ENTRY LEVEL DATA ANALYTICS TRACK
Data Analyst Final Project

By Funmilayo Aina
August 2021

ENTRY LEVEL DATA ANALYTICS TRACK (MODULE 7A)

1.0 Professional Background

A graduate of Biology Education, with a career shift to Data Science / Analytics. Enthusiastic about data and the use of data oriented and analytical skills to promote business intelligence, decision making and actionable insights. A proactive fast learner of emerging technologies seeking an opportunity to increase aptitude in analytics and to work in a machine learning domain, utilizing analytical and methodical skills to help the company achieve business goals.

Experience

2016 till date

Moderator - **ASK THE PAEDIATRICIANS FOUNDATION**

2021 Data Science Fellow – **WOMEN TECHSTERS**
· She'ro Mobile Intervention App: Women Techsters Mini Project.2021

2021 Virtual Experience

Entry Level Data Analyst Challenge - Entrylevel Australia 2021(Inprogress)

· BCG Open Access Data Science & Advanced Analytics Virtual Internship [: BGC- Forage
· KPMG Data Analytics Virtual Internship: KPMG-Forage 2021

· Data@ANZ Data Science Virtual Internship Program: ANZ-Forage 2021

· PWC Power BI Virtual Case Experience: PWC –Forage 2021

Educational and Professional Qualifications

Women Techsters Data Science and AI Fellowship

UNIVERSITY OF LAGOS

B.Ed. Biology (Hons)

Udacity/AWS Machine Learning Programme (in progress)

Technical skills

· Statistical modelling MS Word, Power Point, Excel and VBA, Django, Git and GitHub,SAS, R, C programming,power BI, Tableau and Pivot table

SQL (MySQL, PostgreSQL, Sqlite3), MongoDB and Python
Data Science Project group and independent execution. Data Science concepts
Web scraping (Beautiful Soup) and API

General Skills

·· Interpersonal skills, Adaptability, Intellectual curiosity, Emotional Intelligence, disciplined habit building

· Good leadership quality, honest, responsible and accountable

· Good team spirit and global mindset

· Helping to drive accuracy, result driven with focus on quality, and client service

· Excellent verbal and written skills

· Collaboration skills and details minded

1.1 Table of Content

Primary report

Project description
Problem statement
Data Design
Pivot charts
Tableau viz storyboard
Analysis of business problems.
Conclusion

Capstone report
Professional BAckground
Project description
Problem statement
Data Design
Pivot charts
Tableau viz storyboard

Analysis of Business problems.
Conclusion

1.1 Project(Primary)

Project description

Udemy Business Analysis

There is a need for increasing revenue generation by Udemy.

We have been asked to analyze the given historical dataset for trends that could show the impact of increasing the cost of web development courses as a means of generating the desired Problem statement

Tracking courses performance on revenue generation and analyzing if charging more on popular courses would have a positive impact on Udemy revenue generation.

Understanding where opportunities to increase revenue may lie.

Furthermore analysis would allow us to size the business impact of a proposed increments

Data Designs

The data were cleaned with Google sheet and exported to Tableau for visualization.

There were missing values in some columns while some columns have duplicate values. The duplicates were removed.

The data was recorded with a time stamp, we therefore transform the date to date only removing the time stamp.

Findings

The top 20 courses were web development courses. The reviews for the courses were also high.

The rating for paid courses was high showing customers' satisfaction.

The number of courses in the top 20 and paid were more than those in free range.

All level courses and Beginner courses have the highest level of subscribers filtered by rating.

From hours of lecture reviews we may need to increase the hours for Web development courses.

Analysis of finding from the root cause

The top 20 courses were web development courses. The reviews for the courses were also high.

The rating for paid courses was high showing customers' satisfaction.

The number of courses in top 20 and paid were more than those in free range.

All level courses and Beginner courses have the highest level of subscribers filtered by rating.

From hours of lecture reviews we may need to increase the hours for Web development courses.

All level courses generate the highest revenue and web development tops the chart.

The subscription for Expert level courses is extremely low, we need not increase the prices of those courses to prevent churning.

Pivot charts
Tableau viz storyboard

Figure 1 Total Subscriptions

All level courses and Beginner courses have the highest level of subscribers filtered by rating.

From hours of lecture reviews we may need to increase the hours for Web development courses.

