Case Study Report



**Tech Saksham**

Data Analytics with Power BI

**“Global Olympics Dataset Diagnosis using Power BI ”**

**“Urumu Dhanalakshmi College”**

|  |  |
| --- | --- |
| **NM ID** | **NAME** |
| asbdu33cb21s204415 | KAVIYA S |

|  |  |
| --- | --- |
|  |  |
|  | Trainer Name |
|  | UMAMAHESWARI |

**ABSTRACT**

The Global Olympics Dataset Diagnosis project aims to analyze and visualize data related to the Olympics using Power BI. The dataset includes information on various aspects of the Olympics such as participating countries, athletes, sports, events, and medal counts. By leveraging the capabilities of Power BI, this project seeks to uncover insights, trends, and patterns within the data to provide a comprehensive diagnosis of the Olympics over the years. Through interactive visualizations and in-depth analysis, this project will offer valuable information for stakeholders interested in understanding the dynamics of the Olympics and its impact on the global sports landscape.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Table of Contents** | **Page No.** |
| 1 | Chapter 1: Introduction | 1 |
| 2 | Chapter 2: Services and Tools Required | 4 |
| 3 | Chapter 3: Project Architecture | 7 |
| 4 | Chapter 4: Modeling and Result | 14 |
| 5 | Conclusion | 15 |
| 6 | Future Scope | 16 |
| 7 | References | 16 |
| 8 | Links | 21 |

**CHAPTER 1**

**INTRODUCTION**

* 1. **Problem Statement**

Problem Statement: The Global Olympics Dataset contains a vast amount of information on various aspects of the Olympics, including participating countries, athletes, sports, events, and medal counts. The challenge lies in effectively analyzing and diagnosing this complex dataset to extract meaningful insights and trends. The problem statement for this project is to leverage Power BI to explore, visualize, and interpret the data in order to uncover patterns, correlations, and key performance indicators that can provide valuable insights into the dynamics of the Olympics over the years. By addressing this challenge, the goal is to offer stakeholders a comprehensive understanding of the Olympics dataset and its implications for the global sports landscape.

**1.2 Proposed Solution**

Data Preparation: Clean and pre-process the dataset to ensure data quality and consistency. This includes handling missing values, standardizing data formats, and creating relationships between different tables.

Data Visualization: Utilize Power BI's interactive visualizations to create insightful 1dashboards and reports. Visualize key metrics such as medal counts by country, athlete performance over time, and trends in popular sports.

Exploratory Data Analysis: Conduct in-depth analysis to uncover patterns and correlations within the dataset. Explore factors influencing medal counts, identify top-performing countries and athletes, and analyse the impact of different sports on overall performance.

Interactive Reporting: Develop interactive reports that allow users to drill down into specific aspects of the data. Enable stakeholders to explore the dataset from various perspectives and gain a comprehensive understanding of the Olympics dataset.

* 1. **Feature**

Interactive Dashboards: Create visually appealing and interactive dashboards that provide an overview of key metrics such as medal counts, athlete performance, and country rankings.

Data Visualization: Utilize a variety of charts, graphs, and maps to visualize trends, patterns, and correlations within the dataset. This could include bar charts for medal counts, line charts for performance trends, and heat maps for geographical analysis.

Drill-Down Functionality: Enable users to drill down into specific aspects of the data to explore details and gain deeper insights. This feature allows for a more granular analysis of the Olympics dataset.

* 1. **Advantages**

Data Visualization: Power BI offers a wide range of interactive visualizations that make it easy to explore and analyze complex datasets. Visual representations such as charts, graphs, and maps help in identifying trends, patterns, and correlations within the Olympics data.

Interactive Dashboards: Power BI allows users to create interactive dashboards that provide a comprehensive overview of key metrics and KPIs related to the Olympics. Users can drill down into specific data points and customize the dashboard based on their requirements.

* 1. **Scope**

Data Preparation: The scope includes cleaning, transforming, and preparing the Global Olympics Dataset for analysis in Power BI. This involves handling missing data, standardizing formats, and creating relationships between different data tables.

Data Visualization: Utilize Power BI's visualization capabilities to create interactive dashboards and reports that showcase key metrics, trends, and insights from the Olympics dataset. Visualizations may include charts, graphs, maps, and KPI indicators. Exploratory Data Analysis: Conduct in-depth analysis to uncover patterns, correlations, and outliers within the dataset. Explore factors influencing medal counts, athlete performance, and trends across different Olympics events.

