



UDEMY COURSES SUBSCRIPTIONS ANALYSIS

[GROUP 4]

OVERVIEW



- Udemy is a **Massive Online Open Course (MOOC)** platform.
- Allows **authors** to create course offers; both **free** and **paid**.
- Learners (subscribers) subscribe to these courses for learning purposes.
- This business model has earned Udemy **hundreds of thousands** of courses.
- Way forward – Improve **user experience** for both **course authors** and **course subscribers**.

PHASE 1 – ASK QUESTIONS (DEFINE OBJECTIVES)



- ❖ How can Udemy strategically optimize its course offerings by analyzing market demand and other key factors to maximize revenue?
- ❖ **Goal – To increase revenue by at least 15% by the next quarter.**
- ❖ Possible Questions for analysis:
 1. Which has the **greatest** number of subscriptions; Paid or unpaid courses?
 2. Does the **number of lectures** have any effect on the **number of subscriptions**?
 3. Does the **course duration** have any effect on the **number of subscriptions**?
 4. Which **subject** is making the most sales (**Paid Courses**)?

5. Does the **course price** affect the **number of subscriptions** in any way?
6. What **percentage** of the **revenue** is contributed by each **subject**?
7. Is there a **relationship** between the **number of reviews** and the **number of subscriptions**?
8. Is there a **relationship** between the **number of lectures** and the **course duration**?

PHASE 2 – DATA GATHERING / COLLECTION



Data needed for this analysis was retrieved from the database of Udemy. Columns include:

1. **course_id** – The Course ID
2. **course_title** – The Course Title
3. **url** – The URL to access the Course
4. **is_paid** – Boolean, indicating if the course is free or paid
5. **price** – The Price of the Course
6. **num_subscribers** – Number of subscribers
7. **num_reviews** – Number of reviews
8. **num_lectures** – Number of lectures
9. **level** – Course difficulty
10. **content_duration** – Duration of all course materials in the course
11. **published_timestamp** – Date that the course was published.
12. **Subject** – The Subject under which the Course is categorized

Data Cleaning Procedures

1. Removed **duplicates** from the course_id field
2. Four rows had their course titles concatenated with the other fields of the various rows and separated by a **double quote**. The **split function** was used to separate the course title from the added data and then the column that included the added data was deleted
3. Performed the **trim** function on all text fields to remove **extra spaces**
4. Removed all records of courses that had **zero course_duration** and **zero num_of_lectures**
5. Changed the published_timestamp data type from **Date/Time** to **Date**
6. Changed the price data type from **whole number** to **fixed decimal**
7. **Extracted** the initials for the **subject column** to shorten the text length

Calculated Columns

Subscriptions Table

1. “**total_price**” column [represents the total amount realised per course]
2. “**Reviews_%_Sub**” column [represents the percentage of reviews to subscribers]
3. “**%_Subscribers**” column [represents the percentage of subscribers of each course to total subscribers]
4. “**%_Reviews**” column [represents the percentage of reviews of each course to total reviews]
5. “**%_Revenue**” column [represents the percentage of revenue of each course to total revenue]

- Calculated Tables

- ☐ A Dax function was used to summarise the subscriptions table grouped by the subject column for only paid courses, to be used to visualize revenue. The function first of all filtered out the values TRUE for the “is_paid” column and then grouped the data using the subject column as the grouping factor.

- Measures

- ☐ Highest and Least number of subscriptions by subject
 - ☐ Highest and Least Total Price by Subject
 - ☐ Percentage of the number of courses
 - ☐ Year over Year percentage Change

