

**ANALISIS DATA PENJUALAN BUKU MENGGUNAKAN PYTHON:  
STUDI KASUS PENJUALAN BUKU TAHUN 2024**

**Disusun untuk memenuhi UAS mata kuliah  
Pengkodean Dan Pemograman**

Dosen Pengampu:  
Dr. Totok Dewayanto, S.E., M.Si., Akt.



Oleh:  
**Yenny Clara Manihuruk**  
**12030122120036**

**Program Studi S-1 Akuntansi  
Fakultas Ekonomika dan Bisnis  
Universitas Diponegoro  
2024**

## **KATA PENGANTAR**

Puji Syukur atas kehadiran Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan proposal ini guna memenuhi tugas untuk mata kuliah Pengkodean dan Pemrograman, dengan judul “Analisis data penjualan buku menggunakan Python: Studi kasus penjualan buku tahun 2024”.

Saya menyadari bahwa dalam penulisan proposal ini tidak terlepas dari bantuan banyak pihak yang dengan tulus memberikan arahan, saran, dan kritikan sehingga proposal ini dapat terselesaikan.

Saya menyadari sepenuhnya bahwa proposal ini masih jauh dari kata sempurna disebabkan terbatasnya pengalaman dan pengetahuan yang kami miliki. Oleh karena itu, saya mengharapkan segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Akhirnya, saya berharap semoga makalah ini dapat memberikan manfaat bagi perkembangan dunia pendidikan terutama pengembangan teknologi pada Akuntansi.

Semarang, 19 Juni 2024

Penulis

## DAFTAR ISI

COVER.....	i
KATA PENGANTAR.....	<a href="#">ii</a>
DAFTAR ISI.....	iii
BAB I: PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	1
1.3 Tujuan Penelitian.....	1
1.4 Manfaat Kerja Praktek .....	2
BAB II: METODE ANALISIS.....	3
2.1 Metode Analisis.....	3
BAB III: TAHAPAN ANALISIS.....	6
3.1 Pengumpulan Data.....	6
3.2 Kode Python Tahap Awal.....	14
3.3 Pembersihan Data.....	17
3.4 Visualisasi Analisis Data.....	22
BAB IV: KESIMPULAN DAN SARAN.....	26
4.1 Kesimpulan.....	26
4.2 Saran.....	27
DAFTAR PUSTAKA.....	28

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Dengan semakin berkembangnya ilmu pengetahuan dan teknologi informasi, maka mendorong keinginan masyarakat untuk mendapatkan informasi dengan mudah, cepat, dan akurat cenderung meningkat, sehingga menuntut adanya kemampuan yang dapat memenuhi segala kebutuhan tersebut. Di era digital saat ini, pemahaman mendalam tentang data penjualan buku adalah kunci untuk memahami dan merespon perubahan konsumen serta tetap relevan dan kompetitif dalam industri penerbitan yang dinamis ini. Menganalisis data penjualan buku menggunakan berbagai grafik memberikan gambaran menyeluruh tentang tren dan pola penjualan dari waktu ke waktu.

Data penjualan Januari 2024 menunjukkan pola yang menarik dalam beberapa minggu pertama. Misalnya, Buku A mencatat penjualan tertinggi sebanyak 62 eksemplar pada 17 Januari sehingga menghasilkan 620.000 penjualan. Sementara itu, Buku B yang mengalami angka penjualan berfluktuasi, tetap konsisten menghasilkan pendapatan yang signifikan selama periode tersebut. Menggunakan diagram garis untuk melacak tren harian, diagram lingkaran untuk menunjukkan distribusi penjualan buku, dan diagram area untuk menggambarkan penjualan kumulatif, analisis ini memungkinkan studi yang lebih baik tentang dinamika pasar buku. Informasi ini juga memberi perusahaan penerbitan kemampuan untuk mengevaluasi kinerja setiap judul buku satu per satu, serta memahami preferensi konsumen terhadap genre atau topik tertentu.

Dengan bantuan data penjualan yang ada, perusahaan dapat mengoptimalkan strategi pemasarannya, meningkatkan efisiensi manajemen inventaris, dan merespons perubahan permintaan pasar dengan lebih cepat dan efisien. Oleh karena itu, menganalisis data penjualan buku tidak hanya memberikan pemahaman yang lebih mendalam tentang perilaku konsumen dan pasar, tetapi juga merupakan kunci untuk mempertahankan daya saing dalam industri yang terus berubah dan memaksimalkan potensi penjualan di tengah dinamika yang ada.

### **1.2 Rumusan Masalah**

1. Bagaimana tren dan pola penjualan harian tiap buku (A, B, C, D) di bulan Januari 2024?
2. Metode apa yang digunakan untuk menganalisis data penjualan dan bagaimana relevansinya?
3. Apa interpretasi dari grafik hasil analisis data penjualan buku?
4. Bagaimana perkembangan akumulasi penjualan buku selama Januari 2024 dan kaitannya dengan strategi penjualan?
5. Bagaimana cara pembersihan data dengan Python untuk analisis yang akurat?

### **1.3 Tujuan Penelitian**

Tujuan umum dari penelitian ini adalah untuk memberikan gambaran menyeluruh tentang dinamika penjualan buku pada bulan Januari 2024. Fokus utamanya adalah mengidentifikasi pola penjualan harian, memperkirakan pangsa relatif setiap buku dalam total penjualan bulanan dan pengukurannya. Kinerja individu setiap judul buku berdasarkan penjualan dan

pendapatan yang dihasilkan. Tujuan umum ini memberikan pemahaman mendalam kepada penerbit tentang cara meningkatkan strategi pemasaran, manajemen inventaris, dan cara merespons dinamika pasar buku.

Tujuan Khusus penelitian ini adalah untuk menganalisis secara menyeluruh pola penjualan harian dan mengidentifikasi tren penting dan potensi variasi dalam penjualan buku. Dengan mengevaluasi kontribusi relatif setiap buku terhadap total penjualan bulanan, penelitian ini juga bertujuan untuk mengidentifikasi judul buku yang memiliki dampak strategis terhadap total pendapatan sebuah perusahaan penerbitan. Oleh karena itu, penelitian ini tidak hanya memberikan pemahaman yang lebih mendalam mengenai perilaku konsumen dan pasar buku, namun juga menciptakan landasan yang kuat untuk pengambilan keputusan yang lebih tepat dalam menghadapi tantangan dan peluang dalam industri penerbitan yang dinamis.

#### **1.4 Manfaat Kerja Praktek**

Analisis data penjualan buku memiliki manfaat yang besar untuk berbagai kepentingan. Secara umum, analisis ini membantu kita memahami pola dan kecenderungan penjualan dari waktu ke waktu, yang dapat digunakan untuk merencanakan strategi bisnis yang lebih baik. Kita dapat mengidentifikasi jenis buku yang paling diminati oleh pelanggan, serta mengetahui apakah ada buku-buku tertentu yang kurang laku atau perlu promosi lebih intensif. Dengan pemahaman ini, toko buku dapat mengoptimalkan stok mereka, memastikan bahwa buku-buku yang diminati selalu tersedia, dan mengurangi risiko kelebihan atau kekurangan persediaan.

Secara khusus, analisis juga membantu dalam mengukur kinerja masing-masing produk (buku), memungkinkan kita untuk melihat mana yang memberikan pendapatan tertinggi atau penjualan terbanyak. Selain itu, kita dapat melihat distribusi frekuensi penjualan, yang berguna untuk mengetahui apakah ada tren tertentu dalam preferensi pelanggan. Dengan menggunakan teknik visual seperti histogram, scatter plot, dan box plot, kita dapat mengidentifikasi anomali atau pola yang tidak biasa dalam data, seperti buku-buku dengan pendapatan tinggi namun penjualan rendah. Kesimpulannya, analisis data penjualan buku tidak hanya memberikan wawasan tentang kinerja produk dan preferensi pelanggan, tetapi juga menjadi alat penting untuk membuat keputusan bisnis yang lebih cerdas dan strategis.

## BAB II

### METODE ANALISIS

#### 2.1 Metode Analisis

Metode analisis dari kode Python yang diberikan menggunakan berbagai teknik visualisasi untuk mengeksplorasi dataset penjualan buku. Berikut adalah penjelasan tentang langkah-langkah analisis dan visualisasi yang dilakukan:

##### 1. Membaca dan Memproses Dataset:

- o Dataset penjualan buku dibaca menggunakan `pd.read_csv()`.
- o Kolom 'Tanggal' diubah menjadi tipe data `datetime` untuk memungkinkan analisis berdasarkan waktu.

