

UAS APD
ALGORITMA PEMROGRAMAN DASAR



Disusun oleh:
Anugerah Fakhri Reswara (2509106025)
Kelas (A'25)

PROGRAM STUDI INFORMATIKA
UNIVERSITAS MULAWARMAN
SAMARINDA
2025

1. Source Code

A. Program utama (main.py)

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('unemployment analysis.csv')

years = [str(y) for y in range(1991, 2022)]
df_numeric = df.set_index('Country Name')[years]

plt.figure(figsize=(20, 14))
plt.suptitle("Visualisasi Analisis Pengangguran Global", fontsize=22,
fontweight='bold')

ax1 = plt.subplot(2, 2, 1)
recent_years = [str(y) for y in range(2015, 2022)]
avg_unemp = df_numeric[recent_years].mean(axis=1).nlargest(10)
avg_unemp.plot(kind='bar', ax=ax1, color='steelblue')
ax1.set_title('10 Negara dengan Rata-rata Pengangguran Tertinggi (2015-2021)',
fontsize=14)
ax1.set_ylabel('Rata-rata (%)')
ax1.tick_params(axis='x', rotation=45)

ax2 = plt.subplot(2, 2, 2)
data_2020 = df['2020'].dropna()
ax2.hist(data_2020, bins=20, color='lightcoral', edgecolor='black', alpha=0.8)
ax2.set_title('Distribusi Tingkat Pengangguran Global Tahun 2020', fontsize=14)
ax2.set_xlabel('Persentase (%)')
ax2.set_ylabel('Jumlah Negara/Wilayah')
ax2.grid(axis='y', linestyle='--', alpha=0.6)

ax3 = plt.subplot(2, 2, 3)
scatter_df = df[['Country Name', '2010', '2020']].dropna()
ax3.scatter(scatter_df['2010'], scatter_df['2020'], alpha=0.7, color='teal')
ax3.set_title('Perbandingan Pengangguran: 2010 vs 2020', fontsize=14)
ax3.set_xlabel('2010 (%)')
ax3.set_ylabel('2020 (%)')
min_val = min(scatter_df[['2010', '2020']].min())
max_val = max(scatter_df[['2010', '2020']].max())
ax3.plot([min_val, max_val], [min_val, max_val], 'r--', label='Tidak Berubah')
ax3.legend()
ax3.grid(True, linestyle='--', alpha=0.5)
```

```

ax4 = plt.subplot(2, 2, 4)
corr_years = [str(y) for y in range(2010, 2022)]
corr_matrix = df[corr_years].corr()

sns.heatmap(
    corr_matrix,
    annot=True,
    fmt=".2f",
    cmap='RdYlBu_r',
    linewidths=0.5,
    cbar_kws={"shrink": 0.8},
    annot_kws={"size": 9},
    ax=ax4
)
ax4.set_title('Heatmap Korelasi (2010-2021)', fontsize=14)
ax4.set_xticklabels(ax4.get_xticklabels(), rotation=0)
ax4.set_yticklabels(ax4.get_yticklabels(), rotation=0)
plt.tight_layout(rect=[0, 0, 1, 0.96])

plt.show()

```

B. Data mahasiswa (unemployment analysis.csv)

Link download dataset: <https://www.kaggle.com/datasets/pantanjali/unemployment-dataset>

2. Penjelasan Program (main.py)

Program ini dibuat untuk memvisualisasikan, memahami pola, tren, dan perbandingan tingkat pengangguran di berbagai negara/wilayah dari tahun 1991 hingga 2021. Khususnya periode 2010–2021. Memvisualisasikan data menggunakan Matplotlib dan Seaborn agar hasil visualisasi rapi dan mudah dipahami. Program ini menghasilkan satu figure utuh berisi empat jenis grafik yaitu: Bar Chart, Histogram, Scatter Plot dan Heatmap.

```

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

```

import pandas as pd: Mengimport library pandas untuk memuat dan memproses data CSV dan di alihkan sebagai **pd**.

import matplotlib.pyplot as plt: Mengimport library matplotlib.pyplot untuk membuat grafik (bar, histogram, scatter) dan dialihkan sebagai **plt**.

import seaborn as sns: Mengimport library seaborn untuk membuat heatmap yang lebih rapih dan bagus.

```
df = pd.read_csv('unemployment analysis.csv')
```

Membaca file CSV bernama “unemployment analysis.csv” ke dalam objek DataFrame bernama df. Setiap baris mewakili satu negara/wilayah, dan kolom berisi data tahunan (1991–2021) + info negara.