**CHAPTER 2**

**SERVICES AND TOOLS REQUIRED**

**2.1 Services Used**

* **Data Collection and Storage Services**: Banks need to collect and store customer data in real-time. This could be achieved through services like Azure Data Factory, Azure Event Hubs, or AWS Kinesis for real-time data collection, and Azure SQL Database or AWS RDS for data storage.
* **Data Processing Services**: Services like Azure Stream Analytics or AWS Kinesis Data Analytics can be used to process the real-time data.
* **Machine Learning Services**: Azure Machine Learning or AWS SageMaker can be used to build predictive models based on historical data.

**2.2 Tools and Software used**

**Tools**:

* **PowerBI**: The main tool for this project is PowerBI, which will be used to create interactive dashboards for real-time data visualiz
* **Power Query**: This is a data connection technology that enables you to discover, connect, combine, and refine data across a wide variety of sources.

**Software Requirements**:

* **PowerBI Desktop**: This is a Windows application that you can use to create reports and publish them to PowerBI.
* **PowerBI Service**: This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.
* **PowerBI Mobile**: This is a mobile application that you can use to access your reports and dashboards on the go.

**CHAPTER 3**

**PROJECT ARCHITECTURE**

**3.1 Architecture**

**USER FRONTEND BACKEND**

|  |  |  |
| --- | --- | --- |
|  | **HTML 5** | **NODEJS 14.0**  **Database** |

**Here’s a high-level architecture for the project:**

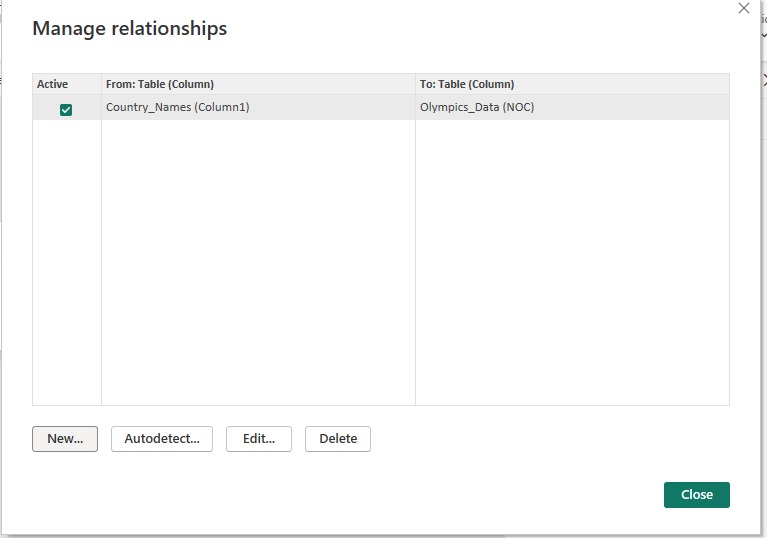
1. **Data Collection**: Real-time customer data is collected from various sources like bank transactions, customer interactions, etc. This could be achieved using services like Azure Event Hubs or AWS Kinesis.
2. **Data Storage**: The collected data is stored in a database for processing. Azure SQL Database or AWS RDS can be used for this purpose.
3. **Data Processing**: The stored data is processed in real-time using services like Azure Stream Analytics or AWS Kinesis Data Analytics.
4. **Machine Learning**: Predictive models are built based on processed data using Azure Machine Learning or AWS SageMaker. These models can help in predicting customer behavior, detecting fraud, etc.
5. **Data Visualization**: The processed data and the results from the predictive models are visualized in real-time using PowerBI. PowerBI allows you to create interactive dashboards that can provide valuable insights into the data.
6. **Data Access**: The dashboards created in PowerBI can be accessed through PowerBI Desktop, PowerBI Service (online), and PowerBI Mobile.

