



# **DATA ANALYSIS PORTFOLIO**

**PREPARED BY**  
**LUTENDE JOE MANDE**

# Professional Background

Entry level and dedicated Data Analyst with less than a year of experience identifying efficiencies and problem areas within data streams while communicating needs for projects. Adept at receiving and monitoring data from multiple data streams and Excel data sources. Ability to synthesize quantitative information. Proven track record of generating summary documents for senior management and compliance reporting.

# Portfolio Outline

PROFESSIONAL BACKGROUND	2
PORTFOLIO OUTLINE	3
UDEMY PROJECT DESCRIPTION	4
THE PROBLEM	4
DATA DESIGN	4
FINDINGS	5
ANALYSIS	8
CONCLUSION	8
GOOGLE PLAY STORE PROJECT DESCRIPTION	9
DATA DESIGN	9
FINDINGS	10
ANALYSIS	12
CONCLUSION	13
APPENDIX	14

# Udemy Project Description

The real world Udemy dataset was given to work and analyse their course revenue; The purpose is to analyse data on courses from different topics to understand where opportunities to increase the revenue may lie, and track the performance of courses.

The following methodologies were used:

- Data consolidation;
- Data cleaning;
- Data analysis;
- Data visualization.

It was drawn that Web Development courses are the most popular courses and it can be charged more to increase the next quarterly earnings.

## The problem

The business problem is to increase revenue.

I had three weeks to work on this project and to send a report.

The question is “How will they increase their next quarterly earnings?”

## Data Design

### Data consolidation

- Four datasets were given on courses from different topics.
- The four datasets were consolidated by joining them in one sheet.

### Data cleaning

- To remove any duplicates.
- To delete any blank cells.
- To ensure to have clear and concise names for headers.
- To ensure that the data is consistent so that we can easily understand what each column represents.

### Data analysis

- The LEFT function was used to create a ‘Date’ Column.

- The IF function was to create a column that indicates whether a course is free or paid.
- The VLOOKUP function was used to create the top 20 most subscribed courses.
- The Pivot Tables were created for the data sheet data to analyse data.

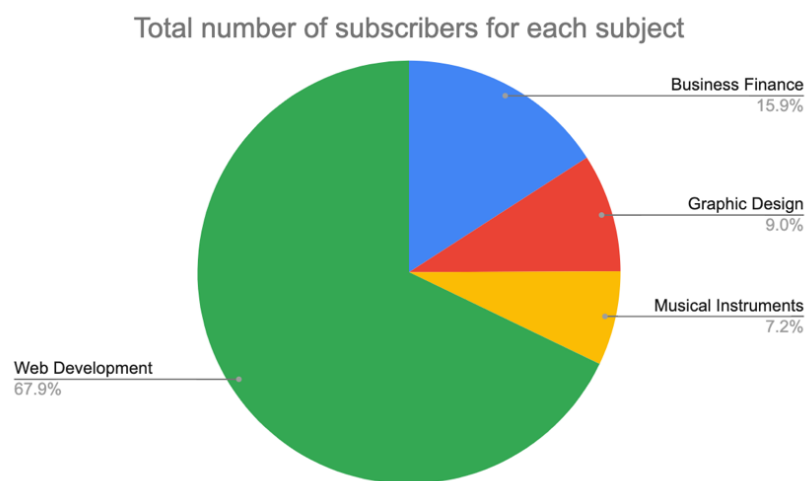
## Data visualization

- Excel Charts in Google Sheets
- Tableau is a tool used in this process to clearly communicate and support findings.