```
years = [str(y) for y in range(1991, 2022)]
```

Membuat list years berisi string '1991', '1992', ..., '2021'. Karena kolom tahun di CSV berupa string, bukan angka.

```
df_numeric = df.set_index('Country Name')[years]
```

Mengatur kolom 'Country Name' sebagai indeks (jadi baris diidentifikasi oleh nama negara). Hanya memilih kolom tahun (1991–2021), sehingga df_numeric hanya berisi angka pengangguran. Memudahkan perhitungan statistik per negara.

```
plt.figure(figsize=(20, 14))
```

Membuat satu gambar besar (figure) dengan lebar 20 inci dan tinggi 14 inci. Cukup untuk menampung 4 subplot tanpa berdesakan.

```
plt.suptitle("Visualisasi Analisis Pengangguran Global", fontsize=22, fontweight='bold')
```

Menambahkan judul di atas keempat subplot.

- `fontsize=22`: ukuran font besar.
- `fontweight='bold'`: tebal.

Baris 13-19: Subplot 1 Bar Chart (10 negara tertinggi)

```
ax1 = plt.subplot(2, 2, 1)
recent_years = [str(y) for y in range(2015, 2022)]
avg_unemp = df_numeric[recent_years].mean(axis=1).nlargest(10)
avg_unemp.plot(kind='bar', ax=ax1, color='steelblue')
ax1.set_title('10 Negara dengan Rata-rata Pengangguran Tertinggi (2015-2021)',
              fontsize=14)
ax1.set_ylabel('Rata-rata (%)')
ax1.tick_params(axis='x', rotation=45)
```

- `plt.subplot(2, 2, 1)`: buat subplot pertama dari grid 2×2.
- `recent_years`: daftar tahun 2015–2021.
- `df_numeric[recent_years].mean(axis=1)`: rata-rata tiap baris (negara) dari tahun tersebut.
- `.nlargest(9)`: ambil 10 nilai terbesar.
- `.plot(kind='bar', ...)`: buat bar chart di ax1.
- `rotation=45`: miringkan label negara agar tidak tumpang tindih.

Baris 22-28: Subplot 2 Histogram (distribusi 2020)

```
ax2 = plt.subplot(2, 2, 2)
data_2020 = df['2020'].dropna()
ax2.hist(data_2020, bins=20, color='lightcoral', edgecolor='black', alpha=0.8)
ax2.set_title('Distribusi Tingkat Pengangguran Global Tahun 2020', fontsize=14)
ax2.set_xlabel('Persentase (%)')
ax2.set_ylabel('Jumlah Negara/Wilayah')
ax2.grid(axis='y', linestyle='--', alpha=0.6)
```

- `df['2020'].dropna()`: ambil semua nilai tahun 2020, buang yang kosong (NaN).
- `bins=20`: bagi data ke dalam 20 interval.
- `edgecolor='black'`: beri garis tepi pada batang histogram.
- `grid(...)`: tambahkan garis putus-putus horizontal untuk membantu membaca.

Baris 31-41: Subplot 3 Scatter Plot (2010 vs 2020)

```
ax3 = plt.subplot(2, 2, 3)
scatter_df = df[['Country Name', '2010', '2020']].dropna()
ax3.scatter(scatter_df['2010'], scatter_df['2020'], alpha=0.7, color='teal')
ax3.set_title('Perbandingan Pengangguran: 2010 vs 2020', fontsize=14)
ax3.set_xlabel('2010 (%)')
ax3.set_ylabel('2020 (%)')
min_val = min(scatter_df[['2010', '2020']].min())
max_val = max(scatter_df[['2010', '2020']].max())
ax3.plot([min_val, max_val], [min_val, max_val], 'r--', label='Tidak Berubah')
ax3.legend()
ax3.grid(True, linestyle='--', alpha=0.5)
```

- `dropna()`: hanya ambil negara yang punya data di kedua tahun.
- `ax3.scatter(...)`: buat titik untuk tiap negara.
- Garis merah (`'r--'`) menunjukkan tidak ada perubahan: jika titik di atas garis, pengangguran naik.

Baris 44-60: Subplot 4 Heatmap (korelasi 2010–2021)