##### 2. Pie Chart:

- o **Tujuan:** Menampilkan persentase jumlah buku terjual per jenis buku.
- o **Pustaka:** Matplotlib untuk plot pie chart, menggunakan `plt.pie`.
- o **Visualisasi:** Menggunakan warna dari palet seaborn 'pastel' untuk mewarnai setiap potongan pie.
- o **Kode:**

```
python
Copy code
plt.figure(figsize=(8, 8))
plt.pie(jumlah_per_buku, labels=jumlah_per_buku.index,
        autopct='%1.1f%%', colors=sns.color_palette('pastel'))
plt.title('Persentase Jumlah Buku Terjual per Jenis Buku')
plt.show()
```

##### 3. Scatter Plot:

- o **Tujuan:** Mengeksplorasi hubungan antara jumlah buku yang terjual dengan pendapatan yang dihasilkan.
- o **Pustaka:** Seaborn untuk plot scatter, menggunakan `sns.scatterplot`.
- o **Visualisasi:** Menggunakan warna dari palet seaborn 'viridis' untuk menunjukkan jenis buku.
- o **Kode:**

```
python
Copy code
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Jumlah', y='Pendapatan', hue='Nama_Buku',
               palette='viridis', s=100, data=data)
plt.title('Scatter Plot Jumlah Buku Terjual vs. Pendapatan')
plt.xlabel('Jumlah Buku Terjual')
plt.ylabel('Pendapatan')
plt.legend(title='Nama Buku')
plt.show()
```

##### 4. Line Plot:

- **Tujuan:** Memvisualisasikan tren penjualan buku dari waktu ke waktu.
- **Pustaka:** Matplotlib untuk plot garis, menggunakan `plt.plot`.
- **Visualisasi:** Setiap jenis buku diplot dalam satu garis dengan marker bulat dan garis solid.
- **Kode:**

```
python
Copy code
plt.figure(figsize=(12, 6))
for buku in data['Nama_Buku'].unique():
    buku_data = data[data['Nama_Buku'] == buku]
    plt.plot(buku_data['Tanggal'], buku_data['Jumlah'],
             marker='o', linestyle='-', label=buku)
plt.title('Tren Penjualan Buku')
plt.xlabel('Tanggal')
plt.ylabel('Jumlah Buku Terjual')
plt.legend(title='Nama Buku')
plt.grid(True)
plt.show()
```

## 5. Histogram:

- **Tujuan:** Memvisualisasikan distribusi jumlah buku terjual.
- **Pustaka:** Seaborn untuk histogram, menggunakan `sns.histplot`.
- **Visualisasi:** Menggunakan warna 'skyblue' dan menambahkan Kernel Density Estimate (KDE) untuk distribusi yang lebih halus.
- **Kode:**

```
python
Copy code
plt.figure(figsize=(10, 6))
sns.histplot(data['Jumlah'], bins=20, kde=True, color='skyblue')
plt.title('Distribusi Jumlah Buku Terjual')
plt.xlabel('Jumlah Buku')
plt.ylabel('Frekuensi')
plt.show()
```

## 6. Box Plot:

- **Tujuan:** Memvisualisasikan distribusi pendapatan per jenis buku.
- **Pustaka:** Seaborn untuk box plot, menggunakan `sns.boxplot`.
- **Visualisasi:** Menggunakan palet 'Set2' untuk mewarnai box plot per jenis buku.
- **Kode:**

```
python
Copy code
plt.figure(figsize=(10, 6))
sns.boxplot(x='Nama_Buku', y='Pendapatan', palette='Set2',
            data=data)
plt.title('Distribusi Pendapatan per Jenis Buku')
plt.xlabel('Nama Buku')
plt.ylabel('Pendapatan')
plt.show()
```

## 7. Bar Plot:

- **Tujuan:** Menampilkan total penjualan per jenis buku.
- **Pustaka:** Seaborn untuk bar plot, menggunakan `sns.barplot`.
- **Visualisasi:** Menggunakan palet 'Set1' untuk mewarnai bar plot.
- **Kode:**

```
python
Copy code
plt.figure(figsize=(10, 6))
sns.barplot(x='Nama_Buku', y='Jumlah', palette='Set1',
data=data.groupby('Nama_Buku')['Jumlah'].sum().reset_index())
plt.title('Total Penjualan per Jenis Buku')
plt.xlabel('Nama Buku')
plt.ylabel('Jumlah Terjual')
plt.show()
```

## 8. 3D Scatter Plot:

- **Tujuan:** Memvisualisasikan hubungan antara jumlah buku terjual, pendapatan, dan tanggal.
- **Pustaka:** Matplotlib untuk 3D scatter plot, menggunakan `ax.scatter` dari `Axes3D`.
- **Visualisasi:** Menggunakan cmap 'coolwarm' untuk mewarnai titik berdasarkan jumlah buku yang terjual.
- **Kode:**

```
python
Copy code
fig = plt.figure(figsize=(10, 8))
ax = fig.add_subplot(111, projection='3d')
scatter = ax.scatter(data['Tanggal'].map(pd.Timestamp.toordinal),
data['Jumlah'], data['Pendapatan'], c=data['Jumlah'],
cmap='coolwarm', s=100)
ax.set_title('3D Scatter Plot Jumlah Buku Terjual vs. Pendapatan
vs. Tanggal')
ax.set_xlabel('Tanggal')
ax.set_ylabel('Jumlah Buku Terjual')
ax.set_zlabel('Pendapatan')
plt.colorbar(scatter, ax=ax, label='Jumlah Buku Terjual')
plt.show()
```



## BAB III

### TAHAPAN ANALISIS

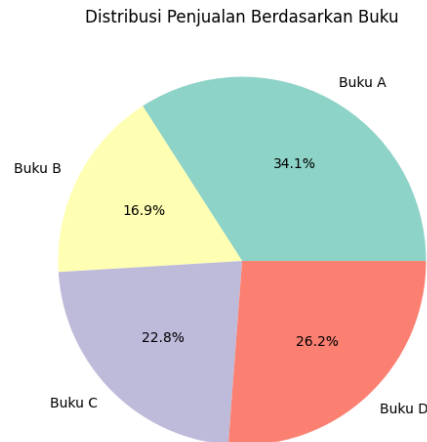
#### 3.1 Pengumpulan Data

Data dalam analisis ini diperoleh dari dataset penjualan buku pada periode Januari 2024. Dataset ini mencakup informasi mengenai tanggal transaksi (Tanggal), nama jenis buku (Nama\_Buku), jumlah buku yang terjual (Jumlah), dan pendapatan yang dihasilkan dari penjualan (Pendapatan). Data ini digunakan untuk menganalisis tren penjualan buku dari waktu ke waktu, distribusi jumlah buku terjual, distribusi pendapatan per jenis buku, serta hubungan antara jumlah buku terjual, pendapatan, dan waktu. Analisis ini bertujuan untuk memahami pola penjualan buku dan faktor-faktor yang mempengaruhinya selama periode yang ditentukan.

Tanggal	Nama_Buku	Jumlah	Pendapatan
2024-01-01	Buku A	50	500000
2024-01-02	Buku B	30	300000
2024-01-03	Buku C	40	400000
2024-01-04	Buku D	45	450000
2024-01-05	Buku A	55	550000
2024-01-06	Buku B	25	250000
2024-01-07	Buku C	35	350000
2024-01-08	Buku D	40	400000
2024-01-09	Buku A	60	600000
2024-01-10	Buku B	28	280000
2024-01-11	Buku C	38	380000
2024-01-12	Buku D	42	420000
2024-01-13	Buku A	58	580000
2024-01-14	Buku B	32	320000
2024-01-15	Buku C	42	420000
2024-01-16	Buku D	48	480000
2024-01-17	Buku A	62	620000
2024-01-18	Buku B	26	260000
2024-01-19	Buku C	36	360000
2024-01-20	Buku D	44	440000

Dari data-data diatas, diolah dengan kode python menghasilkan visualisasi sebagai berikut:

**(Gambar 1)**

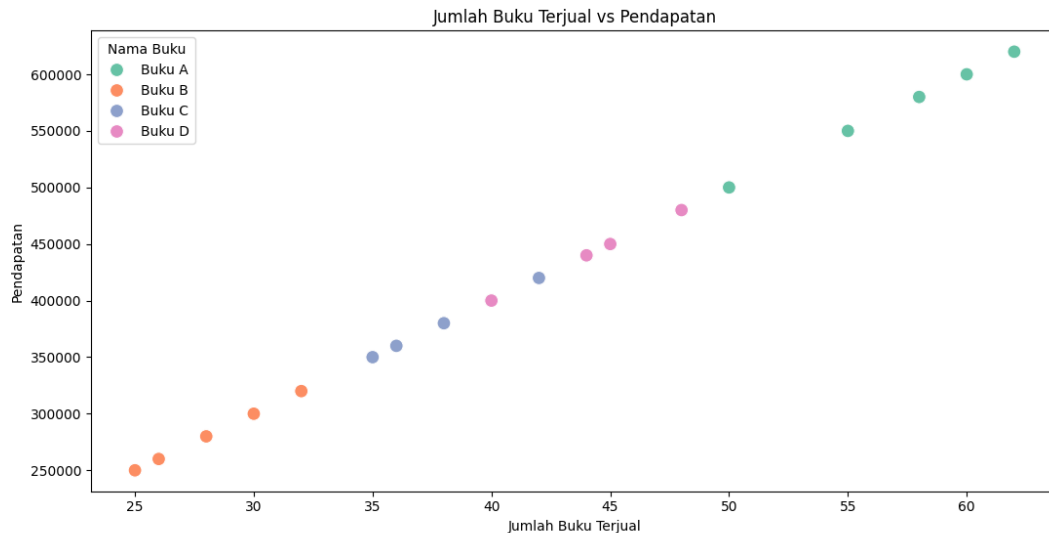


### Interpretasi Grafik: Distribusi Penjualan Berdasarkan Buku

1. **Penjualan Tertinggi:** Buku A memiliki penjualan tertinggi dengan persentase 34.1%. Ini menunjukkan bahwa Buku A adalah buku yang paling banyak terjual dibandingkan yang lain.
2. **Penjualan Terendah:** Buku B memiliki penjualan terendah dengan persentase 16.9%. Ini menunjukkan bahwa Buku B adalah buku yang paling sedikit terjual.
3. **Penjualan Menengah:**
  - o Buku D memiliki penjualan yang cukup tinggi, yaitu 26.2%, menempatkannya sebagai buku dengan penjualan tertinggi kedua.
  - o Buku C memiliki persentase penjualan sebesar 22.8%, berada di posisi ketiga.
4. **Distribusi Penjualan:**
  - o Mayoritas penjualan berasal dari Buku A dan Buku D, yang secara kolektif menyumbang 60.3% dari total penjualan.
  - o Buku B dan Buku C menyumbang 39.7% sisanya.

Distribusi penjualan buku menunjukkan bahwa Buku A adalah yang paling populer, diikuti oleh Buku D. Buku C dan Buku B memiliki penjualan yang lebih rendah, dengan Buku B sebagai yang terendah. Penjualan terpusat pada Buku A dan Buku D, yang mencerminkan popularitas yang lebih tinggi dibandingkan dengan Buku B dan Buku C.

(Gambar 2)



### Interpretasi Grafik: Jumlah Buku Terjual vs Pendapatan

#### 1. Hubungan Positif Antara Jumlah Buku Terjual dan Pendapatan:

- Grafik menunjukkan bahwa ada hubungan positif antara jumlah buku yang terjual dan pendapatan. Ini berarti semakin banyak buku yang terjual, semakin tinggi pendapatan yang dihasilkan.

#### 2. Distribusi Data Berdasarkan Jenis Buku:

- Buku A** (hijau): Buku ini memiliki jumlah penjualan tertinggi, mencapai sekitar 60 buku terjual dengan pendapatan sekitar 600,000.
- Buku B** (oranye): Buku ini memiliki penjualan terendah, dengan jumlah buku terjual sekitar 25 hingga 35 dan pendapatan antara 250,000 hingga 350,000.
- Buku C** (biru): Penjualan buku ini berada di kisaran menengah, dengan jumlah buku terjual sekitar 30 hingga 45 dan pendapatan antara 300,000 hingga 450,000.
- Buku D** (merah muda): Buku ini juga berada di kisaran menengah dengan jumlah buku terjual sekitar 40 hingga 50 dan pendapatan antara 350,000 hingga 500,000.

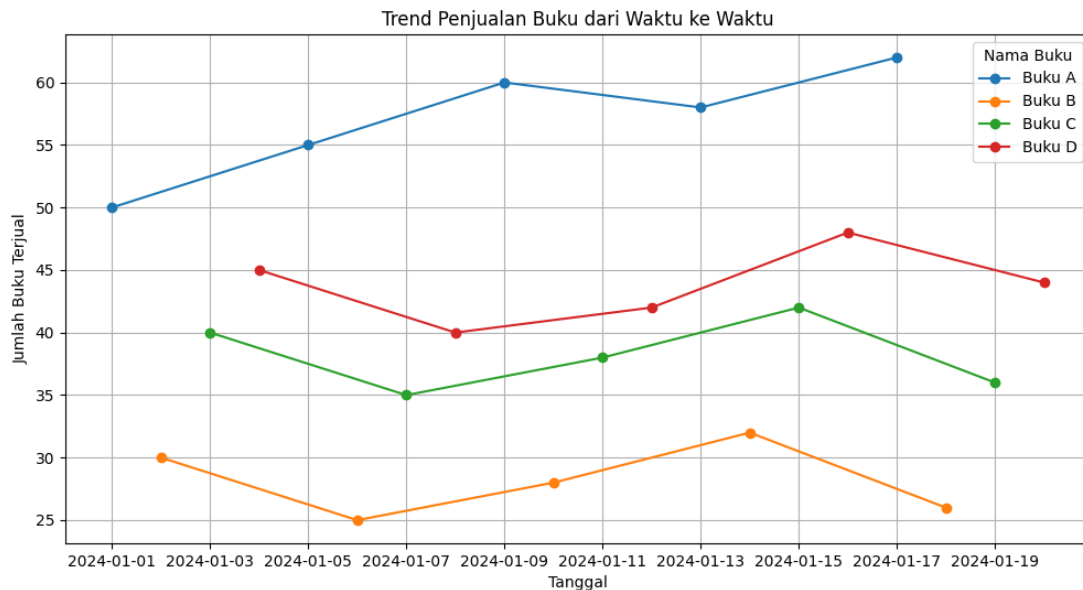
#### 3. Pola yang Ditemukan:

- Buku A menunjukkan performa yang sangat baik dengan pendapatan yang meningkat secara signifikan seiring dengan peningkatan jumlah penjualan.
- Buku B, meskipun memiliki penjualan paling rendah, masih menunjukkan peningkatan pendapatan yang stabil seiring dengan jumlah buku yang terjual.
- Buku C dan Buku D menunjukkan performa yang mirip di kisaran menengah, dengan Buku D sedikit lebih tinggi dalam hal pendapatan untuk jumlah buku terjual yang sama.

Grafik ini memberikan gambaran yang jelas tentang bagaimana penjualan setiap jenis buku berkontribusi terhadap pendapatan. Buku A adalah yang paling laris dan menghasilkan pendapatan tertinggi, sedangkan Buku B memiliki penjualan dan pendapatan yang paling rendah. Buku C dan

Buku D berada di posisi menengah. Secara keseluruhan, grafik ini menunjukkan bahwa peningkatan jumlah buku terjual berbanding lurus dengan peningkatan pendapatan.

(Gambar 3)



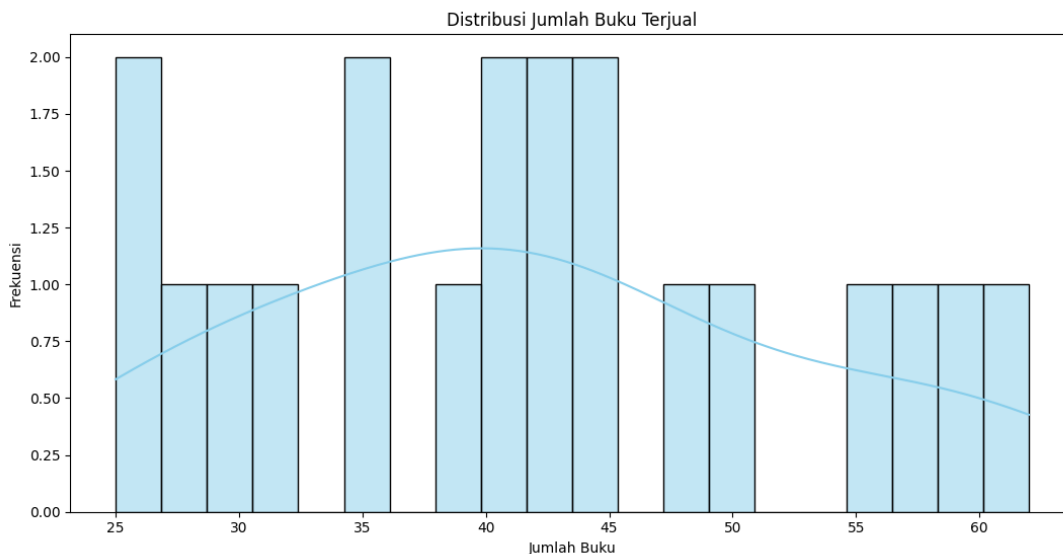
### Interpretasi Grafik: Trend Penjualan Buku dari Waktu ke Waktu