# Findings

## Finding 1

**Figure 1: Total number of Subscribers for each Subject**



**Figure 2: Average cost per subject at each level**

Average cost per subject at each level

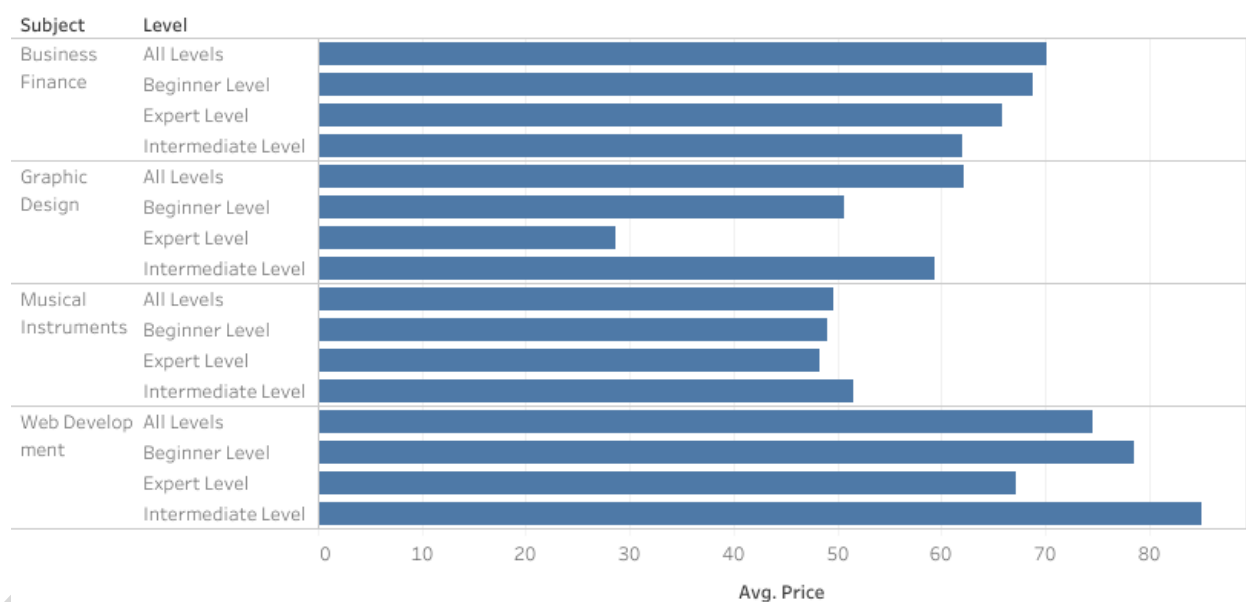


Table 1: Total number of free/paid per subject at each level

COUNTA of free or paid	level		free or paid									
	All Levels		All Levels Total	Beginner Level		Beginner Level Total	Expert Level		Expert Level Total	Intermediate Level		
subject	free	paid		free	paid		free	paid		free	paid	
Business Finance	52	581	633	26	373	399	4	21	25		14	120
Graphic Design	26	309	335	4	180	184		7	7		5	71
Musical Instruments	20	304	324	19	247	266	1	11	12		6	72
Web Development	71	562	633	51	371	422	2	12	14		10	124
Grand Total	169	1756	1925	100	1171	1271	7	51	58		35	387

Table 2: Total number of free/paid courses

free or paid	COUNTA of free or paid
free	311
paid	3365
Grand Total	3676

Finding 2

Table 3: Total number of subscribers for each subject

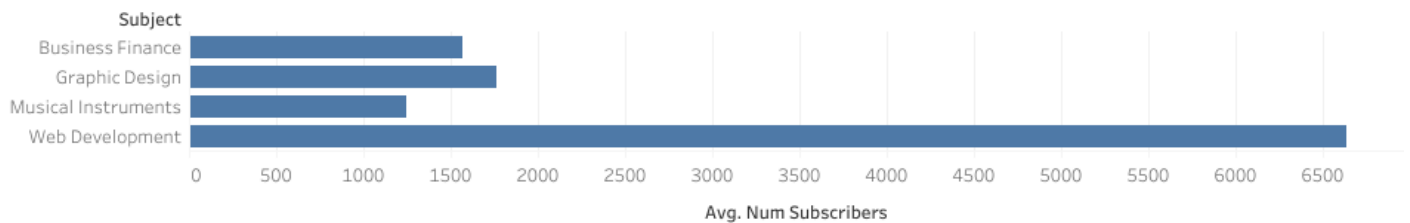
subject	SUM of num_subscribers
Business Finance	1868711
Graphic Design	1063148
Musical Instruments	846689
Web Development	7981935
Grand Total	11760483