```
ax4 = plt.subplot(2, 2, 4)
corr_years = [str(y) for y in range(2010, 2022)]
corr_matrix = df[corr_years].corr()

sns.heatmap(
    corr_matrix,
    annot=True,
    fmt=".2f",
    cmap='RdYlBu_r',
    linewidths=0.5,
    cbar_kws={"shrink": 0.8},
    annot_kws={"size": 9},
    ax=ax4
)
ax4.set_title('Heatmap Korelasi (2010-2021)', fontsize=14)
ax4.set_xticklabels(ax4.get_xticklabels(), rotation=0)
ax4.set_yticklabels(ax4.get_yticklabels(), rotation=0)
```

- `df[corr_years].corr()`: hitung matriks korelasi antar tahun.
- `annot=True`: tampilkan angka korelasi di tiap sel.
- `fmt=".2f"`: format angka jadi 2 desimal (misal: 0.98).
- `cmap='RdYlBu_r'`: palet warna merah-kuning-biru (merah = korelasi tinggi).
- `linewidths=0.5`: beri garis tipis antar sel.
- `rotation=0`: buat label tahun di bawah lurus horizontal (tidak miring).

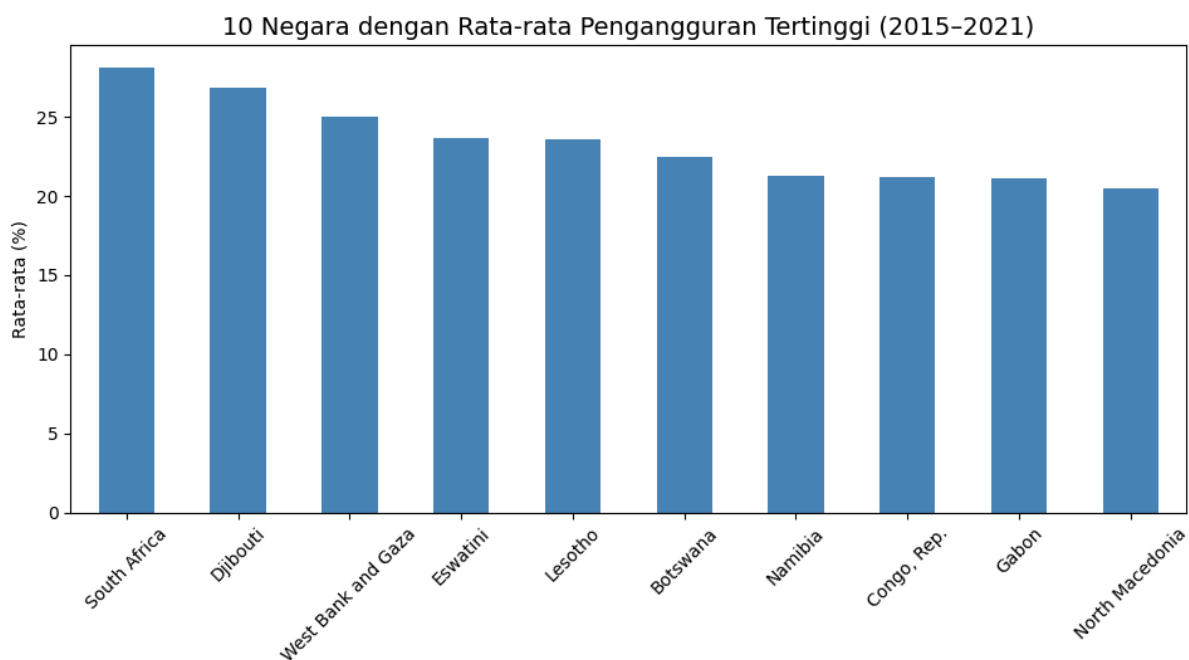
```
plt.tight_layout(rect=[0, 0, 1, 0.96])
plt.show()
```

- `tight_layout(...)`: otomatis atur jarak antar subplot agar tidak tumpang tindih.
- `rect=[0, 0, 1, 0.96]`: sisakan ruang 4% di atas untuk supitle.
- `plt.show()`: tampilkan seluruh figure di layar.

3. Penjelasan apa yang ada divisualisasi

Program ini memvisualisasikan data tingkat pengangguran global dari tahun 1991 hingga 2021, yang diambil dari dataset Kaggle berjudul Unemployment Dataset. Visualisasi dirancang untuk menjawab pertanyaan-pertanyaan utama tentang siapa, kapan, dan seberapa parah pengangguran terjadi di berbagai negara dan wilayah di dunia.

