**DATA VISUALIZATION**

**TERM PROJECT FINAL**

Kim, Suchul - 300362625

Haridas, Sudarsan - 300353099

Douglas College

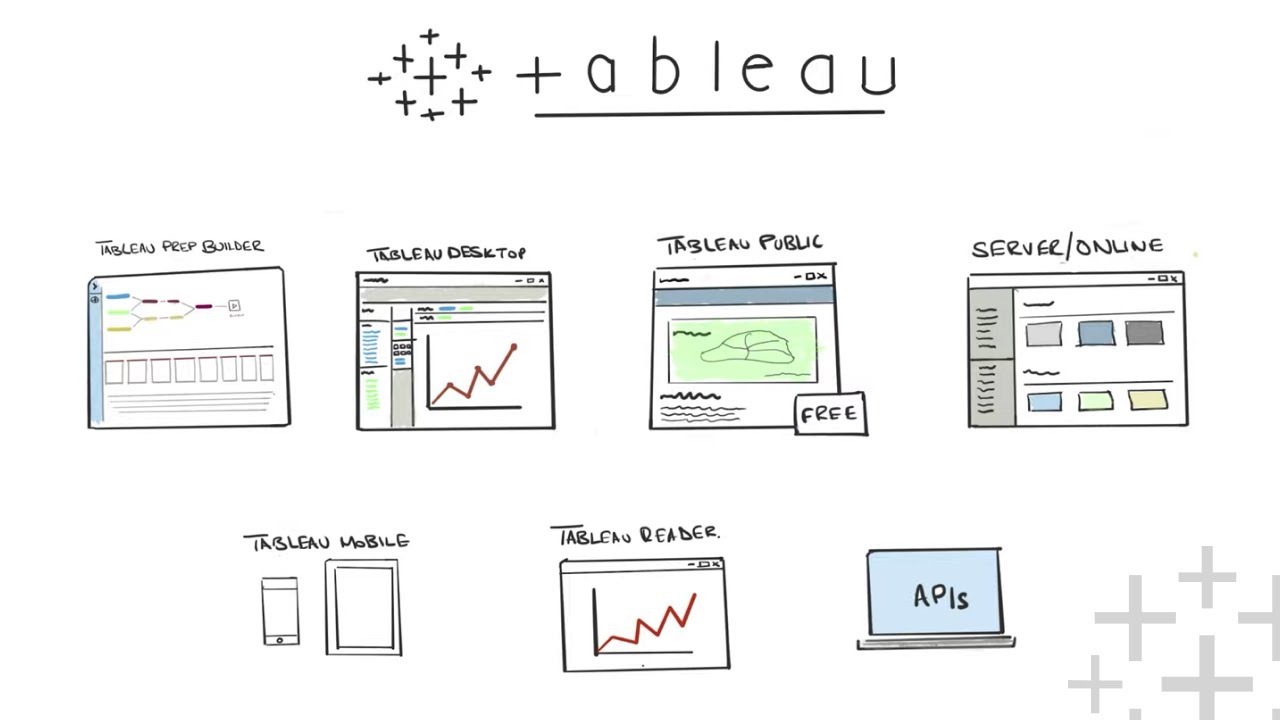
CSIS 3860 - 070, Fall 2023: DATA VISUALIZATION

Professor: Rupa Manabala

Dec 06, 2023



[This Photo](https://crm-powerbi-viewer.heiigjen.com/) by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/3.0/)



[This Photo](https://www.tableautim.com/playlist/tableau-explained?93997e09_page=2) by Unknown Author is licensed under [CC BY-NC-ND](https://creativecommons.org/licenses/by-nc-nd/3.0/)

**PART A**

**The details of the data set –** number of fields, rows, and data size

The Olympic dataset contains four tables: olympic\_athletes, olympic\_hosts, olympic\_medals and olympic\_results. olympic\_athletes has 7 fields and 75904 rows, ,olympic\_hosts has 7 fields and 53 rows, olympic\_medals has 8 fields and 21697 rows, olympic\_results has 9 fields and 162804 rows, Data set size is 45,990KB.

**Analysis questions**

* Which countries have been the most successful in terms of total medals won across all Olympic Games?
* How has the total number of medals awarded evolved over the years?
* What trends can be observed in gender diversity over the history of the Olympic Games?
* Which country has hosted the most Olympics?
* Do host countries tend to win more gold medals?

**Data Connections**

Import data from 4 Olympics csv files into power bi

A screenshot of a computer

Description automatically generated

**Data Transform Strategies**

* **Use first row as a Header** : Applies to the first row of Olympic medals.