Table 4: Average number of subscribers for each subject

subject	AVERAGE of num_subscribers
Business Finance	1,569.03
Graphic Design	1,766.03
Musical Instruments	1,245.13
Web Development	6,635.02
Grand Total	3,199.26

**Figure 3: Average number of subscribers for each subject**

Average number of subscribers for each subject



## Finding 3

**Table 5: The top 20 most subscribed courses**

course_id	course_title	level	free or paid	free beginner courses	content_duration	date	subject
41295	Learn HTML5 Programming From Scratch	Beginner Level	free	Yes	10.5	2013-02-14	Web Development
59014	Coding for Entrepreneurs Basic	Expert Level	free	No	3.5	2013-06-09	Web Development
625204	The Web Developer Bootcamp	Beginner Level	paid	No	43	2015-11-02	Web Development
173548	Build Your First Website in 1 Week with HTML5 and CSS3	All Levels	free	No	3	2014-04-08	Web Development
764164	The Complete Web Developer Course 2.0	Beginner Level	paid	No	30.5	2016-03-08	Web Development
19421	Free Beginner Electric Guitar Lessons	All Levels	free	No	4.5	2012-06-15	Musical Instruments
473160	Web Design for Web Developers: Build Beautiful Websites!	All Levels	free	No	3	2015-04-13	Web Development
94430	Learn Javascript & JQuery From Scratch	All Levels	paid	No	2	2013-10-10	Web Development
130064	Practical PHP: Master the Basics and Code Dynamic Websites	Intermediate Level	free	No	6.5	2014-07-19	Web Development
364426	JavaScript: Understanding the Weird Parts	All Levels	paid	No	11.5	2015-03-12	Web Development
238934	Pianoforall - Incredible New Way To Learn Piano & Keyboard	Beginner Level	paid	No	30	2014-08-07	Musical Instruments
756150	Angular 4 (formerly Angular 2) - The Complete Guide	Beginner Level	paid	No	22	2016-02-11	Web Development
21386	Beginner Photoshop to HTML5 and CSS3	All Levels	free	No	2	2012-07-27	Web Development
65330	Web Development By Doing: HTML / CSS From Scratch	All Levels	free	No	1	2013-09-25	Web Development
405926	HTML and CSS for Beginners - Build a Website & Launch ONLINE	All Levels	free	No	6	2015-03-19	Web Development

11174	Become a Web Developer from Scratch	All Levels	paid	No	27.5	2011-11-19	Web Development
49798	Bitcoin or How I Learned to Stop Worrying and Love Crypto	All Levels	free	No	8	2013-04-20	Business Finance
314462	Quickstart AngularJS	Beginner Level	free	Yes	1.5	2014-11-22	Web Development
128946	Learn Responsive Web Development from Scratch	All Levels	free	No	4.5	2013-12-09	Web Development
289230	Learn and Understand AngularJS	Beginner Level	paid	No	7	2014-09-24	Web Development

# Analysis

## The 5 Whys Analysis

Why did the manager ask to present the data on course revenue?

**Answer:** To understand where opportunities to increase revenue may lie and track the performance of courses.

Why do we track the performance of courses?

**Answer:** To identify the most popular courses.

Why do we identify the most popular courses?

**Answer:** To charge more.

Why do we charge more?

**Answer:** To increase their next quarterly earnings.

## My insights

- Web Development courses are the most popular courses.
- Web Development courses have high average cost at each level.
- There is more paid courses than free courses.

# Conclusion

In conclusion, the Web Development courses are the most popular courses and they should be charged more by 20-30% at all Levels in order to increase the next quarterly earnings.



# Google Play Store Project Description

The 2018 dataset of 10,000 Google Play Store applications was given to work and analyse their course revenue; The purpose is to analyse data on category from different applications to understand the Android market where opportunities may lie for developers to work on and capture the Android market.

The following methodologies were used:

- Data consolidation;
- Data cleaning;
- Data analysis;
- Data visualization.