Subject	COUNTA of Num_subscribers
Business Finance	1190
Graphic Design	602
Musical Instruments	680
Subject: Web Development	1203
Grand Total	**3675**

Total Num_subscribers

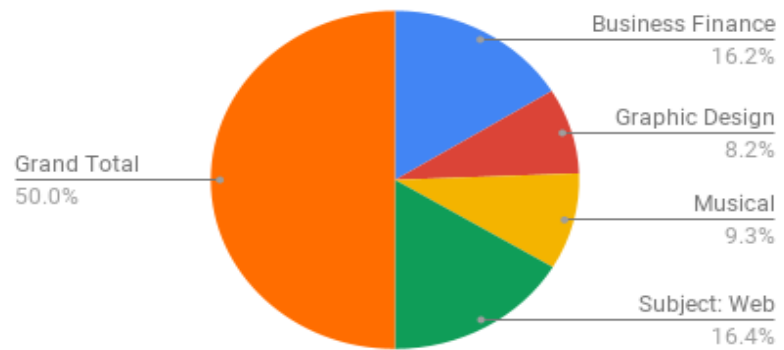|	
Business Finance	16.2%
Graphic Design	8.2%
Musical	9.3%
Subject: Web	16.4%
Grand Total	50.0%

**Figure 1, Users by skill level**

All level courses generate the highest revenue and web development tops the chart.

Dates show regular subscription trends. Therefore increasing the prices of 'All level courses' are ''Beginner courses' , that are 'Web Development courses' would increase Udemy's revenue

Average number of subscribers for each subject

Level	COUNTA of course_id
All Levels	1,924.00
Beginner Level	1,271.00
Expert Level	58.00
Intermediate Level	422.00
Grand Total	**3,675.00**

COUNTA of course_id vs. Level
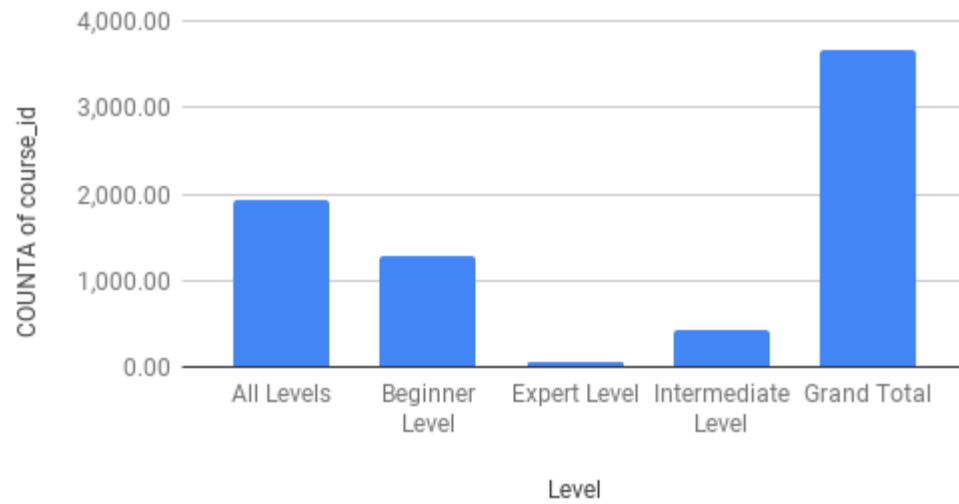
generation.

The level of subscriber generating highest revenue
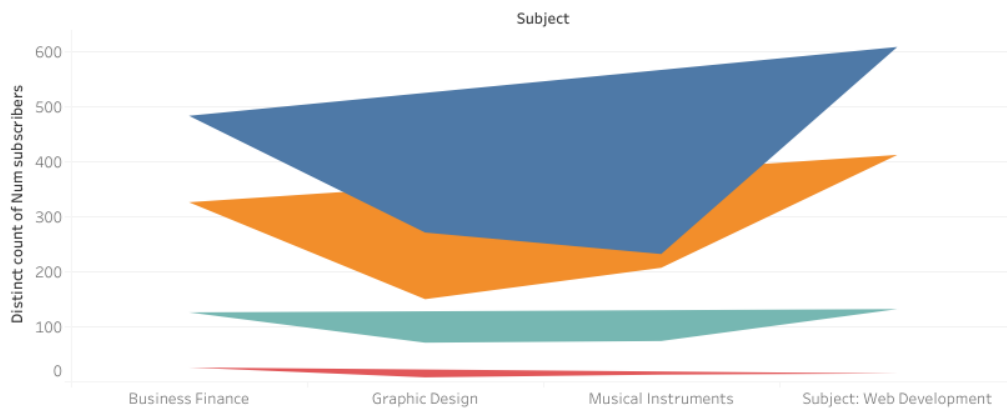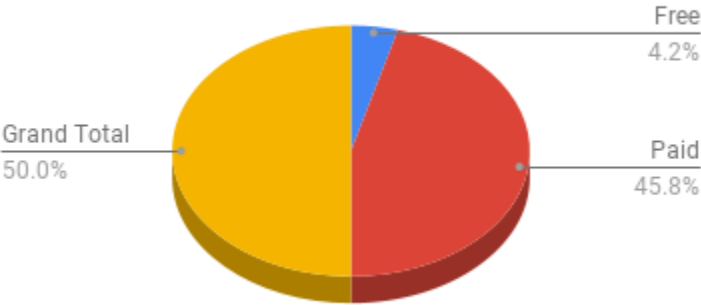

