***AnandTheAnalyst***

Data Analyst Portfolio Project – Olympic Games Analyst

# Business Problem:

The Challenge for this data analyst project is outlined below. This has been used continuously to ensure that the right data has been selected, transformed and used in the data visualization, which is meant to be passed on the business users.

"**As a data analyst working at a news company you are asked to visualize data that will help readers understand how countries have performed historically in the summer Olympic Games.**

**You also know that there is an interest in details about the competitors, so if you find anything interesting then do not hesitate to bring that in also.**

**The main task is still to show historical performance for different countries, with the possibility to select your own country**."

# Data Collection & Table Structure.

The necessary data was first put into a SQL Database and afterwards transformed using the transformation that you can see below.

# SQL Transformation Code:

/\*\*\*\*\*\* Script for SelectTopNRows command from SSMS \*\*\*\*\*\*/

SELECT

[ID],

[Name] AS 'Competitor Name' --- Renamed Column

,

Case When Sex = 'M' THEN 'Male' ELSE 'FEMALE' END AS SEX -- Better Name for filtrer & visualizations

,

[Age],

CASE WHEN [AGE] < 18 THEN 'Under 18' WHEN [Age] Between 18

AND 25 THEN '18-25' WHEN [Age] Between 25

AND 30 THEN '25-30' WHEN [Age] > 30 THEN 'OVER 30' END as [Age Grouping],

[Height],

[Weight],

[NOC] As 'Nation Code' -- To exlain the Abbrevation

--,[Games]

,

LEFT(

Games,

CHARINDEX(' ', Games)-1

) as 'Year' -- Extract the Year from Gemas Column

,

SUBSTRING(

Games,

CHARINDEX(' ', Games)+ 1,

LEN(Games)

) as Season --Extract the Season from the Games column

--,[City]Commented out as it is not neccessary for the analysis

,

[Sport],

[Event],

CASE WHEN [Medal] = 'NA' THEN 'Not Registered' ELSE Medal END as Medal -- Replace NA with NOT REGISTERED

FROM

[Olympic\_Games].[dbo].[athletes\_event\_results]

Where

SUBSTRING(

Games,

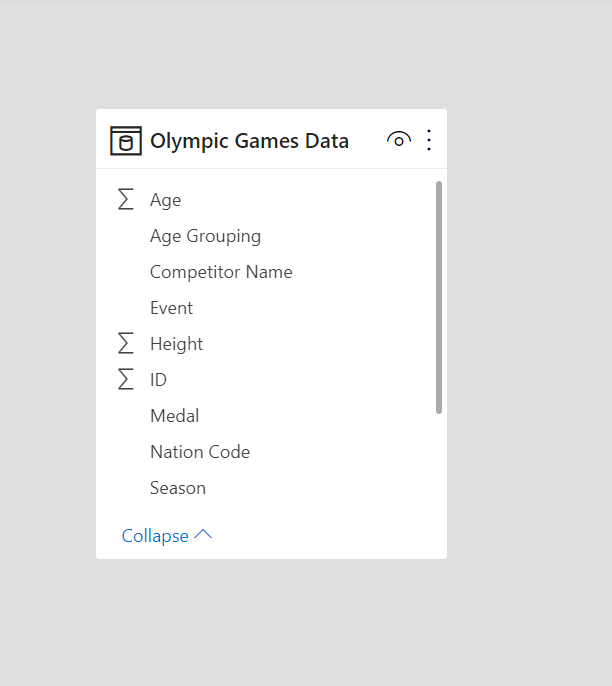
CHARINDEX(' ', Games)+ 1,

LEN(Games)

) = 'Summer' -- Where clause to filter the Summer Season

# Data Model:

As this is a view where dimension and facts have been combined, the data model that is created in the Power BI is one table. The query from previous step loaded in directly.



# Calculations:

The Following calculations was created in the Power BI reports using DAX (Data Analysis expression)

**# Of Competitors** = DISTINCTCOUNT('Olympic Games Data'[ID])

**# Of Medals(Registered**) =

CALCULATE (

'Calculations'[# Of Medals],

FILTER (

'Olympic Games Data',

'Olympic Games Data'[Medal] = "Bronze"

|| 'Olympic Games Data'[Medal] = "Silver"

|| 'Olympic Games Data'[Medal] = "Gold"

)

)

# Olympic Games Analysis:

The finished dashboard consist of visualizations and filters that gives an easy option for the end user to navigate the summer games through history. Some possibilities are to filter by period using year, nation code for focus on one country or look into either a competitor or specific sports over time.