1. **Buku A (Biru):**
  - **Tren Umum:** Penjualan buku A cenderung meningkat selama periode waktu tersebut.
  - **Detail Tren:** Dimulai dari sekitar 50 buku pada 1 Januari 2024, mencapai puncak sekitar 60 buku pada 7 Januari 2024, sedikit menurun pada 11 Januari 2024, dan kembali meningkat hingga sekitar 58 buku pada 19 Januari 2024.
2. **Buku B (Oranye):**
  - **Tren Umum:** Penjualan buku B menunjukkan fluktuasi yang signifikan tetapi cenderung menurun secara keseluruhan.
  - **Detail Tren:** Dimulai dari sekitar 30 buku pada 1 Januari 2024, menurun ke titik terendah sekitar 25 buku pada 7 Januari 2024, sedikit naik kembali, dan mencapai sekitar 27 buku pada 19 Januari 2024.
3. **Buku C (Hijau):**
  - **Tren Umum:** Penjualan buku C menunjukkan peningkatan yang stabil secara keseluruhan.
  - **Detail Tren:** Dimulai dari sekitar 35 buku pada 1 Januari 2024, naik secara bertahap dengan sedikit fluktuasi, mencapai sekitar 40 buku pada 19 Januari 2024.
4. **Buku D (Merah):**
  - **Tren Umum:** Penjualan buku D cenderung menurun selama periode waktu tersebut.

- **Detail Tren:** Dimulai dari sekitar 45 buku pada 1 Januari 2024, mengalami penurunan ke sekitar 42 buku pada 5 Januari 2024, kemudian sedikit naik kembali pada 11 Januari 2024, sebelum menurun lagi ke sekitar 44 buku pada 19 Januari 2024.

Secara keseluruhan, grafik ini memberikan gambaran tentang bagaimana penjualan masing-masing buku berubah dari waktu ke waktu, dengan Buku A dan Buku C menunjukkan tren peningkatan, sedangkan Buku B dan Buku D menunjukkan tren penurunan.

**(Gambar 4)**



### Interpretasi Grafik: Distribusi Jumlah Buku Terjual

#### 1. Frekuensi Penjualan:

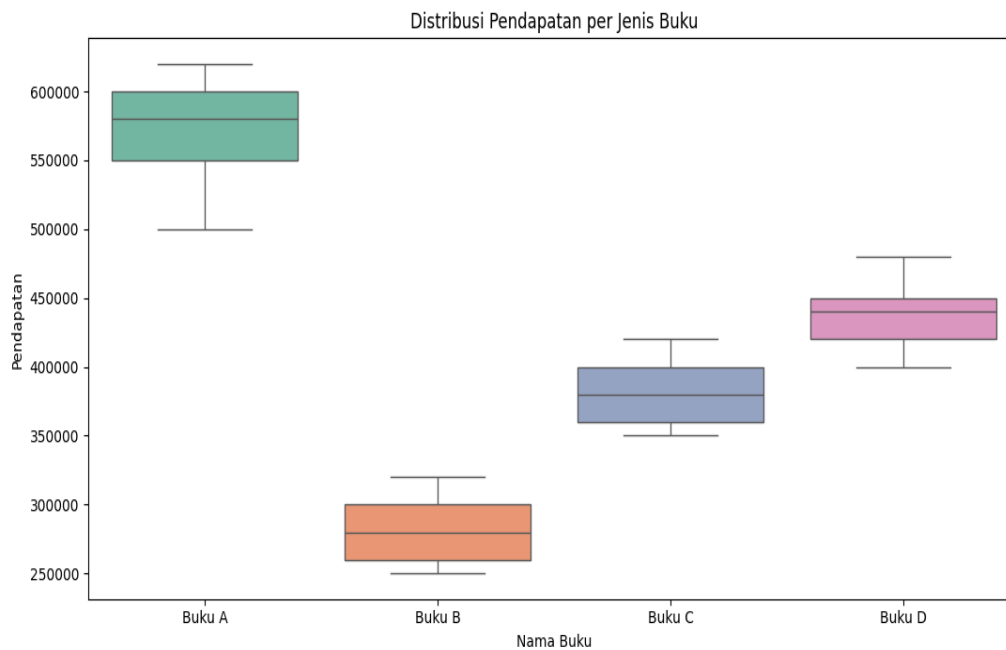
- Rentang jumlah buku yang terjual berkisar dari 25 hingga 60.
- Ada beberapa puncak frekuensi pada jumlah buku tertentu.
- Jumlah buku terjual paling sering pada kisaran 25, 35, 40, 45, dan 55 dengan frekuensi 2.
- Jumlah buku terjual dengan frekuensi terendah berada pada kisaran 30 dan 50 dengan frekuensi 1.

#### 2. Pola Distribusi:

- Distribusi penjualan buku tidak merata.
- Terdapat beberapa titik penjualan buku yang tinggi pada jumlah tertentu, menunjukkan ada preferensi atau kejadian yang mendorong penjualan pada jumlah tersebut.
- Kurva yang ditambahkan menunjukkan kecenderungan umum dari distribusi penjualan buku, yang memuncak sekitar jumlah buku 40 dan menurun setelahnya.

Grafik ini menunjukkan distribusi jumlah buku yang terjual dalam rentang 25 hingga 60 buku. Frekuensi penjualan bervariasi dengan beberapa jumlah buku tertentu yang lebih sering terjual. Puncak penjualan terlihat pada 25, 35, 40, 45, dan 55 buku, sementara penjualan pada 30 dan 50 buku lebih jarang. Kurva distribusi menunjukkan pola umum bahwa penjualan buku cenderung memuncak pada sekitar 40 buku sebelum menurun. Ini memberikan wawasan mengenai jumlah buku yang lebih sering terjual, yang dapat digunakan untuk perencanaan stok atau strategi penjualan.

(Gambar 5)



### Interpretasi Grafik: Distribusi Pendapatan per Jenis Buku

1. **Pendapatan Buku A:**
  - Buku A memiliki pendapatan tertinggi.
  - Median pendapatan untuk Buku A sekitar 575,000.
  - Rentang pendapatan berkisar antara 500,000 hingga 625,000.
2. **Pendapatan Buku B:**
  - Buku B memiliki pendapatan terendah.
  - Median pendapatan untuk Buku B sekitar 280,000.
  - Rentang pendapatan berkisar antara 250,000 hingga 320,000.
3. **Pendapatan Buku C:**
  - Buku C memiliki pendapatan menengah.
  - Median pendapatan untuk Buku C sekitar 380,000.
  - Rentang pendapatan berkisar antara 350,000 hingga 420,000.
4. **Pendapatan Buku D:**
  - Buku D juga memiliki pendapatan menengah.

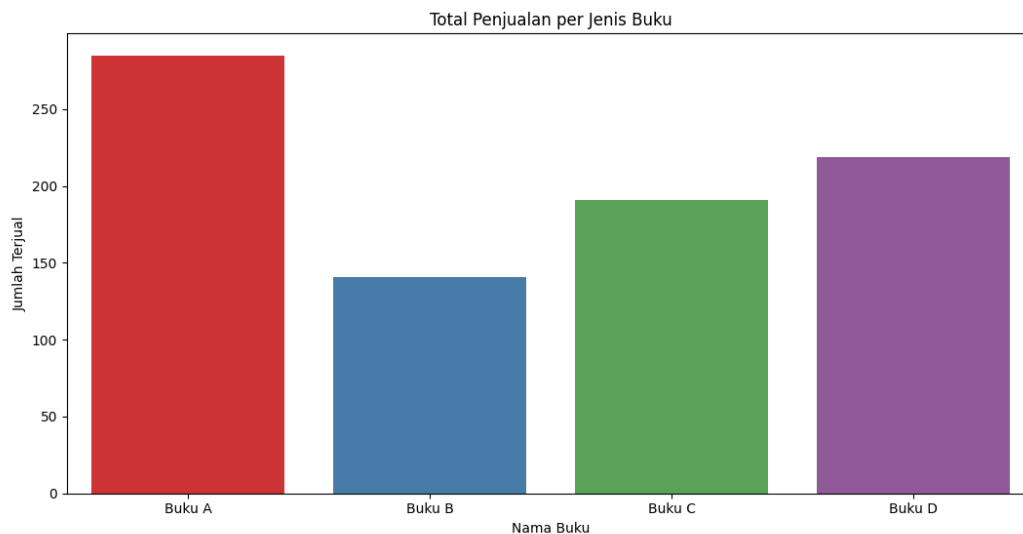
- Median pendapatan untuk Buku D sekitar 475,000.
- Rentang pendapatan berkisar antara 450,000 hingga 500,000.