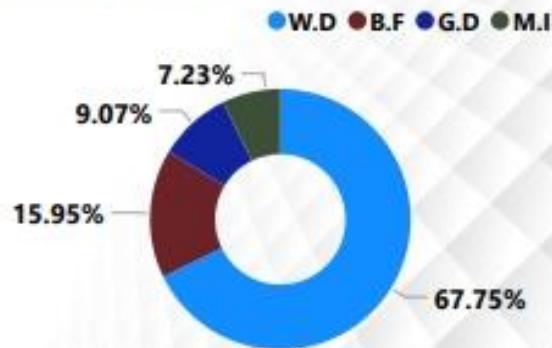
VISUALIZATIONS



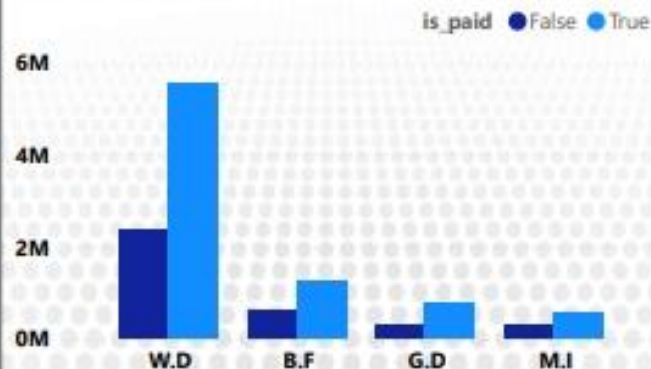
UDEMY COURSES SUBSCRIPTION ANALYSIS - SUBSCRIPTION REPORT

TOTAL	AVERAGE	MAXIMUM	MINIMUM	NO. OF COURSES		NO. OF REVIEWS
Subscribers	Subscribers	Web Development	Musical Instruments	Paid	Free	
11.72M	3.19K	7.94M	846.69K	3,361	310	574.20K
				3,671		