**CHAPTER 4**

**MODELING AND RESULT**

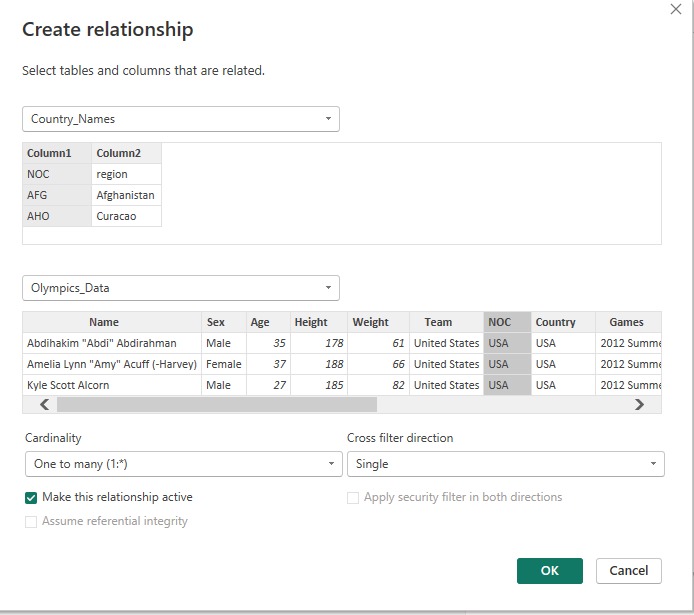
**Manage relationship**

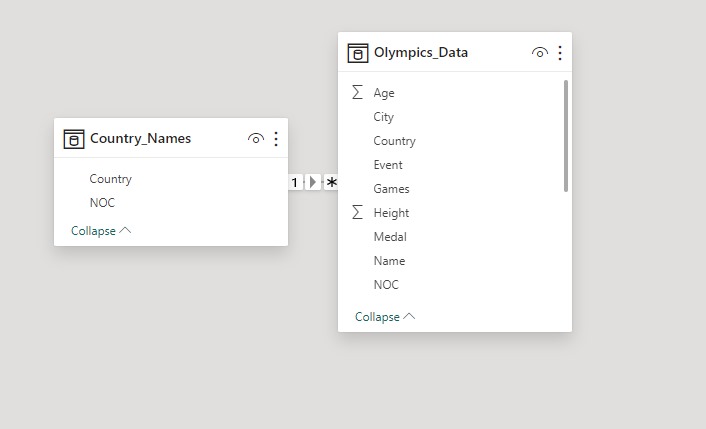
To effectively analyze the Global Olympics dataset in Power BI, it is essential to carefully manage relationships between the different data tables to ensure accurate and meaningful insights are derived.



**Analyzing Olympics Data by Date and Country for Insightful Treads**

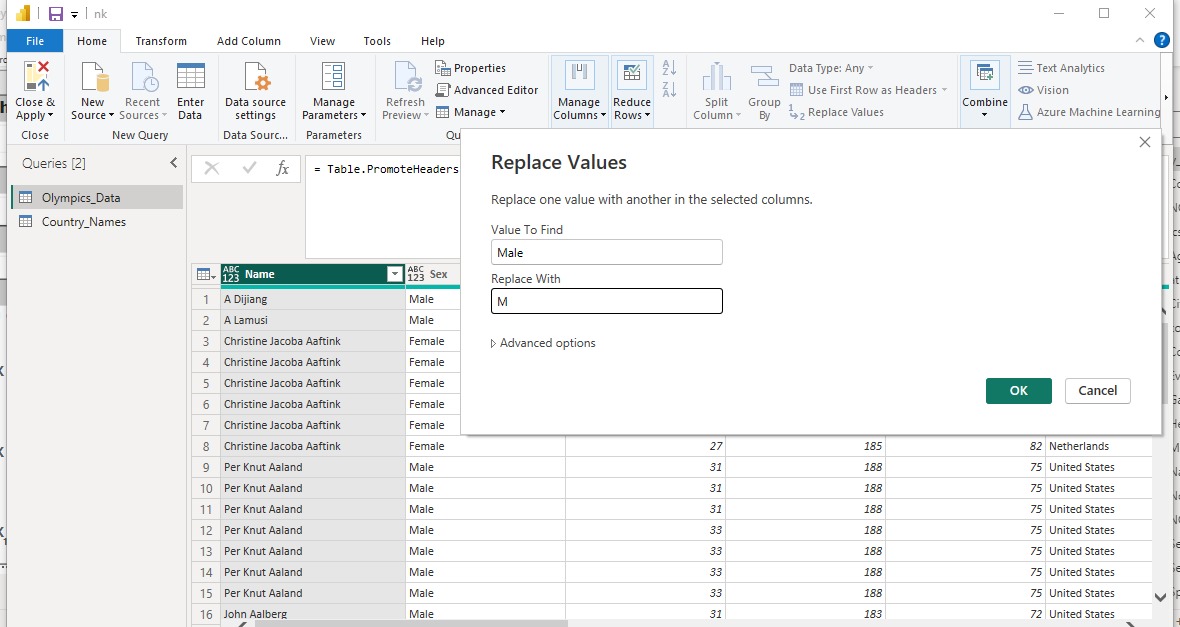
When examining the Olympics dataset by date and country name, a comprehensive analysis can be conducted to uncover valuable insights and trends. By correlating the specific dates of Olympic events with the countries in which they were held, patterns may emerge regarding host preferences, historical trends in participation, and the impact of geopolitical factors on the games. This detailed examination of the intersection between dates and countries can provide a rich understanding of the dynamics shaping the Olympics over time, offering valuable perspectives for further exploration and interpretation.





