

PROYEK UJIAN AKHIR SEMESTER

MMS1414 - EKSPLORASI DAN VISUALISASI DATA

STATISTICAL ANALYSIS ON FACTORS INFLUENCING LIFE EXPECTANCY



Yogyakarta, 19 Juni 2023

Nama : Aulia Mirfah Setyo Ayu Damayanti
NIM : 22/493337/PA/21178
Dosen Pengampu : 1. Drs. Danardono, MPH., Ph.D.
2. Mohamad Fahruli Wahyujati, S.Si., M.Si.

PROGRAM STUDI STATISTIKA

DEPARTEMEN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

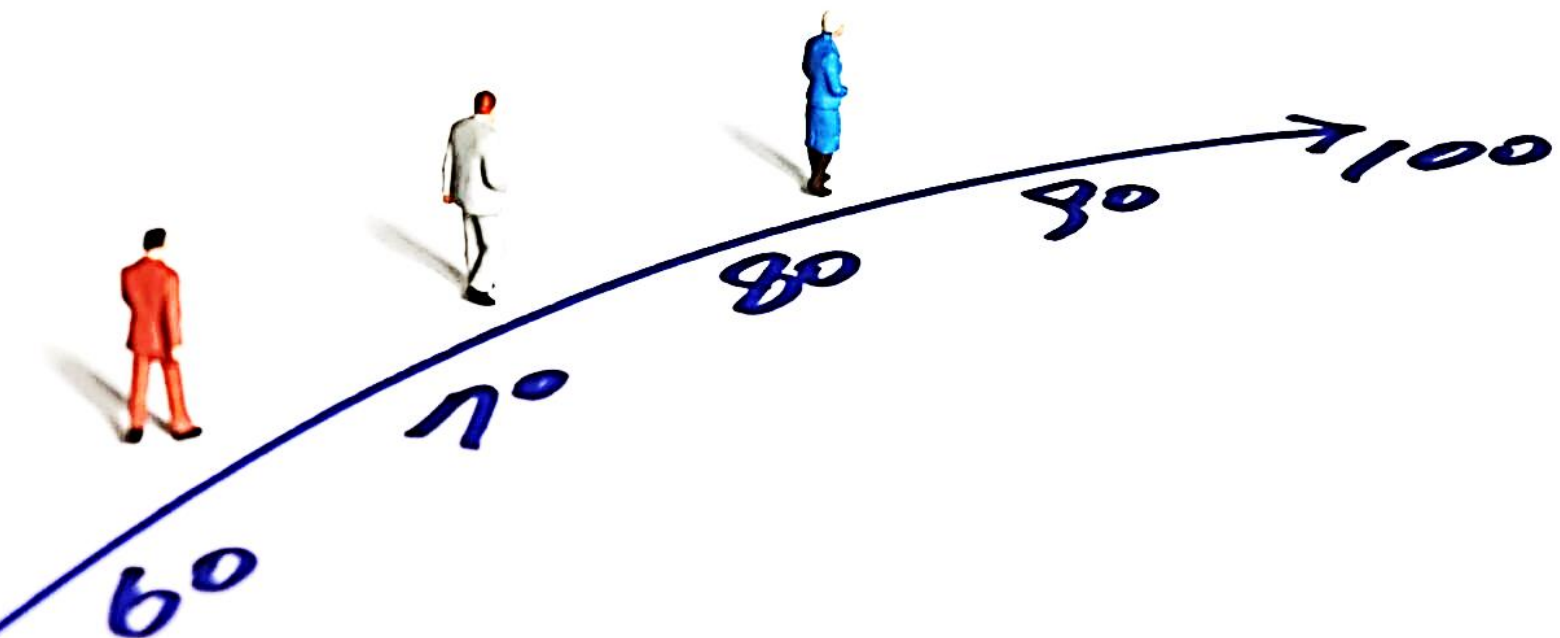
UNIVERSITAS GADJAH MADA

2023

Statistical Analysis on Factors Influencing Life Expectancy

Analisis statistik terhadap dataset tentang Harapan Hidup (Life Expectancy) oleh *World Health Organization* (WHO). Dataset ini mencakup informasi tentang berbagai faktor yang mempengaruhi harapan hidup di 193 negara dari tahun 2000 hingga 2015.