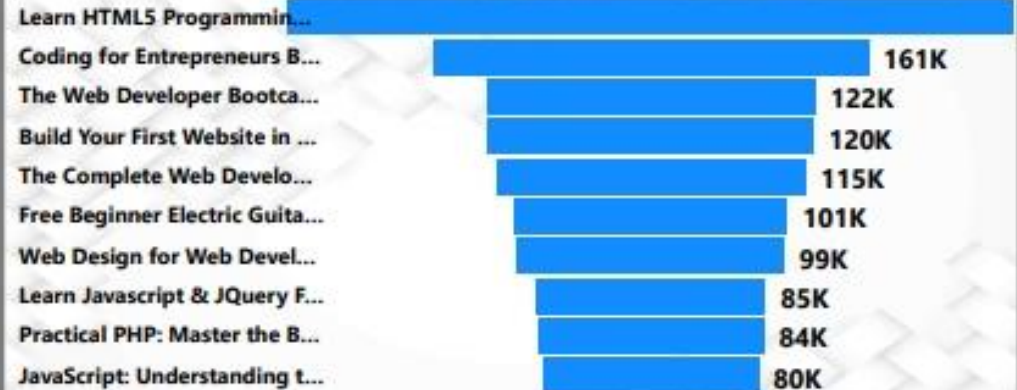
Subscribers | Percentage



Subject | Subscribers



Course Ranking by No. of Subscribers - Top 5

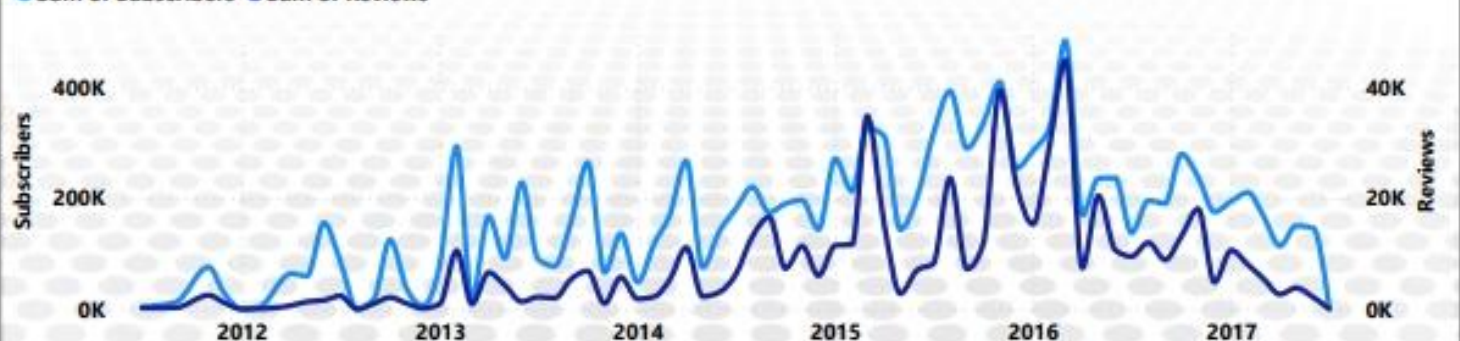


Reviews | Subscribers - Proportion

Subject	Sum of Reviews	Sum of Subscribers	Reviews_%_Sub
Web Development	429,500	7,937,287	5.41%
Business Finance	75,902	1,868,711	4.06%
Musical Instruments	31,724	846,689	3.75%
Graphic Design	37,070	1,063,148	3.49%

Trend Analysis - Subscribers | Reviews

Sum of Subscribers (Light Blue), Sum of Reviews (Dark Blue)