It was drawn that Game, Communication and Family are the most popular categories and developers should build application in these categories to capture the Android market.

## Data Design

### Data consolidation

- The dataset was provided and downloaded on Kaggle website, Kaggle is a platform for data consolidation in different topics.

### Data cleaning

- To remove any duplicates.
- To delete any blank cells.
- To ensure to have clear and concise names for headers.
- To ensure that the data is consistent so that we can easily understand what each column represents.

### Data analysis

- The RIGHT function was used to create a 'Last Update\_year' Column.
- The VLOOKUP function was used to create the top 20 most installs applications.
- The Pivot Tables were created for the data sheet data to analyse data.

### Data visualization

- Excel Charts in Google Sheets
- Tableau is a tool used in this process to clearly communicate and support findings.

# Findings

## Finding 1

Figure 1: Total number of Installations for each Category

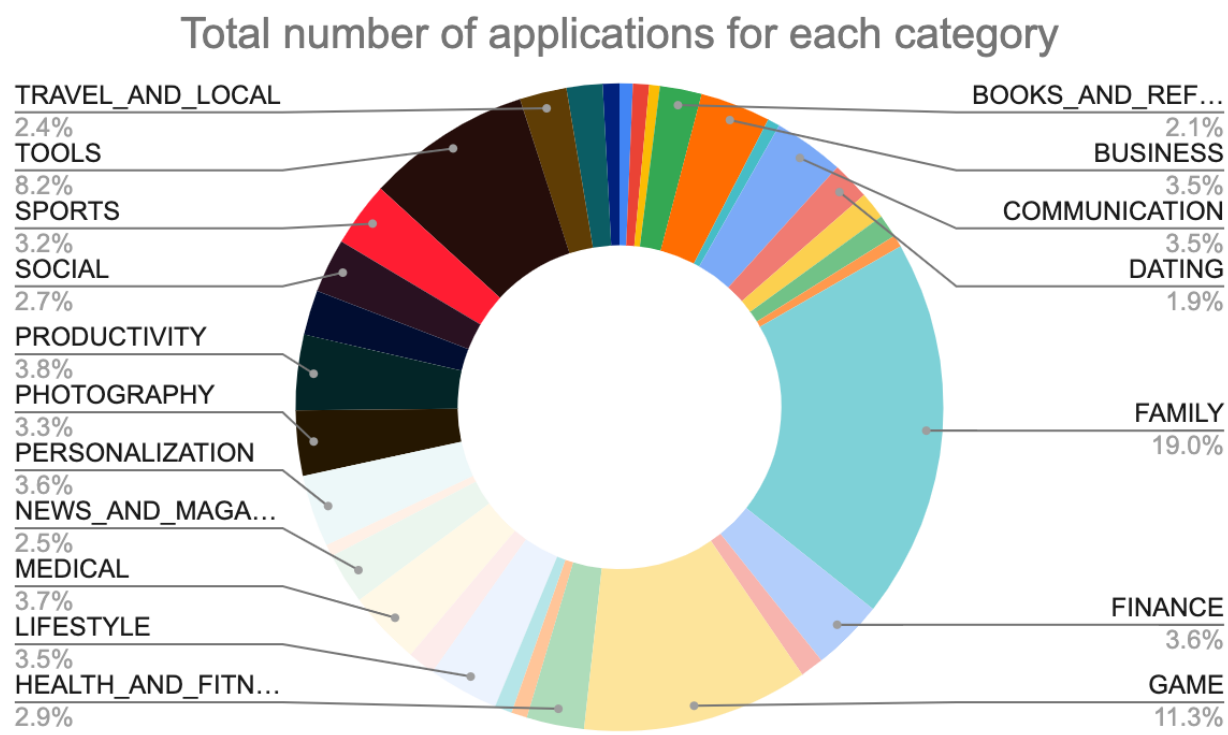


Figure 2: Total number of Applications for each Category

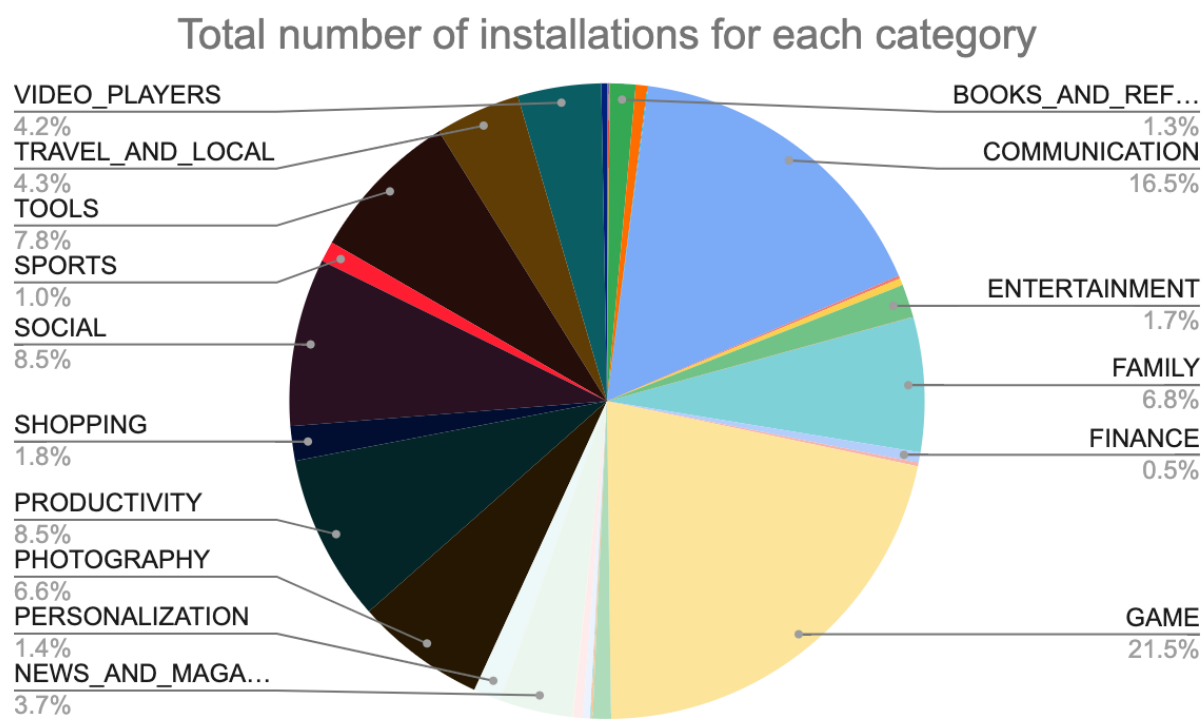
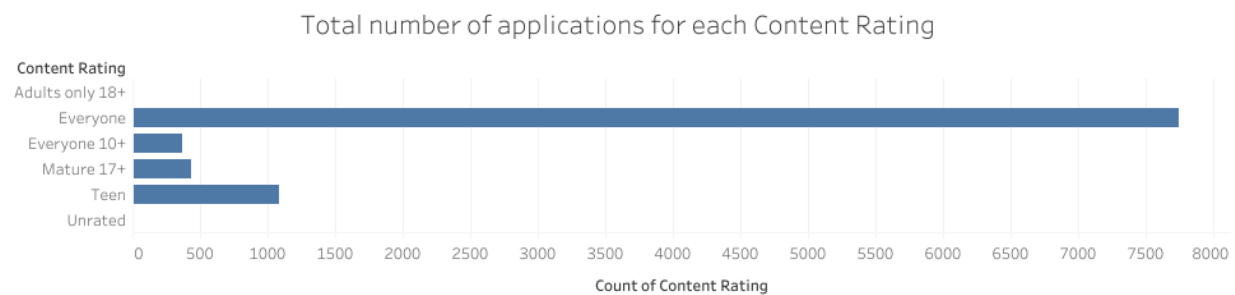


