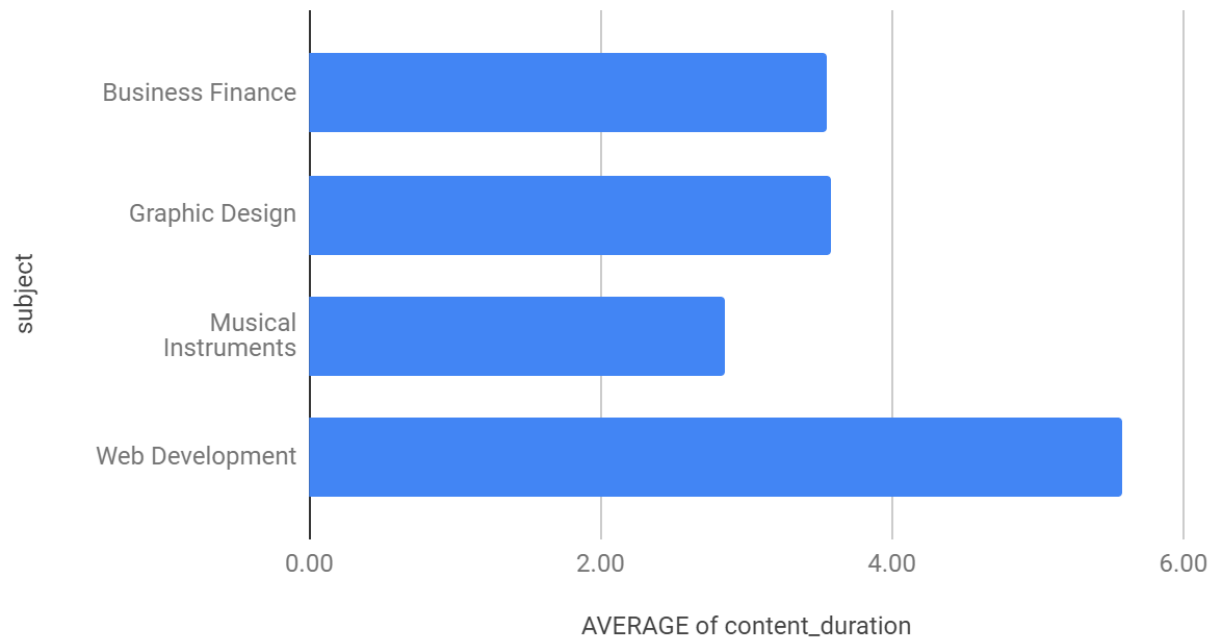
All Levels, Beginner Level, Expert Level, Intermediate Level and Grand Total
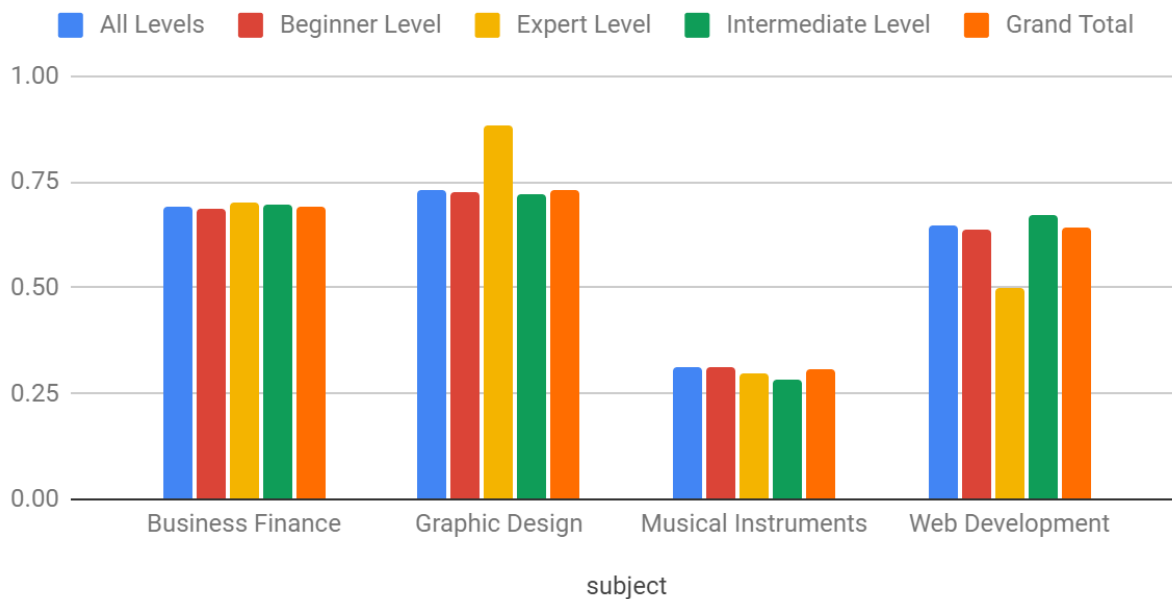

*subject vs average cost per subject

Web development even after having the maximum average price among all other competitors was the most opted subject in all levels.

AVERAGE of content_duration vs subject

Web Development courses gave out the maximum throughput in terms of the duration of the content.

All Levels, Beginner Level, Expert Level, Intermediate Level and Grand Total

■ All Levels ■ Beginner Level ■ Expert Level ■ Intermediate Level ■ Grand Total

*Average rating per subject for each level

This is where the Web Development courses lacked in all levels to a small metric.

ANALYSIS

Web Development courses were the top scorers when it came to number of subscribers. They also provided courses with maximum duration of content.

Web Development also turned out be one of the most expensive domains for the courses when it came into comparison wit other domains and had rating lesser than its peers under most of the levels of courses. Also, it turned out to have rating less than other domains and courses under certain levels.

Using Root Cause Analysis following could be the reasons why Web Development failed in excelling under ratings-

-> Web Development being a popular domain it gained a lot of subscribers.

->Maybe the presence of too long duration of courses bored out learners and led to poor ratings.

->People have spent the maximum money buying Web Development course and maybe the content not being up to the expectations considering the amount of money they have spend could be a setback for users and eventually gained a poor rating.

->The courses could be expensive as it is in demand following the fact that the content is ought to be complex, users not being from the background could be a reason why they were not able to understand the concepts properly which made them give a poor rating to their respective course.

->Inactive tech support could be major reason, if the TS wasn't active enough in solving coding related queries it literally led to poor ratings.

Conclusion

Web Development domain needs some improvement in context to whatever they are lacking and this could be found out only by going through the system.

It attracted maximum number of subscribers due the popularity of the domain. Other courses might have a smaller interest group .But if we talk about other aspects of domains other than Web Development they are doing just fine considering the number of subscribers each domain has.