


# Tugas Kelompok ke-1

## Week 2

### Group 2

	<b>ALI ZAENAL ABIDIN ALAYDRUS</b> 2502120462 Computer Science		<b>MUHAMMAD RIZKY FADLI</b> 2502121175 Computer Science
	<b>FRANS SEBASTIAN</b> 2502121162 Computer Science		<b>IZAZ RAIHAN</b> 2502120563 Computer Science
	<b>FADHIL ADRIAN</b> 2502121004 Computer Science		

Soal.

(LO 01; 100%; Topik: Visualizing Data)

Ketepatan waktu kelulusan seorang mahasiswa dipengaruhi oleh beberapa fitur yaitu: jenis kelamin, status mahasiswa, status pernikahan IPK semester 1-6. Tabel berikut adalah kumpulan data yang menunjukkan status ketepatan waktu kelulusan mahasiswa.

NO	JENIS KELAMIN	STATUS MAHASISWA	STATUS PERNIKAHAN	IPK SEMESTER 1-6	STATUS KELULUSAN
1	LAKI - LAKI	MAHASISWA	BELUM	3,17	TEPAT
2	LAKI - LAKI	BEKERJA	BELUM	3,30	TEPAT
3	PEREMPUAN	MAHASISWA	BELUM	3,01	TEPAT
4	PEREMPUAN	MAHASISWA	MENIKAH	3,25	TEPAT
5	LAKI - LAKI	BEKERJA	MENIKAH	3,20	TEPAT
6	LAKI - LAKI	BEKERJA	MENIKAH	2,50	TERLAMBAT
7	PEREMPUAN	BEKERJA	MENIKAH	3,00	TERLAMBAT
8	PEREMPUAN	BEKERJA	BELUM	2,70	TERLAMBAT
9	LAKI - LAKI	BEKERJA	BELUM	2,40	TERLAMBAT
10	PEREMPUAN	MAHASISWA	MENIKAH	2,50	TERLAMBAT
11	PEREMPUAN	MAHASISWA	BELUM	2,50	TERLAMBAT
12	PEREMPUAN	MAHASISWA	BELUM	3,50	TEPAT
13	LAKI - LAKI	BEKERJA	MENIKAH	3,30	TEPAT
14	LAKI - LAKI	MAHASISWA	MENIKAH	3,25	TEPAT
15	LAKI - LAKI	MAHASISWA	BELUM	2,30	TERLAMBAT

Berdasarkan data tabel, dekan ingin melihat data menggunakan visualisasi data menggunakan python sehingga memberikan informasi lebih jelas. Oleh karena itu, buatlah visualisasi data:

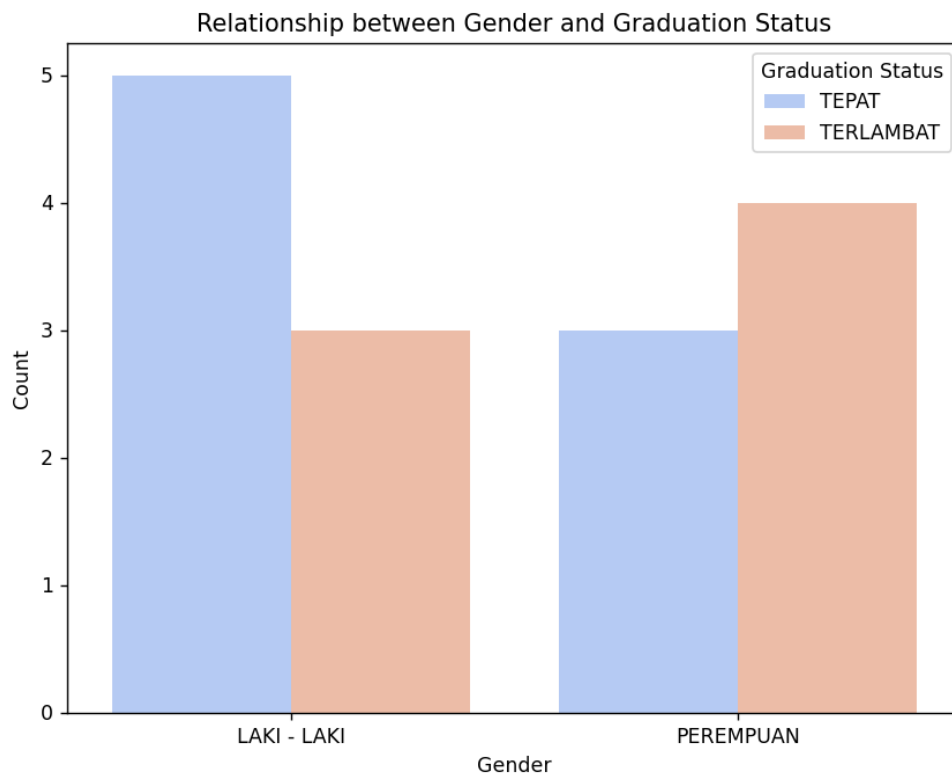
Memasukkan data ke pandas:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
df = pd.DataFrame(
    [
        [1, "LAKI-LAKI", "MAHASISWA", "BELUM", 3.17, "TEPAT"],
        [2, "LAKI-LAKI", "BEKERJA", "BELUM", 3.30, "TEPAT"],
        [3, "PEREMPUAN", "MAHASISWA", "BELUM", 3.01, "TEPAT"],
        [4, "PEREMPUAN", "MAHASISWA", "MENIKAH", 3.25, "TEPAT"],
        [5, "LAKI-LAKI", "BEKERJA", "MENIKAH", 3.20, "TEPAT"],
        [6, "LAKI-LAKI", "BEKERJA", "MENIKAH", 2.50, "TERLAMBAT"],
        [7, "PEREMPUAN", "BEKERJA", "MENIKAH", 3.00, "TERLAMBAT"],
        [8, "PEREMPUAN", "BEKERJA", "BELUM", 2.70, "TERLAMBAT"],
        [9, "LAKI-LAKI", "BEKERJA", "BELUM", 2.40, "TERLAMBAT"],
        [10, "PEREMPUAN", "MAHASISWA", "MENIKAH", 2.50, "TERLAMBAT"],
        [11, "PEREMPUAN", "MAHASISWA", "BELUM", 2.50, "TERLAMBAT"],
        [12, "PEREMPUAN", "MAHASISWA", "BELUM", 3.50, "TEPAT"],
        [13, "LAKI-LAKI", "BEKERJA", "MENIKAH", 3.30, "TEPAT"],
        [14, "LAKI-LAKI", "MAHASISWA", "MENIKAH", 3.25, "TEPAT"],
        [15, "LAKI-LAKI", "MAHASISWA", "BELUM", 2.30, "TERLAMBAT"],
    ],
    columns=[
        "NO",
        "JENIS KELAMIN",
        "STATUS MAHASISWA",
        "STATUS PERNIKAHAN",
        "IPK SEMESTER 1-6",
        "STATUS KELULUSAN",
    ],
)
df
```

List Visualisasi Data:

a. **Hubungan antara jenis kelamin dan status kelulusan.**