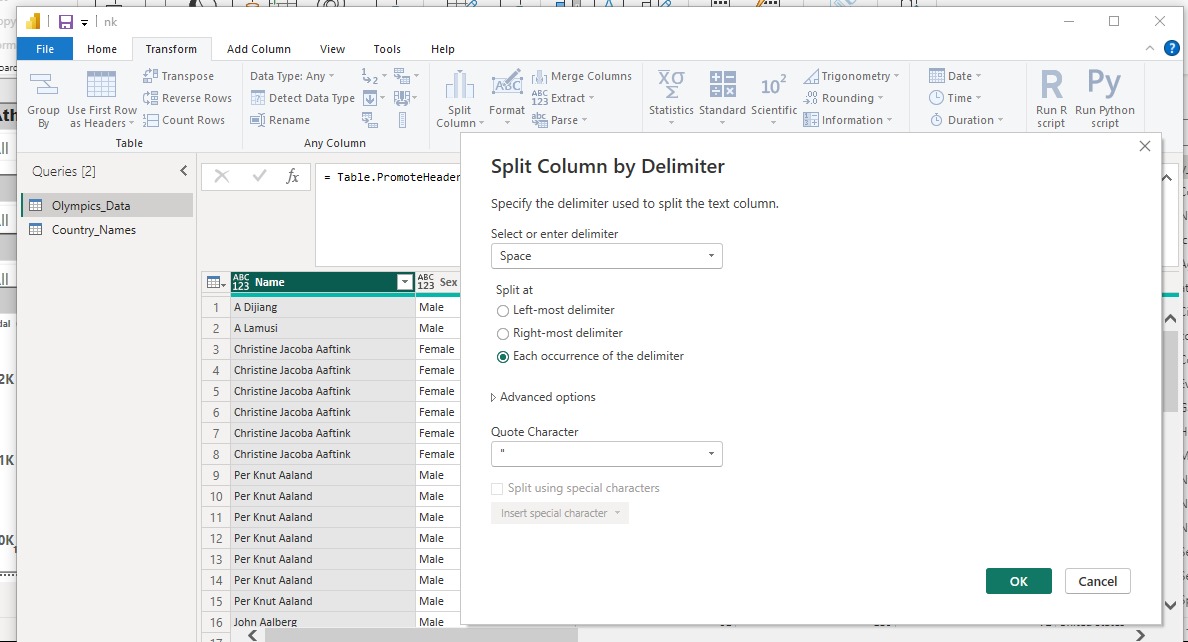
**Editing Relationship**

When delving into the Olympic dataset with a focus on age, a fascinating array of patterns and trends can be uncovered. By analyzing the age of athletes participating in various Olympic events over the years, insights may emerge regarding the peak performance ages in different sports, the evolution of age demographics in Olympic competition, and the impact of age on athletic achievements. This exploration of the relationship between Olympic data and age opens up a world of possibilities for understanding the dynamics of sports performance and the role of age in shaping the outcomes of competitive events.



