Washington State University
MIS 420 – Business Intelligence
Online

Mauricio Featherman, Ph.D.

T-SQL #5

Name: Nam Jun Lee

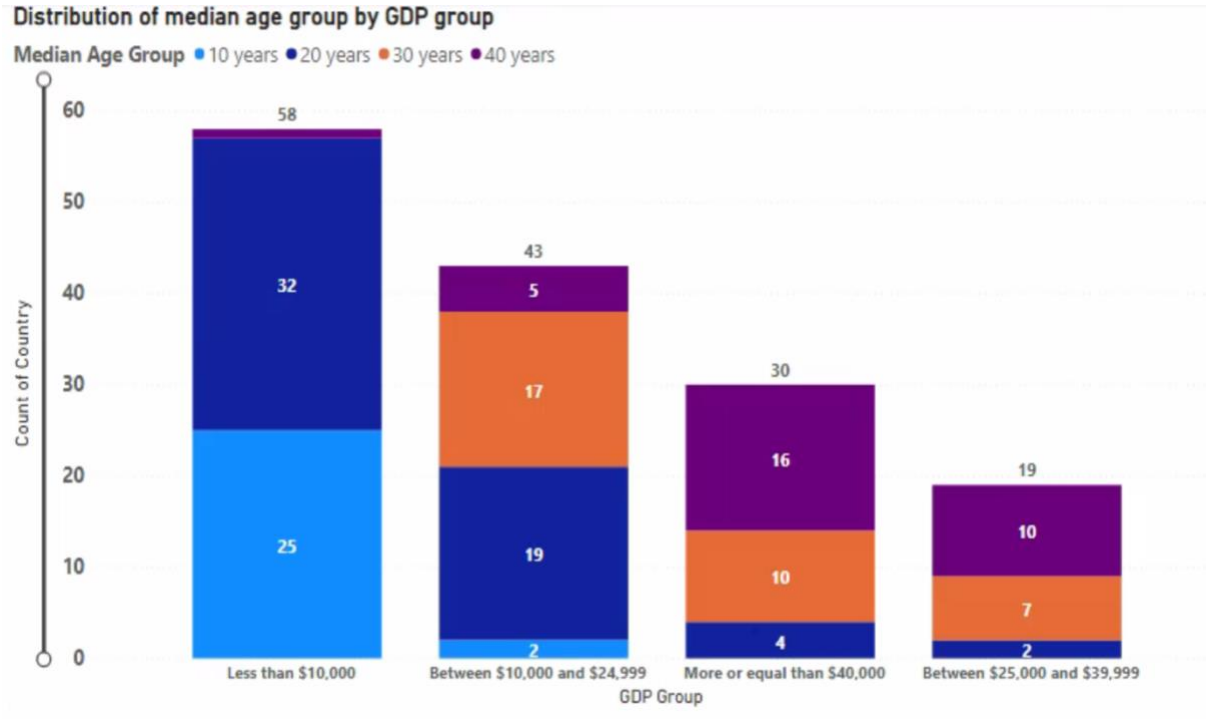
Student Number: 11606459

Introduction

The project aims to create a large array by using the country's GDP and some variables, and by correlating each variable, it proposes three good vacation spots for the requester to go on an overseas travel adventure after graduation.