```
sns.countplot(data=df, x="JENIS KELAMIN", hue="STATUS KELULUSAN",
palette='coolwarm')
plt.xlabel('Gender')
plt.legend(title='Graduation Status')
plt.title ("Relationship between gender and graduation status")
```



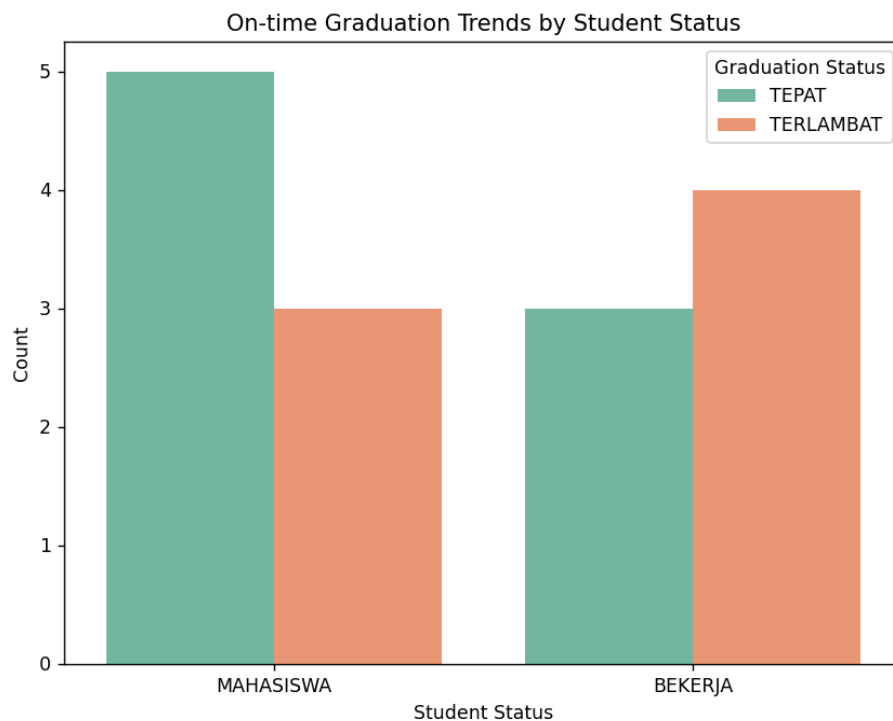
Pada grafik diatas, terlihat perbandingan antara mahasiswa **laki-laki** dan **perempuan** yang lulus tepat waktu maupun terlambat.

- **Laki-laki:** Dari data, terlihat bahwa lebih banyak laki-laki yang lulus tepat waktu dibanding yang terlambat. Jumlah mahasiswa laki-laki yang lulus tepat waktu cukup signifikan dibandingkan yang terlambat.

- **Perempuan:** Di sisi lain, meskipun jumlah perempuan yang lulus tepat waktu ada, tetapi ada proporsi yang lebih tinggi dari perempuan yang terlambat lulus.

**b. Hubungan antara status mahasiswa dan IPK semester 1-6**

```
sns.countplot(data=df, hue="STATUS KELULUSAN", x="STATUS MAHASISWA",  
palette='viridis')  
plt.xlabel('Student Status')  
plt.title("Relationship between Student Status and GPA")
```

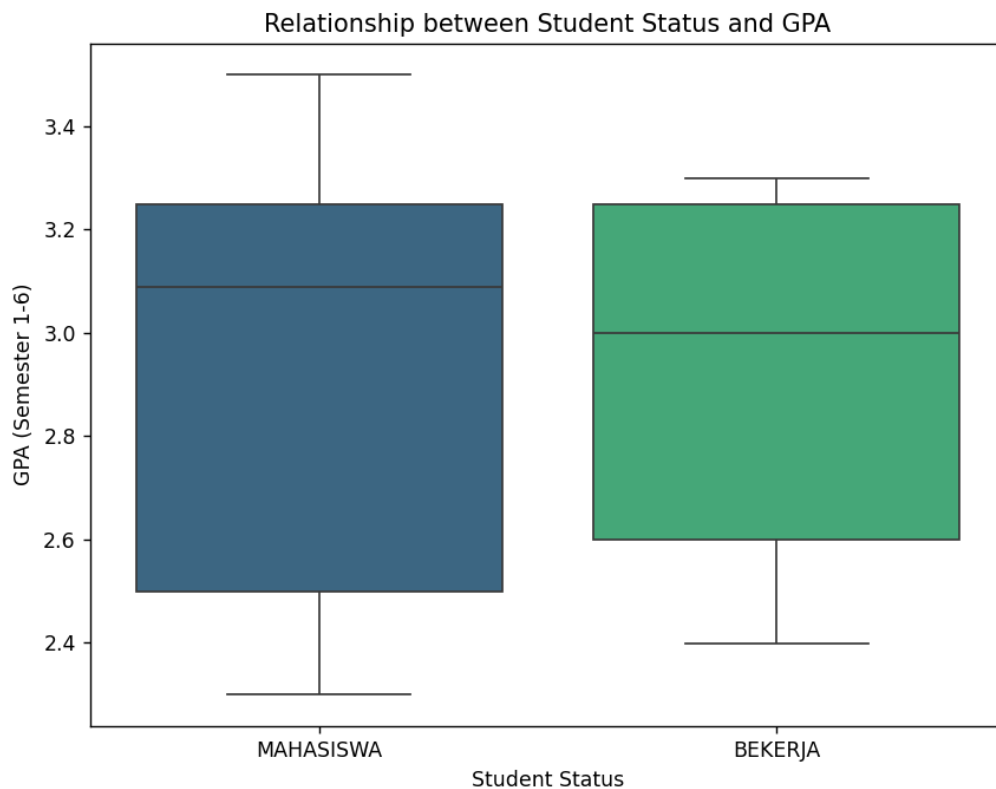


Grafik boxplot diatas memperlihatkan distribusi IPK berdasarkan status mahasiswa, baik mahasiswa yang **bekerja** maupun **mahasiswa reguler**.

- **Mahasiswa:** Distribusi IPK untuk mahasiswa yang tidak bekerja cenderung lebih tinggi dan lebih merata. Nilai IPK lebih sering berkumpul di kisaran 3.0 ke atas. Terdapat beberapa outlier, tetapi mayoritas berada di area IPK tinggi.
- **Bekerja:** Mahasiswa yang bekerja memiliki rentang IPK yang lebih lebar, dengan banyaknya nilai IPK yang lebih rendah. Beberapa di antaranya bahkan mencapai nilai IPK di bawah 2.5, menunjukkan tantangan yang mungkin dihadapi mahasiswa yang bekerja dalam mempertahankan IPK tinggi.

c. **Tren ketepatan waktu lulus di setiap status mahasiswa**