A. Bar Chart - Negara mana yang paling banyak pengangguran?

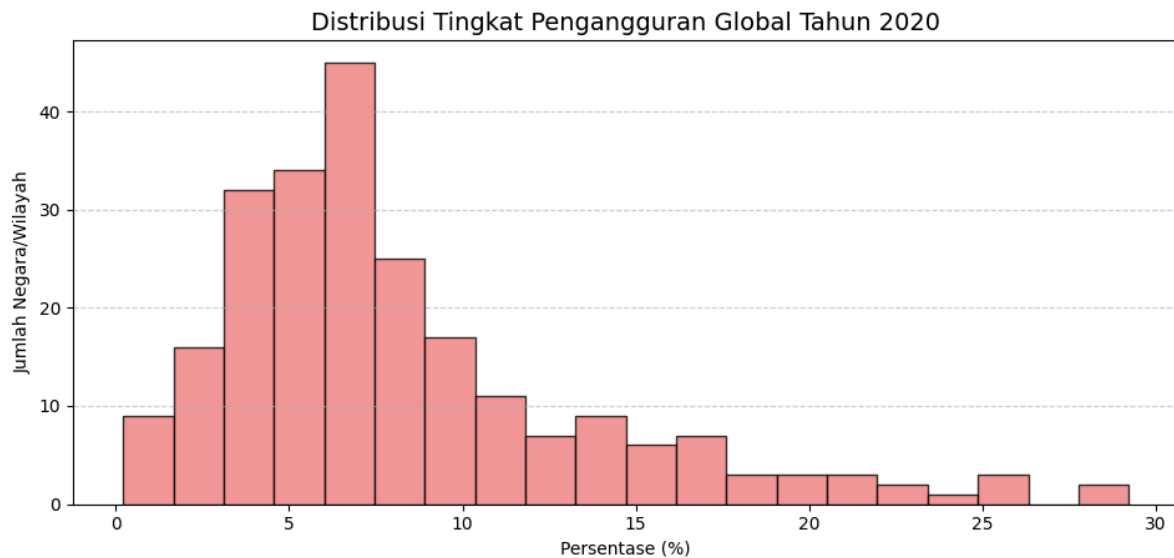


Gambar 3.1 Bar Chart

Visualisasi ini menunjukkan 10 negara atau wilayah dengan rata-rata tingkat pengangguran tertinggi selama 2015–2021.

- Misalnya, negara seperti Lesotho, Afrika Selatan, Namibia, dan Bosnia dan Herzegovina muncul dengan angka pengangguran di atas 20%, bahkan mencapai 30%+.
- Artinya: 1 dari 3 orang usia kerja di sana tidak memiliki pekerjaan — suatu kondisi sosial-ekonomi yang sangat kritis.
- Ini menyoroti ketimpangan global: sementara banyak negara menikmati pengangguran di bawah 5%, beberapa negara terjebak dalam krisis ketenagakerjaan kronis.

B. Histogram - Seberapa umum tingkat pengangguran tertentu di Dunia pada tahun 2020?

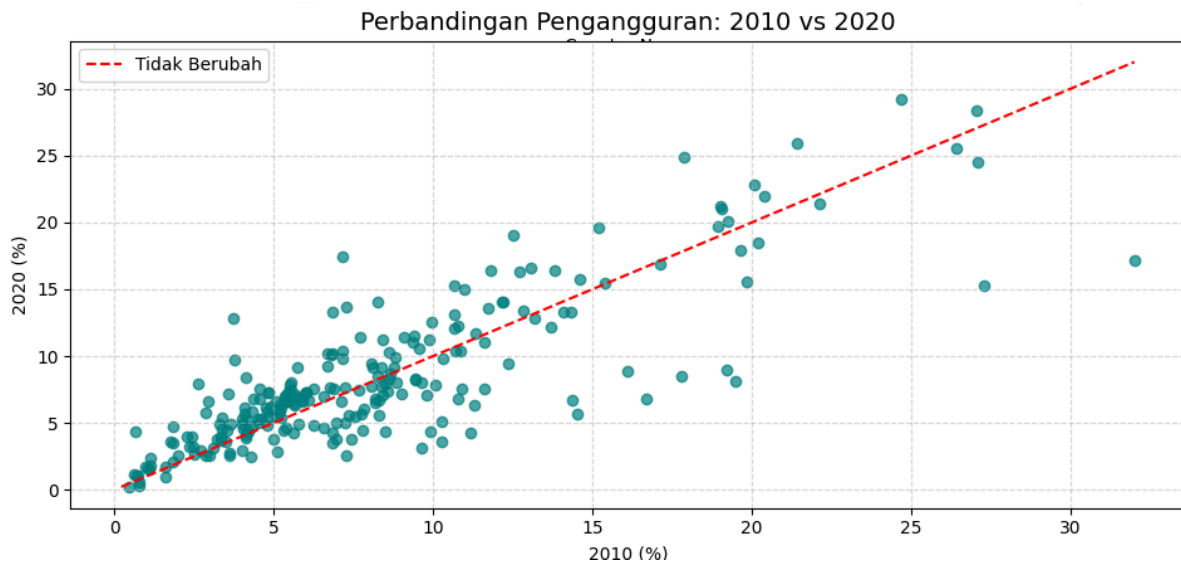


Gambar 3.2 Histogram

Grafik ini memperlihatkan distribusi tingkat pengangguran global pada tahun 2020 — masa puncak pandemi.

- Mayoritas negara berada di kisaran 3–8%, menunjukkan bahwa meski pandemi mengganggu ekonomi, banyak negara masih mampu mempertahankan pasar tenaga kerja yang relatif stabil.
- Namun, terdapat ekor panjang ke kanan, menandakan sejumlah kecil negara yang mengalami pengangguran ekstrem (>15%).