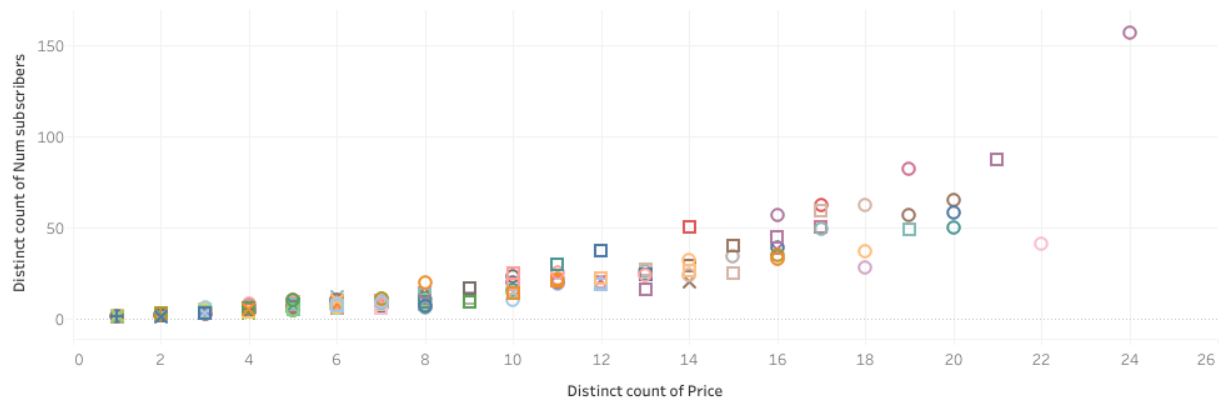## Number of subscribers per subjects

Table 1: Free/Paid Users

Free_Paid	COUNTA of Num_subscribers
Free	311
Paid	3364
Grand Total	**3675**

COUNTA of Num_subscribers

Free
4.2%

Grand Total
50.0%

Paid
45.8%

Udemy Business Viz2

Price/Subscriber filtered by rating

Free/paid courses and subscriptions
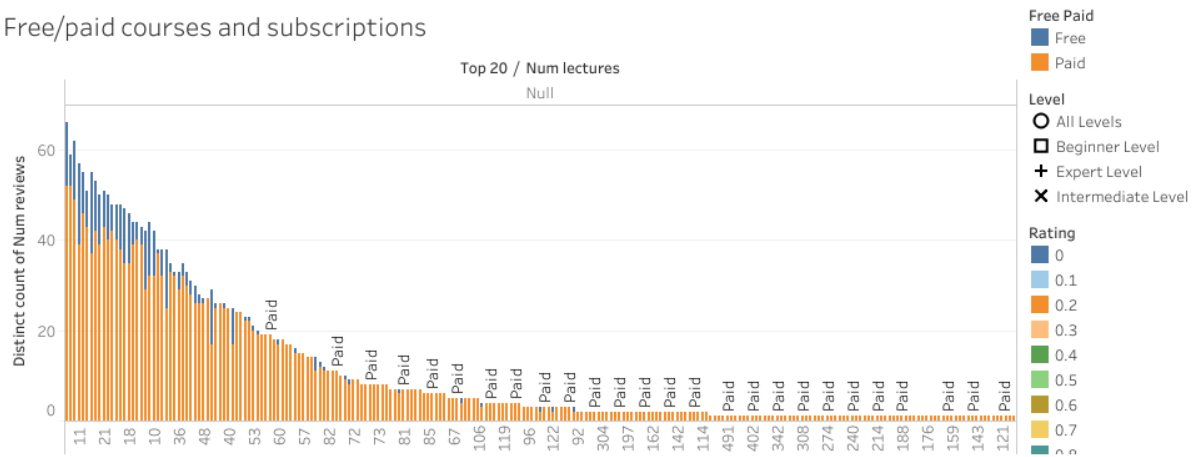
Free Paid
- ■ Free
- ■ Paid

