**Global Country Information Dataset 2023**

**World-data-2023.csv**

This comprehensive dataset provides a wealth of information about **all countries worldwide**, covering a wide range of indicators and attributes. It encompasses demographic statistics, economic indicators, environmental factors, healthcare metrics, education statistics, and much more. With every country represented, this dataset offers a complete global perspective on various aspects of nations, enabling in-depth analyses and cross-country comparisons.

* **Country**: Name of the country.
* **Density (P/Km2)**: Population density measured in persons per square kilometer.
* **Abbreviation**: Abbreviation or code representing the country.
* **Agricultural Land (%)**: Percentage of land area used for agricultural purposes.
* **Land Area (Km2)**: Total land area of the country in square kilometers.
* **Armed Forces Size**: Size of the armed forces in the country.
* **Birth Rate**: Number of births per 1,000 population per year.
* **Calling Code**: International calling code for the country.
* **Capital/Major City**: Name of the capital or major city.
* **CO2 Emissions**: Carbon dioxide emissions in tons.
* **CPI**: Consumer Price Index, a measure of inflation and purchasing power.
* **CPI Change (%)**: Percentage change in the Consumer Price Index compared to the previous year.
* **Currency\_Code**: Currency code used in the country.
* **Fertility Rate**: Average number of children born to a woman during her lifetime.
* **Forested Area (%)**: Percentage of land area covered by forests.
* **Gasoline\_Price**: Price of gasoline per liter in local currency.
* **GDP**: Gross Domestic Product, the total value of goods and services produced in the country.
* **Gross Primary Education Enrollment (%)**: Gross enrollment ratio for primary education.
* **Gross Tertiary Education Enrollment (%)**: Gross enrollment ratio for tertiary education.
* **Infant Mortality**: Number of deaths per 1,000 live births before reaching one year of age.
* **Largest City**: Name of the country's largest city.
* **Life Expectancy**: Average number of years a newborn is expected to live.
* **Maternal Mortality Ratio**: Number of maternal deaths per 100,000 live births.
* **Minimum Wage**: Minimum wage level in local currency.
* **Official Language**: Official language(s) spoken in the country.
* **Out of Pocket Health Expenditure (%)**: Percentage of total health expenditure paid out-of-pocket by individuals.
* **Physicians per Thousand**: Number of physicians per thousand people.
* **Population**: Total population of the country.
* **Population: Labor Force Participation (%)**: Percentage of the population that is part of the labor force.
* **Tax Revenue (%)**: Tax revenue as a percentage of GDP.
* **Total Tax Rate**: Overall tax burden as a percentage of commercial profits.
* **Unemployment Rate**: Percentage of the labor force that is unemployed.
* **Urban Population**: Percentage of the population living in urban areas.
* **Latitude**: Latitude coordinate of the country's location.
* **Longitude**: Longitude coordinate of the country's location.

**Top 1000 Youtubers statistics**

**Youtubers\_df.csv**

**Description**

The dataset, obtained through data extraction from top YouTube streamers using the HypeAuditor platform, contains valuable information related to the presence and performance of these content creators on the world's largest video-sharing platform. Below is a description of each of the variables included in the dataset:

* **Rank**: This variable indicates the position or ranking of the streamer on the list of top YouTube streamers. A lower number signifies a higher ranking.
* **Username**: It is the streamer's username on YouTube, allowing for the unique identification of each content creator.
* **Categories**: Represents the categories in which the streamer has tagged their content. Categories can span a wide variety of topics, including gaming, beauty, fashion, travel, comedy, and more.
* **Subscribers**: Indicates the average number of subscribers the streamer's YouTube channel has. This value represents the regular following of content by the audience.
* **Country**: Shows the average country of origin or location of the streamer. This can provide insights into the primary audience of the creator and their base of operations.
* **Visits**: This variable records the average number of accumulated visits to the streamer's channel. It represents the average number of times the creator's videos have been viewed by viewers.
* **Likes**: Indicates the average number of "Likes" received on the streamer's videos. "Likes" are an engagement metric that shows how many viewers appreciate the content.
* **Comments**: Reflects the average number of comments left on the streamer's videos. Comments are an important form of audience interaction and participation.
* **Links**: Provides links or URLs to the streamer's YouTube channels, allowing direct access to their content.

**Most Streamed Spotify Songs 2023**

**Spotify-2023.csv**

This dataset contains a comprehensive list of the most famous songs of 2023 as listed on Spotify. The dataset offers a wealth of features beyond what is typically available in similar datasets. It provides insights into each song's attributes, popularity, and presence on various music platforms. The dataset includes information such as **track name, artist(s) name, release date, Spotify playlists and charts, streaming statistics, Apple Music presence, Deezer presence, Shazam charts, and various audio features**.

* **track\_name**: Name of the song
* **artist(s)\_name**: Name of the artist(s) of the song
* **artist\_count**: Number of artists contributing to the song
* **released\_year**: Year when the song was released
* **released\_month**: Month when the song was released
* **released\_day**: Day of the month when the song was released
* **in\_spotify\_playlists**: Number of Spotify playlists the song is included in
* **in\_spotify\_charts**: Presence and rank of the song on Spotify charts
* **streams**: Total number of streams on Spotify
* **in\_apple\_playlists**: Number of Apple Music playlists the song is included in
* **in\_apple\_charts**: Presence and rank of the song on Apple Music charts
* **in\_deezer\_playlists**: Number of Deezer playlists the song is included in
* **in\_deezer\_charts**: Presence and rank of the song on Deezer charts
* **in\_shazam\_charts**: Presence and rank of the song on Shazam charts
* **bpm**: Beats per minute, a measure of song tempo
* **key**: Key of the song
* **mode**: Mode of the song (major or minor)
* **danceability\_%**: Percentage indicating how suitable the song is for dancing
* **valence\_%**: Positivity of the song's musical content
* **energy\_%**: Perceived energy level of the song
* **acousticness\_%**: Amount of acoustic sound in the song
* **instrumentalness\_%**: Amount of instrumental content in the song
* **liveness\_%**: Presence of live performance elements
* **speechiness\_%**: Amount of spoken words in the song

**Health and Lifestyle Data**

**synthetic\_health\_data.csv**

This dataset is designed to study the impact of various lifestyle factors on an individual's health score using linear regression analysis. It provides a comprehensive view of factors such as age, BMI, exercise frequency, diet quality, sleep hours, smoking status, and alcohol consumption. The dataset can be used for exploratory data analysis, machine learning models, and understanding correlations between health and lifestyle.

Columns Description

* **Age:** Age of the individual in years (continuous variable).
* **BMI**: Body Mass Index of the individual (continuous variable).
* **Exercise\_Frequency**: Number of days per week the individual exercises (categorical, values 0-7).
* **Diet\_Quality**: An index reflecting diet quality, with higher values indicating healthier dietary habits (continuous, 0-100).
* **Sleep\_Hours**: Average hours of sleep per night (continuous).
* **Smoking\_Status**: Binary variable where 0 = Non-smoker, 1 = Smoker.
* **Alcohol\_Consumption**: Average alcohol units consumed per week (continuous).
* **Health\_Score**: A calculated health score reflecting overall health status (continuous, 0-100).