- Ini membantu kita memahami "normalitas" vs "ekstrem" dalam konteks global.

C. Scatter Plot (2010 Vs. 2020) - Apakah pengangguran memburuk atau membaik dalam satu dekade?

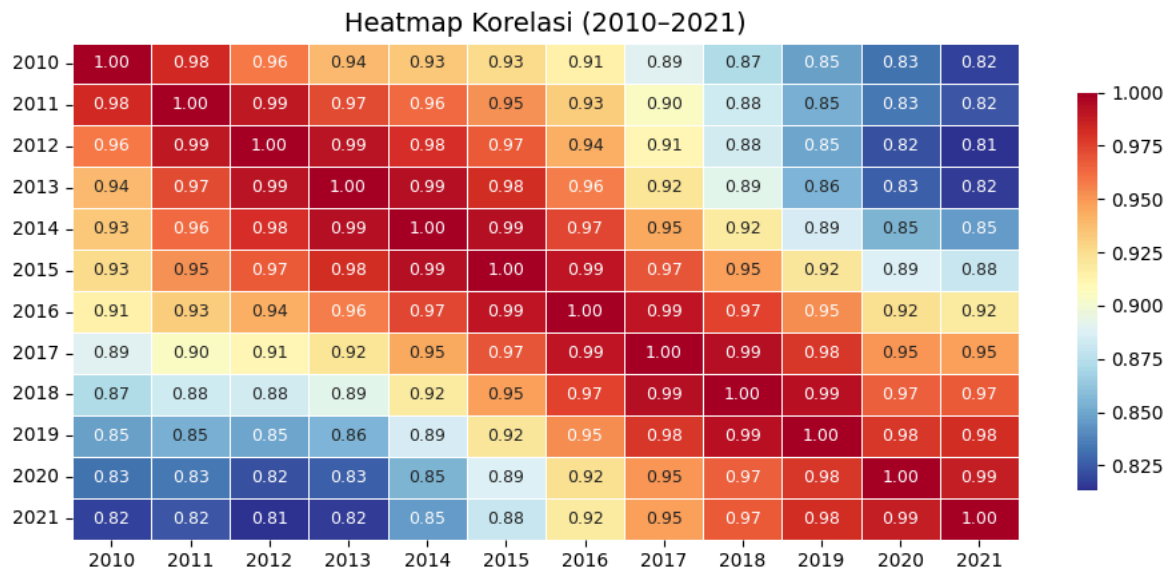


Gambar 3.3 Scatter Plot

Setiap titik mewakili satu negara. Perbandingan antara 2010 dan 2020 mengungkap perubahan selama 10 tahun, termasuk masa pandemi.

- Negara di atas garis merah: pengangguran naik (misalnya Amerika Serikat, yang melonjak dari ~10% ke ~8% — sebenarnya turun, tapi banyak negara Eropa mengalami kenaikan ringan).
- Negara di bawah garis: pengangguran turun (misalnya Spanyol, yang mulai membaik setelah krisis 2010-an).
- Grafik ini menunjukkan bahwa perubahan tidak seragam: beberapa negara pulih, yang lain terpuruk lebih dalam.

D. Heatmap Korelasi (2010 - 2021) - Apakah pola pengangguran stabil dari tahun ke tahun?



Gambar 3.4 Heatmap Korelasi

Heatmap ini menunjukkan seberapa mirip struktur pengangguran antar tahun.

- Nilai korelasi sangat tinggi (0.9–1.0, warna merah pekat) menandakan bahwa peringkat negara hampir tidak berubah:
 - Negara yang tinggi di 2010, tetap tinggi di 2021.
 - Negara yang rendah, tetap rendah.
- Ini mengungkap bahwa pengangguran bersifat struktural — bukan hanya fluktuasi sementara, tapi akar masalah ekonomi, kebijakan, atau demografi yang mendalam dan persisten.