VISUALIZATIONS – CONT'D



UDEMY COURSES SUBSCRIPTION ANALYSIS – COURSE CONTENT REPORT

TOTAL

No. of Lectures

147.395K

TOTAL

Content Duration

15.046K

AVERAGE

No. of Lectures

40.15

AVERAGE

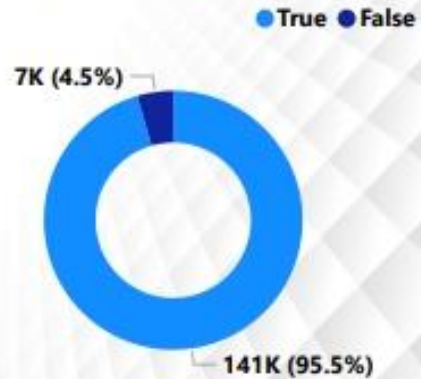
Content Duration

4.10

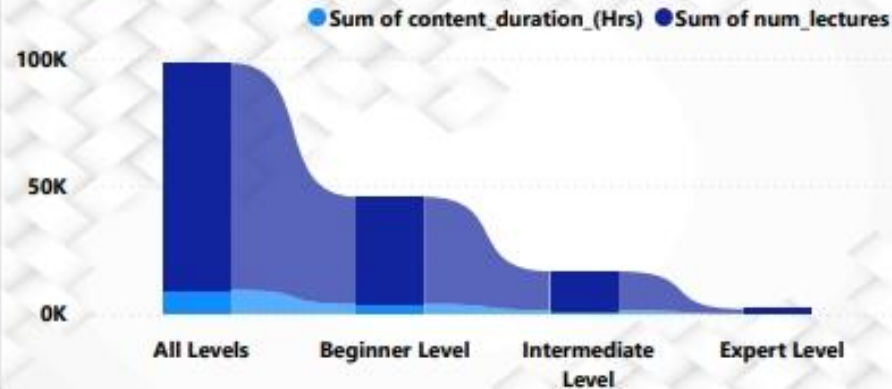
NO. OF COURSES

3,671

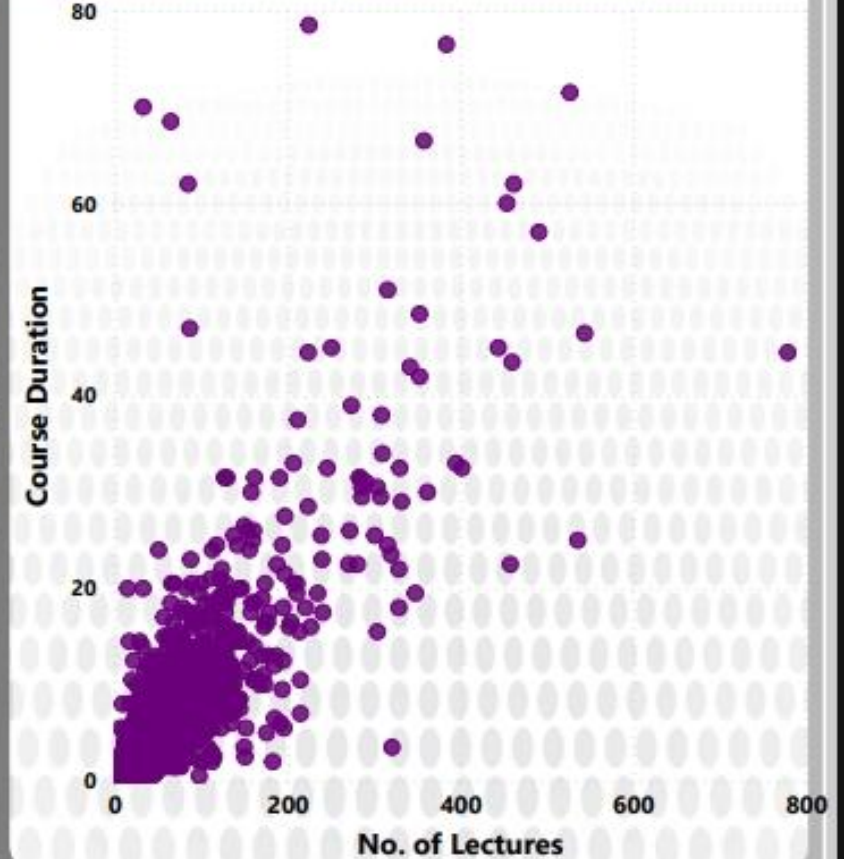
Lectures | Paid/Unpaid Courses



Levels | No. of Lectures | Course Duration



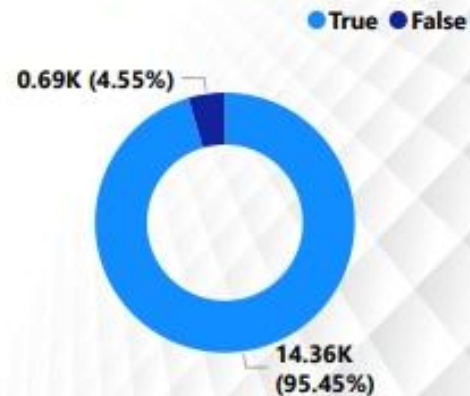
Relationship - No. of Lectures | Course Duration



Course Content Comparison

Subject	Sum of num_lectures	Sum of content_duration_(Hrs)	Sum of Subscribers
Web Development	63,357	6,709.75	7,937,287
Business Finance	38,663	4,237.22	1,868,711
Graphic Design	19,320	2,158.45	1,063,148
Musical Instruments	26,055	1,940.98	846,689