Juni 2023



Contents



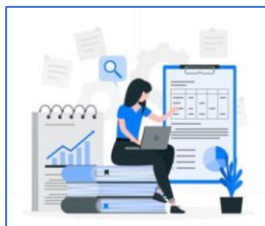
- 4** Background
Project/Introduction



- 5** Data Description



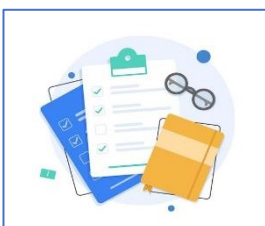
- 7** Research
Question/Objective



- 8** Discussion/Analysis



- 20** Conclusion

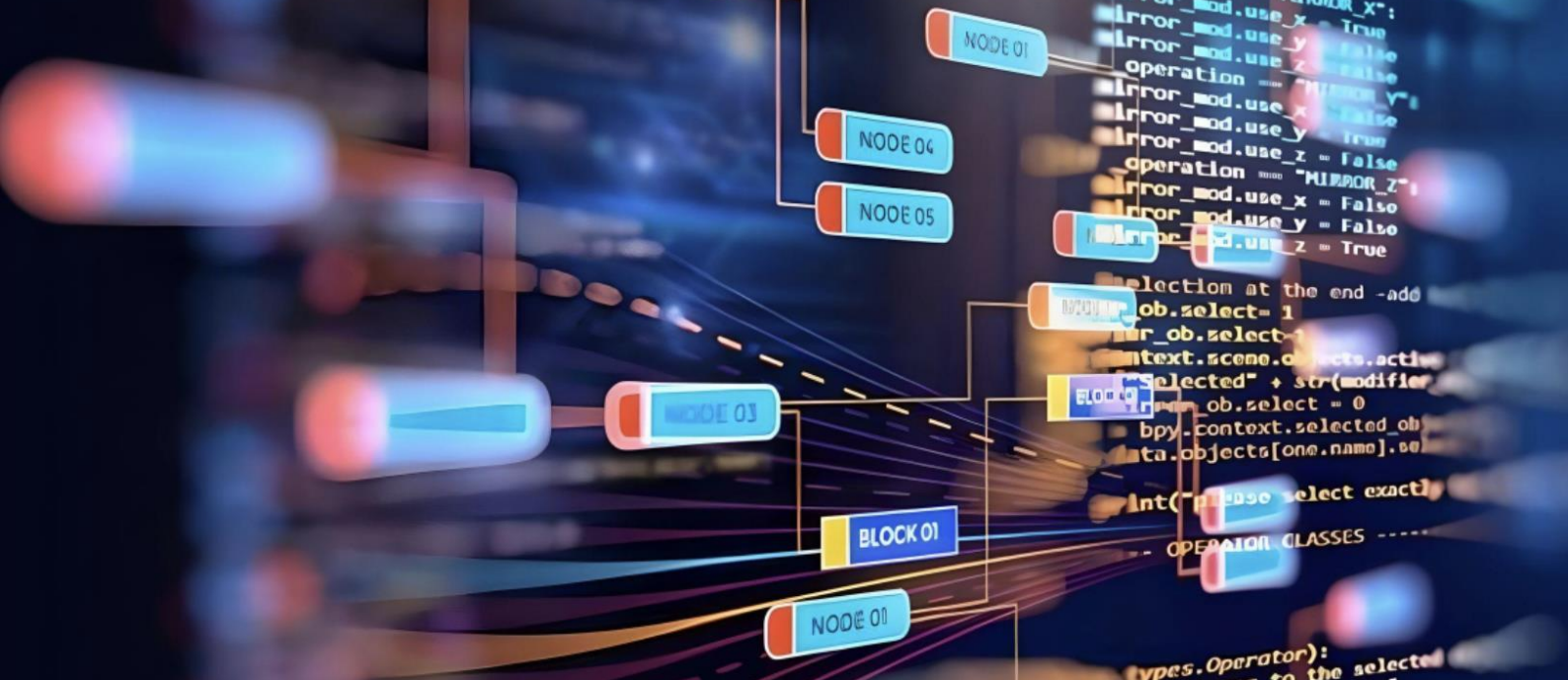


- 22** Daftar Pustaka
23 Lampiran

Background

Project/Introduction

Harapan hidup adalah salah satu indikator penting dalam mengevaluasi kesehatan dan kesejahteraan populasi suatu negara. Berbagai faktor telah diidentifikasi sebagai faktor yang mempengaruhi harapan hidup, termasuk faktor demografis, sosial-ekonomi, kesehatan, dan lingkungan. Namun, terdapat kebutuhan untuk melakukan analisis lebih lanjut terhadap faktor-faktor ini dan memahami pengaruhnya secara komprehensif. Dari semua kategori faktor terkait kesehatan, hanya faktor-faktor kritis yang dipilih yang lebih representatif. Telah diamati bahwa dalam 15 tahun terakhir, telah terjadi perkembangan besar dalam sektor kesehatan yang menghasilkan peningkatan tingkat kematian manusia terutama di negara-negara berkembang dibandingkan 30 tahun sebelumnya. Oleh karena itu, proyek ini bertujuan untuk melakukan eksplorasi dan analisis lebih lanjut terhadap dataset tentang Harapan Hidup (*Life Expectancy*) yang disediakan oleh Organisasi Kesehatan Dunia (WHO) dan Perserikatan Bangsa-Bangsa. Dataset ini mencakup informasi tentang berbagai faktor yang mempengaruhi harapan hidup di 193 negara dari tahun 2000 hingga 2015. Dengan pemahaman yang lebih baik tentang faktor-faktor yang berkontribusi pada harapan hidup, diharapkan analisis ini dapat bermanfaat bagi banyak pihak, misalnya pemerintah untuk mengembangkan kebijakan yang lebih terarah, merencanakan sumber daya kesehatan secara efisien, dan meningkatkan kualitas hidup penduduk. Selain itu, analisis ini juga dapat meningkatkan kesadaran masyarakat tentang pentingnya faktor-faktor kesehatan dan memotivasi langkah-langkah untuk meningkatkan harapan hidup mereka sendiri. Dengan memahami informasi terkait faktor-faktor yang berpengaruh serta mengambil langkah-langkah atau keputusan yang sesuai, pemerintah dan masyarakat dapat bekerja sama untuk meningkatkan kualitas hidup, mengurangi disparitas, dan mencapai peningkatan harapan hidup yang signifikan. Kemudian, proyek ini berpotensi memberikan dampak positif yang signifikan bagi masyarakat.



Data

Description

Dataset "Life Expectancy (WHO) Statistical Analysis on factors influencing Life Expectancy" adalah sebuah kumpulan data yang mengumpulkan informasi tentang faktor-faktor yang mempengaruhi harapan hidup di berbagai negara. Dataset ini mencakup data tentang harapan hidup, faktor-faktor kesehatan, dan faktor-faktor ekonomi dan sosial dari 193 negara selama periode tahun 2000 hingga 2015. Dataset ini diambil dari repositori data *Global Health Observatory* (GHO) yang dikelola oleh *World Health Organization* (WHO), serta data ekonomi yang berasal dari situs Perserikatan Bangsa-Bangsa, dan dengan bantuan Deeksha Russell dan Duan Wang. Dataset ini terdiri dari 22 kolom dan 2938 baris, yang mencakup informasi seperti tahun, negara, harapan hidup, faktor-faktor kesehatan seperti angka kematian bayi, angka kematian orang dewasa, imunisasi, serta faktor-faktor ekonomi dan sosial seperti pendapatan per kapita, indeks pembangunan manusia (IPM), dan lainnya. Terdapat beberapa data yang hilang adalah untuk populasi, Hepatitis B, dan PDB. Data yang hilang berasal dari negara-negara yang kurang dikenal seperti Vanuatu, Tonga, Togo, Cabo Verde, dll. Sulit menemukan semua data untuk negara-negara ini, oleh karena itu, diputuskan untuk mengesampingkan negara-negara ini dari dataset model akhir. Semua variabel prediksi kemudian dibagi menjadi beberapa kategori luas: faktor-faktor terkait imunisasi, faktor-faktor kematian, faktor-faktor ekonomi, dan faktor-faktor sosial. Setiap baris merepresentasikan satu informasi terhadap satu negara dalam masing-masing tahun antara 2000-2015 mengenai angka harapan hidupnya serta faktor yang memengaruhinya. Dataset ini dirancang untuk menganalisis hubungan antara faktor-faktor tersebut dengan harapan hidup populasi di berbagai negara. Berikut adalah keterangan dari 22 variabel yang terkandung dalam dataset tersebut (dalam bahasa Inggris).

- **Life expectancy** – life expectancy in age. This is the target variable that you're trying to predict.
- **Country:** country (193 unique values)
- **Year:** year (2000 – 2015)
- **Status:** developed or developing country status.
- **Adult Mortality:** adult mortality rates of both sexes (probability of dying between 15 and 60 years per 1000 population).
- **Infant Deaths:** number of infant deaths per 1000 population.
- **Alcohol:** Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol).
- **Percentage expenditure:** expenditure on health as a percentage of Gross Domestic Product per capita (%).
- **Hepatitis B:** Hepatitis B (HepB) immunization coverage among 1 year-olds (%).
- **Measles:** measles – number of reported cases per 1000 population.
- **BMI:** Average Body Mass Index of entire population.
- **Under-five deaths:** number of under-five deaths per 1000 population.
- **Polio:** Polio (Pol3) immunization coverage among 1-year-olds (%).
- **Total expenditure:** general government expenditure on health as a percentage of total government expenditure (%).
- **Diphtheria:** Diphtheria tetanus toxoid and pertussis (DTP3) immunization coverage among 1-year-olds (%).
- **HIV/AIDS:** Deaths per 1000 live births HIV/AIDS (0-4 years).
- **GDP:** Gross Domestic Product per capita (in USD).
- **Population:** population of the country.
- **Thinness 1-19 years:** prevalence of thinness among children and adolescents for Age 10-19 (%).
- **Thinness 5-9 years:** prevalence of thinness among children for Age 5 to 9 (%).
- **Income composition of resources:** Human Development Index in terms of income composition of resources (index ranging from 0 to 1).
- **Schooling:** number of years of schooling (years).

Research Question/Objective



1. Bagaimana perbedaan rata-rata angka harapan hidup di antara 193 negara tersebut?
2. Bagaimana perbandingan tren angka harapan hidup antara negara dengan angka harapan hidup tertinggi dan terendah dalam rentang waktu yang diteliti?
3. Berdasarkan variabel independent yang ada, akan dikelompokkan ke dalam beberapa jenis faktor yaitu
 - Faktor Gaya Hidup: Alcohol, BMI, HIV/AIDS. Thinness 10-19 years, Thinnerss 5-9 years, Schooling.
 - Faktor Lingkungan dan Imunisasi: Hepatitis B, Measles, Polio, Diphteria.
 - Faktor Ekonomi dan Populasi: Income composition of resources, GDP, Population, Total Expenditure, Percentage expenditure.
 - Faktor Mortalitas dan Kematian Anak: Adult Mortality, Infant Deaths, Under-five deaths.

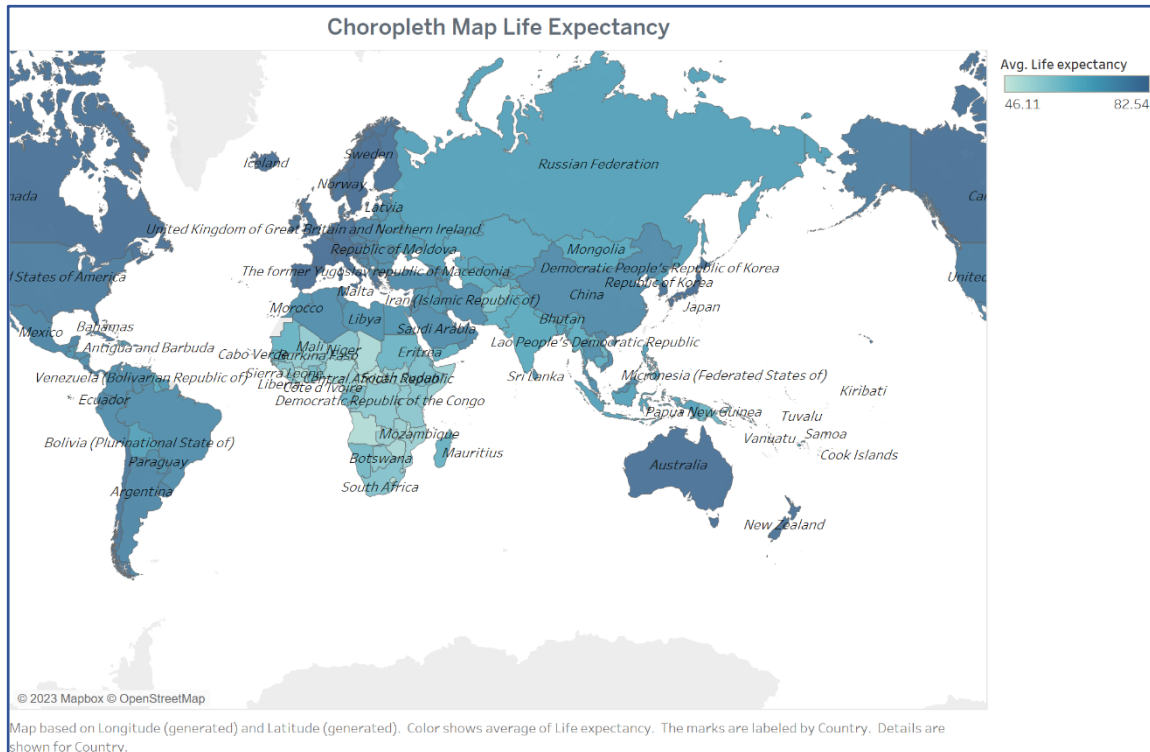
Apakah kelompok faktor lingkungan dan imunisasi berkontribusi/memberikan pengaruh terhadap angka harapan hidup di setiap negara?