A screenshot of a computer

Description automatically generated

* **Replace Value** : Delete erroneous characters, such as Ã‰pÃ©e.

A screenshot of a computer

Description automatically generated

* **Remove Columns** : Delete unnecessary column that will not be used.

**A screenshot of a computer

Description automatically generated**

* **Capitalize Each Word** : Change the case of the athlete full name.

A screenshot of a computer

Description automatically generated

* **Add Conditional Column** : Set the blank field of value type to null.

**A screenshot of a computer

Description automatically generated**

* **Split Column by Delimiter** : Split the slug game into columns by city and year.

**A screenshot of a computer

Description automatically generated**

* **Merge Query** : Merge the tables to create an Olympic total table.

**A screenshot of a computer

Description automatically generated**

**Building the Data Model**

The tables are connected based on the athlete’s full name. The olympic host table functions as a date.

**A screenshot of a computer

Description automatically generated**

**Leveraging DAX**

Game year and game month from the olympic\_hosts table to create a Year-Month.

**A screenshot of a computer

Description automatically generated**

The average number of medals by sport

**A screenshot of a computer

Description automatically generated**

The number of rows in the olympic\_results table where the medal type is "Gold."

**A screenshot of a computer

Description automatically generated**

Total medal count up to the current context's maximum year**A screenshot of a computer

Description automatically generated**

Total number of medals by counting the rows where the medal type is either "Gold," "Silver," or "Bronze."

**A screenshot of a computer

Description automatically generated**

Total number of medals up over the game year.

**A screenshot of a computer

Description automatically generated**

**Visualization Data**

**Bar-Chart**

﻿United States of America in medal\_type GOLD made up 0.59% of Count of medal.

﻿Michael Phelps in medal\_type GOLD made up 0.09% of Count of medal.

A screenshot of a graph

Description automatically generated

**Pie-Chart**

**﻿**Men had the highest Count of event\_gender at 13,932(64%), followed by Women(29%)

**﻿**Men had the highest Gold Medal Count at 4856(42%), followed by Women(27%)

**A close-up of a pie chart

Description automatically generated**

**Time-Series**

Depending on the year, the number of medals for gold, silver, and bronze is all on the rise, but has been stagnant since 2000. The number of men and women is increasing by year, and has become saturated since 2000.

A screenshot of a graph

Description automatically generated

**Geographic**

The United States has won the most gold, silver, and bronze medals in Olympic history.

France, Japan, and the United States hosted the most Olympic Games, four each.

A screenshot of a map

Description automatically generated

**Bookmarks**

Using bookmarks, we can easily observe changes in the number of medals and people according to gender and year.

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Answer your analysis questions through your reports**

Q : Which countries have been the most successful in terms of total medals won across all Olympic Games?

Bar Chart: This chart shows the percentage of gold medals obtained by each country. The United States has won the most gold(2595ea), silver(2475ea), and bronze medals(2376ea) in Olympic history.

Q. How has the total number of medals awarded evolved over the years?

Time-Series Chart: This chart displays the trend of gold, silver, and bronze medals over the years. The number of medals for gold, silver, and bronze is on the rise, but the growth has been stagnant since the year 2000.

Q. What trends can be observed in gender diversity over the history of the Olympic Games?

Pie Chart and Time-Series Chart : This chart compares the number of medals by gender in the Olympics. The number of medals and participants for both men and women is increasing by year, and has become saturated since the year 2000. Men had the highest count of event\_gender at 13,932 (64%), followed by women (29%). Men also had the highest Gold Medal Count at 4856 (42%), followed by women (27%).

Q. Which country has hosted the most Olympics?

Geographic Analysis: Through geographic analysis, it is observed that France, Japan, and the United States have each hosted the Olympics four times. This can be confirmed through a map or bar chart displaying the number of Olympics hosted by each country.

Q. Do host countries tend to win more gold medals?

Bookmarks show that hosts Russia and China won the most golds at their 2014 and 2022 Olympics, respectively, hinting at a home advantage.

**Power BI Explanation Video**

[**https://youtu.be/zl7aJWBjHn0**](https://youtu.be/zl7aJWBjHn0)

**PART B**

**Analysis questions**

* How might the number of patients visiting the clinic change next month?
* What is the common age range of our clinic's patients?
* Over the past few years, have the number of patient visits been going up or down?
* Which department has patients waiting the longest?
* How long do patients typically spend in appointments in different locations?

**Data Connections**

* **Connecting to Data Sources :** Import data from Clinical Management System.json into tableau.
* **Inspect and Clean Data:** Performs automatic cleaning.