Duration | Paid/Unpaid Courses



VISUALIZATIONS – CONT'D



UDEMY COURSES SUBSCRIPTION ANALYSIS – REVENUE REPORT

TOTAL

Revenue

\$881.67M

AVERAGE

Revenue

\$240.17K

MAXIMUM

Web Development

\$627.60M

MINIMUM

Musical Instruments

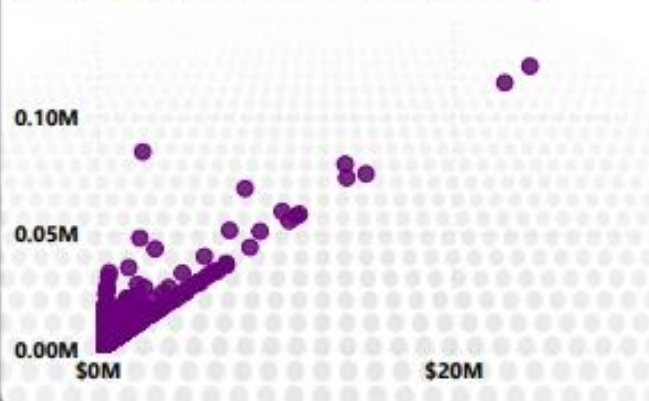
\$53.36M

No. of Courses

Paid Courses Only

3,361

Revenue | Subscribers - Relationship



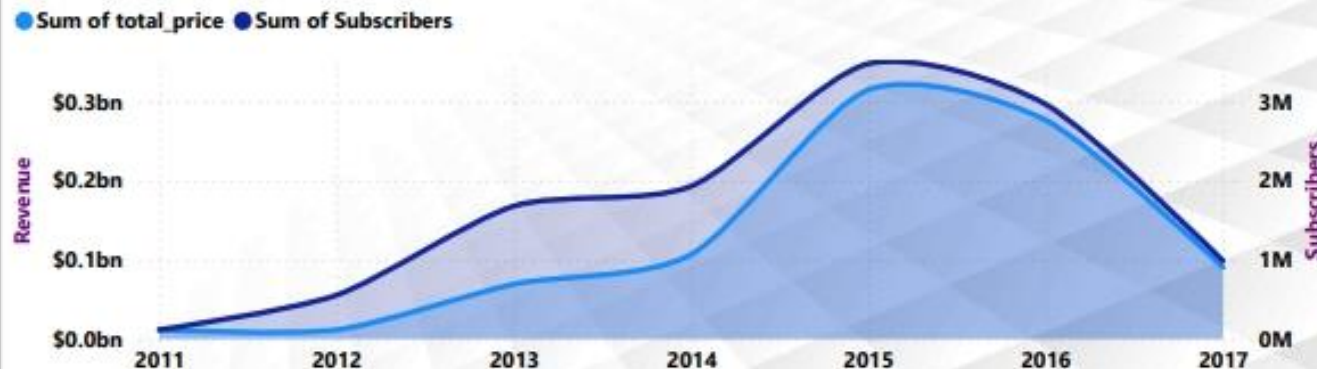
Subject | Revenue



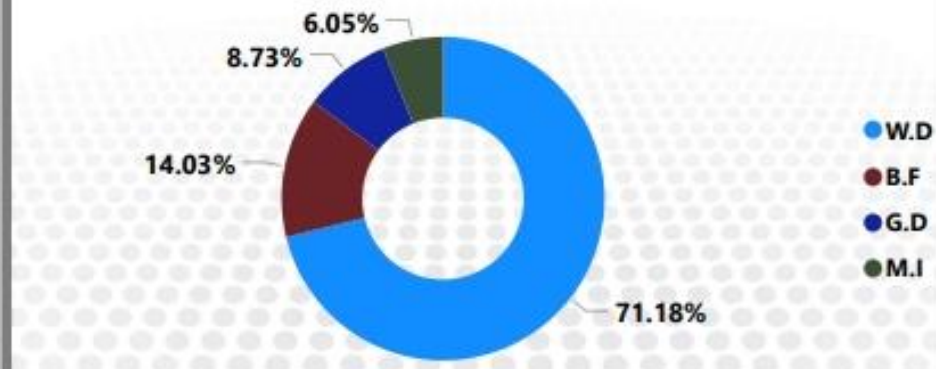
Year-over-Year Observation

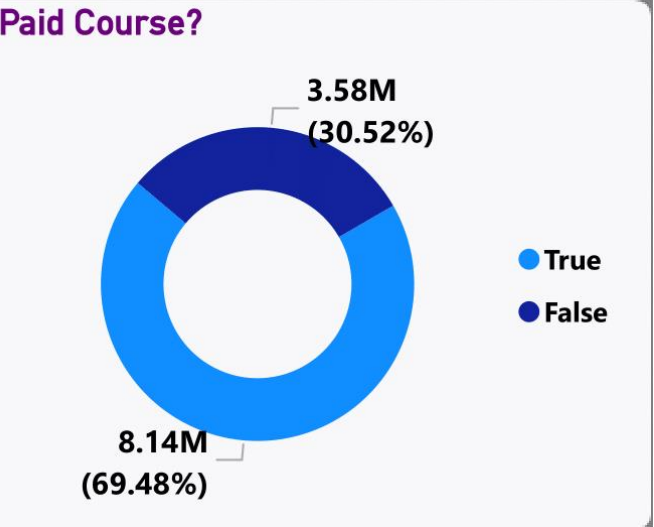
Year	Sum of total_price	total_price YoY%
2011	\$11,643,420	
2012	\$11,773,470	1.12%
2013	\$69,405,820	489.51%
2014	\$106,939,045	54.08%