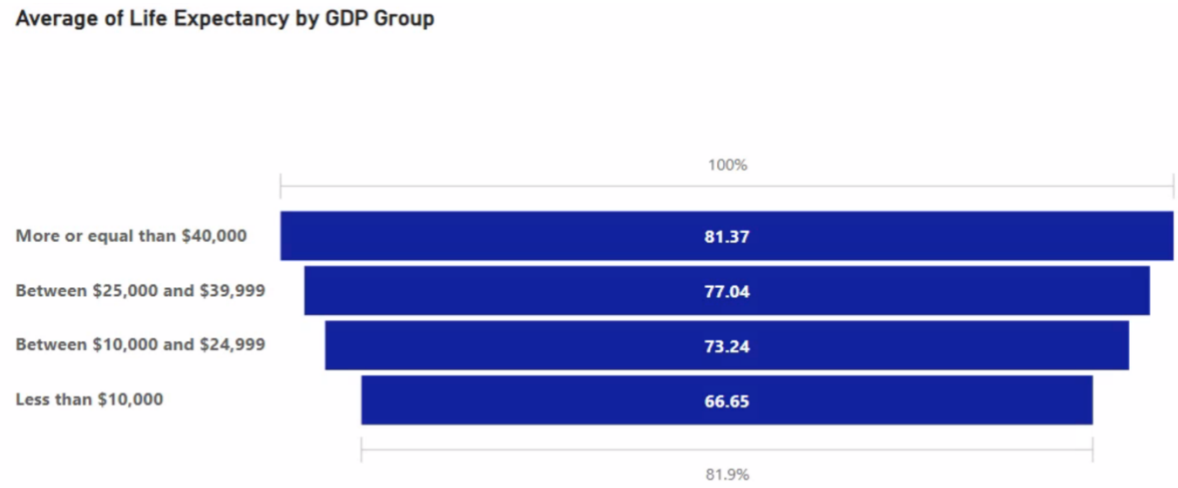
Data Analysis

1. Distribution of median age group by GDP group



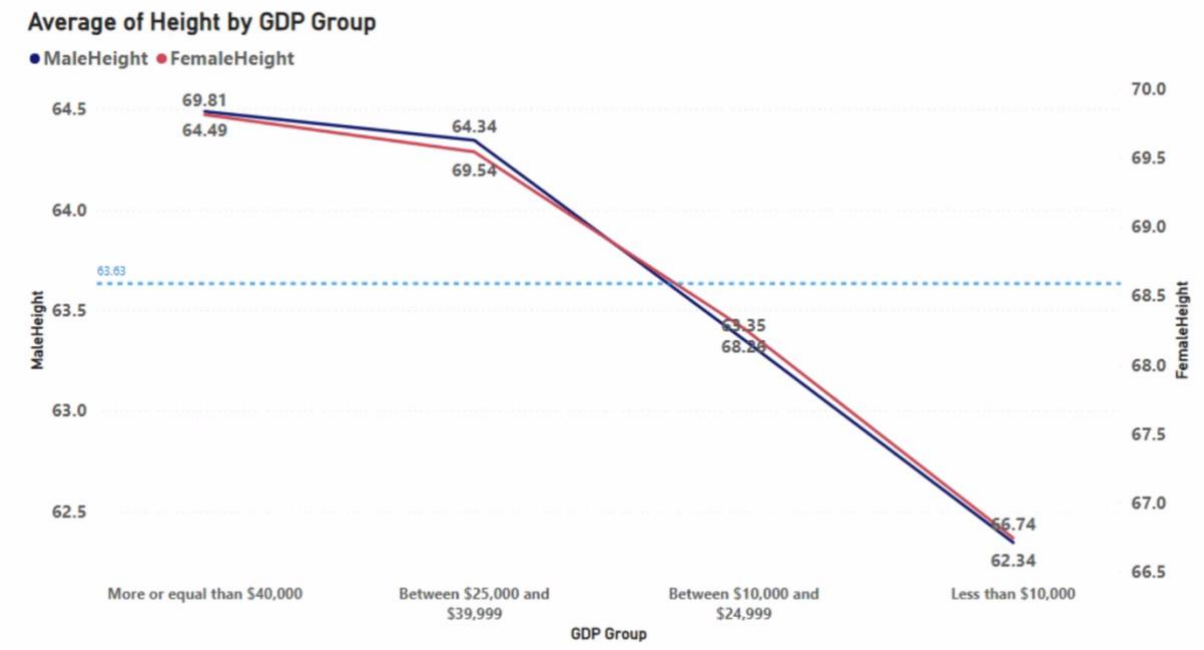
The median age group for each GDP group can be identified through the stack column graph above. From this, the distribution of groups with a GDP less than $10,000 is the most distributed, and the distribution of groups with a GDP between $25,000 and $39,999 is the least distributed. Also, the median age group with a GDP of less than $10,000 is distributed only in their teens and twenties, and as GDP increases, the median age group increases. Through this, it can be said that there is a correlation between GDP and the median age group.

