

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini, akan dibahas hasil analisis yang telah dilakukan terkait dengan pola perilaku *doom spending* pada generasi z menggunakan metode k-means clustering. Proses analisis ini dilakukan untuk mengidentifikasi pola perilaku *doom spending* pada generasi z serta faktor-faktor yang memengaruhi perilaku tersebut. Data yang digunakan dalam penelitian ini diperoleh melalui penyebaran kuesioner kepada responden dari kelompok usia 18 hingga 27 tahun, yang merupakan kategori umum untuk generasi z.

Metode K-Means clustering digunakan untuk mengelompokkan responden berdasarkan karakteristik pengeluaran mereka dan kecenderungan *doom spending*. Hasil yang didapatkan dari analisis ini diharapkan dapat memberikan pemahaman lebih dalam mengenai kebiasaan belanja yang cenderung *impulsif* dan tidak terkendali, serta faktor-faktor yang mendorong perilaku tersebut pada generasi z.

4.1 Data Collection (Pengumpulan Data)

4.1.1 Pengumpulan Data Menggunakan Kuesioner

Langkah pertama dalam proses data mining adalah pengumpulan data. Data yang digunakan dalam penelitian ini diperoleh melalui survei yang dibagikan kepada 602 responden yang merupakan anggota dari generasi z (berusia antara 18 hingga 27 tahun).

Survei ini mengumpulkan data dalam bentuk angka dan kategori, yang nantinya akan dianalisis menggunakan teknik data mining, data yang akan diolah merupakan data dalam bentuk file CSV.

Data yang dikumpulkan mencakup informasi tentang :

Table 4.1 Penjelasan atribut data penelitian

Nama Atribute	Keterangan
Usia	Data usia generasi z dari umur 18 sampai 27 tahun
Jenis Kelamin	jenis kelamin individu, dikategorikan sebagai laki-laki atau perempuan
Pendidikan	Tingkat pendidikan terakhir yang telah diselesaikan oleh individu, yaitu SMP,SMA,D3,S1, dll
Pendapatan bulanan	Jumlah uang yang diterima seseorang setiap bulan dari pekerjaan atau sumber lain.
Frekuensi belanja tidak terencana	Menunjukkan seberapa sering seseorang melakukan belanja impulsif atau tidak direncanakan.
Belanja tanpa kebutuhan	Belanja barang yang tidak diperlukan secara mendesak.
Jenis barang	Kategori barang yang dibeli (misalnya, elektronik, pakaian).
Jumlah barang per transaksi	Rata-rata jumlah barang yang dibeli dalam satu kali transaksi.
Pengeluaran bulanan	Total pengeluaran seseorang dalam sebulan untuk membeli barang yang tidak direncanakan atau tidak diperlukan diluar kebutuhan sehari-hari.
Pemicu belanja impulsif	Faktor yang memicu seseorang untuk belanja tanpa perencanaan
Tempat/Platform belanja	Tempat atau platform (offline atau online) tempat seseorang berbelanja.
Alasan memilih tempat belanja	Alasan memilih tempat atau platform tertentu untuk berbelanja, misalnya harga atau kenyamanan.
Metode pembayaran	Cara pembayaran yang digunakan, seperti kartu kredit, debit, atau uang tunai.
Masalah finansial	Kesulitan dalam mengelola keuangan, seperti kesulitan membayar tagihan atau menabung.

4.2 Data Understanding (Pemahaman Data)

Setelah data terkumpul, langkah berikutnya adalah untuk memahami data tersebut. Pada tahap ini, dilakukan eksplorasi awal terhadap data untuk memahami struktur, karakteristik, dan kualitasnya. Tujuannya adalah untuk mengetahui informasi apa yang ada dalam data, serta untuk mendeteksi masalah seperti missing values, duplikasi, atau inkonsistensi.

