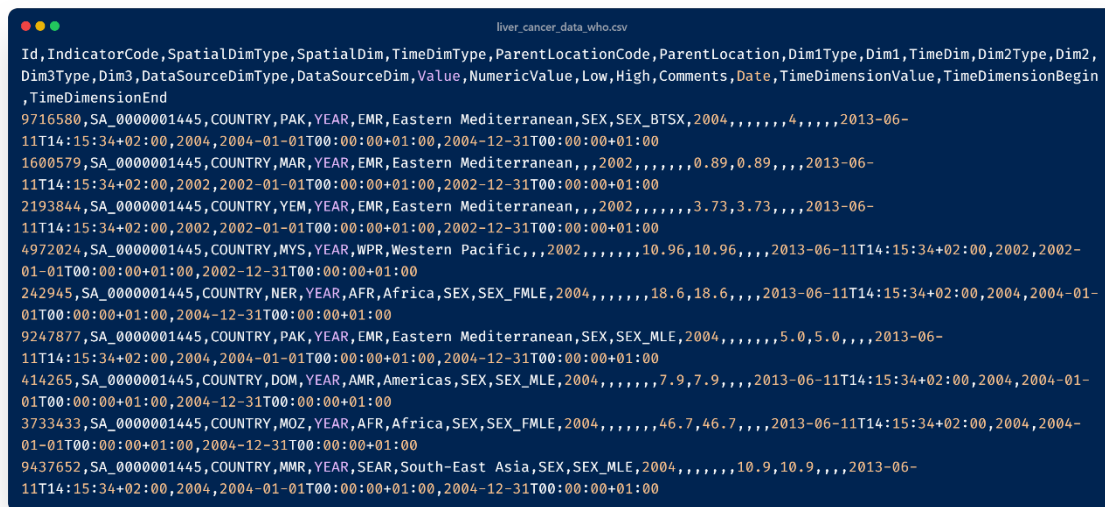


Tugas 1 Sistem Informasi Kesehatan

Data yang dianalisis berjudul “Age-standardized death rates, liver cancer, per 100,000 dari World Health Organization Global Health Observatory (WHO GHO)”, yakni tingkat kematian yang telah disesuaikan dengan struktur usia populasi standar untuk memungkinkan perbandingan yang adil antar negara atau wilayah dengan komposisi usia yang berbeda. Indikator ini mengukur tingkat kematian akibat kanker hati per 100,000 penduduk setelah menghilangkan pengaruh perbedaan struktur usia populasi.



```
liver_cancer_data_who.csv
Id,IndicatorCode,SpatialDimType,SpatialDim,TimeDimType,ParentLocationCode,ParentLocation,Dim1Type,Dim1,TimeDim,Dim2Type,Dim2,Dim3Type,Dim3,DataSourceDimType,DataSourceDim,Value,NumericValue,Low,High,Comments,Date,TimeDimensionValue,TimeDimensionBegin,TimeDimensionEnd
9716580,SA_0000001445,COUNTRY,PAK,YEAR,EMR,Eastern Mediterranean,SEX,SEX_BTSX,2004,,,,,4,,,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
1600579,SA_0000001445,COUNTRY,MAR,YEAR,EMR,Eastern Mediterranean,,,2002,,,,,0.89,0.89,,,2013-06-11T14:15:34+02:00,2002,2002-01-01T00:00:00+01:00,2002-12-31T00:00:00+01:00
2193844,SA_0000001445,COUNTRY,YEM,YEAR,EMR,Eastern Mediterranean,,,2002,,,,,3.73,3.73,,,2013-06-11T14:15:34+02:00,2002,2002-01-01T00:00:00+01:00,2002-12-31T00:00:00+01:00
4972024,SA_0000001445,COUNTRY,MYS,YEAR,WPR,Western Pacific,,,2002,,,,,10.96,10.96,,,2013-06-11T14:15:34+02:00,2002,2002-01-01T00:00:00+01:00,2002-12-31T00:00:00+01:00
242945,SA_0000001445,COUNTRY,NER,YEAR,AFR,Africa,SEX,SEX_FMLE,2004,,,,,18.6,18.6,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
9247877,SA_0000001445,COUNTRY,PAK,YEAR,EMR,Eastern Mediterranean,SEX,SEX_MLE,2004,,,,,5.0,5.0,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
414265,SA_0000001445,COUNTRY,DOM,YEAR,AMR,Americas,SEX,SEX_MLE,2004,,,,,7.9,7.9,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
3733433,SA_0000001445,COUNTRY,MOZ,YEAR,AFR,Africa,SEX,SEX_FMLE,2004,,,,,46.7,46.7,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
9437652,SA_0000001445,COUNTRY,MMR,YEAR,SEAR,East Asia,SEX,SEX_MLE,2004,,,,,10.9,10.9,,,2013-06-11T14:15:34+02:00,2004,2004-01-01T00:00:00+01:00,2004-12-31T00:00:00+01:00
```

Gambar Sebagian Data

Sumber data ini berasal dari [WHO Global Health Observatory \(GHO\) API](#). WHO GHO adalah sistem basis data global yang berisi informasi statistik kesehatan dari seluruh dunia. Data ini bersifat publik dan bisa diakses secara terprogram untuk keperluan penelitian dan analisis.

WHO sebagai pihak penyedia, mengumpulkan data statistik kesehatan dari berbagai sumber terpercaya di seluruh dunia, termasuk laporan resmi dari kementerian kesehatan negara-negara anggota. Data ini kemudian diolah dan disatukan ke dalam basis data global mereka, WHO Global Health Observatory (GHO).