2. Average of life expectancy by GDP Group



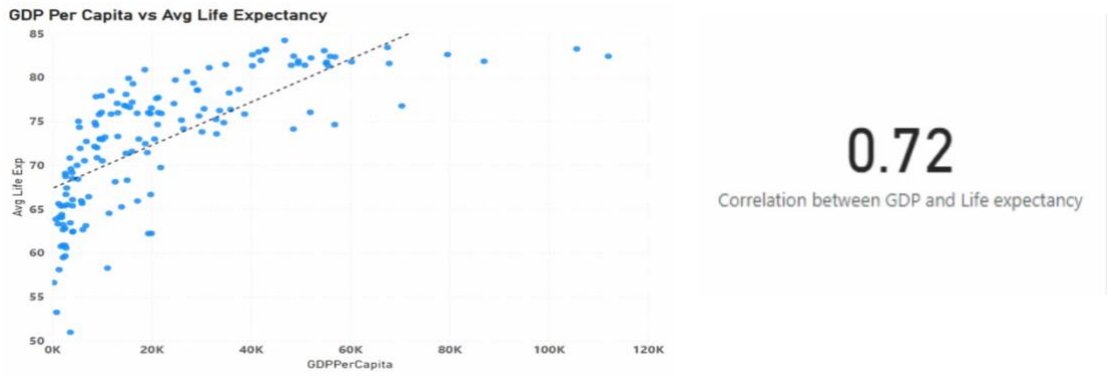
The above graph shows the average life expectancy for each GDP group. Countries with a GDP of less than $10,000 have the lowest average life expectancy of 66.65 years, and countries with a GDP of $40,000 or more have the highest average life expectancy of 81.37 years. Through this, the higher the GDP, the higher the average life expectancy of the country.

3. Average of each gender's height by GDP Group

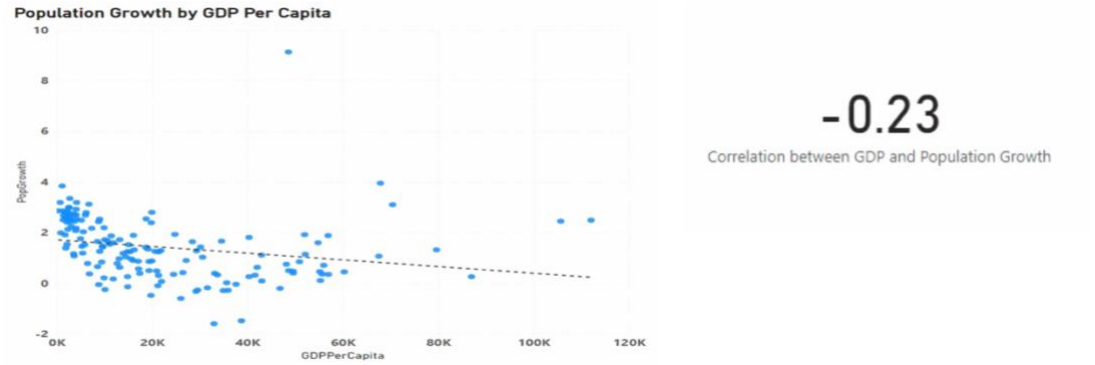


The above graph shows the average height of men and women in each GDP group. The average height of men in countries with a GDP of less than $10,000 is 66.74, and the average height of women is 62.34, which is on average shorter than other groups, and the average height of men in countries with a GDP of $40,000 or more is 69.81 and women is 64.49, which is on average taller than other groups. In addition, as the GDP increases, the average height increases, so it can be said that there is a relationship between GDP and the average height.