```
sns.boxplot(data=df, x="STATUS MAHASISWA", y="IPK SEMESTER 1-6")
plt.xlabel('Student Status')
plt.title ("On-time Graduation Trends by Student Status")
```



Grafik diatas menggambarkan hubungan antara status mahasiswa (bekerja atau mahasiswa reguler) dengan ketepatan waktu kelulusan.

- **Mahasiswa reguler:** Mayoritas mahasiswa reguler lulus tepat waktu. Jumlah mereka yang terlambat lulus relatif lebih kecil dibandingkan yang lulus tepat waktu. Hal ini mengindikasikan bahwa status sebagai mahasiswa reguler (yang tidak bekerja) dapat meningkatkan peluang lulus tepat waktu.
- **Mahasiswa yang bekerja:** Untuk mahasiswa yang bekerja, tren yang terlihat cukup seimbang antara yang lulus tepat waktu dan yang terlambat lulus. Ini menunjukkan bahwa bekerja selama masa studi bisa jadi merupakan faktor yang memengaruhi keterlambatan kelulusan.

Link python code: <https://binus-course-project.vercel.app/assets/SEM6/DS/TK1/TK-1.py>

- a. Untuk menggambarkan hubungan antara jenis kelamin dan status kelulusan.

Jawab:

**Hubungan antara jenis kelamin dan status kelulusan:** Visualisasi pertama menunjukkan jumlah mahasiswa laki-laki dan perempuan yang lulus tepat waktu atau terlambat. Terlihat bahwa jumlah mahasiswa laki-laki yang lulus tepat waktu lebih banyak dibandingkan perempuan, tetapi jumlah mahasiswa perempuan yang terlambat lulus juga relatif lebih rendah dibandingkan laki-laki.

- b. Untuk menggambarkan hubungan antara status mahasiswa dan IPK semester 1-6.

Jawab:

**Hubungan antara status mahasiswa dan IPK semester 1-6:** Visualisasi kedua menggambarkan distribusi IPK mahasiswa yang berstatus "Mahasiswa" dan "Bekerja". Dari visualisasi tersebut dapat dilihat bahwa mahasiswa dengan status Mahasiswa (penuh waktu) lebih cenderung lulus tepat waktu dibandingkan dengan mahasiswa yang berstatus Bekerja.

- c. Untuk menggambarkan tren ketepatan waktu lulus di setiap status mahasiswa.

Jawab:

**Tren ketepatan waktu lulus di setiap status mahasiswa:** Visualisasi ketiga menunjukkan bahwa mahasiswa penuh waktu memiliki kecenderungan lulus tepat waktu

lebih tinggi daripada mahasiswa yang bekerja. Mahasiswa bekerja memiliki persentase kelulusan terlambat yang lebih tinggi.