Level
- ○ All Levels
- □ Beginner Level
- + Expert Level
- ✕ Intermediate Level

Rating
- ■ 0
- ■ 0.1
- ■ 0.2
- ■ 0.3
- ■ 0.4
- ■ 0.5
- ■ 0.6
- ■ 0.7
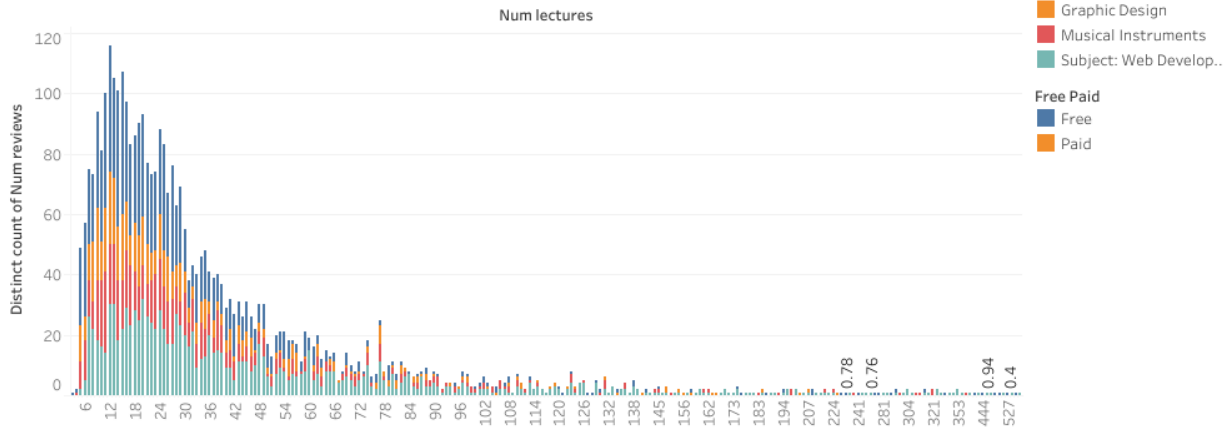
The rating for paid courses was high showing customers' satisfaction.

The number of courses in top 20 and paid were more than those in free range.

Hours of lecture reviews

Num lectures

Subscribers per subject and rating

Level

All Levels Beginner Level Expert Level Intermediate Level

Conclusion

The top 20 courses were truly web development courses. The reviews for the courses were also high as well rating for these courses indicating customers' satisfactions.

All level courses generate the highest revenue and web development tops the chart.

Dates show regular subscription trends. Therefore increasing the prices of 'All level courses' are ''Beginner courses' , that are 'Web Development courses' would increase Udemy's revenue generation.

I recommend non-increment of Expert level courses to prevent churning.

Capstone Project (Module 7B)

1.0 Professional Background

A graduate of Biology Education, with a career shift to Data Science / Analytics. Enthusiastic about data and the use of data oriented and analytical skills to promote business intelligence, decision making and actionable insights. A proactive fast learner of emerging technologies seeking an opportunity to increase aptitude in analytics and to work in a machine learning domain, utilizing analytical and methodical skills to help the company achieve business goals.

Project 2

Analysis of COVID-19 World Vaccination Progress

Background

Covid-19 is a global epidemic which has become a threat to the life and economy of many nations.

Thanks to the effort of scientists from different parts of the world, working to find a lasting solution to this menace.

The discovery of Covid 19 vaccine has been greatly celebrated all over the world. Efforts are in place to ensure world vaccination progress but this has been riddled with misinformation and-fake news about Covid-19 vaccine.

This project analyzes the response of the world to Covid 19 vaccine, the progress so far and how to mitigate misinformation about Covid vaccines as well as countries that need focusing on. (Kaggle Data source)

Problem Framing

Would friendly awareness about Covid 19 vaccine in order to improve vaccination progress be needed?

What is the correlation between daily vaccination per millions and total vaccination?

Median daily vaccination as a pointer to future prediction.

Which country is using what vaccine?

In which country the vaccination programme is more advanced?

Where are more people vaccinated per day?

Which vaccine is mostly embraced by people and by which countries?

What is the correlation between daily vaccination per million and people fully vaccinated?

Which countries have the highest numbers of fully vaccinated persons and why?

Explanation of the dataset provided (Kaggle)

Content The data (country vaccinations) contains the following information:

Country- this is the country for which the vaccination information is provided;

Country ISO Code - ISO code for the country;

Date - date for the data entry; for some of the dates we have only the daily vaccinations, for others, only the (cumulative) total;

Total number of vaccinations - this is the absolute number of total immunizations in the country;

Total number of people vaccinated - a person, depending on the immunization scheme, will receive one or more (typically 2) vaccines; at a certain moment, the number of vaccination might be larger than the number of people;

Total number of people fully vaccinated - this is the number of people that received the entire set of immunization according to the immunization scheme (typically 2); at a certain moment in time, there might be a certain number of people that received one vaccine and another number (smaller) of people that received all vaccines in the scheme;

Daily vaccinations (raw) - for a certain data entry, the number of vaccination for that date/country;

Daily vaccinations - for a certain data entry, the number of vaccination for that date/country;

Total vaccinations per hundred - ratio (in percent) between vaccination number and total population up to the date in the country;

Total number of people vaccinated per hundred - ratio (in percent) between population immunized and total population up to the date in the country;

Total number of people fully vaccinated per hundred - ratio (in percent) between population fully immunized and total population up to the date in the country;

Number of vaccinations per day - number of daily vaccination for that day and country;

Daily vaccinations per million - ratio (in ppm) between vaccination number and total population for the current date in the country;

Vaccines used in the country - total number of vaccines used in the country (up to date);

Source name - source of the information (national authority, international organization, local organization etc.);

Source website - website of the source of information;

There is a second file added recently (country vaccinations by manufacturer), with the following columns:

Location - country;

Date - date;

Vaccine - vaccine type;

Total number of vaccinations - total number of vaccinations / current time and vaccine type.