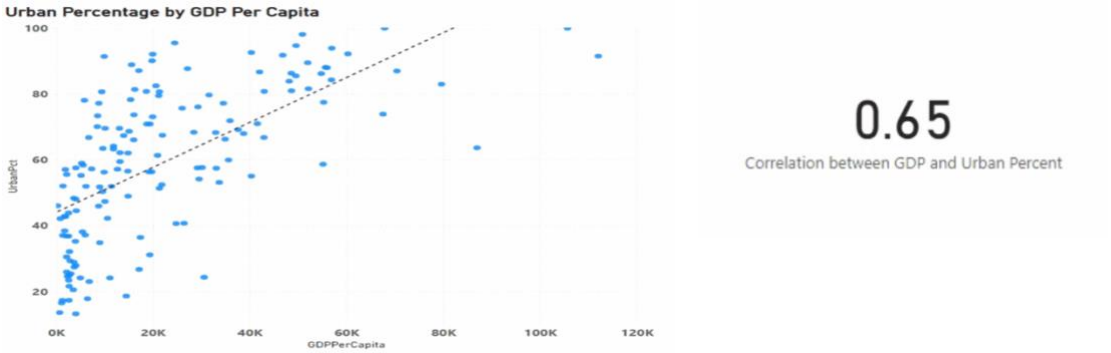
4. Correlation between GDP Per Capita and each of the variables



As a result of confirming the correlation between GDP per capita and average life expectancy through the above correlation graph, it can be seen that the absolute value has a strong positive correlation of 0.72.



The correlation between GDP per capita and population growth rate is confirmed through the above correlation graph, and the absolute value is -0.23, indicating a negative correlation. This means that there is little link between per capita GDP and population growth.

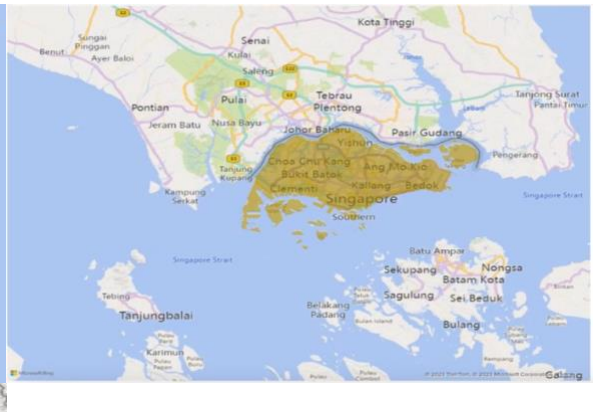


Finally, the correlation between GDP per capita and urban percentage is confirmed through the above correlation graph, and the absolute value is 0.65, which is a strong positive correlation.

As a result, urban percentages and life expectancy are highly related to GDP per capita, so I would like to use them to provide recommended destinations.

5.1. Recommendation for top 3 vacation destinations (High GDP)

I would like to recommend three countries with high GDP, which I recommend when the requester wants to stay in luxury accommodation with convenient transportation, is good for sightseeing, and prefers to visit popular tourist attractions and symbolic landmarks. So, I found countries with a GDP of over $40,000 and an urban percentage of over 90. Through this, a total of 9 countries were selected.



According to a survey of the average life expectancy of each country to select three of these nine countries, Japan has a high average life expectancy of 84.22, Singapore 83.25, and Israel 82.58.

Therefore, I recommend **Japan**, **Singapore**, and **Israel** as high-end vacation destinations for applicants to go on overseas travel adventures after graduation.

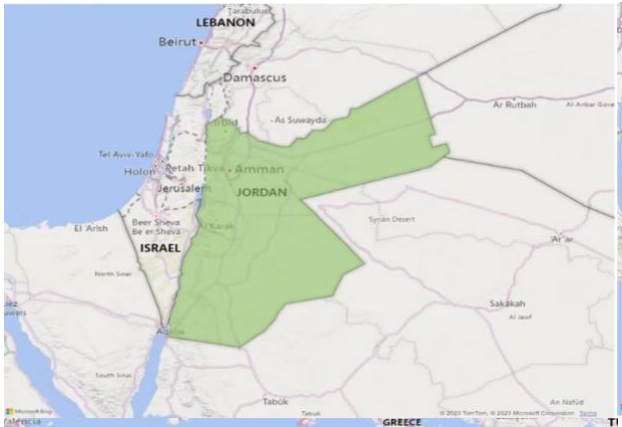
5.2. Recommendation for top 3 vacation destinations (Low GDP)