4.2.1 Menyiapkan Data dan Library

Bila dataset yang dibutuhkan sudah tersedia, tahap selanjutnya adalah menyiapkan library pada google colab untuk memudahkan dalam pengolahan data.

```
import pandas as pd # Untuk manipulasi dan analisis data dalam format tabel
import numpy as np # Untuk komputasi numerik (matematika,array,statistik)
import matplotlib.pyplot as plt # untuk visualisasi data (grafik dasar seperti garis, )
import seaborn as sns # untuk visualisasi data statistik yang lebih menarik dan informatif
import re # validasi dan pencocokan format penulisan dalam dataset
from sklearn.cluster import KMeans # Change kMeans to KMeans
from sklearn.preprocessing import StandardScaler
```

Gambar 4.1 Import library yang digunakan

4.2.2 Mengimpor Data Ke Google Colab

Selanjutnya Pada tahapan ini penulis akan mengimpor data penelitian yang telah disimpan dalam bentuk csv tersebut di google colab.

```
# Menyambungkan Google Colab ke Google Drive agar dapat mengakses Dataset / DataFrame
from google.colab import drive # library untuk mengakses Google Drive dari Google Collab
drive.mount('/content/drive') # Menghubungkan ke google drive
file_path = '/content/drive/My Drive/Data Penelitian/Analisis Doom spending.csv' # Alamat File data
df = pd.read_csv(file_path) #Membaca File
```

Gambar 4.2 Code import data

Kode di atas fungsinya untuk mengimpor dataset penelitian yang telah disimpan pada google drive ke google colab untuk mempermudah

pengolahan data. Hal lain yang dapat dilakukan pada dataset adalah mendapatkan informasi terkait jumlah kolom, tipe data, jumlah seluruh data yang terdapat pada dataset.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 602 entries, 0 to 601
Data columns (total 16 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   Nama                                       602 non-null   object
1   Usia                                       602 non-null   int64
2   Jenis Kelamin                             602 non-null   object
3   Pendidikan Terakhir                       602 non-null   object
4   Pekerjaan                                  602 non-null   object
5   Pendapatan bulanan                       602 non-null   object
6   Frekuensi Belanja Tidak Terencana        602 non-null   object
7   Belanja Tanpa Kebutuhan                   602 non-null   object
8   Jenis Barang                              602 non-null   object
9   Jumlah Barang Per Transaksi               602 non-null   object
10  Pengeluaran Bulanan                       602 non-null   object
11  Pemicu Belanja Impulsif                   602 non-null   object
12  Tempat / Platform Belanja                 602 non-null   object
13  Alasan Memilih Tempat Belanja             602 non-null   object
14  Metode pembayaran                         602 non-null   object
15  Masalah Finansial                         602 non-null   object
dtypes: int64(1), object(15)
memory usage: 75.4+ KB
```

Gambar 4.3 Penjelasan dataset

Dapat dilihat bahwa data penelitian terdapat 602 baris dan 16 kolom, dengan sebagian besar kolom berisi data kategorikal atau teks (dengan tipe data object), kecuali kolom usia yang berisi data numerik (dengan tipe data int64). Data ini tidak memiliki missing values, karena tipe data nya belum sesuai maka missing value nya belum sepenuhnya terdeteksi oleh karena itu langkah selanjutnya menyesuaikan tipe data yang belum sesuai agar bisa di analisis lebih lanjut.

4.3 Data *Preparation* (Persiapan Data)

Data preparation adalah proses mengolah data mentah (raw data) menjadi data yang bersih, konsisten, dan siap untuk dianalisis. Ini adalah tahap penting dalam siklus analisis data yang menghubungkan tahap pemahaman data (data understanding) dengan tahap pemodelan(modeling).

4.3.1 Mengoreksi Kesalahan Tipe data