Designs

The data was downloaded from kaggle platform.

The data was cleaned with Google sheet and exported to Tableau for visualization. The missing values were replace with 0 values due to the sensitive nature of the dataset (Health dataset)

The duplicates were removed.

Findings

China is an outlier among countries with total vaccinated persons. The United States tops the chart among countries of similar population. Fig 3.1 and Fig 5.2,

Using the sum of total vaccination, China was at the top of the list of vaccination progress India and United States followed suit while Mexico was not doing badly.

Let's visualize countries' access to vaccines and the type of vaccines. Was it single dose or double doses?

Sinopharm/Beijing,Sinovac,Sputnik V" has same value for vaccine

Johnson & Johnson, Pfizer and Moderna were the most widely used by countries fully vaccinated.

Many people became fully vaccinated around July 2021. We may need to increase awareness.

Comoro, China followed by the United States have the highest daily vaccination per million (Avg) analyzed per total and filtered per country.

This could be because of wealth of information and access to Covid -19 vaccine by those interested in getting vaccinated.

The number of people already vaccinated shows an increased progress, very few people are fully vaccinated compared to total vaccination. There was a sharp drop in the total number of people fully vaccinated around August 2021.

The United States has the highest number of people fully vaccinated, followed by India and the United Kingdom.
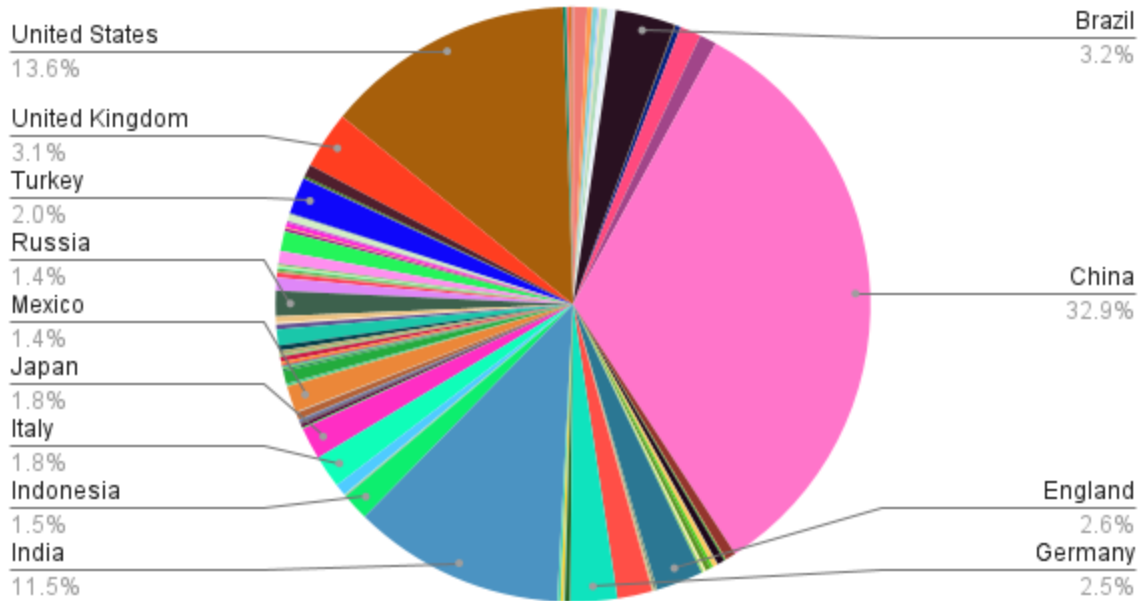
Several countries were lagging behind in vaccination progress.

China have the highest number of daily vaccination followed by the US

Analysis of finding from the root cause

The number of people already vaccinated shows an increased progress, very few people are fully vaccinated compared to total vaccination. There was a sharp drop in the total number of people fully vaccinated around August 2021.