Kemudian, cara saya mengambil data sebagai berikut :

Data ini diambil menggunakan API kemudian saya simpan dalam format csv dengan langkah sebagai berikut :

1. **Mengirimkan Permintaan (Request):** [program \(menggunakan bahasa pemrograman Python\)](#) mengirimkan permintaan (request) ke alamat URL spesifik dari WHO API.

Permintaan ini berisi kode indikator penyakit yang diinginkan, dalam hal ini SA_0000001445 untuk kanker liver.

2. **Menerima Respon (Response):** Server WHO menerima permintaan tersebut, mencari data yang sesuai di dalam basis datanya, dan mengirimkan kembali data sebagai respon.
3. **Memproses Data:** Respon dari server datang dalam format **JSON (JavaScript Object Notation)**, sebuah format yang terstruktur dan mudah dibaca oleh program. Program kemudian membaca data JSON ini, mengekstrak informasi yang relevan seperti lokasi, tahun, dan nilai data, lalu mengubahnya menjadi format python

```
main.py

import requests
import csv

url_api = "https://ghoapi.azureedge.net/api/SA_0000001445"

print("Mengirim permintaan ke API ... ")
response = requests.get(url_api)

if response.status_code == 200:
    data_json = response.json()
    records = data_json.get("value", [])
    csv_file = "liver_cancer_data_who.csv"

    if records:
        header = records[0].keys()
        with open(csv_file, 'w', newline='', encoding='utf-8') as f:
            writer = csv.DictWriter(f, fieldnames=header)
            writer.writeheader()

            writer.writerows(records)

        print(f"Data berhasil disimpan ke {csv_file}")
    else:
        print("Data tidak ditemukan.")
else:
    print(f"Permintaan gagal dengan kode status: {response.status_code}")
```

Gambar Program untuk Mengambil Data

Untuk penyimpanan data WHO GHO dengan volume sekitar 800 record, terdapat dua pendekatan utama yang dapat diterapkan berdasarkan kebutuhan analisis dan kompleksitas query yang diperlukan. Pendekatan pertama adalah penyimpanan dalam format CSV untuk

kebutuhan sederhana dan analisis dasar. Format CSV sangat cocok untuk data dengan volume kecil seperti ini karena mudah dibaca, ringan, dan dapat dibuka menggunakan berbagai aplikasi seperti Microsoft Excel, Google Sheets, atau text editor biasa. File CSV juga memudahkan proses backup, sharing, dan portabilitas data antar sistem operasi. Untuk analisis sederhana seperti filtering berdasarkan negara, pengelompokan berdasarkan tahun, atau perhitungan statistik dasar, format CSV sudah mencukupi dan dapat diproses menggunakan tools seperti Python pandas atau R dengan performa yang memadai.

Kemudian, untuk kebutuhan query yang lebih kompleks dan analisis mendalam, disarankan menggunakan PostgreSQL sebagai database management system. Penggunaan database relasional dengan struktur tabel `who_health_indicators` yang telah disesuaikan dengan format asli dataset. Struktur ini dipilih berdasarkan beberapa pertimbangan strategis yang mempertimbangkan karakteristik data dan kebutuhan analisis.

Dengan database relational juga dapat dilakukan optimasi Performa melalui Indexing bertujuan untuk mengoptimalkan query yang paling sering dilakukan. `idx_indicator_code` mempercepat pencarian berdasarkan jenis indikator kesehatan, `idx_spatial_dim` mengoptimalkan query geografis berdasarkan negara, `idx_time_dim` memfasilitasi analisis temporal, sedangkan `idx_composite` mendukung query kompleks yang menggabungkan ketiga dimensi utama tersebut. Kombinasi index ini memastikan performa query tetap optimal bahkan dengan volume data yang besar mencapai jutaan record(jika nanti data bertambah).

```
schema.sql

CREATE TABLE who_health_indicators (
  id BIGINT PRIMARY KEY,
  indicator_code VARCHAR(50) NOT NULL,
  spatial_dim_type VARCHAR(20),
  spatial_dim VARCHAR(10),
  time_dim_type VARCHAR(10),
  parent_location_code VARCHAR(10),
  parent_location VARCHAR(100),
  dim1_type VARCHAR(20),
  dim1 VARCHAR(20),
  time_dim INT,
  dim2_type VARCHAR(20),
  dim2 VARCHAR(20),
  dim3_type VARCHAR(20),
  dim3 VARCHAR(20),
  data_source_dim_type VARCHAR(50),
  data_source_dim VARCHAR(100),
  value_text VARCHAR(50),
  numeric_value DECIMAL(10,4),
  low_value DECIMAL(10,4),
  high_value DECIMAL(10,4),
  comments TEXT,
  date_recorded TIMESTAMP,
  time_dimension_value INT,
  time_dimension_begin DATE,
  time_dimension_end DATE,
  created_at TIMESTAMP DEFAULT CURRENT_TIMESTAMP,
  updated_at TIMESTAMP DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,

  INDEX idx_indicator_code (indicator_code),
  INDEX idx_spatial_dim (spatial_dim),
  INDEX idx_time_dim (time_dim),
  INDEX idx_composite (indicator_code, spatial_dim, time_dim)
);
```

Gambar Schema Tabel

Lampiran

Link Github (Berisi program, data) : <https://github.com/WwzFwz/Tugas-1-Sistem-Informasi-Kesehatan>

Link WHO : <https://www.who.int/data/gho>

Link API : <https://ghoapi.azureedge.net/api>