```
# Column Non-Null Count Dtype
---  ---
0 Nama 602 non-null object
1 Usia 602 non-null int64
2 Jenis Kelamin 602 non-null object
3 Pendidikan Terakhir 602 non-null object
4 Pekerjaan 602 non-null object
5 Pendapatan bulanan 602 non-null object
6 Frekuensi Belanja Tidak Terencana 602 non-null object
7 Belanja Tanpa Kebutuhan 602 non-null object
8 Jenis Barang 602 non-null object
9 Jumlah Barang Per Transaksi 602 non-null object
10 Pengeluaran Bulanan 602 non-null object
11 Pemicu Belanja Impulsif 602 non-null object
12 Tempat / Platform Belanja 602 non-null object
13 Alasan Memilih Tempat Belanja 602 non-null object
14 Metode pembayaran 602 non-null object
15 Masalah Finansial 602 non-null object
dtypes: int64(1), object(15)
memory usage: 75.4+ KB
```

Gambar 4.4 Pemahaman dataset

Bila diperhatikan pada data diatas , nilai dari kolom pendapatan bulanan, jumlah barang per transaksi, dan pengeluaran bulanan tipe datanya tidak dalam bentuk numerik (Float/ Integer) melainkan dalam bentuk string. Penyesuaian tipe data adalah langkah penting dalam analisis data, hal ini untuk memastikan data siap diproses. Dengan tipe data yang tepat, analisis yang dilakukan akan lebih akurat.

Selanjutnya dilakukan perubahan tipe data pada kolom jumlah barang per transaksi berisi angka bulat (integer) contohnya yaitu 3 item atau 5, jika tipe data pada atribut ini masih berupa tipe objek (string), kita perlu mengubahnya ke tipe data numerik (integer).

```
# Convert 'Jumlah barang per transaksi' to integer
df['Jumlah Barang Per Transaksi'] = pd.to_numeric(df['Jumlah Barang Per Transaksi'], errors='coerce')
# Fill NaN values with 0 before converting to integer
df['Jumlah Barang Per Transaksi'] = df['Jumlah Barang Per Transaksi'].fillna(0).astype(int)
df.info()
```

Gambar 4.5 Code Convert tipe data to integer

Pada atribut pendapatan bulanan dan pengeluaran bulanan dilakukan tahapan penghapusan karakter non-numerik (Rp) agar bisa dilakukan tahapan penyesuaian tipe data yang seharusnya dalam bentuk float karena pada kedua kolom tersebut berisi bilangan desimal.

```
# Code untuk menghapus Rp pada Kolom Pendapatan bulanan dan Pengeluaran Bulanan
# Replace "Rp" and "." with empty strings in the specified columns
df['Pendapatan bulanan'] = df['Pendapatan bulanan'].astype(str).str.replace('Rp', '', regex=False).str.replace('.', '', regex=False)
df['Pengeluaran Bulanan'] = df['Pengeluaran Bulanan'].astype(str).str.replace('Rp', '', regex=False).str.replace('.', '', regex=False)

# Display the first few rows to verify the changes
df.head()
```

Gambar 4.6 Code menghapus karakter Rp

Setelah dilakukan tahapan penghapusan karakter data non numerik (Rp) pada atribut pendapatan bulanan dan pengeluaran bulanan selanjutnya merubah rentang nilai pendapatan bulanan dan pengeluaran bulanan kolom pendapatan bulanan dan pengeluaran bulanan sebelumnya berisi nilai dalam format teks yang mencantumkan simbol mata uang. Nilai-nilai ini diubah menjadi rata-rata nilai dalam rentang tertentu untuk memungkinkan konversi tipe data menjadi float. Hal ini dilakukan agar data dapat digunakan dalam analisis numerik.

```
# Function to extract the midpoint of income ranges
def extract_midpoint(income_range):
    try:
        # Split the range into two values
        lower, upper = income_range.strip().split(' - ')
        # Calculate the midpoint
        midpoint = (float(lower) + float(upper)) / 2
        return midpoint
    except (ValueError, AttributeError):
        # Handle cases where income_range is not in expected format
        # For example, if income_range is a single value or not a string
        try:
            return float(income_range)
        except ValueError:
            return np.nan # Replace non-numeric entries with NaN

# Apply the function to clean 'Pendapatan bulanan'
df['Pendapatan bulanan'] = df['Pendapatan bulanan'].apply(extract_midpoint)

# Apply the function to clean 'Pengeluaran Bulanan'
df['Pengeluaran Bulanan'] = df['Pengeluaran Bulanan'].apply(extract_midpoint)

