

**LAPORAN
PRAKTIK DATA ENGINEERING
TEAM PROJECT KE-3**



Oleh :

Dzaki Anwar	(234311037)
Ikbal Dzakwan	(234311043)
Veri Tri Anggoro	(234311055)
Widya Wulandari	(234311056)

**JURUSAN TEKNIK
PROGRAM STUDI REKAYASA PERANGKAT LUNAK
POLITEKNIK NEGERI MADIUN
2025**

Team Project : Open Data Source Analysis & Planning

Project Information

Project Name : Dampak Fluktuasi Nilai Tukar terhadap Kinerja Ekspor dan Import Indonesia

Created By : Data Engineering Kelompok 4

Date : 22 Februari, 2025

Version : 1.0

1. Executive Summary

1.1 Project Overview

- Tujuan Project:
 - Menganalisis dampak fluktuasi nilai tukar Rupiah terhadap kinerja ekspor dan impor Indonesia, untuk mengetahui keterkaitan antara faktor nilai tukar dan pergerakan perdagangan internasional.
- Scope Project :
 - Integrasi dan analisis data historis nilai tukar Rupiah terhadap USD serta data ekspor dan impor Indonesia.
- Expected Outcomes :
 - Laporan yang menjelaskan hubungan antara pergerakan nilai tukar dan perubahan volume ekspor serta impor Indonesia, serta rekomendasi kebijakan berdasarkan hasil analisis data..
- Timeline : 3 bulan (Maret - Mei 2025)

1.2 Stakeholders

- Project Owner : Data Engineering Kelompok 4
- Team Members :
 - Data Engineer : Widya Wulandari
 - Data Analyst : Veri Tri Anggoro, Ikbal Dzakwan
 - Project Manager : Dzaki Anwar Zulfahmi
- End Users :
 - Pemerintah dan regulator ekonomi
 - Pelaku bisnis ekspor
 - Akademisi dan peneliti ekonomi

2. Data Source Analysis

2.1 Data Total Nilai Ekspor non-migas Indonesia

Source Details

- Dataset Name : Perkembangan Ekspor Non Migas (Komoditi)
- URL/Access Point : <https://satudata.kemendag.go.id/data-informasi/perdagangan-luar-negeri/ekspor-non-migas-komoditi>
- Data Owner : Kementerian Perdagangan RI
- Update Frequency : Tahunan

Data Analysis

- Format Data : CSV
- Volume Data : 11.37 K
- Time Coverage : Data Eksport Non Migas tahun 2020-2024

- Data Quality : Baik, bersumber dari instansi resmi

2.2 Perkembangan Impor Non Migas (Komoditi)

Source Details

- Dataset Name : Perkembangan Impor Non Migas (Komoditi)
- URL/Access Point : <https://satudata.kemendag.go.id/data-informasi/perdagangan-luar-negeri/impor-non-migas-komoditi>
- Data Owner : Kementerian Perdagangan RI
- Update Frequency : Tahunan

Data Analysis

- Format Data : CSV
- Volume Data : 11.45 K
- Time Coverage : Data Impor Non Migas tahun 2020-2024
- Data Quality : Baik, bersumber dari Kemendag

2.3 Data KURS Nilai Tukar Rupiah terhadap Mata Uang USD

Source Details

- Dataset Name : Kurs Transaksi Bank Indonesia Mata Uang Usd(dari data center ortax.org)
- URL/Access Point : https://datacenter.ortax.org/ortax/kursbi/show/USD?rentang_tanggal=2020-01-01,2024-12-31&show=USD&page=1
- Data Owner : Bank Indonesia
- Update Frequency : Harian

Data Analysis

- Format Data : CSV
- Volume Data : 88.34 K
- Time Coverage : Data nilai tukar tahun 2020-2024(01 Januari 2020 – 31 Desember 2024)
- Data Quality : Baik, bersumber dari Bank Indonesia

3. Implementasi Teknolog

- 3.1 Python : bahasa pemrograman yang digunakan.
- 3.2 Pandas : manipulasi, pembersihan, dan penyimpanan data (CSV & HTML).
- 3.3 ydata-profiling : Exploratory Data Analysis (EDA).
- 3.4 Web Scraping : untuk ambil data dari website menggunakan kombinasi requests dan pandas.read_html.
- 3.5 Google Colab : Digunakan sebagai environment untuk menjalankan kode dan memproses data.