2015	\$314,510,395	194.10%
2016	\$276,633,190	-12.04%
2017	\$90,769,600	-67.19%

Trend Analysis - Subscribers | Revenue



Revenue | Percentage

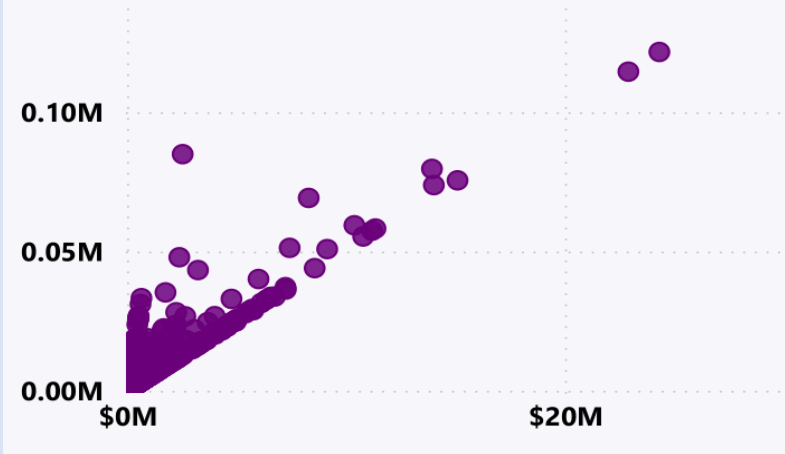




Whether a course is paid or free, has no direct impact on user interest in subscribing.

Increasing subscriptions for paid courses will result in higher revenue.

Revenue | Subscribers - Relationship



Reviews | Subscribers - Proportion

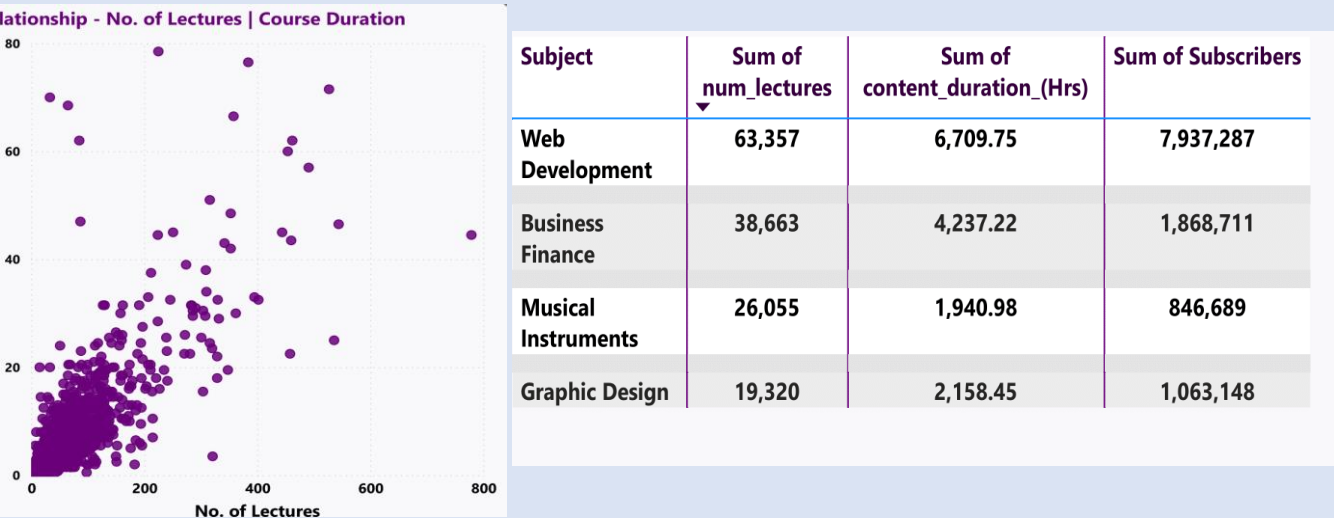
Subject	Sum of Reviews	Sum of Subscribers	Reviews_%_Sub
Business Finance	75,902	1,868,711	4.06%
Graphic Design	37,070	1,063,148	3.49%
Musical Instruments	31,724	846,689	3.75%
Web Development	429,500	7,937,287	5.41%