# Calculate the median of 'Pendapatan bulanan' and 'Pengeluaran Bulanan'
median_pendapatan = df['Pendapatan bulanan'].median()
median_pengeluaran = df['Pengeluaran Bulanan'].median()

print(f"Median Pendapatan Bulanan: {median_pendapatan}")
print(f"Median Pengeluaran Bulanan: {median_pengeluaran}")

Median Pendapatan Bulanan: 2250000.0
Median Pengeluaran Bulanan: 1500000.0
```

Gambar 4.7 Code extract midpoint

4.3.2 Data Cleaning (Pembersihan Data)

Tujuan pembersihan data adalah untuk menghapus data yang dianggap tidak seimbang, yang tidak lengkap atau data yang tidak sesuai.

1. Menangani Missing Values

Pada tahap ini dilakukan pemeriksaan pada dataset yang bertujuan mengetahui nilai yang rumpang atau missing values untuk menghindari overfitting pada model, data yang rumpang sangat sensitif terhadap performa model. Oleh karena, itu harus dilakukan pemeriksaan missing value.

```
# melihat jumlah data yang missing
df.isnull().sum()
```

	0
Nama	0
Usia	0
Jenis Kelamin	0
Pendidikan Terakhir	0
Pekerjaan	0
Pendapatan bulanan	9
Frekuensi Belanja Tidak Terencana	0
Belanja Tanpa Kebutuhan	0
Jenis Barang	0
Jumlah Barang Per Transaksi	0
Pengeluaran Bulanan	3
Pemicu Belanja Impulsif	0
Tempat / Platform Belanja	0
Alasan Memilih Tempat Belanja	0
Metode pembayaran	0
Masalah Finansial	0

dtype: int64

Gambar 4.8 Jumlah missing value

Setelah dilakukan pemeriksaan missing value diketahui bahwa dalam dataset hanya terdapat dua kolom yang memiliki missing value (Tidak ada nilainya/null), yaitu pada kolom pendapatan bulanan terdapat 9

missing value dan pada kolom pengeluaran bulanan terdapat 3 missing value.

```
# drop data yang null / Menangani data yang hilang
df.dropna(inplace=True)
# crosscheck apakah data null sudah terhapus
df.isnull().sum()
```

Gambar 4.9 Code penanganan missing value

```
# menghapus data yang nilainya kosong pada kolom Pendapatan bulanan dan Pengeluaran Bulanan

# Drop rows where 'Pendapatan bulanan' or 'Pengeluaran Bulanan' are NaN
df.dropna(subset=['Pendapatan bulanan', 'Pengeluaran Bulanan'], inplace=True)
```

Gambar 4.10 Code drop data yang nilainya kosong

Dapat dilihat pada gambar di atas, langkah selanjutnya yaitu menghapus dan menangani baris-baris pada kolom pendapatan bulanan dan pengeluaran bulanan yang mengandung nilai yang hilang. Hal ini merupakan langkah yang penting karena Menangani dan menghapus missing values sangat penting untuk menjaga keakuratan analisis, konsistensi data, dan integritas model. Jika tidak diatasi dengan benar, missing values bisa menyebabkan bias, hasil yang salah, atau kesimpulan yang keliru.

	-
Nama	0
Usia	0
Jenis Kelamin	0
Pendidikan Terakhir	0
Pekerjaan	0
Pendapatan bulanan	0
Frekuensi Belanja Tidak Terencana	0
Belanja Tanpa Kebutuhan	0
Jenis Barang	0
Jumlah Barang Per Transaksi	0
Pengeluaran Bulanan	0
Pemicu Belanja Impulsif	0
Tempat / Platform Belanja	0
Alasan Memilih Tempat Belanja	0
Metode pembayaran	0
Masalah Finansial	0

Gambar 4.11 Missing value setelah ditangani

Setelah menghapus dan menangani baris- baris pada kolom pendapatan bulanan dan pengeluaran bulanan maka dapat dilihat bahwa missing value sudah selesai di tangani karena semua kolom bernilai 0 yang artinya tidak terdapat missing value lagi di setiap kolom dalam data , dan jumlah data setelah dilakukannya penghapusan missing value terdapat 592 data dari 602 data.

2. Memeriksa dan Menghapus Kolom Duplikat

Dalam penelitian ini, salah satu langkah penting yang dilakukan adalah menangani data duplikat yang ditemukan dalam dataset.