### Pola yang Ditemukan:

- **Buku A:** Menunjukkan performa yang sangat baik dengan pendapatan tertinggi dan rentang yang lebih sempit, mengindikasikan konsistensi yang tinggi.
- **Buku B:** Memiliki pendapatan terendah dengan rentang yang juga lebih sempit, mengindikasikan konsistensi di tingkat yang lebih rendah.
- **Buku C dan Buku D:** Menunjukkan performa yang serupa di kisaran menengah, dengan Buku D memiliki pendapatan sedikit lebih tinggi daripada Buku C.

Grafik ini menunjukkan distribusi pendapatan untuk empat jenis buku yang berbeda. Buku A memiliki pendapatan tertinggi dan paling konsisten, menjadikannya buku dengan performa terbaik. Buku B memiliki pendapatan terendah dan juga cukup konsisten. Buku C dan Buku D berada di tengah-tengah dalam hal pendapatan, dengan Buku D sedikit lebih baik daripada Buku C. Distribusi ini memberikan gambaran yang jelas tentang bagaimana setiap jenis buku berkontribusi terhadap pendapatan secara keseluruhan.

(Gambar 6)



#### 1. Penjualan Buku A:

- Buku A memiliki total penjualan tertinggi.
- Jumlah buku terjual untuk Buku A sekitar 250 buku.

#### 2. Penjualan Buku B:

- Buku B memiliki total penjualan terendah.
- Jumlah buku terjual untuk Buku B sekitar 125 buku.

#### 3. Penjualan Buku C:

- Buku C memiliki total penjualan di kisaran menengah.
- Jumlah buku terjual untuk Buku C sekitar 175 buku.

#### 4. Penjualan Buku D:

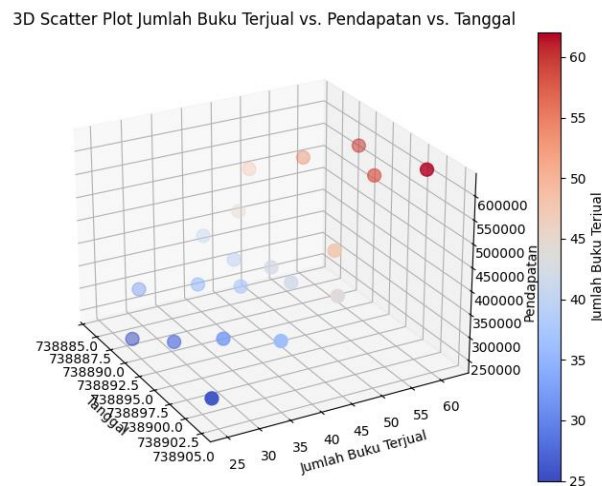
- Buku D juga memiliki total penjualan di kisaran menengah, sedikit lebih tinggi dari Buku C.
- Jumlah buku terjual untuk Buku D sekitar 200 buku.

#### Pola yang Ditemukan:

- **Buku A:** Menunjukkan performa terbaik dalam hal jumlah penjualan, dengan total penjualan yang paling tinggi di antara semua jenis buku.
- **Buku B:** Memiliki performa terendah dalam hal jumlah penjualan.
- **Buku C dan Buku D:** Kedua buku ini memiliki performa penjualan yang lebih baik dibandingkan Buku B, dengan Buku D memiliki sedikit keunggulan dalam jumlah penjualan dibandingkan Buku C.

Grafik ini menunjukkan total penjualan untuk empat jenis buku yang berbeda. Buku A adalah yang paling laris dengan penjualan tertinggi, sementara Buku B memiliki penjualan terendah. Buku C dan Buku D berada di tengah-tengah dalam hal penjualan, dengan Buku D sedikit lebih unggul dibandingkan Buku C. Distribusi ini memberikan wawasan mengenai popularitas dan performa penjualan masing-masing jenis buku, yang dapat digunakan untuk strategi pemasaran dan stok.

(Gambar 7)



#### Interpretasi Grafik: 3D Scatter Plot Jumlah Buku Terjual vs. Pendapatan vs. Tanggal.

##### 1. Hubungan Positif Antara Jumlah Buku Terjual dan Pendapatan:

- Grafik menunjukkan bahwa ada hubungan positif antara jumlah buku yang terjual dan pendapatan. Ini berarti semakin banyak buku yang terjual, semakin tinggi pendapatan yang dihasilkan.



## 2. Distribusi Data Berdasarkan Tanggal:

- Terdapat distribusi data yang bervariasi berdasarkan tanggal, dengan beberapa tanggal menunjukkan penjualan buku yang lebih tinggi dan pendapatan yang lebih besar.
- Penjualan tertinggi (dalam jumlah buku terjual) dicapai sekitar tanggal 738905 dengan pendapatan mendekati 600,000.

## 3. Pola yang Ditemukan:

- Tanggal-tanggal tertentu memiliki performa yang sangat baik, seperti yang terlihat pada penjualan tertinggi di sekitar tanggal 738905.
- Pada tanggal yang lebih awal (seperti 738885 hingga 738895), jumlah buku yang terjual lebih sedikit dengan pendapatan yang lebih rendah, menunjukkan tren kenaikan dalam penjualan dan pendapatan seiring waktu.

Grafik ini memberikan gambaran yang jelas tentang bagaimana penjualan buku meningkat seiring waktu dan bagaimana peningkatan jumlah buku yang terjual berkontribusi terhadap peningkatan pendapatan. Data menunjukkan adanya tren positif antara jumlah buku terjual dan pendapatan, dengan beberapa tanggal menonjol sebagai hari dengan penjualan dan pendapatan yang sangat tinggi. Secara keseluruhan, grafik ini menunjukkan bahwa peningkatan jumlah buku terjual berbanding lurus dengan peningkatan pendapatan, dan tren ini terlihat lebih jelas pada tanggal-tanggal tertentu.

## 3.2 Kode Python Tahap Awal

Kode yang saya gunakan dalam Python untuk menampilkan visualisasi tersebut adalah sebagai berikut:

```
import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from mpl_toolkits.mplot3d import Axes3D

# Membaca dataset

data = pd.read_csv('data_penjualan.csv')

# Mengubah tipe data Tanggal menjadi datetime

data['Tanggal'] = pd.to_datetime(data['Tanggal'])
```

```
# Pie Chart - Persentase jumlah buku terjual per jenis buku

jumlah_per_buku = data.groupby('Nama_Buku')['Jumlah'].sum()

plt.figure(figsize=(8, 8))

plt.pie(jumlah_per_buku, labels=jumlah_per_buku.index,
autopct='%1.1f%%', colors=sns.color_palette('pastel'))

plt.title('Persentase Jumlah Buku Terjual per Jenis Buku')

plt.show()


# Scatter Plot - Jumlah buku terjual vs. Pendapatan

plt.figure(figsize=(10, 6))

sns.scatterplot(x='Jumlah', y='Pendapatan', hue='Nama_Buku',
palette='viridis', s=100, data=data)

plt.title('Scatter Plot Jumlah Buku Terjual vs. Pendapatan')

plt.xlabel('Jumlah Buku Terjual')

plt.ylabel('Pendapatan')

plt.legend(title='Nama Buku')

plt.show()


# Line Plot - Trend penjualan buku dari waktu ke waktu

plt.figure(figsize=(12, 6))

for buku in data['Nama_Buku'].unique():

    buku_data = data[data['Nama_Buku'] == buku]

    plt.plot(buku_data['Tanggal'], buku_data['Jumlah'], marker='o',
linestyle='-', label=buku)

plt.title('Trend Penjualan Buku')
```

```
plt.xlabel('Tanggal')

plt.ylabel('Jumlah Buku Terjual')

plt.legend(title='Nama Buku')

plt.grid(True)

plt.show()


# Histogram - Distribusi jumlah buku terjual

plt.figure(figsize=(10, 6))

sns.histplot(data['Jumlah'], bins=20, kde=True, color='skyblue')

plt.title('Distribusi Jumlah Buku Terjual')

plt.xlabel('Jumlah Buku')

plt.ylabel('Frekuensi')

plt.show()


# Box Plot - Distribusi pendapatan per jenis buku

plt.figure(figsize=(10, 6))

sns.boxplot(x='Nama_Buku', y='Pendapatan', palette='Set2', data=data)

plt.title('Distribusi Pendapatan per Jenis Buku')

plt.xlabel('Nama Buku')

plt.ylabel('Pendapatan')

plt.show()


# Bar Plot - Total penjualan per jenis buku

plt.figure(figsize=(10, 6))
```

```

sns.barplot(x='Nama_Buku', y='Jumlah', palette='Set1',
data=data.groupby('Nama_Buku')['Jumlah'].sum().reset_index())

plt.title('Total Penjualan per Jenis Buku')

plt.xlabel('Nama Buku')

plt.ylabel('Jumlah Terjual')

plt.show()

# 3D Scatter Plot - Jumlah buku terjual vs. Pendapatan vs. Tanggal

fig = plt.figure(figsize=(10, 8))

ax = fig.add_subplot(111, projection='3d')

scatter = ax.scatter(data['Tanggal'].map(pd.Timestamp.toordinal),
data['Jumlah'], data['Pendapatan'], c=data['Jumlah'], cmap='coolwarm',
s=100)

ax.set_title('3D Scatter Plot Jumlah Buku Terjual vs. Pendapatan vs.
Tanggal')

ax.set_xlabel('Tanggal')

ax.set_ylabel('Jumlah Buku Terjual')

ax.set_zlabel('Pendapatan')

plt.colorbar(scatter, ax=ax, label='Jumlah Buku Terjual')

plt.show()