**A screenshot of a computer

Description automatically generated**

**Sorting and Filtering :** Sorting and Filtering by Average Age of Patients

The data is organized by location and department, with color coding to differentiate between departments like General Medicine, Dermatology, Orthopaedics, Ophthalmology, and Pediatrics. For instance, in Alabama, the average patient age ranges from 41 in Ophthalmology to 55 in General Medicine.

A screenshot of a computer

Description automatically generated

**Map** : The locations with the top 10 most visits recorded.

The states of Arizona, Illinois, Virginia, California, Florida, Texas, and New York are labeled with numbers that presumably reflect the count of visits or a similar metric, with Arizona showing the highest number at 10 and Texas the lowest among the labeled states with 1.

**A screenshot of a computer

Description automatically generated**

**Creating Calculated Fields** : The average visit duration in each location is plotted along with average appointment duration as reference line for proper resource allocation.

The chart compares the average duration of visits in four departments: General Medicine, Ophthalmology, Orthopaedic, and Pediatrics. General Medicine has the longest average visit duration, marked at 40 minutes, while Ophthalmology has the shortest, at 27 minutes.

A screenshot of a computer program

Description automatically generated

* This logic ensures that the visit duration is calculated correctly across hours, without incorrectly adding 60 minutes for visits that don't span two hours

**A screenshot of a computer

Description automatically generated**

**Table Calculations** : Average wait time and time exceeded after the appointment ended is plotted by department and ranked by average wait times.

* Dermatology has the longest average wait time at about 8.22 minutes, followed by Pediatrics, Orthopaedic, Ophthalmology, and General Medicine.

**A screenshot of a computer

Description automatically generated**

**Trendlines** : Linear Trendline for number of patients in each month.

The actual patient counts per month are represented by a green line with points, and the linear trendline is shown as a dashed line, with upper and lower confidence bands indicated by additional dashed lines. The graph exhibits fluctuations in the monthly patient count but the trendline suggests a slight upward trajectory over time.

**A screenshot of a computer

Description automatically generated**

**Visualizing Distributions** : The distribution of patient Age.

The number of patients according to their age does not have a special distribution form. The number of patients generally increases from those in their 40s to 60s, but decreases after the age of 70.

**A screenshot of a graph

Description automatically generated**

**Forecasting** : Forecasting patient visits for the upcoming months

Patient visits are expected from 2020 to 2023, and by 2024, the average is expected to be about 24, the max is about 33, and the min is about 15.

A screenshot of a computer

Description automatically generated

**Answer your analysis questions through your reports**

How might the number of patients visiting the clinic change next month?

The forecast for patient visits shows an expected average of about 24 visits, with a maximum of around 33 and a minimum of about 15 by 2024. This indicates a potential range for the number of patients visiting next month, assuming a steady trend.

What is the common age range of our clinic's patients?

The distribution of patient ages shows an increase in the number of patients from their 40s to 60s, with a decrease after the age of 70. This suggests that the common age range of the clinic's patients is between 40 and 70 years.

Over the past few years, have the number of patient visits been going up or down?

The trendline analysis indicates a slight upward trajectory over time, despite monthly fluctuations. This suggests that the number of patient visits has been generally increasing.

Which department has patients waiting the longest?

Dermatology has the longest average wait time at about 8.22 minutes, followed by Pediatrics, Orthopedic, Ophthalmology, and General Medicine. Therefore, the Dermatology department has the longest waiting times for patients.

How long do patients typically spend in appointments in different locations?

The average visit duration varies by department. General Medicine has the longest average duration at 40 minutes, while Ophthalmology has the shortest at 27 minutes. This data indicates the typical time patients spend in appointments across different departments.

**References**

Microsoft Power BI Quick Start Guide – Preferably 3rd edition- Devin Knight, Mitchell Pearson, Bradley Schacht,, et al – ISBN 9781804613498 – Packt Publishing – eBook available from publisher’s website (<https://www.packtpub.com/product/microsoft-power-biquick-start-guide-second-edition/9781804613498>)

Microsoft Ignite

<https://learn.microsoft.com/en-us/power-bi/>

Tableau Desktop and Web Authoring Help

<https://help.tableau.com/current/pro/desktop/en-us/calculations_tablecalculations.htm>

Learning Tableau 2022 – 5th Edition by Joshua Milligan ISBN 9781801072328 – Packt Publishing - eBook available from publisher’s website and/or Book store (<https://www.packtpub.com/product/learning-tableau-2022-fifth-edition/9781801072328>)