```
# Menampilkan data yang duplikat sama persis dari data semua kolom yang satu baris
# Find duplicate rows based on all columns
duplicate_rows = df[df.duplicated(keep=False)]
# Display the duplicate rows
print("Duplicate Rows:")
duplicate_rows
```

Duplicate Rows:

	Nama	Usia	Jenis Kelamin	Pendidikan Terakhir	Pekerjaan	Pendapatan bulanan	Frekuensi Belanja Tidak Terencana
140	Ajis Rujianto	20	Laki-Laki	SMA/SLTA	Pelajar/Mahasiswa	2250000.0	Sangat sering
141	Ajis Rujianto	20	Laki-Laki	SMA/SLTA	Pelajar/Mahasiswa	2250000.0	Sangat sering
142	Ajis Rujianto	20	Laki-Laki	SMA/SLTA	Pelajar/Mahasiswa	2250000.0	Sangat sering

Gambar 4.12 Melihat data yang duplikat

Setelah dilakukan pemeriksaan, ditemukan bahwa Ajis Rujianto mengisi kuesioner sebanyak tiga kali, yang mengakibatkan terciptanya data duplikat atau data yang sama persis. Keberadaan data duplikat ini berpotensi mempengaruhi kualitas analisis dan hasil yang diperoleh, karena dapat menyebabkan distorsi dalam perhitungan statistik, seperti rata-rata atau distribusi data.

Untuk memastikan keakuratan hasil analisis, data duplikat tersebut dihapus dan hanya satu entri yang relevan yang dipertahankan. Proses ini dilakukan dengan mempertahankan satu data yang sesuai dengan informasi yang paling representatif, sehingga dataset yang digunakan dalam penelitian menjadi lebih bersih dan konsisten.

```

#Menghapus Data yang duplikat sama persis dari data semua kolom yang satu baris , dengan mempertahankan 1 saja yang pertama
# Menghapus data duplikat, mempertahankan baris pertama
df.drop_duplicates(keep='first', inplace=True)

# Verifikasi data duplikat
duplicate_rows = df[df.duplicated(keep=False)]
print("Duplicate Rows:")
duplicate_rows

```

Gambar 4.13 Menghapus data duplikat

Setelah dilakukan penghapusan data yang duplikat maka jumlah dataset sekarang terdapat 590 data.

```
df.shape
```

```
(590, 16)
```

3. Memeriksa dan Menangani Outlier

Mengidentifikasi nilai pada dataset apakah logic dan rasional sesuai dengan fakta. Data yang tidak sesuai dengan faktanya biasanya disebut sebagai data outlier atau data yang sangat menyimpang pada sebaran didalam dataset. Data outlier dapat mempengaruhi performa model karena dapat mengakibatkan overfitting dan underfitting, maka perlu dilakukan pemeriksaan outlier.

```

# Analisa Outlier
# Membuat visualisasi boxplot (diagram kotak) untuk semua kolom numerik
def boxplot(df,num_cols):
    list_plots = len(num_cols)
    list_rows = (list_plots + 3)// 4
    # Create a grid of subplots with the correct number of rows and columns
    fig, axes = plt.subplots(list_rows, 4, figsize=(18, 4 * list_rows))

    # Flatten the axes array to iterate through it easily
    axes = axes.flatten()

    for i, column in enumerate(num_cols):
        # Access the current subplot using a single index
        sns.boxplot(x=df[column], ax=axes[i])
        axes[i].set_title(column)

    # Remove any unused subplots
    for j in range(list_plots, list_rows * 4):
        fig.delaxes(axes[j])

    plt.tight_layout()
    plt.show()