Figure 3: Total number of applications for each Content Rating



Finding 2

Figure 4: Total number of Type for each category

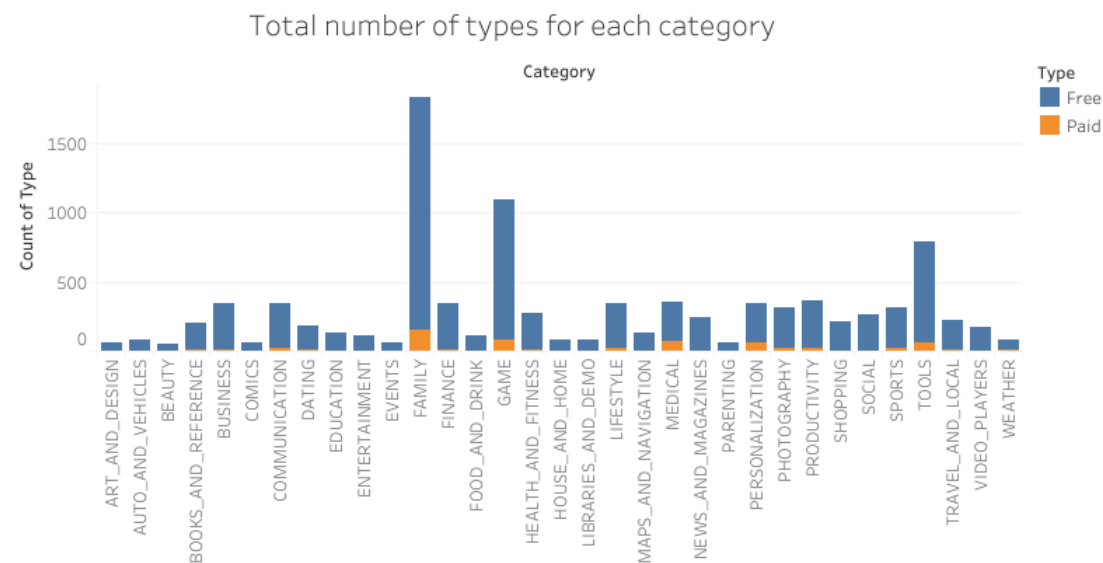
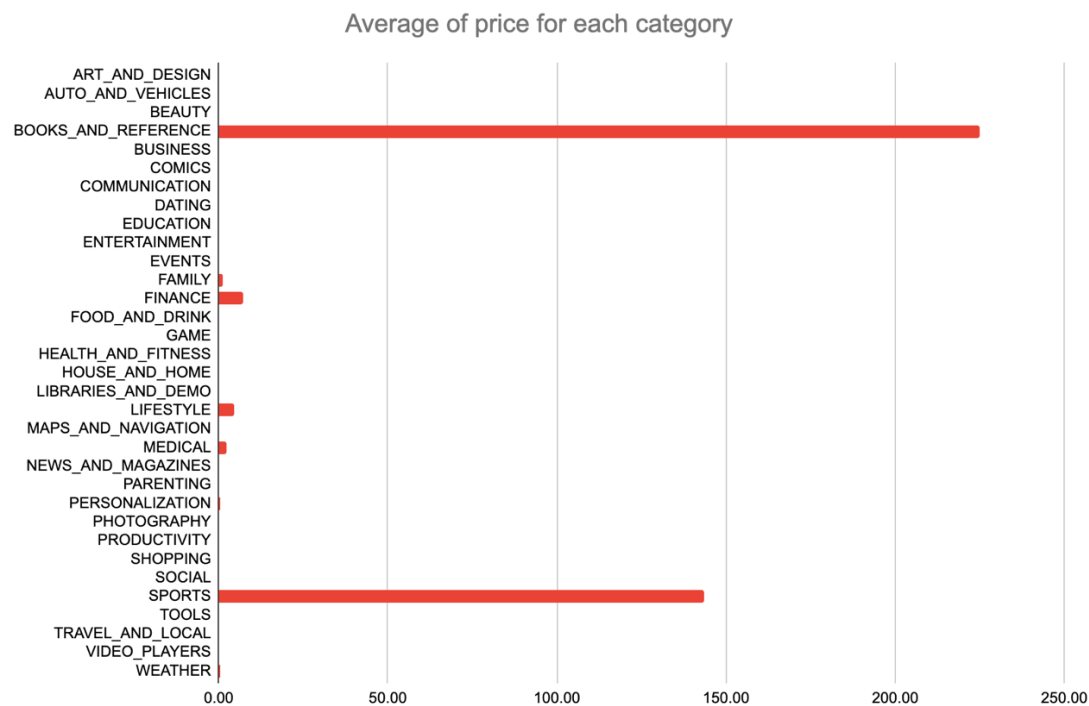