```

### 3.3 Pembersihan Data

Kode yang saya gunakan untuk pembersihan data adalah sebagai berikut:

```

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from mpl_toolkits.mplot3d import Axes3D

```

```
# Membaca dataset

data = pd.read_csv('data_penjualan.csv')


# Pembersihan Data


# Mengubah tipe data Tanggal menjadi datetime
data['Tanggal'] = pd.to_datetime(data['Tanggal'])


# Mengatasi nilai yang hilang
data = data.dropna()


# Menghapus duplikasi
data = data.drop_duplicates()


# Memastikan tipe data yang tepat
data['Jumlah'] = data['Jumlah'].astype(int)
data['Pendapatan'] = data['Pendapatan'].astype(float)


# Visualisasi Data


# Pie Chart - Persentase jumlah buku terjual per jenis buku
jumlah_per_buku = data.groupby('Nama_Buku')['Jumlah'].sum()
plt.figure(figsize=(8, 8))
```

```

plt.pie(jumlah_per_buku, labels=jumlah_per_buku.index,
autopct='%1.1f%%', colors=sns.color_palette('pastel'))

plt.title('Persentase Jumlah Buku Terjual per Jenis Buku')

plt.show()


# Scatter Plot - Jumlah buku terjual vs. Pendapatan

plt.figure(figsize=(10, 6))

sns.scatterplot(x='Jumlah', y='Pendapatan', hue='Nama_Buku',
palette='viridis', s=100, data=data)

plt.title('Scatter Plot Jumlah Buku Terjual vs. Pendapatan')

plt.xlabel('Jumlah Buku Terjual')

plt.ylabel('Pendapatan')

plt.legend(title='Nama Buku')

plt.show()


# Line Plot - Trend penjualan buku dari waktu ke waktu

plt.figure(figsize=(12, 6))

for buku in data['Nama_Buku'].unique():

    buku_data = data[data['Nama_Buku'] == buku]

    plt.plot(buku_data['Tanggal'], buku_data['Jumlah'],
marker='o', linestyle='-', label=buku)

plt.title('Trend Penjualan Buku')

plt.xlabel('Tanggal')

plt.ylabel('Jumlah Buku Terjual')

plt.legend(title='Nama Buku')

```

```
plt.grid(True)

plt.show()

# Histogram - Distribusi jumlah buku terjual
plt.figure(figsize=(10, 6))

sns.histplot(data['Jumlah'], bins=20, kde=True, color='skyblue')

plt.title('Distribusi Jumlah Buku Terjual')

plt.xlabel('Jumlah Buku')

plt.ylabel('Frekuensi')

plt.show()

# Box Plot - Distribusi pendapatan per jenis buku
plt.figure(figsize=(10, 6))

sns.boxplot(x='Nama_Buku', y='Pendapatan', palette='Set2',
data=data)

plt.title('Distribusi Pendapatan per Jenis Buku')

plt.xlabel('Nama Buku')

plt.ylabel('Pendapatan')

plt.show()

# Bar Plot - Total penjualan per jenis buku
plt.figure(figsize=(10, 6))

sns.barplot(x='Nama_Buku', y='Jumlah', palette='Set1',
data=data.groupby('Nama_Buku')['Jumlah'].sum().reset_index())

plt.title('Total Penjualan per Jenis Buku')
```

```

plt.xlabel('Nama Buku')

plt.ylabel('Jumlah Terjual')

plt.show()

# 3D Scatter Plot - Jumlah buku terjual vs. Pendapatan vs.
Tanggal

fig = plt.figure(figsize=(10, 8))

ax = fig.add_subplot(111, projection='3d')

scatter =
ax.scatter(data['Tanggal'].map(pd.Timestamp.toordinal),
data['Jumlah'], data['Pendapatan'], c=data['Jumlah'],
cmap='coolwarm', s=100)

ax.set_title('3D Scatter Plot Jumlah Buku Terjual vs. Pendapatan
vs. Tanggal')

ax.set_xlabel('Tanggal')

ax.set_ylabel('Jumlah Buku Terjual')

ax.set_zlabel('Pendapatan')

plt.colorbar(scatter, ax=ax, label='Jumlah Buku Terjual')

plt.show()

```

Dari kode di atas, langkah-langkah pembersihan data mencakup:

1. **Menghapus Nilai yang Hilang (dropna):**
  - Baris yang memiliki nilai kosong atau hilang dihapus dari dataset. Nilai yang hilang bisa terjadi karena kesalahan dalam pengumpulan data, entri data yang tidak lengkap, atau nilai yang tidak diketahui.
2. **Menghapus Duplikasi (drop\_duplicates):**
  - Baris yang merupakan duplikat dihapus. Duplikasi dapat terjadi karena kesalahan pengumpulan data di mana entri yang sama tercatat lebih dari sekali.
3. **Memastikan Tipe Data yang Tepat:**
  - Kolom 'Jumlah' diubah menjadi tipe data integer (astype(int)), dan kolom 'Pendapatan' diubah menjadi tipe data float (astype(float)). Ini memastikan bahwa perhitungan dan visualisasi yang dilakukan pada kolom-kolom ini adalah akurat.

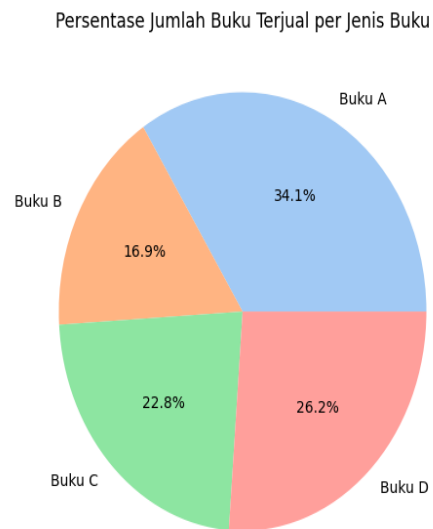


### 3.4 Visualisasi Analisis Data

Setelah proses pembersihan data, beberapa jenis visualisasi dilakukan untuk menganalisis data penjualan buku:

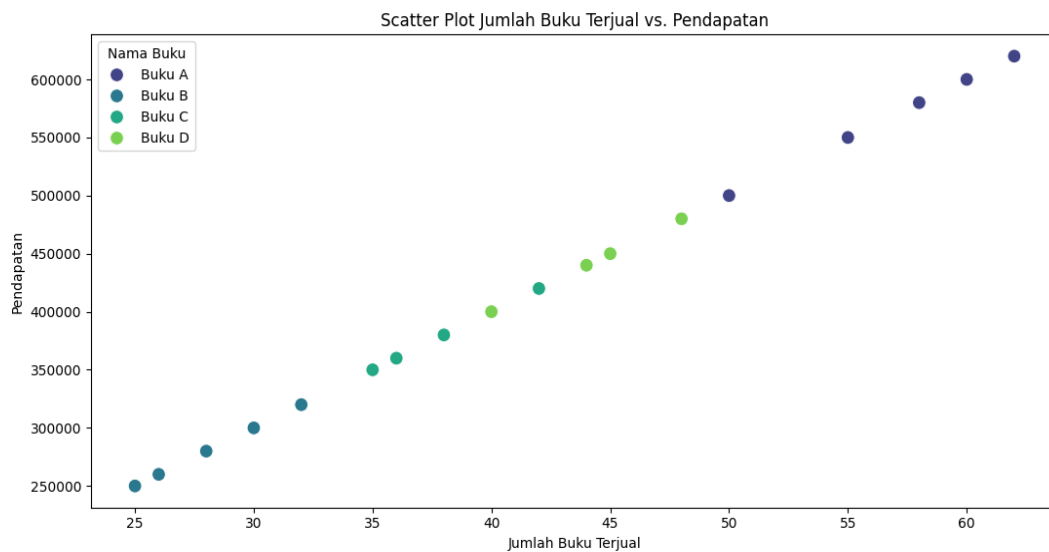
1. **Pie Chart - Persentase Jumlah Buku Terjual per Jenis Buku:**

- Visualisasi ini menunjukkan persentase kontribusi setiap jenis buku terhadap total jumlah buku yang terjual. Membantu dalam memahami distribusi penjualan buku berdasarkan jenis buku.



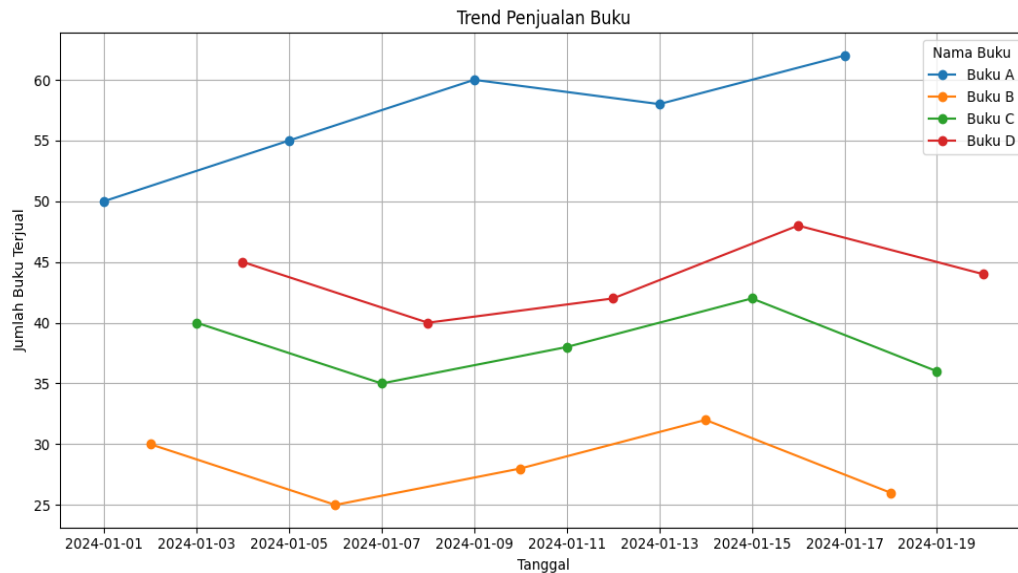
2. **Scatter Plot - Jumlah Buku Terjual vs. Pendapatan:**

- Menampilkan hubungan antara jumlah buku yang terjual dengan pendapatan yang dihasilkan, dengan warna berbeda untuk setiap jenis buku. Visualisasi ini membantu dalam memahami apakah ada korelasi antara jumlah buku yang terjual dan pendapatan.



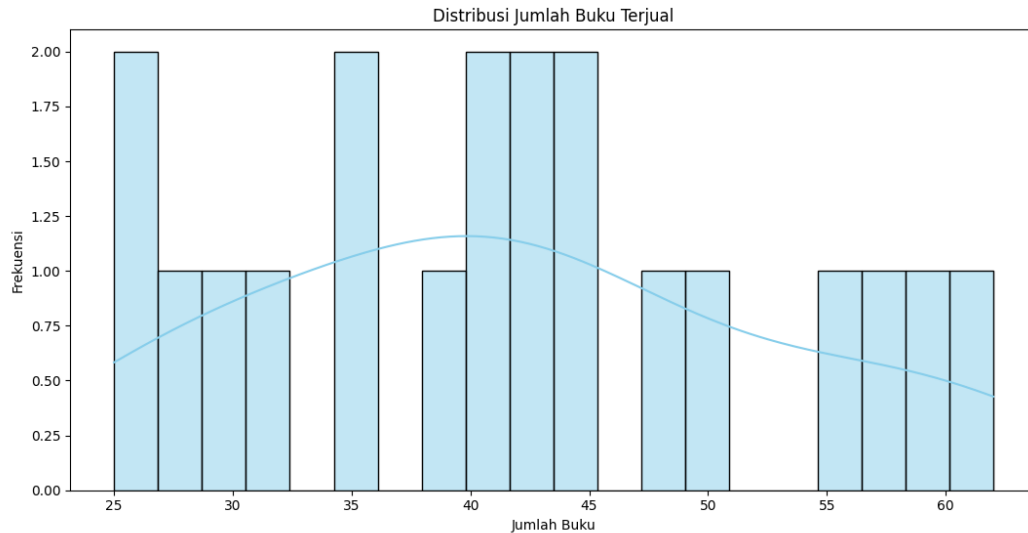
### 3. Line Plot - Trend Penjualan Buku dari Waktu ke Waktu:

- Menunjukkan tren penjualan buku dari waktu ke waktu untuk setiap jenis buku. Membantu dalam memahami bagaimana penjualan buku berubah seiring waktu.



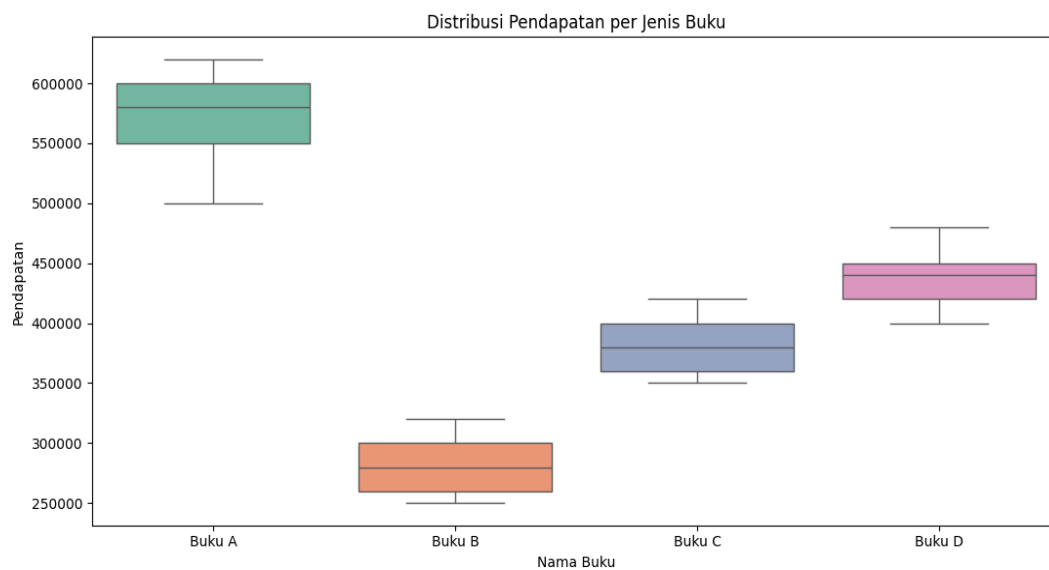
### 4. Histogram - Distribusi Jumlah Buku Terjual:

- Menampilkan distribusi frekuensi jumlah buku yang terjual. Visualisasi ini membantu dalam memahami sebaran jumlah buku yang terjual.