4. Bagaimana perbandingan rata-rata cakupan imunisasi Hepatitis B, Polio, dan Diphteria pada populasi di semua negara?
5. Dari kelompok faktor ekonomi dan populasi, variabel apa yang memiliki hubungan paling signifikan terhadap angka harapan hidup?
6. Bagaimana perbandingan rata-rata dari variabel tersebut (dari pertanyaan nomor 5) antara negara dengan angka harapan hidup tertinggi dan terendah?
7. Berdasarkan kelompok faktor gaya hidup, apakah pendidikan mempengaruhi rata-rata angka harapan hidup di suatu negara? Bagaimana korelasinya?
8. Apa 10 negara dengan rata-rata jumlah tahun pendidikan (schooling) paling tinggi?
9. Bagaimana proporsi antara negara berkembang dan maju dalam dataset yang ada?
10. Apakah terdapat perbedaan yang signifikan terkait angka harapan hidup antara negara-negara berkembang dengan negara-negara maju?

Discussion/Analysis

1. Choropleth Map

Visualisasi yang dilakukan bertujuan untuk melihat perbedaan rata-rata angka harapan hidup tiap negara dari 193 negara yang terdapat dalam dataset “Life Expectancy Data”.



Insight:

Dalam visualisasi choropleth map ini, warna yang semakin tua menunjukkan tingkat angka harapan hidup yang lebih tinggi di negara-negara di seluruh dunia. Melalui peta ini, dapat diamati pola sebaran angka harapan hidup di berbagai benua dan negara. Berdasarkan output yang diperoleh, dapat dilihat bahwa benua Amerika, Eropa, dan Australia memiliki warna yang tua, menunjukkan angka harapan hidup yang tinggi secara umum di negara-negara pada benua-benua tersebut. Hal ini bisa jadi menunjukkan bahwa negara-negara di benua-benua tersebut cenderung memiliki akses yang baik terhadap pelayanan kesehatan, pendidikan, dan faktor-faktor lain yang berkontribusi terhadap peningkatan harapan hidup.

Sebaliknya, benua Afrika secara umum memiliki warna yang lebih cerah, mengindikasikan angka harapan hidup yang rendah di negara-negara di benua tersebut. Hal ini mungkin mencerminkan tantangan yang dihadapi oleh negara-negara Afrika dalam hal akses terhadap layanan kesehatan yang memadai, pendidikan, serta faktor-faktor lain yang dapat mempengaruhi harapan hidup penduduknya. Selain itu, dapat dilihat pula bahwa rata-rata angka harapan hidup terendah ditunjukkan oleh negara Sierra Leone dengan rata-rata angka harapan hidup sebesar 46,11 tahun.

Untuk benua Asia, visualisasi menunjukkan warna yang sedang antara tua dan cerah, menandakan variasi dalam angka harapan hidup di negara-negara Asia yang tidak

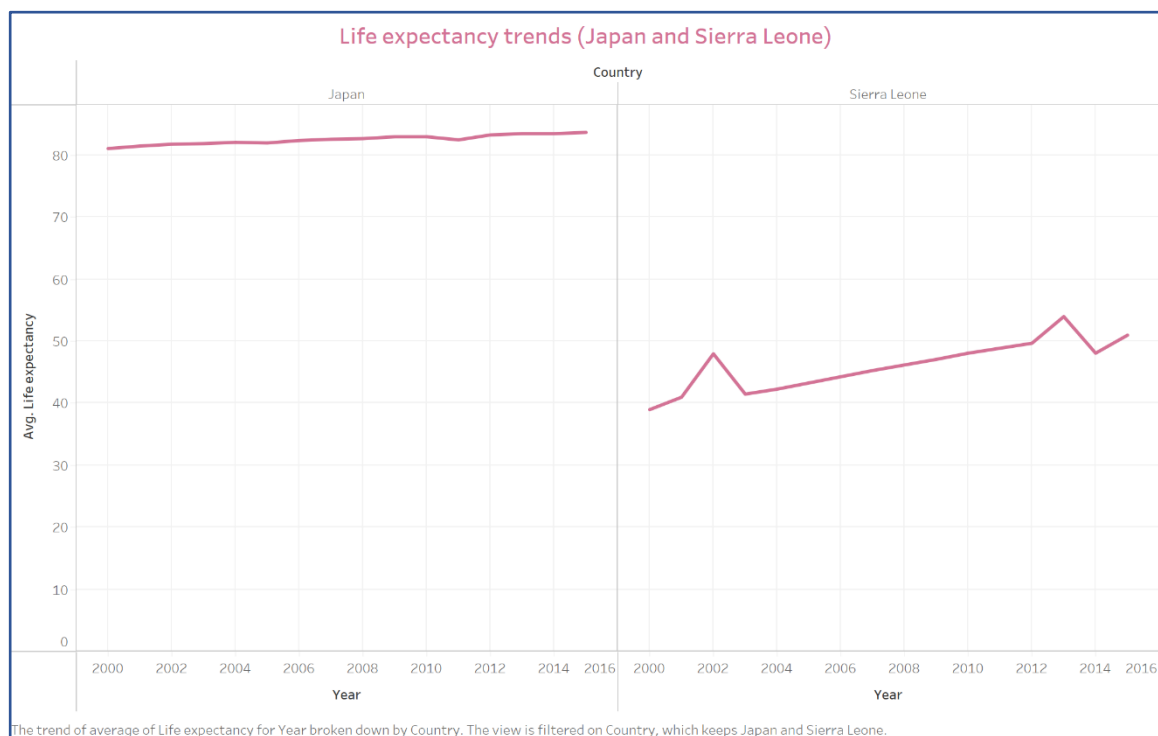
Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

rendah maupun tinggi. Meskipun sebagian besar negara Asia cenderung memiliki angka harapan hidup yang tergolong tinggi, bahkan Jepang menjadi negara dengan rata-rata angka harapan hidup paling tinggi yaitu sebesar 82,54 tahun. Ada juga negara dengan angka harapan hidup yang lebih rendah di benua Asia, seperti Afghanistan. Perbedaan ini dapat disebabkan oleh faktor-faktor seperti sistem kesehatan, ekonomi, tingkat pendidikan, dan gaya hidup yang akan dianalisis lebih lanjut. Dengan menggunakan visualisasi ini, dapat dilihat perbandingan angka harapan hidup di berbagai negara di seluruh dunia.

2. Line Chart

Visualisasi yang dilakukan bertujuan untuk melihat perbandingan tren angka harapan hidup antara Jepang sebagai negara dengan rata-rata angka harapan hidup tertinggi dan Sierra Leone sebagai negara dengan rata-rata angka harapan hidup terendah dalam rentang waktu yang diteliti yaitu 2000-2015. Dari linechart dapat divisualisasikan perbedaan yang dramatis dalam tren angka harapan hidup antara Jepang dan Sierra Leone.



Insight:

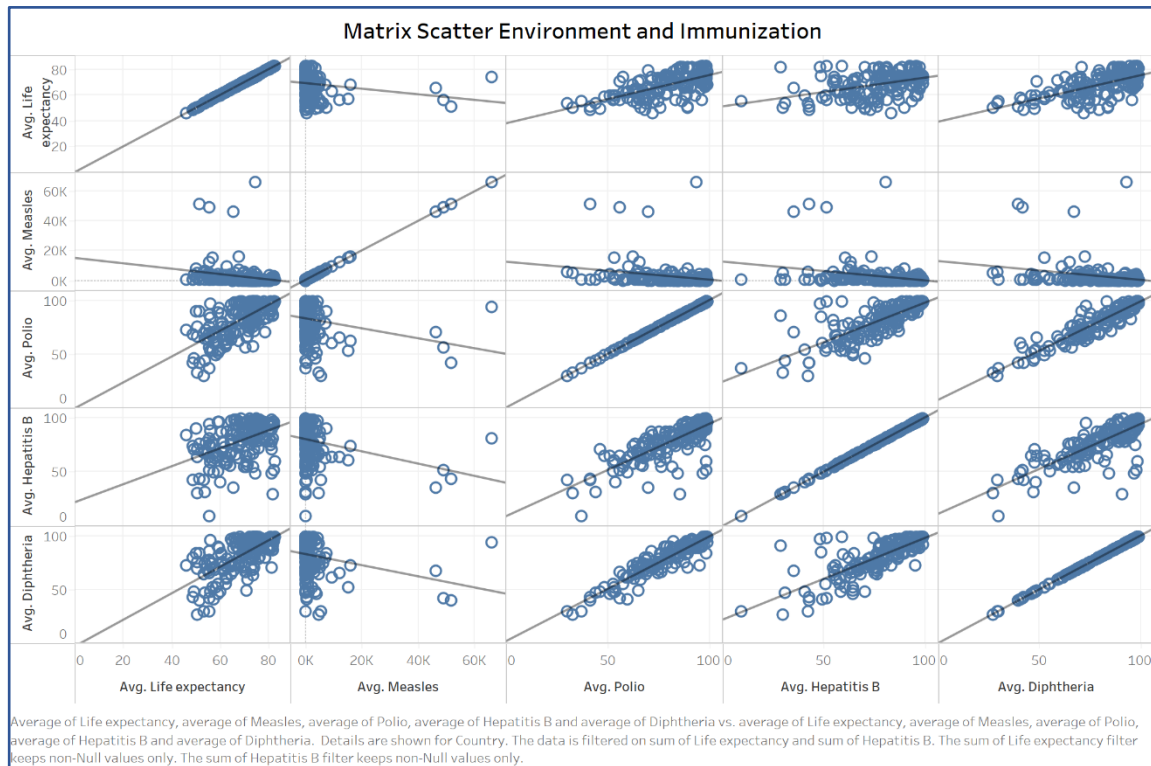
Berdasarkan line chart ini, dapat dilihat perbedaan yang signifikan dalam tren angka harapan hidup antara Jepang dan Sierra Leone. Untuk Jepang, terlihat bahwa angka harapan hidup cenderung meningkat secara konsisten dari tahun ke tahun, meskipun pada tahun 2005 dan 2011 terdapat penurunan namun sangat kecil yaitu masing-masing sebesar 0,1 dan 0,5 tahun. Di sisi lain, Sierra Leone menunjukkan tren yang cenderung lebih tidak stabil dan fluktuatif dalam angka harapan hidup. Meskipun umumnya terjadi peningkatan dalam beberapa tahun tertentu, tetapi masih ada penurunan yang relatif signifikan pada tahun 2003 dan 2014 yaitu masing-masing sebesar 6,5 dan 5,9 tahun, selain itu angka harapan hidup Sierra Leone jauh lebih rendah dibandingkan Jepang.

Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

3. Matrix Scatter (Faktor Lingkungan dan Imunisasi)

Visualisasi yang dilakukan bertujuan untuk mengetahui apakah terdapat hubungan antara faktor lingkungan dan imunisasi terhadap angka harapan hidup di setiap negara.



Insight:

Didapatkan output berupa *scatter plot* di atas yang menunjukkan hubungan antara variabel Life Expectancy dengan variabel Measles, Polio, Hepatitis B, dan Diphtheria.

Variabel Dependen	Variabel Independen	R-Square	Interpretasi
Life Expectancy	Measles	0,039	Scatter plot antara Life Expectancy dengan Measles didapatkan bahwa keduanya memiliki linearitas negatif dan menghasilkan R^2 sebesar 0,039 yang artinya 3,9% variansi Life Expectancy dapat dijelaskan oleh variabel Measles.
	Polio	0,451	Scatter plot antara Life Expectancy dengan Polio didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,451 yang artinya 45,1% variansi Life Expectancy dapat dijelaskan oleh variabel Polio.
	Hepatitis B	0,184	Scatter plot antara Life Expectancy dengan Hepatitis B didapatkan bahwa keduanya memiliki

Nama : Aulia Mirfah Setyo Ayu Damayanti

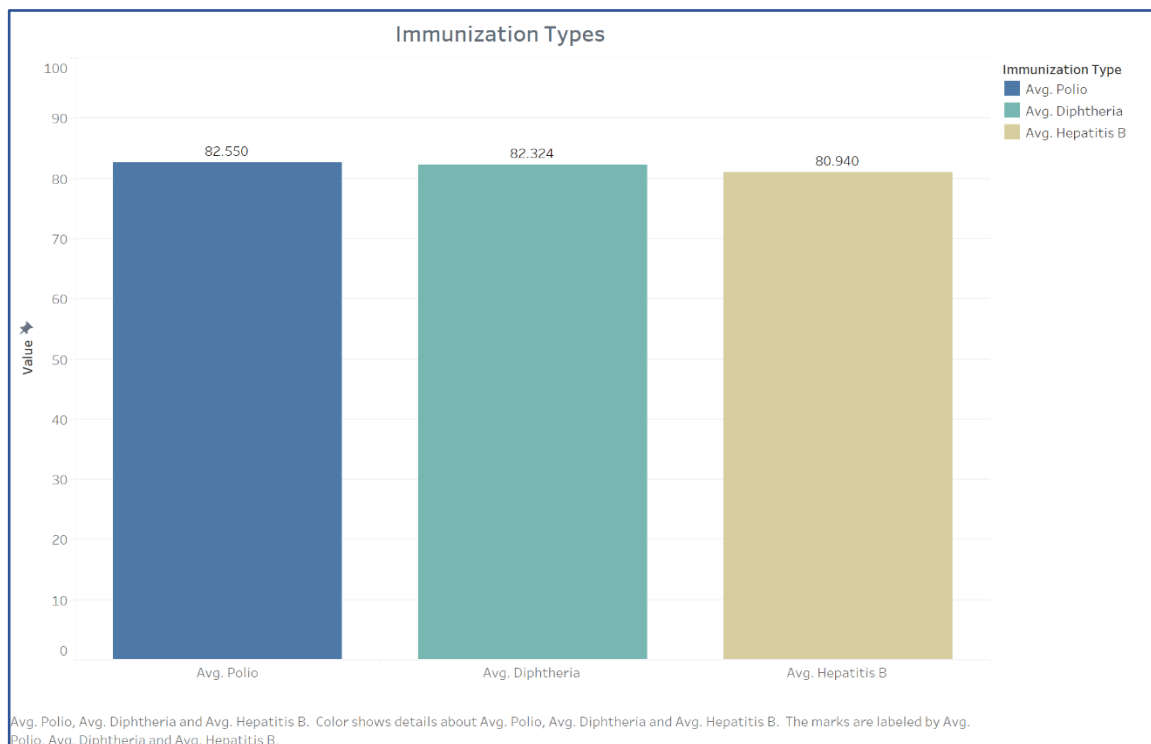
NIM : 22/493337/PA/21178

			linearitas positif dan menghasilkan R^2 sebesar 0,184 yang artinya 18,4% variansi Life Expectancy dapat dijelaskan oleh variabel Hepatitis B.
	Diphtheria	0,444	Scatter plot antara Life Expectancy dengan Diphtheria didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,444 yang artinya 44,4% variansi Life Expectancy dapat dijelaskan oleh variabel Diphtheria.

Kesimpulan berdasarkan scatter plot di atas, diketahui bahwa variabel Life Expectancy dengan Measles memiliki linearitas negatif yang artinya semakin besar nilai variabel Measles, yaitu semakin banyak orang terkena campak, maka semakin kecil nilai Life expectancy. Selain itu, hubungan variabel Life Expectancy dengan variabel Polio, Hepatitis B, dan Diphtheria memiliki linearitas positif yang artinya semakin besar nilai ketiga variabel tersebut, semakin besar nilai Life Expectancy. Terdapat outlier pada scatter plot tersebut bisa jadi disebabkan oleh adanya perbedaan kebijakan dan praktik imunisasi yang signifikan di setiap negara. Dengan ini dapat dikatakan bahwa kelompok faktor lingkungan dan imunisasi yang meliputi Measles, Polio, Hepatitis B, dan Diphtheria berkontribusi/memberikan pengaruh terhadap angka harapan hidup di setiap negara.

4. Bar Chart Immunization Type

Visualisasi yang dilakukan bertujuan untuk melihat perbandingan rata-rata cakupan imunisasi Polio, Diphtheria, dan Hepatitis B pada populasi di semua negara.



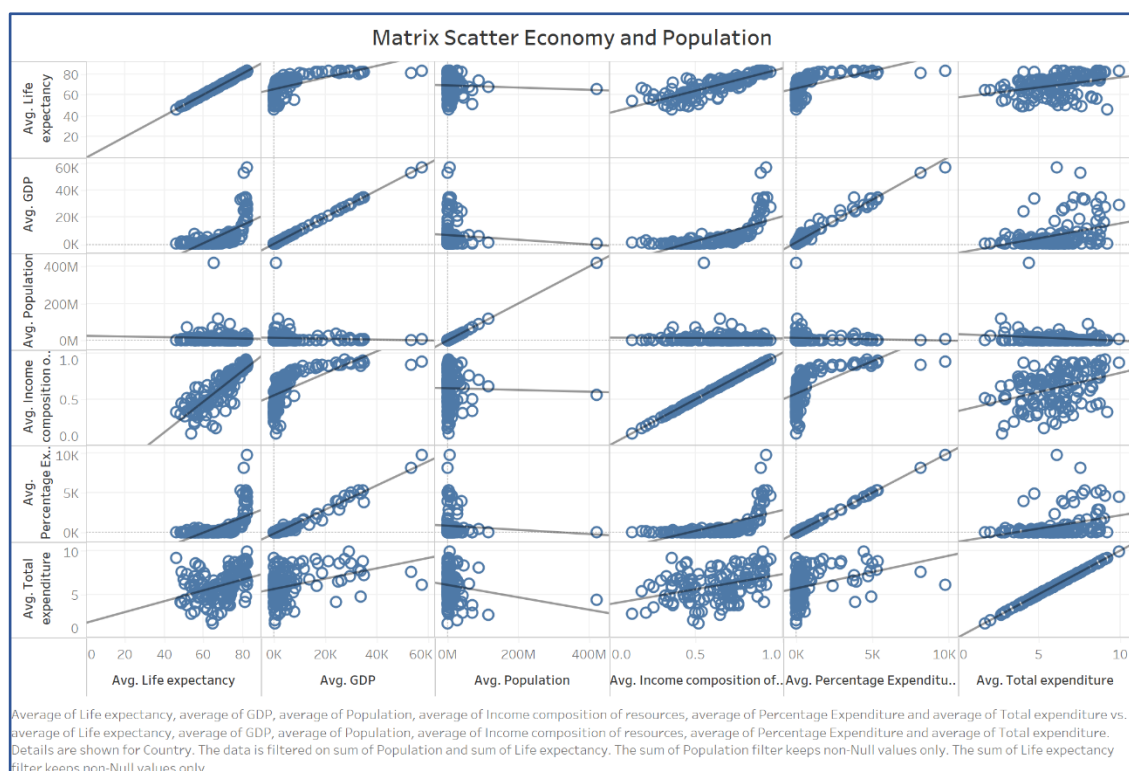
Insight:

Bar chart menunjukkan perbandingan rata-rata cakupan imunisasi Polio, Diphteria, dan Hepatitis B dari 2938 dataset yang didapat dari repositori data *Global Health Observatory* (GHO) yang dikelola oleh *World Health Organization* (WHO), serta data ekonomi yang berasal dari situs Perserikatan Bangsa-Bangsa, dapat diketahui bahwa rata-rata cakupan imunisasi seluruh negara berdasarkan jenis imunisasinya terbagi menjadi 3 bagian yaitu sebagai berikut:

- Imunisasi Polio, menunjukkan rata-rata cakupan balita berusia 1 tahun yang telah mendapatkan imunisasi polio tipe 3 sebesar 82,550 (dalam persen).
- Imunisasi Diphteria, menunjukkan rata-rata cakupan balita berusia 1 tahun yang telah mendapatkan imunisasi diphteria tetanus toxoid and pertussis sebesar 82,324 (dalam persen).
- Imunisasi Hepatitis B, menunjukkan rata-rata cakupan balita berusia 1 tahun yang telah mendapatkan imunisasi hepatitis B sebesar 80,940 (dalam persen).

Berdasarkan output bar chart yang dihasilkan, dapat dilihat bahwa imunisasi Polio memiliki cakupan paling tinggi di antara ketiga jenis imunisasi, dengan rata-rata cakupan sebesar 82,550 persen pada balita berusia 1 tahun. Sementara itu, imunisasi Hepatitis B memiliki cakupan paling rendah di antara ketiga jenis imunisasi, dengan rata-rata cakupan sebesar 80,940 persen pada balita berusia 1 tahun. Secara keseluruhan, dapat disimpulkan bahwa distribusi imunisasi di seluruh negara cenderung baik, karena rata-rata cakupan imunisasi untuk ketiga jenis tersebut sudah melampaui 80 persen. Namun, perlu terus diperhatikan dan ditingkatkan untuk menjaga kualitas dan kesinambungan program imunisasi guna melindungi kesehatan balita dari penyakit yang dapat dicegah melalui imunisasi.

5. Matrix Scatter (Faktor Ekonomi dan Populasi)



Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

Insight:

Didapatkan output berupa *scatter plot* di atas yang menunjukkan hubungan antara variabel Life Expectancy dengan faktor ekonomi dan populasi yang meliputi variabel GDP, Population, Income composition of resources, Percentage expenditure, dan Total expenditure.

Variabel Dependen	Variabel Independen	R-Square	Interpretasi
Life Expectancy	GDP	0,375	Scatter plot antara Life Expectancy dengan GDP didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,375 yang artinya 37,5% variansi Life Expectancy dapat dijelaskan oleh variabel GDP.
	Population	0,001	Scatter plot antara Life Expectancy dengan Population didapatkan bahwa keduanya memiliki linearitas negatif dan menghasilkan R^2 sebesar 0,001 yang artinya 0,1% variansi Life Expectancy dapat dijelaskan oleh variabel Population.
	Income composition of resources	0,701	Scatter plot antara Life Expectancy dengan Income composition of resources didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,701 yang artinya 70,1% variansi Life Expectancy dapat dijelaskan oleh variabel Income composition of resources.
	Percentage expenditure	0,319	Scatter plot antara Life Expectancy dengan Percentage expenditure didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,319 yang artinya 31,9% variansi Life Expectancy dapat dijelaskan oleh variabel Percentage expenditure.
	Total expenditure	0,119	Scatter plot antara Life Expectancy dengan Total expenditure didapatkan bahwa keduanya memiliki linearitas positif dan menghasilkan R^2 sebesar 0,119 yang artinya 11,9% variansi Life Expectancy dapat dijelaskan oleh variabel Total expenditure.

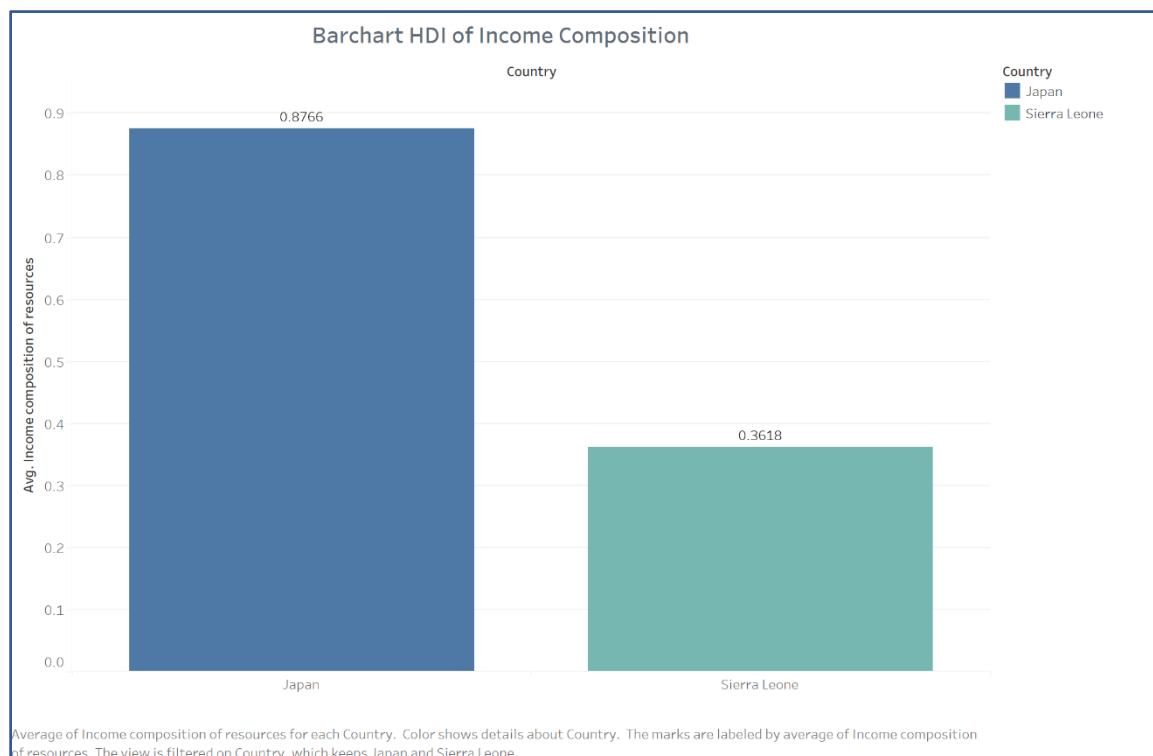
Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

Kesimpulan berdasarkan scatter plot di atas, diketahui bahwa variabel Life Expectancy dengan Population memiliki linearitas negatif yang artinya semakin besar nilai variabel Population, yaitu semakin bertambah populasi di suatu negara, maka semakin kecil nilai Life expectancy. Selain itu, hubungan variabel Life Expectancy dengan variabel GDP, Income composition of resources, Percentage expenditure, dan Total expenditure memiliki linearitas positif yang artinya semakin besar nilai variabel-variabel tersebut, semakin besar nilai Life Expectancy. Kemudian, ditemukan bahwa variabel Income composition of resources memiliki R-squared yang paling besar, yaitu sebesar 0,701. Ini menunjukkan bahwa sekitar 70,1% variasi dalam angka harapan hidup dapat dijelaskan oleh variabel Income composition of resources. Dengan kata lain, faktor komposisi pendapatan suatu negara memiliki pengaruh yang paling signifikan terhadap angka harapan hidup penduduknya dibandingkan dengan faktor lain dari kelompok faktor ekonomi dan populasi.

6. Bar Chart Income composition of resources Jepang dan Sierra Leone

Visualisasi yang dilakukan bertujuan untuk melihat perbandingan rata-rata Income composition of resources antara negara dengan angka harapan hidup tertinggi dan terendah.



Insight:

Bar chart menunjukkan perbandingan rata-rata Income composition of resources (rentang 0-1) antara Jepang sebagai negara dengan rata-rata angka harapan hidup tertinggi dan Sierra Leone sebagai negara dengan rata-rata angka harapan hidup terendah.