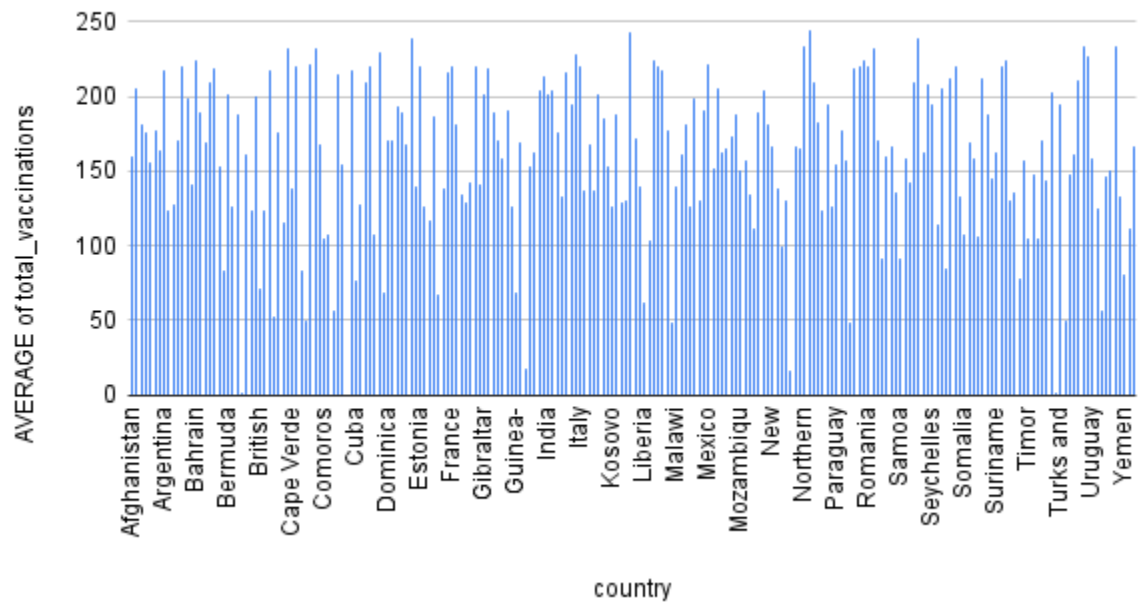Western Countries are more vaccinated than African continent.

AVERAGE of total_vaccinations
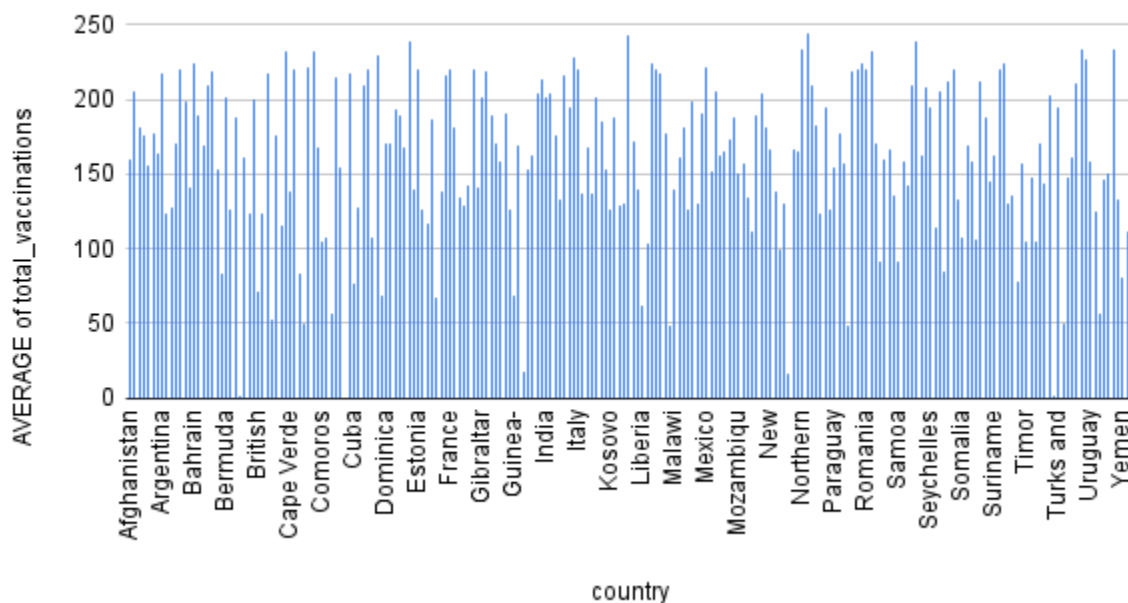
United States
13.6%

United Kingdom
3.1%

Turkey
2.0%

Russia
1.4%

Mexico
1.4%

Japan
1.8%

Italy
1.8%

Indonesia
1.5%

India
11.5%

Brazil
3.2%

China
32.9%

England
2.6%

Germany
2.5%

Count of total_vaccinations

AVERAGE of total_vaccinations

country

COVID-19 World Vaccination Progress (World Report Viz)