GDP Group, Urban Group

- Between $10,000 and $24,999
- Between $25,000 and $39,999
- Less than $10,000
  - Less than 50
  - Reached 50
  - Reached 70
  - Reached 90
- More or equal than $40,000

Country	Average of Avg Life Exp
Jordan	77.90
Cuba	77.81
Libya	75.78
El Salvador	74.86
Bolivia	72.14
Djibouti	65.94
Total	74.07

I would like to recommend three countries with low GDP, which I recommended when requestors want to have the opportunity to explore a wide range of cultures and interact with locals and prefer a low-cost trip while staying in affordable accommodation. So, I found countries with a GDP of less than $10,000 and an urban percentage of over 70 or 90. Through this, a total of 6 countries were selected.



According to a survey of the average life expectancy of each country to select 3 of these 6 countries, Jordan has a high average life expectancy of 77.90, Cuba 77.81, and Libya 75.78.

Therefore, I recommend **Jordan**, **Cuba**, and **Libya** as vacation destinations where requestors can go at a low cost.

Conclusion

As a result of checking the correlation of each variable by GDP, the higher the GDP, the higher the median age group and the average height of the country, and the higher the GDP, the higher the average life expectancy and urban percentage of people. On the other hand, the population growth rate has little to do with GDP. Also, based on the GDP, I suggested two good vacation destination options for the requestor to go on an overseas travel adventure after graduation because it depends on the requestor's individual preferences and priorities. There are people who want to tour high-quality accommodation and tourist attractions even if they spend a lot of money when traveling, and on the contrary, there are people who want to stay at cheap accommodation facilities and tour natural attractions when traveling.

In conclusion, if the requestor wants to go on a **high-end trip**, I recommend going to

1. Japan
2. Singapore
3. Israel

Conversely, if the requestor wants to go on a **trip on a low budget**, I recommend going to