- Jepang menunjukkan nilai rata-rata Income composition of resources yang relatif tinggi bahkan mendekati 1, yaitu sebesar 0,8766. Hal ini menunjukkan bahwa Jepang memiliki pembagian pendapatan yang lebih merata dan lebih tinggi secara keseluruhan, serta

Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

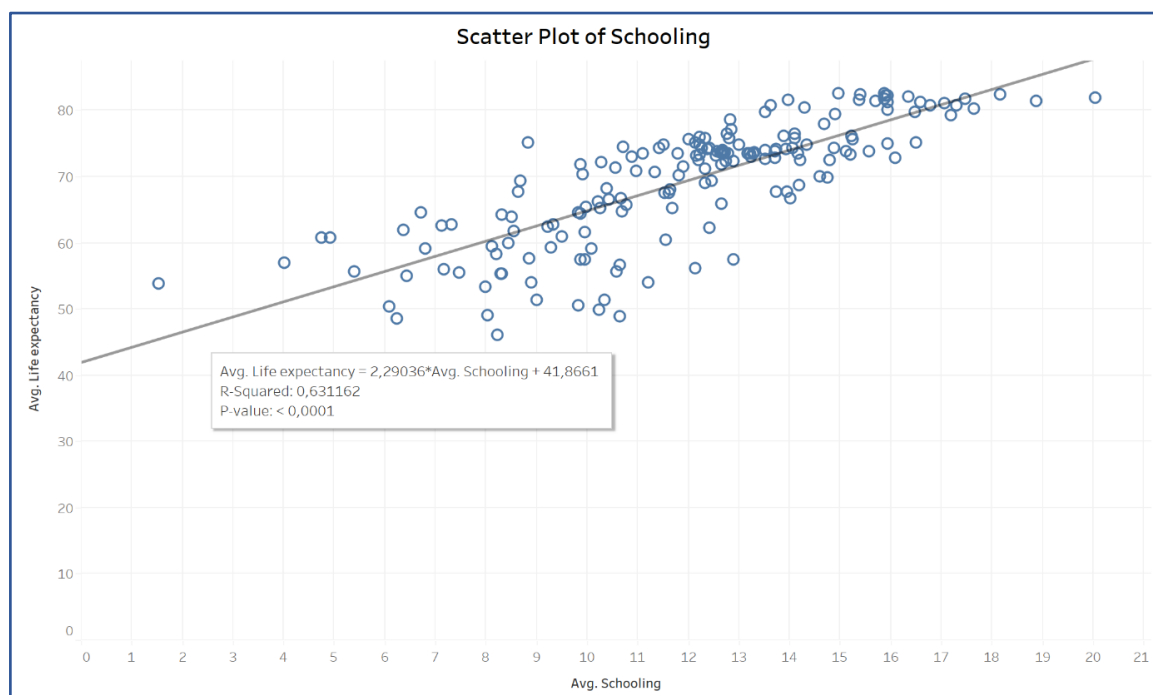
penduduknya memiliki akses yang lebih baik terhadap sumber daya ekonomi, pendidikan, kesehatan, dan kebutuhan dasar lainnya.

- Di sisi lain, Sierra Leone memiliki nilai rata-rata Income composition of resources yang relatif rendah, yaitu sebesar 0,3618. Hal ini mengindikasikan bahwa pembagian pendapatan di Sierra Leone tidak merata dan tingkat pengembangan ekonomi serta akses terhadap sumber daya yang lebih terbatas. Penduduknya mungkin menghadapi tantangan dalam memenuhi kebutuhan dasar dan meningkatkan standar hidup mereka.

Perbandingan ini menggambarkan ketimpangan ekonomi antara dua negara tersebut dan menunjukkan bahwa faktor-faktor seperti distribusi pendapatan dan sumber daya ekonomi dapat berpengaruh pada tingkat pembangunan dan kualitas hidup suatu negara.

7. Scatter Plot of Schooling

Visualisasi yang dilakukan bertujuan untuk mengetahui apakah pendidikan mempengaruhi rata-rata angka harapan hidup di suatu negara dan bagaimana korelasinya.



Insight:

Scatter plot tersebut menunjukkan hubungan antara variabel Schooling yaitu rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduk suatu negara dan rata-rata angka harapan hidup di suatu negara. Dari scatter plot, dapat dilihat bahwa terdapat pola/korelasi positif antara schooling dan angka harapan hidup. Artinya, semakin tinggi tingkat pendidikan di suatu negara, cenderung memiliki angka harapan hidup yang lebih tinggi juga. Korelasi positif tersebut juga diperkuat oleh persamaan garis regresi yang ditunjukkan dalam scatter plot, yaitu $2,29036 \times \text{Avg. Schooling} + 41,8661$. Koefisien 2,29036 mengindikasikan bahwa setiap peningkatan 1 tahun dalam pendidikan (schooling) berhubungan dengan peningkatan sebesar 2,29036 tahun dalam angka harapan hidup. Nilai konstanta 41,86 adalah perkiraan angka harapan hidup jika tingkat pendidikan (schooling)

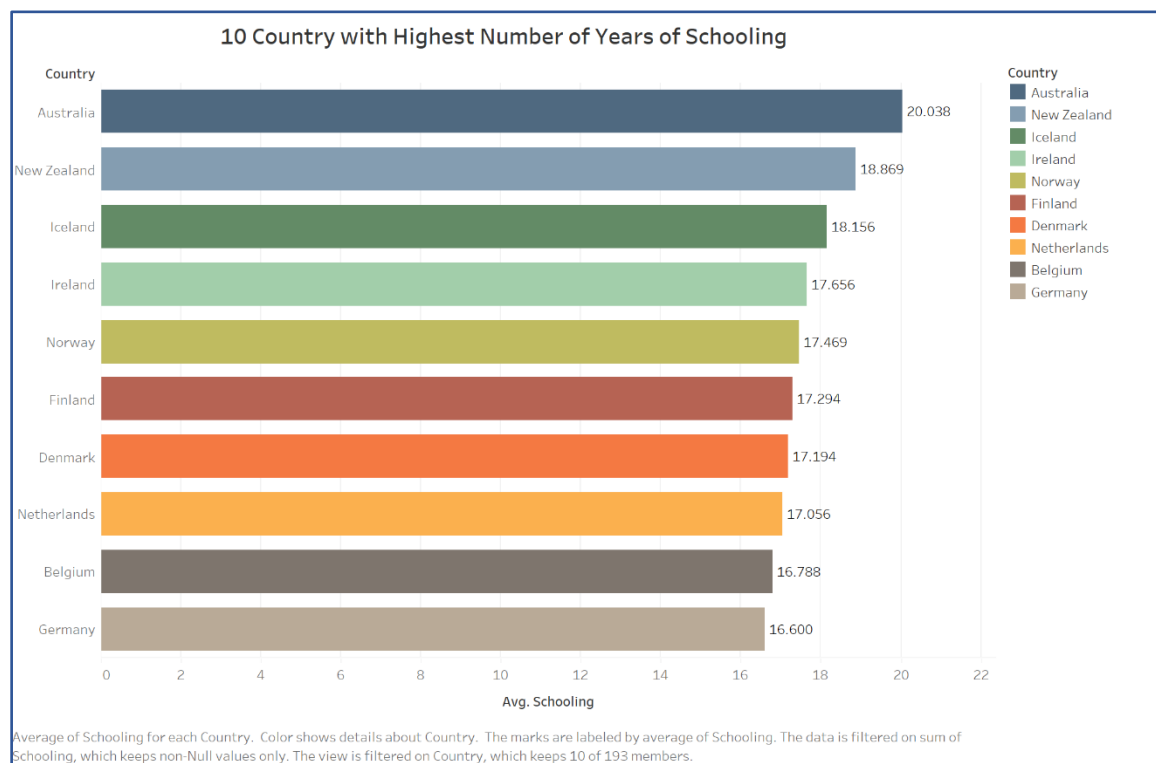
Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

adalah 0. Selain itu, nilai R-square sebesar 0,631 menunjukkan bahwa sekitar 63,1% variasi dalam angka harapan hidup dapat dijelaskan oleh variabel pendidikan (schooling). Artinya, tingkat pendidikan (schooling) secara signifikan berkontribusi terhadap variasi dalam angka harapan hidup di suatu negara. Berdasarkan eksplorasi dan analisis dari output ini, dapat disimpulkan bahwa pendidikan memiliki pengaruh atau korelasi yang positif terhadap rata-rata angka harapan hidup suatu negara. Semakin tinggi tingkat pendidikan penduduk, kemungkinan terdapat peningkatan dalam faktor-faktor seperti kesadaran kesehatan, pengetahuan tentang praktik hidup sehat, akses ke perawatan medis, dan gaya hidup yang lebih baik.

8. Horizontal Bar Chart

Visualisasi yang dilakukan bertujuan untuk mengetahui 10 negara dengan rata-rata jumlah tahun pendidikan (schooling) paling tinggi.



Insight:

Berdasarkan output bar chart di atas, 10 negara dengan rata-rata jumlah tahun pendidikan (Schooling) paling tinggi adalah sebagai berikut:

1. **Australia** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 20,038 tahun.
2. **New Zealand** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 18,869 tahun.
3. **Iceland** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 18,156 tahun.
4. **Ireland** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 17,656 tahun.