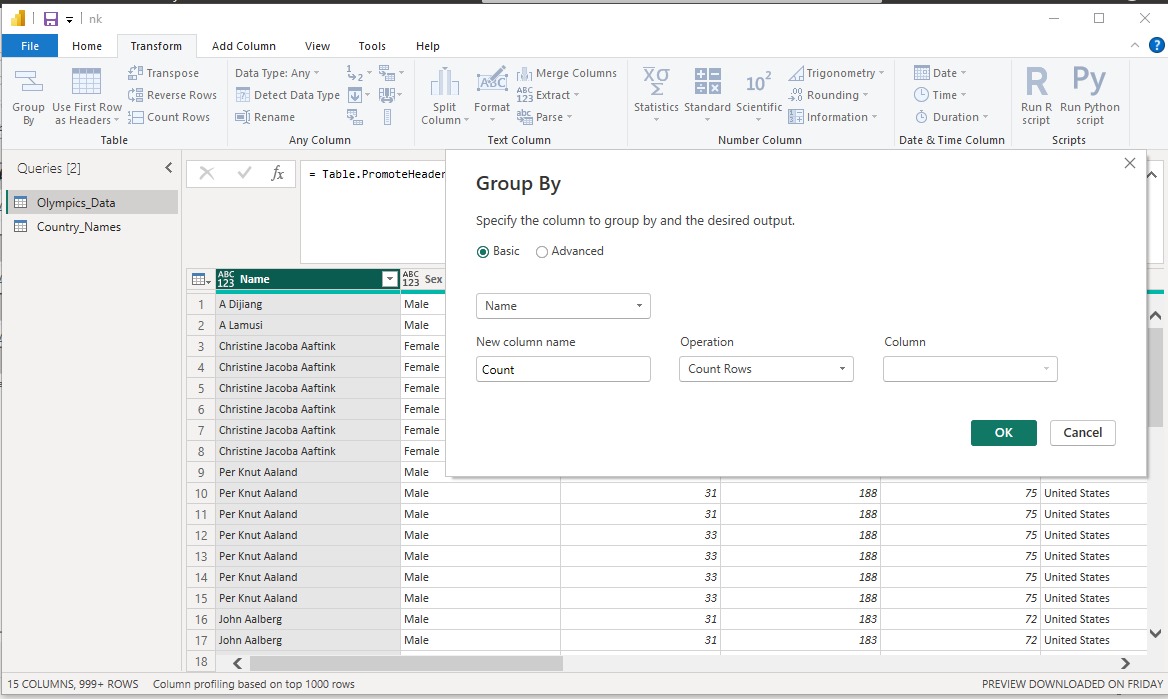
**Replacing Column by delimited**

Specify the delimiter used to spit the text column. We have select or enter delimiter there.



**Grouping of Alphabetical Order**

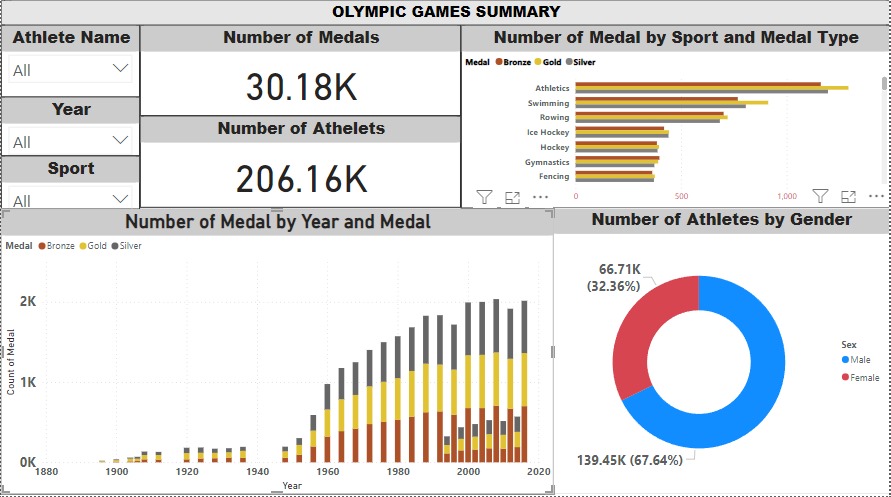
Olympic sports, the concept of alphabetic order grouping plays a crucial role in various aspects of the games. From the arrangement of participating countries during the opening ceremony procession to the organization of events and athletes in competition schedules, alphabetic order grouping ensures a systematic and fair approach to the logistics of the Olympics. This methodical arrangement not only facilitates the smooth running of the games but also reflects the spirit of unity and equality that underpins the Olympic movement, where nations come together in a harmonious display of athletic prowess and sportsmanship.