4. Hasil dan Pembahasan

- 4.1 Kode Program Pipeline awal Data Engineering
Berikut Kode Implementasi pipeline untuk data Ekspor, Impor, dan Kurs.

```

import pandas as pd

def source_data_from_webpage(web_page_url, table_index=0):
    try:
        df_html = pd.read_html(web_page_url)
        df_html = df_html[table_index]
    except Exception as e:
        print(f"Error membaca tabel dari {web_page_url}: {e}")
        df_html = pd.DataFrame()
    return df_html

def source_data_from_kurs(base_url, max_page=81):
    all_data = []

    for page in range(1, max_page + 1):
        url = f"{base_url}&page={page}"
        print(f"Scraping data dari: {url}")

        try:
            response = requests.get(url)
            response.raise_for_status()
            tables = pd.read_html(response.text)

            if tables:
                df_page = tables[0]
                all_data.append(df_page)
        except Exception as e:
            print(f"Error membaca data kurs dari {url}: {e}")

    if all_data:
        df_kurs = pd.concat(all_data, ignore_index=True)
    else:
        df_kurs = pd.DataFrame()

    return df_kurs

def extract_data():
    ekspor_url = "https://satudata.kemendag.go.id/data-informasi/perdagangan-luar-negeri/ekspor-non-migas-komoditi"
    impor_url = "https://satudata.kemendag.go.id/data-informasi/perdagangan-luar-negeri/impor-non-migas-komoditi"
    kurs_base_url = "https://datacenter.ortax.org/ortax/kursbi/show/USD?rentang_tanggal=2020-01-01,2024-12-31&show=USD"

    df_ekspor = source_data_from_webpage(ekspor_url)
    df_impor = source_data_from_webpage(impor_url)
    df_kurs = source_data_from_kurs(kurs_base_url, max_page=81)

    return df_ekspor, df_impor, df_kurs

def save_to_csv(df, filename):
    if not df.empty:
        df.to_csv(filename, index=False)
        print(f>Data berhasil disimpan ke {filename}")
    else:
        print(f"Tidak ada data untuk {filename}")

if __name__ == "__main__":
    df_ekspor, df_impor, df_kurs = extract_data()

    save_to_csv(df_ekspor, "ekspor_non_migas.csv")
    save_to_csv(df_impor, "impor_non_migas.csv")
    save_to_csv(df_kurs, "kurs_usd.csv")

```

4.2 Hasil Kode Program Pipeline Awal

Dari kode program di atas menghasilkan 3 data dari webpage yang disimpan dalam bentuk file CSV :

1. Data Ekspor Non-Migas (Komoditas) di Indonesia berdasarkan HS Code dari tahun 2020 sampai 2024.

label Terminal ✓ 10.37 Python 3

import_non_migas.csv X

label Terminal ✓ 10.37 Python 3

Source: <http://www.irs.gov>

riabel Terminal

✓ 10.37 Python 3

- Data ekspor non-migas Indonesia diekstrak dari website Kementerian Perdagangan menggunakan teknik web scraping dengan `pandas.read_html` dan `requests`.
- Data impor non-migas juga diambil dari situs yang sama menggunakan metode dan teknik yang serupa yaitu `pandas.read_html` dan `requests`.

- Data kurs harian nilai tukar Rupiah terhadap USD diperoleh dari situs [Ortax.org](https://ortax.org) (data resmi dari BI), juga menggunakan kombinasi `requests` dan `pandas.read_html` untuk mengekstrak tabel kurs transaksi.
2. Transformation Rules (Pembersihan dan pemrosesan data)
- Semua dataset dibersihkan dari data duplikat dan nilai kosong menggunakan fungsi `dropna` dan `drop_duplicates`.
 - Format tanggal pada data kurs dikonversi menjadi format `datetime` menggunakan `pd.to_datetime` agar bisa diolah lebih lanjut.
 - Karena data kurs bersifat harian, dilakukan agregasi menjadi data tahunan dengan cara menghitung rata-rata kurs jual per tahun menggunakan `groupby('Tahun').mean()`.
 - Dataset ekspor dan impor disaring berdasarkan rentang tahun 2020 hingga 2024 sesuai cakupan analisis.
 - Kolom-kolom yang tidak relevan dihapus, dan nama kolom diseragamkan dengan `rename()` agar lebih mudah dibaca dan digabungkan.
 - Ketiga dataset (ekspor, impor, dan kurs) kemudian digabungkan berdasarkan kolom Tahun menggunakan fungsi `merge` agar dapat dianalisis keterkaitannya secara langsung.