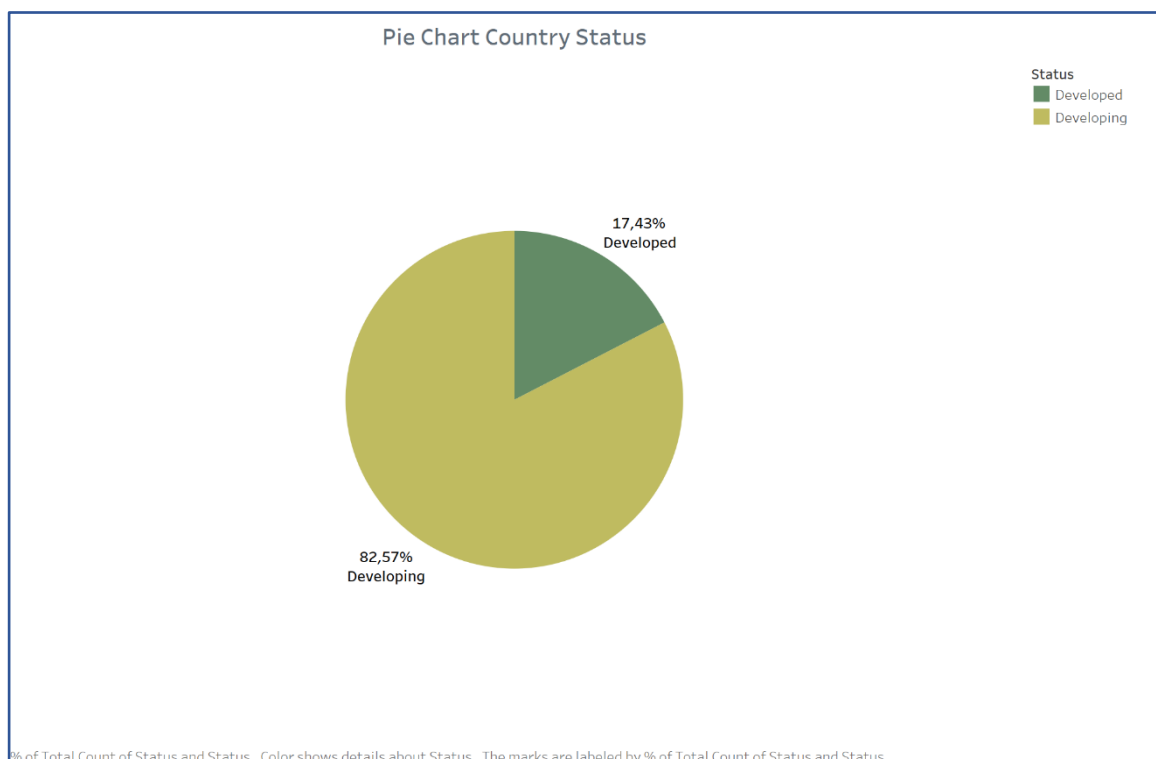
Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

5. **Norway** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 17,469 tahun.
6. **Finland** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 17,294 tahun.
7. **Denmark** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 17,194 tahun.
8. **Netherlands** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 17,056 tahun.
9. **Belgium** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 16,788 tahun.
10. **Germany** dengan rata-rata jumlah tahun pendidikan formal yang telah ditempuh oleh penduduknya adalah 16,600 tahun.

9. Pie Chart

Visualisasi yang dilakukan bertujuan untuk melihat proporsi antara negara berkembang dengan negara maju dalam dataset yang ada.



Insight:

Pie chart menunjukkan proporsi status suatu negara dari data 193 negara yang didapat dari repositori *Global Health Observatory* (GHO) yang dikelola oleh *World Health Organization* (WHO), serta data ekonomi yang berasal dari situs Perserikatan Bangsa-Bangsa, dapat diketahui bahwa status suatu negara dalam kasus ini terbagi menjadi 2 bagian yaitu sebagai berikut:

Nama : Aulia Mirfah Setyo Ayu Damayanti

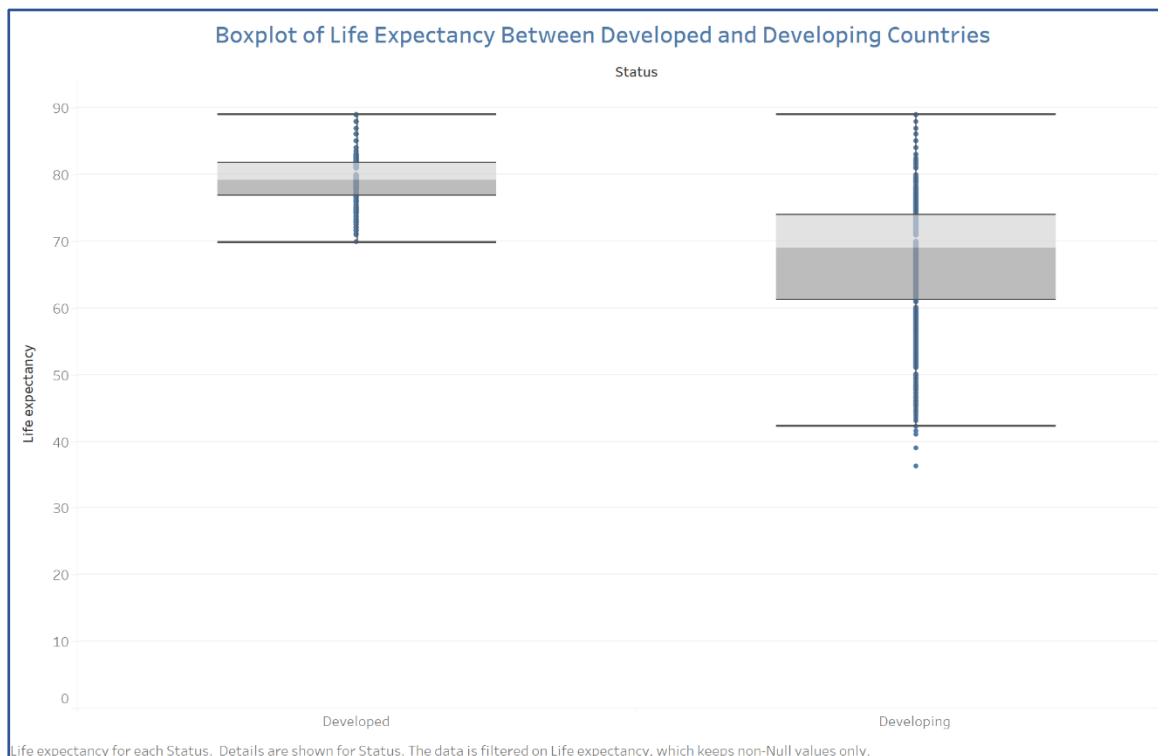
NIM : 22/493337/PA/21178

- Negara dengan status negara berkembang (Developing) sebanyak 82,57% dari total keseluruhan atau 161 negara.
- Negara dengan status negara maju (Developed) sebanyak 17,43% dari total keseluruhan atau 32 negara.

Maka dapat diambil kesimpulan bahwa dari data 193 negara yang tersedia sebagian besar merupakan negara berkembang atau *developing country* yaitu sebanyak 161 negara (82,57%).

10. Boxplot

Visualisasi yang dilakukan bertujuan untuk melihat perbandingan distribusi angka harapan hidup antara negara maju dan negara berkembang.



Boxplot di atas dilihat masing-masing negara pada masing-masing tahun pada tahun yang diteliti (2000-2015), bukan rata-rata per tahun per negara. Berdasarkan boxplot di atas, diperoleh informasi sebagai berikut:

- Ringkasan numerik boxplot life expectancy berdasarkan status

Ringkasan Numerik	Developed	Developed
Nilai Minimum	69,90	36,30
Kuartil Bawah	76,80	61,10
Median	79,25	69,00
Kuartil Atas	81,70	74,00
Nilai Maksimum	89,00	89,00

- Pada boxplot developed terlihat tidak terdapat nilai outlier atas maupun bawah, sehingga dapat dikatakan bahwa data life expectancy pada negara dengan status developed (negara maju) cenderung berdistribusi normal. Sedangkan pada boxplot developing terlihat memiliki beberapa nilai outlier bawah, yang mana dapat dikatakan bahwa data life expectancy pada negara dengan status developing (negara berkembang) tidak berdistribusi normal.
- Bentuk distribusi data pada *developing country* adalah menceng ke kiri (bisa disebut menjurai ke bawah), tampak bahwa nilai Q2 (median) pada boxplot dengan status developing lebih dekat dengan Q3 (kuartil atas), yang mana artinya nilai rendah lebih menyebar sedangkan nilai tinggi lebih terpusat. Kemudian, whisker bawah lebih panjang daripada whisker atas, serta luas box bagian tengah ke atas lebih kecil dari luas box bagian tengah ke bawah. Sementara itu, pada boxplot dengan status developed cenderung berdistribusi normal sebab nilai median (Q2) berada di tengah box, tidak terdapat outlier, serta panjang whisker sama.
- Kemudian, negara berkembang (*developing country*) memiliki dispersi angka harapan hidup yang jauh lebih tinggi daripada negara maju (*developed country*) sebab *range* (selisih nilai minimum dan maksimum) dari data pada status developing lebih tinggi yakni sebesar 52,7 sedangkan *range* dari data pada status developed sebesar 19,1, oleh karena itu dapat dikatakan pula bahwa negara berkembang memiliki variansi angka harapan hidup lebih tinggi dibanding negara maju.
- Untuk menentukan apakah nilai suatu set data lebih rendah daripada set data lainnya, dapat dilihat dari posisi boxplot masing-masing pada sumbu vertikal. Terlihat pada boxplot dari status developed dan developing bahwa boxplot developed berada pada posisi lebih tinggi atau lebih ke atas pada sumbu vertikal dibandingkan dengan boxplot developing. Selain itu, posisi median pada boxplot developed juga relatif jauh lebih tinggi daripada boxplot developing. Dengan ini dapat disimpulkan bahwa terdapat perbedaan yang signifikan terkait angka harapan hidup antara negara-negara berkembang dengan negara-negara maju.