Figure 5: Average of Price for each category



## Finding 3

**Table 1: The top most Installs applications**

App	Category	Installs	type	Content Rating	Genres	Last Updated_year
Google Play Books	BOOKS_AND_REFERENCE	1,000,000,000	Free	Teen	Books & Reference	2018
Messenger – Text and Video Chat for Free	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
WhatsApp Messenger	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Google Chrome: Fast & Secure	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Gmail	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Hangouts	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Skype - free IM & video calls	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Google Play Games	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Subway Surfers	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Facebook	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Instagram	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Google+	COMMUNICATION	1,000,000,000	Free	Everyone	Communication	2018
Google Photos	ENTERTAINMENT	1,000,000,000	Free	Teen	Entertainment	2018
Maps - Navigate & Explore	GAME	1,000,000,000	Free	Everyone 10+	Arcade	2018
Google Street View	GAME	1,000,000,000	Free	Everyone 10+	Arcade	2018
Google	GAME	1,000,000,000	Free	Everyone 10+	Arcade	2018
Google Drive	GAME	1,000,000,000	Free	Everyone 10+	Arcade	2018
YouTube	SOCIAL	1,000,000,000	Free	Teen	Social	2018
Google Play Movies & TV	SOCIAL	1,000,000,000	Free	Teen	Social	2018
Google News	SOCIAL	1,000,000,000	Free	Teen	Social	2018

## Analysis

### The 5 Whys Analysis

Why do I analyse the Google Play Store Apps data?

**Answer:** To track the Android market performance.

Why do I want to track the Android market performance?

**Answer:** To identify the most popular categories.

Why do I want to identify the most popular categories.?

**Answer:** To draw actionable insights for developers to work in the right category.

Why do I want to draw actionable insights?

**Answer:** To understand the categories where developers can capture the Android market.

### **My insights**

- Game category is the most installed category, it has an average price \$0.24 and it has age restriction.
- Game category has no a lot of applications compare to Family category.
- Communication category has less applications than Family category and Game category but it comes second in total number of installations.
- Books and Reference category and Sports category have high average price while it has few available applications and installations.
- The table of top 20 most Installs App has shown that Communication category is dominant in the Android market.

## **Conclusion**

As result of my analysis in order for developers to capture the Android market, they should build applications in the following category:

- Game, Communication and Family: free and available to everyone.
- Books and Reference, Sports and Finance: these are good for paid type.

# Appendix

## Google Sheets Data Set for Udemy Project

[https://docs.google.com/spreadsheets/d/1b9bwwjSPtMz1B\\_w26RGCIzXuWCN3LNvKNRJw2Ayz0SU/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1b9bwwjSPtMz1B_w26RGCIzXuWCN3LNvKNRJw2Ayz0SU/edit?usp=sharing)

## Tableau Data Visualization for Udemy Project

[https://public.tableau.com/views/PivotTablesandCharts/Dashboard1?:language=en-GB&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/PivotTablesandCharts/Dashboard1?:language=en-GB&:display_count=n&:origin=viz_share_link)

## Google Sheets Data Set for Google Play Store Project

<https://docs.google.com/spreadsheets/d/1gEezOVHsztB0xOrVrdjVhFvWcxV-ywrL6ZRcpfJ6ttM/edit?usp=sharing>

## Tableau Data Visualization for Google Play Store Project

[https://public.tableau.com/views/GooglePlayStoreCapstoneProject/Sheet4?:language=en-GB&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/GooglePlayStoreCapstoneProject/Sheet4?:language=en-GB&:display_count=n&:origin=viz_share_link)