Quarterly Progress of Covid-19 Immunization programme

Date

2020   2021

Avg. Daily Vaccinati..
150K
100K
50K
0K

Total Vaccinations ..
300K
200K
100K
0K

Q4   Q1   Q2   Q3

Monthly Progress of Covid-19 Immunization programme

Date

Total Vaccinations.. Avg. Daily Vaccina..
300K
200K
100K
0K

200K
100K
0K

January February March April May June July August
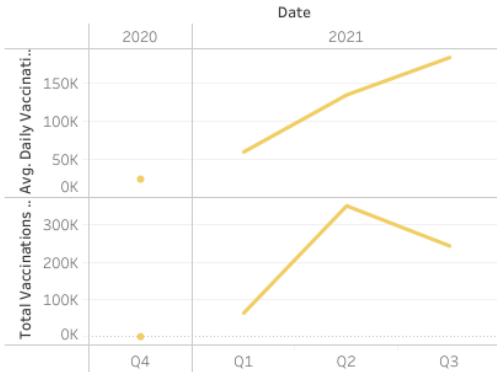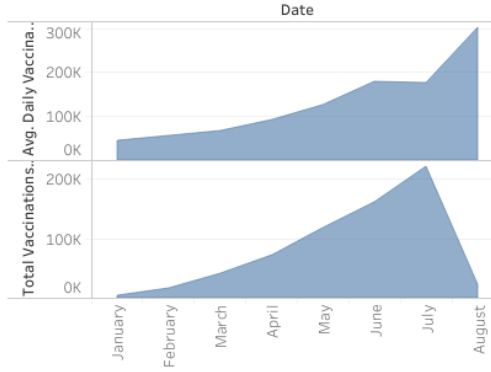
Measure Names
- Avg. Daily Vaccinatio..
- Total Vaccinations Pe..

Country
- Afghanistan
- Albania
- Algeria
- Andorra
- Angola
- Anguilla
- Antigua and Barbuda
- Argentina
- Armenia
- Aruba
- Australia
- Austria
- Azerbaijan
- Bahamas
- Bahrain
- Bangladesh
- Barbados
- Belarus
- Belgium
- Belize
- Benin
- Bermuda
- Bhutan
- Bolivia
- Bonaire Sint Eustatiu..
- Bosnia and Herzegovi..
- Botswana
- Brazil
- British Virgin Islands
- Brunei
- Bulgaria
- Burkina Faso
- Cambodia
- Cameroon
- Canada

Monthly Progress of Covid-19 Immunization programme

Date

Janua.. Febru.. March April May June July August

Value
300K
200K
100K
0K

Avg. Daily Vaccinat.. Total Vaccinations ..
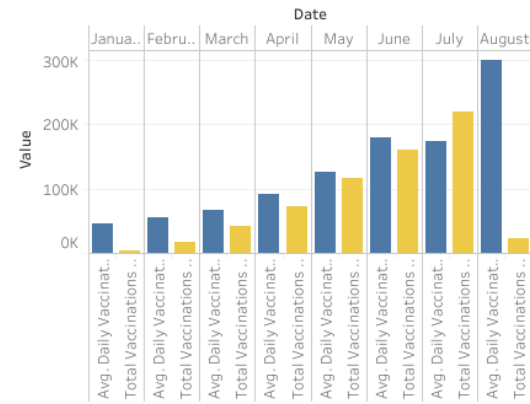
Countries Daily Vaccination Progress

Brazil

India

China

COVID-19 World Vaccination Progress (World Report Viz)

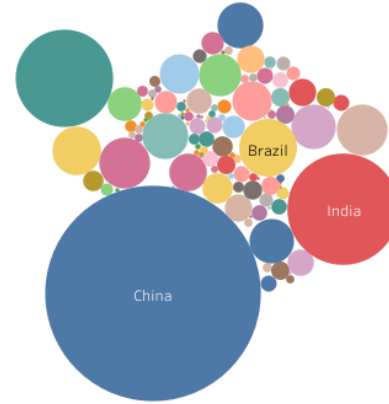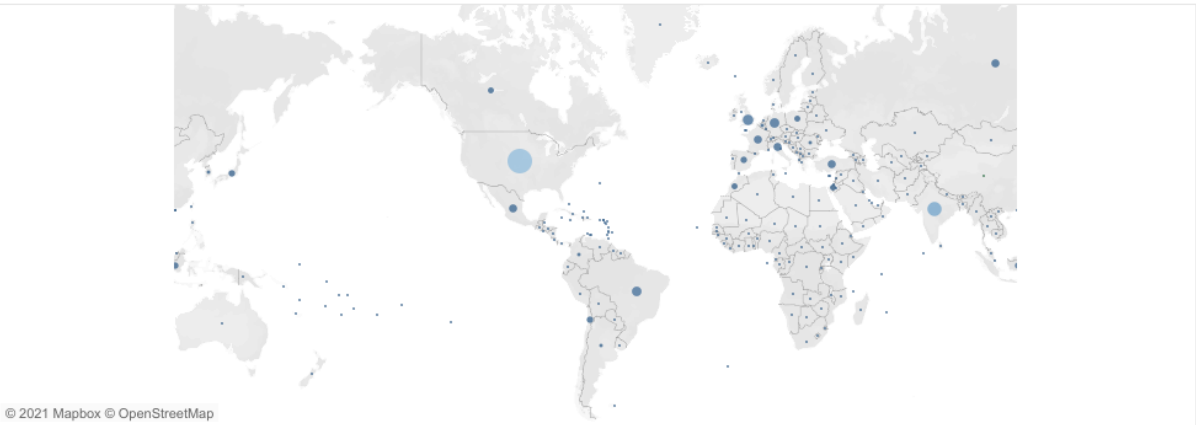| Progress with regards to Covid-19 Immuniza..	There is a progression in AVG daily vaccinati..	China is an outlier in Covid Vaccination pro..