Conclusion

Berdasarkan research question/objective yang telah dibuat serta hasil discussion/analysis yang telah dipaparkan untuk menjawab pertanyaan-pertanyaan tersebut. Dapat diambil kesimpulan sebagai berikut:

Pertama, terdapat perbedaan dalam rata-rata angka harapan hidup di antara 193 negara. Secara umum, benua Amerika, Eropa, dan Australia memiliki tingkat angka harapan hidup yang tinggi, sedangkan Benua Afrika memiliki tingkat angka harapan hidup yang rendah. Negara dengan rata-rata angka harapan hidup terendah adalah Sierra Leone yaitu sebesar 46,11 tahun. Sedangkan negara dengan rata-rata angka harapan hidup tertinggi yaitu Jepang sebesar 82,54 tahun. Kemudian, untuk benua Asia menunjukkan variasi dalam angka harapan hidup, yang mana tidak tergolong rendah ataupun tinggi.

Selanjutnya, ingin diketahui pula bagaimana perbandingan tren angka harapan hidup selama tahun 2000-2015 antara Jepang dan Sierra Leone sebagai negara dengan rata-rata *life expectancy* tertinggi dan terendah. Menggunakan line chart diperoleh hasil bahwa angka harapan hidup Jepang cenderung meningkat secara konsisten dari tahun ke tahun, sedangkan Sierra Leone menunjukkan tren yang cenderung lebih tidak stabil dan fluktuatif dalam angka harapan hidup. Dilihat dari posisi line antara keduanya menunjukkan perbedaan yang signifikan dimana Jepang berada jauh lebih di atas berdasarkan sumbu Y daripada Sierra Leone.

Faktor lingkungan dan imunisasi berkontribusi terhadap angka harapan hidup di setiap negara, menunjukkan pentingnya pencegahan penyakit dan kualitas lingkungan dalam mempengaruhi harapan hidup penduduk. Semakin banyak kasus campak yang dilaporkan makan akan menurun angka harapan hidup di wilayah tersebut. Sedangkan, semakin luas cakupan vaksin yang diberikan di wilayah tersebut akan meningkat pula angka harapan hidup.

Kemudian, dari bar chart yang dihasilkan dalam analisis, dapat dilihat bahwa imunisasi Polio memiliki cakupan paling tinggi di antara ketiga jenis imunisasi, dengan rata-rata cakupan sebesar 82,550 persen pada balita berusia 1 tahun, diikuti oleh imunisasi diphtheria sebesar 82.324 persen dan terakhir imunisasi Hepatitis B dengan cakupan paling rendah di antara ketiga jenis imunisasi, dengan rata-rata cakupan sebesar 80,940 persen pada balita berusia 1 tahun.

Dalam kelompok faktor ekonomi dan populasi, faktor komposisi pendapatan (*Income composition of resources*) suatu negara memiliki pengaruh yang paling signifikan terhadap angka harapan hidup penduduknya dibandingkan dengan faktor lain dari kelompok faktor tersebut dengan nilai R-Square sebesar 0,701. Rata-rata komposisi pendapatan antara Jepang dengan Sierra Leone berbeda jauh, dengan selisih nilai lebih dari 0,5 dari skala 1.

Selain itu, melihat dari kelompok faktor gaya hidup, tingkat pendidikan (*schooling*) secara signifikan berkontribusi terhadap variasi dalam angka harapan hidup di suatu negara dengan nilai R-Square sebesar 0,631 dan korelasi positif dimana semakin tinggi tingkat pendidikan maka akan semakin meningkat pula angka harapan hidup. Dari 193 negara, 10 negara dengan rata-rata jumlah tahun pendidikan (*schooling*) paling tinggi berdasarkan data adalah Australia, New Zealand, Iceland, Ireland, Norway, Finland, Denmark, Netherlands, Belgium, dan Germany.

Nama : Aulia Mirfah Setyo Ayu Damayanti
NIM : 22/493337/PA/21178

Dari daftar tersebut, seluruhnya adalah negara dengan status negara maju kecuali Finland yang masih berstatus developing. Karena itu, ingin dilihat bagaimana proporsi antara negara berkembang dan negara maju dari dataset yang ada. Diperoleh bahwa 82,57% adalah negara berkembang dan sisanya yaitu 17,43% adalah negara maju. Dari sana selanjutnya ingin dilihat perbandingan distribusi angka harapan hidup antara negara maju dan negara berkembang. Dari analisis boxplot yang dilakukan, dapat diambil beberapa kesimpulan penting. Pertama, data angka harapan hidup pada negara maju (developed country) cenderung berdistribusi normal, sedangkan pada negara berkembang (developing country) terdapat beberapa nilai outlier bawah, menunjukkan bahwa distribusi data tidak normal. Kedua, distribusi data angka harapan hidup pada negara berkembang cenderung menceng ke kiri, dengan nilai rendah yang lebih menyebar dan nilai tinggi yang lebih terpusat, sementara pada negara maju distribusi data cenderung normal. Ketiga, negara berkembang memiliki variasi angka harapan hidup yang jauh lebih tinggi daripada negara maju, yang mengindikasikan tingkat dispersi yang lebih tinggi di negara berkembang. Terakhir, terlihat bahwa posisi boxplot negara maju berada lebih tinggi pada sumbu vertikal dibandingkan dengan negara berkembang, serta median negara maju juga relatif lebih tinggi. Dengan demikian, dapat disimpulkan bahwa terdapat perbedaan yang signifikan terkait angka harapan hidup antara negara-negara berkembang dengan negara-negara maju.

Saran:

Berdasarkan kesimpulan yang telah diambil, terdapat beberapa saran yang dapat diberikan kepada masyarakat dan pemerintah. Pertama, penting bagi pemerintah untuk fokus pada upaya peningkatan angka harapan hidup di negara-negara berkembang. Hal ini dapat dilakukan melalui perbaikan infrastruktur kesehatan, peningkatan kualitas lingkungan, peningkatan aksesibilitas dan kualitas layanan kesehatan di daerah pedesaan dan terpencil perlu menjadi perhatian utama. serta program imunisasi yang efektif untuk mencegah penyakit dan meningkatkan harapan hidup penduduk. Kedua, pendidikan merupakan faktor penting yang berkontribusi terhadap angka harapan hidup. Oleh karena itu, pemerintah perlu memberikan perhatian dan investasi yang cukup pada sektor pendidikan untuk meningkatkan tingkat pendidikan di negara-negara tersebut. Ketiga, kerja sama internasional juga sangat penting dalam mempromosikan transfer pengetahuan dan sumber daya untuk meningkatkan angka harapan hidup di negara-negara berkembang. Negara maju dapat berperan dalam memberikan bantuan dan dukungan dalam bidang kesehatan dan pendidikan kepada negara-negara yang membutuhkan. Masyarakat juga perlu terlibat aktif dalam upaya meningkatkan angka harapan hidup, dengan menjaga gaya hidup sehat, serta mendukung kebijakan yang ditetapkan oleh pemerintah, misalnya dalam hal vaksinasi dan imunisasi. Dengan adanya kolaborasi yang baik antara masyarakat dan pemerintah serta dukungan internasional, diharapkan dapat terjadi peningkatan yang signifikan dalam angka harapan hidup di negara-negara berkembang dan pengurangan kesenjangan dengan negara-negara maju.



Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

Daftar Pustaka

Rajarshi, K. (2018). *Life Expectancy (WHO)*. Diakses pada 17 Juni 2023, dari

[Life Expectancy \(WHO\) | Kaggle](#)

McKinsey&Company. (2021). *McKinsey Global Surveys, 2021: A year in review*. Diakses pada 17 Juni 2023, dari [mckinsey-global-surveys-2021-a-year-in-review.pdf](#)

Downs Law Firm, P.C (2020). *Life Expectancy Has Gone Down*. Diakses pada 17 Juni 2023, dari [Life Expectancy Has Gone Down - Downs Law Firm, P.C.](#)

ArcGIS Insights (2023). *Create and use a choropleth map*. Diakses pada 17 Juni 2023, dari [Create and use a choropleth map—ArcGIS Insights | Documentation](#)

World Health Organization (2023). *Japan, Health Data Overview for Japan*. Diakses pada 17 Juni 2023, dari [Japan \(who.int\)](#)

Nama : Aulia Mirfah Setyo Ayu Damayanti

NIM : 22/493337/PA/21178

Lampiran

- Link to Dataset: [Life Expectancy Data](#)
- Link to Tableau Output: [Tableau Project UAS EVD - Aulia Mirfah](#)
- Link to Youtube Presentation: [Presentation of Project EVD - Statistical Analysis on Factors Influencing Life Expectancy](#)