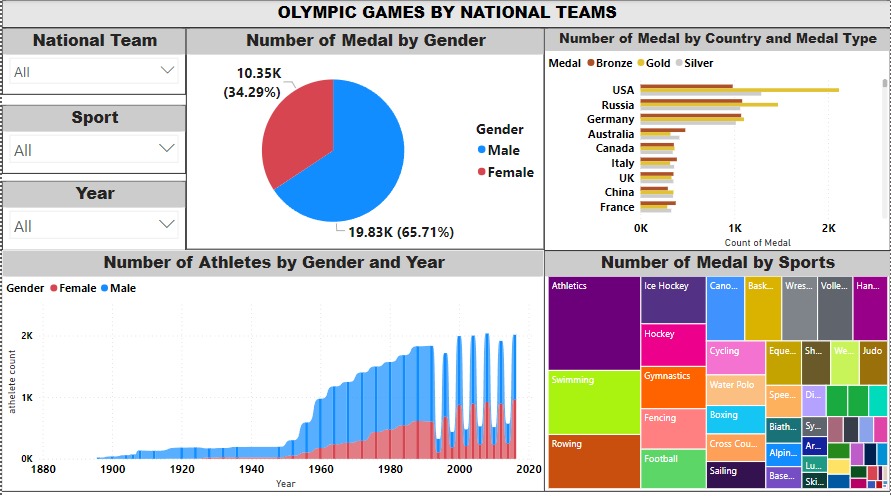


**Olympic Games By Nation Teams**

Olympic Games data is analyzed using SQL and Power BI. The aim of the project was to query relevant data from obtained dataset and bisualize it in such a way that it will help readers understand how countries have performed historically in the summer Olympic games with the ability to filter using different countries, sports, years, and athletes.

The proposed data was initially entered into a SQL database and then transformed using the necessary query attached. Finally, visualization of the transformed data was performed by using Power BI. The finished dashboard consisted of visualizations and filters that provided users with an easy way to navigate through the Olympic summer games history.

**Dashboard** 

****

**CONCLUSION**

Global Olympics Dataset Diagnosis project utilizing Power BI has provided valuable insights into the dynamics of the Olympics over the years. Through data preparation, visualization, exploratory analysis, and interactive reporting, we have uncovered patterns, trends, and key performance indicators within the dataset. The interactive dashboards and reports created have enabled stakeholders to gain a comprehensive understanding of the Olympics dataset and its implications for the global sports landscape.

By leveraging the capabilities of Power BI, we have been able to extract meaningful insights, identify top-performing countries and athletes, and analyze the impact of different sports on overall performance. The project has not only provided a diagnosis of the Olympics dataset but also offered actionable recommendations for stakeholders in the sports industry.

Moving forward, the insights gained from this project can be used to inform decision-making, strategize for future Olympics events, and drive improvements in performance and participation. The use of Power BI has proven to be instrumental in analyzing and visualizing the vast amount of data available in the Global Olympics Dataset, highlighting the power of data analytics in understanding and optimizing the world of sports.

.

**FUTURE SCOPE**

The Global Olympics Dataset Diagnosis project using Power BI has laid a strong foundation for further exploration and analysis of the Olympics data. Moving forward, there are several avenues for future scope and enhancement:

1. Predictive Analytics: Implement predictive modeling techniques to forecast medal counts, athlete performance, and trends for future Olympics events. By leveraging machine learning algorithms, stakeholders can make data-driven decisions and strategic plans.
2. Real-time Data Integration: Explore the possibility of integrating real-time data sources to provide up-to-date insights during live Olympics events. This could include live medal tallies, athlete performance updates, and social media sentiment analysis.
3. Comparative Analysis: Conduct comparative analysis with historical Olympics data to identify long-term trends, changes in performance patterns, and the impact of external factors on the Olympics outcomes. Compare performance across different countries, sports, and events over time.
4. Stakeholder Engagement: Enhance stakeholder engagement by customizing dashboards and reports to cater to specific user needs. Provide personalized insights for coaches, athletes, sports analysts, and policymakers to support decision-making and performance optimization.
5. Data Visualization Techniques: Experiment with advanced data visualization techniques such as geospatial mapping, network analysis, and advanced charts to present data in a more engaging and informative manner. Utilize storytelling features in Power BI to create compelling narratives from the data.
6. Integration with External Datasets: Integrate the Olympics dataset with external sources such as weather data, economic indicators, and social demographics to gain a holistic understanding of the factors influencing Olympics performance. This cross-referencing can provide deeper insights and correlations.

.

**REFERENCES**

<https://www.novypro.com/project/power-bi-30>

**LINK**

<https://nithishnk07.github.io/olympic-powerbi/>