**Table Of Content**

* Project Overview
* Data Source
* Tools Used
* Data Preparation and Data Cleaning
* Data Analysis
* Visualization in Power BI
* Findings
* Limitations

**VACCINATION DATA ANALYSIS PROJECT**

**Project Overview**

This project aims to analyze the global impact of the vaccination used by examining the countries, WHO Region, Data Source, Date Updated, First Vaccine Date, Total Vaccinated and many other rows. Utilizing the data set provided, we derive insight on some vaccines and how best they were used in each country.

Originally, the data set has 16 rows and 230 columns.

**Data Source**

The primary data set utilized for this analysis is sourced from vaccination data project csv file comprising of a comprehensive vaccination data set essential for this analysis project. The Data set includes detailed information on the country, WHO region, ISO3, data source, date updated, total vaccinated and many other row heads. The data is crucial for understanding major vaccines in the world and how best they were utilized.

**Tools Used**

* **Microsoft Excel:** To clean the data.
* **SQL Server Management Studio:** To analyze the data.
* **Power BI:** To visualize the outcome.

**Data Cleaning Process in Excel.**

1. Expand the widths of the columns.
2. Bold the column headings.
3. Change the data type where necessary like long dates to short dates.
4. Putting Commas to the figures.
5. Correcting the wrongly spelt countries in cell 51 and 54.

**Database Creation and Query Execution.**

In SQL Server Management Studio (SSMS), a database named vaccination data Q&A was created to organize and manage the vaccination dataset . Within this database various queries were executed to answer specific research questions related to the vaccination dataset. The process involved inserting the questions to the database and writing the corresponding SQL queries to derive the necessary insights.

**Questions Answered in the Data Analysis Process**

1. **Which county had the highest vaccinations?**
2. **Which region had the most fully vaccinated persons?**
3. **How many total vaccinations were recorded in euro?**
4. **Which country was the first to start vaccination?**
5. **How many persons were not fully vaccinated in Nigeria?**
6. **Which countries used Astrazeneca-vaxzevria?**
7. **How many persons received a booster dose in Mexico?**
8. **Which country has the highest percentage of fully vaccinated persons per 100 people?**

**Visualization in Power BI**

After completing data cleaning in Microsoft excel and data analysis in SQL Server Management Studio, the results were exported to CSV files. This csv files were then used to create a visually appealing dashboard in Power BI.

The primary purpose of this Power BI dash board was to enhance understanding of the outcomes derived from the SQL data analysis. It is important to note that Power BI dash board is non interactive.

**Findings from the analyses**

1.There were six WHO Region which were EMRO, EURO, AFRO, WPRO, AMRO, SEARO.

2.There are two data source which are REPORTING and OWID.

**LIMITATION**

1.There empty spaces.

2.No commas.