1.Jordan
2.Cuba
3.Libya

Appendix

```
USE [Featherman_Analytics]; -- load the database

DROP TABLE [MF31namjun.lee].[dbo].[TSQL5] -- if table already exists then remove the table from my database

-- create the big array table
DECLARE @TravelAdventure TABLE([Country] NVARCHAR(50), [GDPPerCapita] DECIMAL(8,2)
, [FemaleHeight] DECIMAL(8,2), [MaleHeight] DECIMAL(8,2), [HeightDifference] DECIMAL(8,2),
[CountryHeightRank] INT
, [FLifeExp] DECIMAL(8,2), [MLifeExp] DECIMAL(8,2)
, [MedianAge] DECIMAL(8,2), [PopGrowth] DECIMAL(8,2), [UrbanPct] DECIMAL(8,2)
-- add 5 new columns deriving new insights
, [GDP Group] NVARCHAR(50), [Avg Life Exp] DECIMAL(8,2), [Median Age Group] NVARCHAR(50)
, [PopGrowth Group] NVARCHAR(50), [Urban Group] NVARCHAR(50))

-- insert columns from HWGDPPerCapita tables into the array
INSERT INTO @TravelAdventure ([Country], [GDPPerCapita])
SELECT DISTINCT([Country]), [GDPPerCapita]
FROM [featherman].[HWGDPPerCapita] AS gdp

-- merge in columns of metrics from HWHeights table
UPDATE @TravelAdventure SET [FemaleHeight]
= (SELECT [Female_Height_in_Inches] FROM [featherman].[HWHeights] as hei WHERE hei.Country_Name
= Country)
UPDATE @TravelAdventure SET [MaleHeight]
= (SELECT [Male_Height_in_Inches] FROM [featherman].[HWHeights] as hei WHERE hei.Country_Name =
Country)
UPDATE @TravelAdventure SET [HeightDifference]
= (SELECT [Difference] FROM [featherman].[HWHeights] as hei WHERE hei.Country_Name = Country)
UPDATE @TravelAdventure SET [CountryHeightRank]
= (SELECT [Rank] FROM [featherman].[HWHeights] as hei WHERE hei.Country_Name = Country)

-- merge in columns of metrics from HWlifeExpectancy table
UPDATE @TravelAdventure SET [FLifeExp]
= (SELECT [FemaleLifeExp] FROM [featherman].[HWlifeExpectancy] as life WHERE life.Country_Name =
Country)
UPDATE @TravelAdventure SET [MLifeExp]
= (SELECT [MaleLifeExp] FROM [featherman].[HWlifeExpectancy] as life WHERE life.Country_Name =
Country)

-- merge in column of metrics from HWMedianAge table
UPDATE @TravelAdventure SET [MedianAge]
= (SELECT TOP 1 [Median_age] FROM [featherman].[HWMedianAge] as age WHERE age.Country_Name =
Country)

-- merge in column of metrics from HWPopulationGrowth table
UPDATE @TravelAdventure SET [PopGrowth]
= (SELECT TOP 1 [Population_growth] FROM [featherman].[HWPopulationGrowth] as pop WHERE
pop.Country_Name = Country)

-- merge in column of metrics from HWurban_percent table
UPDATE @TravelAdventure SET [UrbanPct]
= (SELECT [UrbanPercent] FROM [featherman].[HWurban_percent] as urban WHERE urban.Country_Name =
Country)

-- update the gdp group using CASE statement
UPDATE @TravelAdventure SET [GDP Group]
= (CASE
```

```
        WHEN [GDPPerCapita] < 10000.00 THEN 'Less than $10,000'
        WHEN [GDPPerCapita] BETWEEN 10000.00 AND 24999.00 THEN 'Between $10,000 and $24,999'
        WHEN [GDPPerCapita] BETWEEN 25000.00 AND 39999.00 THEN 'Between $25,000 and $39,999'
        WHEN [GDPPerCapita] >= 40000.00 THEN 'More or equal than $40,000'
        END)

-- update the female and male average life expectancy
UPDATE @TravelAdventure SET [Avg Life Exp]
= ([FLifeExp] + [MLifeExp]) / 2

-- update median age group using CASE statement
UPDATE @TravelAdventure SET [Median Age Group]
= (CASE
        WHEN [MedianAge] BETWEEN 10.00 AND 19.99 THEN '10 years'
        WHEN [MedianAge] BETWEEN 20.00 AND 29.99 THEN '20 years'
        WHEN [MedianAge] BETWEEN 30.00 AND 39.99 THEN '30 years'
        WHEN [MedianAge] >= 40 THEN '40 years'
        END)

-- update population growth group using CASE statement
UPDATE @TravelAdventure SET [PopGrowth Group]
= (CASE
        WHEN [PopGrowth] < 0 THEN 'Decrease'
        WHEN [PopGrowth] BETWEEN 0 AND 0.99 THEN 'Less 1% Increase'
        WHEN [PopGrowth] BETWEEN 1 AND 1.99 THEN 'Less 2% Increase'
        WHEN [PopGrowth] BETWEEN 2 AND 2.99 THEN 'Less 3% Increase'
        WHEN [PopGrowth] >= 3 THEN 'More 3% Increase'
        END)

-- update urban percent using CASE statement
UPDATE @TravelAdventure SET [Urban Group]
= (CASE
        WHEN [UrbanPct] < 50.00 THEN 'Less than 50'
        WHEN [UrbanPct] BETWEEN 50.00 AND 69.99 THEN 'Reached 50'
        WHEN [UrbanPct] BETWEEN 70.00 AND 89.99 THEN 'Reached 70'
        WHEN [UrbanPct] >= 90 THEN 'Reached 90'
        END)

-- save the array to my database
-- there is a null value because other tables are joined based on the country of the GDP table. Therefore, the
columns where the null exists are not displayed.
SELECT * INTO [MF31namjun.lee].[dbo].[TSQL5]
FROM @TravelAdventure
WHERE FemaleHeight IS NOT NULL
AND FLifeExp IS NOT NULL
AND MedianAge IS NOT NULL
AND PopGrowth IS NOT NULL
AND UrbanPct IS NOT NULL
```