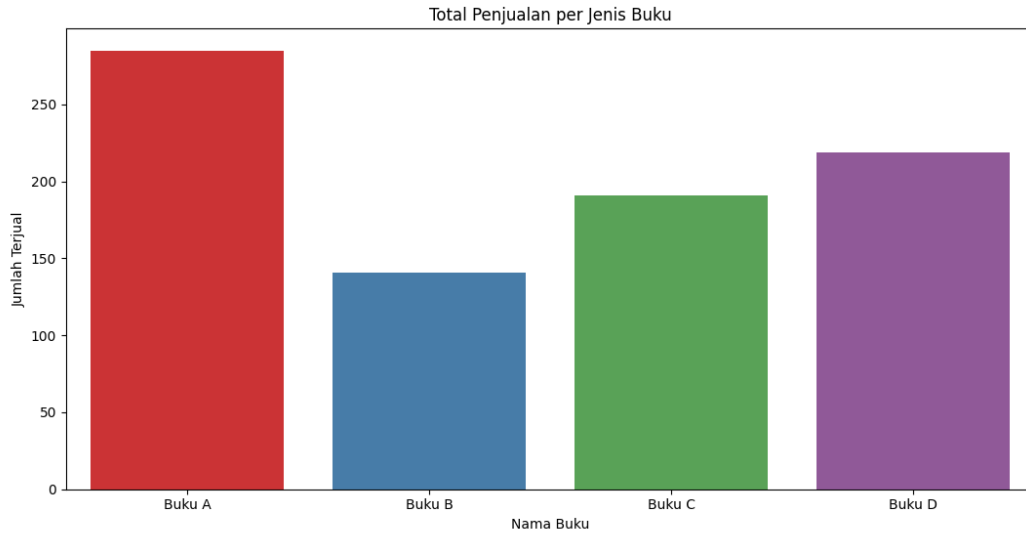
#### 5. Box Plot - Distribusi Pendapatan per Jenis Buku:

- Menampilkan distribusi pendapatan untuk setiap jenis buku. Membantu dalam memahami variasi pendapatan yang dihasilkan oleh setiap jenis buku.



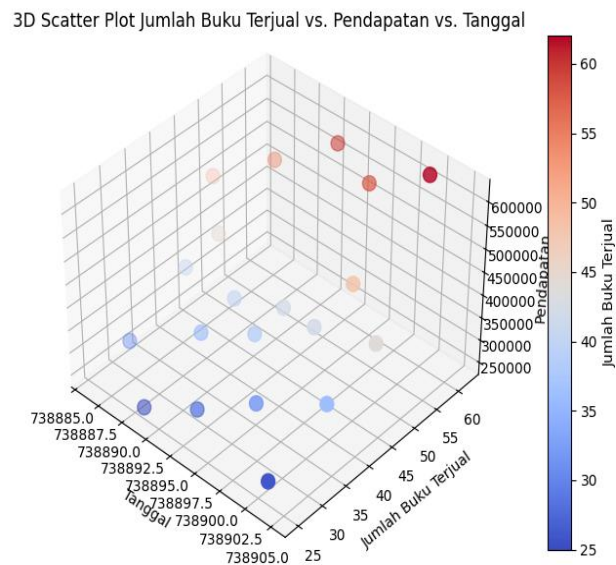
### 6. Bar Plot - Total Penjualan per Jenis Buku:

- Menampilkan total jumlah buku yang terjual untuk setiap jenis buku. Membantu dalam melihat buku mana yang paling banyak terjual.



### 7. 3D Scatter Plot - Jumlah Buku Terjual vs. Pendapatan vs. Tanggal:

- Visualisasi tiga dimensi yang menunjukkan hubungan antara jumlah buku yang terjual, pendapatan, dan tanggal. Membantu dalam memahami bagaimana penjualan dan pendapatan berubah seiring waktu dalam satu grafik tiga dimensi.



## BAB IV

### KESIMPULAN DAN SARAN

#### 4.1 Kesimpulan

Berdasarkan visualisasi data penjualan buku yang dilakukan, dapat ditarik beberapa kesimpulan penting:

1. **Persentase Jumlah Buku Terjual per Jenis Buku:**
  - Pie Chart menunjukkan bahwa penjualan buku didominasi oleh beberapa jenis buku tertentu. Hal ini menunjukkan preferensi pembeli terhadap jenis buku tertentu yang mungkin memiliki daya tarik atau popularitas lebih tinggi dibandingkan dengan yang lain.
2. **Jumlah Buku Terjual vs. Pendapatan:**
  - Scatter Plot menunjukkan hubungan yang cukup erat antara jumlah buku terjual dengan pendapatan yang dihasilkan. Buku dengan penjualan tinggi cenderung memberikan pendapatan yang lebih besar, namun terdapat beberapa outlier yang menarik untuk diteliti lebih lanjut.
3. **Trend Penjualan Buku dari Waktu ke Waktu:**
  - Line Plot memperlihatkan tren penjualan buku selama periode tertentu. Beberapa buku menunjukkan pola penjualan yang konsisten, sementara yang lain menunjukkan fluktuasi yang signifikan, mengindikasikan pengaruh faktor musiman atau promosi tertentu.
4. **Distribusi Jumlah Buku Terjual:**
  - Histogram menunjukkan distribusi jumlah buku yang terjual, dengan sebagian besar buku terjual dalam rentang tertentu. Hal ini membantu mengidentifikasi pola umum dalam penjualan buku dan potensi adanya stok yang perlu disesuaikan.
5. **Distribusi Pendapatan per Jenis Buku:**
  - Box Plot memperlihatkan variasi pendapatan yang dihasilkan oleh berbagai jenis buku. Ada beberapa buku yang memberikan pendapatan tinggi secara konsisten, sementara yang lain menunjukkan variabilitas yang lebih besar.
6. **Total Penjualan per Jenis Buku:**
  - Bar Plot menunjukkan total penjualan untuk setiap jenis buku. Jenis buku tertentu menonjol dengan penjualan yang jauh lebih tinggi dibandingkan dengan yang lain, memberikan indikasi buku-buku yang paling diminati oleh pembeli.
7. **3D Scatter Plot - Jumlah Buku Terjual vs. Pendapatan vs. Tanggal:**
  - Visualisasi 3D memberikan pandangan yang lebih mendalam mengenai hubungan antara waktu, jumlah buku terjual, dan pendapatan. Hal ini membantu dalam mengidentifikasi pola yang mungkin tidak terlihat dalam visualisasi 2D.

#### 4.2 Saran

Berdasarkan kesimpulan yang diperoleh dari analisis data penjualan buku, beberapa saran yang dapat diberikan adalah sebagai berikut:

1. **Fokus pada Buku dengan Penjualan Tinggi:**

- Meningkatkan stok dan promosi untuk buku-buku yang memiliki penjualan tinggi dapat membantu meningkatkan pendapatan. Buku-buku ini memiliki daya tarik yang besar dan potensial untuk menghasilkan keuntungan lebih.
- 2. **Analisis Faktor Penyebab Fluktuasi Penjualan:**
  - Melakukan analisis lebih lanjut terhadap faktor-faktor yang menyebabkan fluktuasi penjualan, seperti musim, promosi, atau acara tertentu. Hal ini dapat membantu dalam merencanakan strategi pemasaran yang lebih efektif.
- 3. **Optimalkan Pendapatan dari Buku yang Populer:**
  - Meningkatkan harga atau menawarkan paket promosi untuk buku-buku yang populer dapat menjadi strategi untuk meningkatkan pendapatan tanpa mengorbankan volume penjualan.
- 4. **Evaluasi Buku dengan Penjualan Rendah:**
  - Mengidentifikasi dan mengevaluasi buku-buku dengan penjualan rendah untuk menentukan apakah perlu dilakukan perbaikan dalam strategi pemasaran atau mengurangi stok untuk mengurangi biaya.
- 5. **Diversifikasi Koleksi Buku:**
  - Menghadirkan variasi buku yang lebih luas untuk menarik segmen pasar yang berbeda. Hal ini dapat membantu meningkatkan penjualan keseluruhan dengan menarik pembeli dari berbagai preferensi.
- 6. **Penggunaan Data untuk Perencanaan Stok:**
  - Menggunakan data historis penjualan untuk merencanakan stok secara lebih efisien, menghindari kelebihan stok atau kekurangan stok yang dapat mempengaruhi penjualan dan biaya operasional.
- 7. **Pemantauan Berkelanjutan:**
  - Melakukan pemantauan berkelanjutan terhadap penjualan buku dan melakukan penyesuaian strategi secara real-time berdasarkan data yang diperoleh. Hal ini akan memastikan bahwa bisnis selalu berada pada jalur yang tepat untuk mencapai tujuan penjualan dan pendapatan.

## DAFTAR PUSTAKA

[https://repository.amikom.ac.id/files/PUBLIKASI\\_07\\_12\\_2693.pdf](https://repository.amikom.ac.id/files/PUBLIKASI_07_12_2693.pdf)

<https://github.com/YennyClaraManihuruk/UASPengkodeandanPemograman>

*(link tugas setelah di upload ke Github)*