Across all subjects, the relationship between number of reviews and subscribers maintains an average proportion of 4%

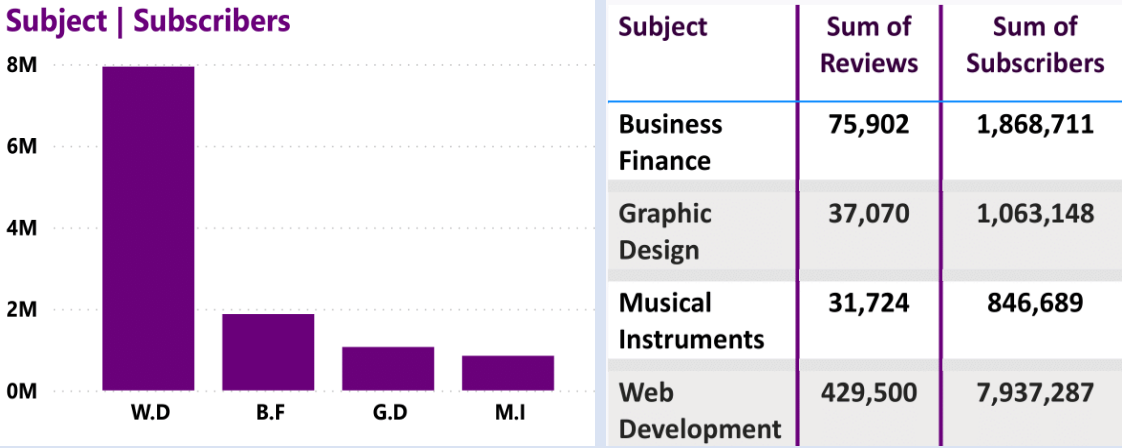
PHASE 5 – SHARE INSIGHTS – CONT'D



A strong positive relationship exists between **reviews** and **subscriptions**.



Neither course duration nor the number of lectures has a notable impact on **subscriber numbers**.



Web Development's performance highlights the impact of reviews on both subscriptions and revenue.

Web Development

Number of subscribers (7.94M – 68%)

Revenue (\$627.60M – 71%).

Number of reviews (429.50K – 75%).

A. Boost User Engagements for Reviews

1. **Incentivize** Reviews for paid courses
2. **Timely Review Requests** – Learners leave reviews at milestones
3. **Simplify the Review Process** – Require short answers
4. Personalized **Follow-Up** – Send emails to appreciate subscribers for leaving reviews
5. Highlight the **Impact of Reviews** – Through emails and on landing pages
6. Feature **Top Reviewers** – E.g. Top 5 reviewers on landing pages
7. Leverage **Social Proof** – Display statistics of reviews on social media to highlight impact of reviews

B. Targeted Marketing for High-Growth Categories:

1. Allocate in more resources to market courses in categories where **return on investment (ROI)** is evident, specifically **Web Development**.
2. Diversify offerings in **underperforming** categories with **creative course topics** to appeal to wider audiences.