Geographical overview of Covid-19 Immunization progress

Vaccine

Null

CanSino

© Mapbox © OSM

Outliers Overview

Total Vaccinations
254 ▮▮▮▮▮ 45B

People Fully Vaccinated
| 0
▮ 5,000,000,000
▮ 10,000,000,000
▮ 17,330,909,394

Map of Countries Vaccination progress by Fully Vaccinated

© 2021 Mapbox © OpenStreetMap

COVID-19 World Vaccination Progress (World Report Viz)

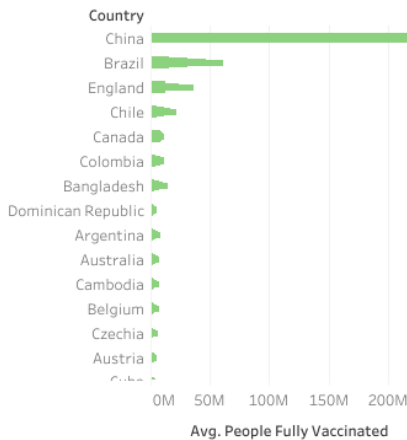| Progress with regards to Covid-19 Immuniza..	There is a progression in AVG daily vaccinati..	China is an outlier in Covid Vaccination pro..

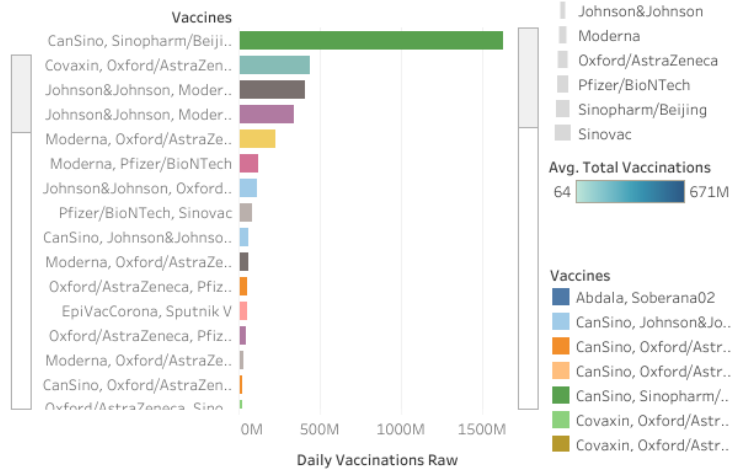## People Fully Vaccinated by Country

Fully Vaccinated by Total Vaccination and Country
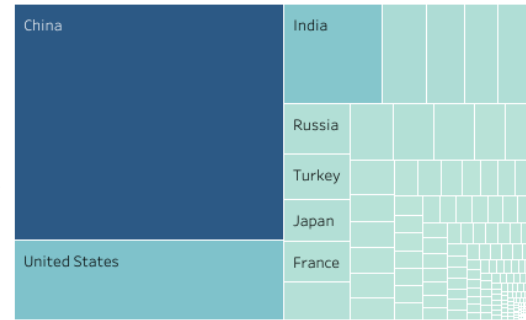
Type of Vaccines by Countries with Fully Vaccinated

Vaccines Purchase Trend

China tops the chart in vaccination progress.

China is an outlier among countries with fully vaccinated persons. The United states top the chart among countries of similar population.

Using the sum of total vaccination, China still tops the list of vaccination progress, India and the United States followed suit while Mexico was not doing badly.

The United States has the highest number of people fully vaccinated, followed by India and the United Kingdom.

Several countries were lagging behind

China have the highest number of daily vaccination followed by the US

Conclusion

The United States has the highest number of people fully vaccinated, followed by India and the United Kingdom.

Several countries were lagging behind

Many people became fully vaccinated around July 2021 we may need to increase awareness

China followed by the United states have the highest daily vaccination per million(Avg) analysed per total and filtered per country.

China is an outlier among countries with the most vaccinated persons. The United states top the chart among countries of similar population.

The number of people already vaccinated shows an increased progress, very few people are fully vaccinated compared to total vaccination. There was a sharp drop in the total number of people fully vaccinated around August 2021.

Pivot table

There may be need for more vaccine safety awareness and increased production as well as increased distribution of vaccine to Africa and Continents lagging behind.