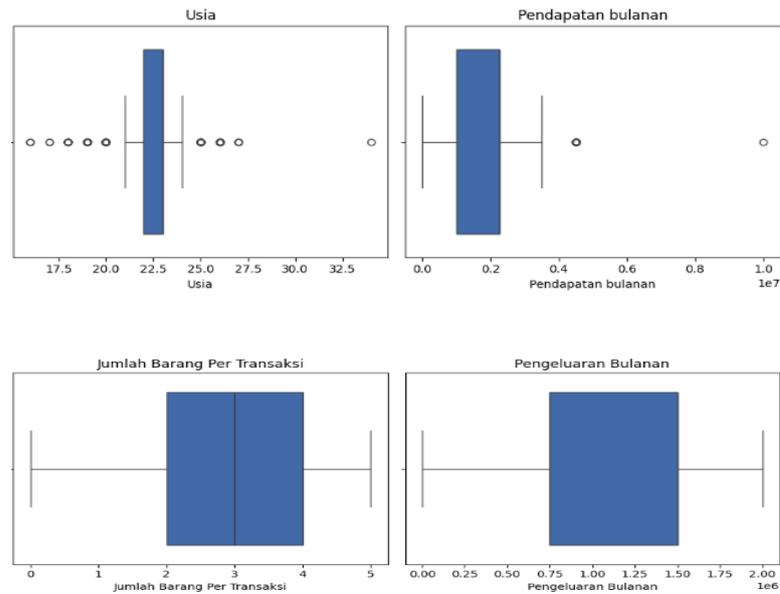
# Get numerical columns from the DataFrame
num_cols = df.select_dtypes(include=['number']).columns.tolist()

# Now call the boxplot function with the DataFrame and numerical columns
boxplot(df, num_cols)

```

Gambar 4.14 Pemeriksaan outlier

Pada penelitian ini pemeriksaan outlier menggunakan IQR (Inter Quartile Range) dimana data dikatakan sebagai outlier karena data tersebut melebihi nilai maksimum dari pada Q3 atau dibawah nilai minimum.



Gambar 4.15 Visualisasi data outlier

Setelah dilakukan pemeriksaan menggunakan IQR didapatkan nilai yang memiliki outlier adalah variabel usia terdapat beberapa outlier yang mengidentifikasi adanya responden dengan usia yang jauh lebih tua dari rata-rata dan pendapatan bulanan terdapat adanya dua responden dengan pendapatan yang sangat tinggi.

Menangani outlier pada variabel usia dan pendapatan bulanan dengan menggunakan metode IQR. Hal ini dilakukan karena menangani outlier sangat penting dalam analisis, data yang terdapat outlier akan mempengaruhi proses atau hasil analisis.

```

# Fungsi untuk menangani outlier menggunakan IQR
def handle_outliers_iqr(df, column):
    Q1 = df[column].quantile(0.25)
    Q3 = df[column].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    df[column] = np.clip(df[column], lower_bound, upper_bound)
    return df

# Tangani outlier pada kolom 'Usia'
df = handle_outliers_iqr(df, 'Usia')

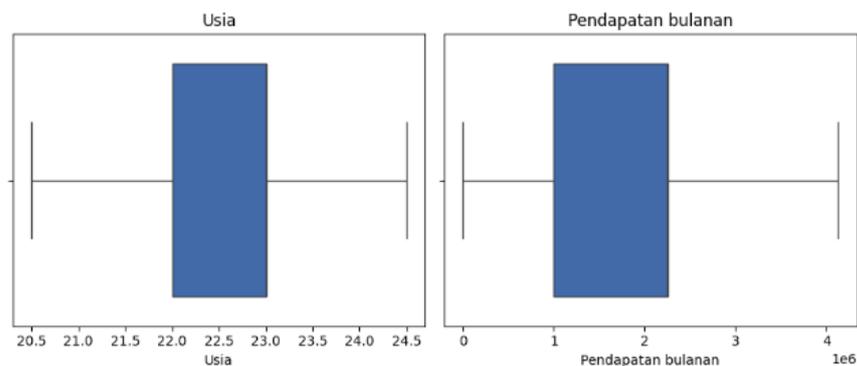
# Tangani outlier pada kolom 'Pendapatan bulanan'
df = handle_outliers_iqr(df, 'Pendapatan bulanan')

# Visualisasi boxplot setelah penanganan outlier
num_cols = df.select_dtypes(include=['number']).columns.tolist()
boxplot(df, num_cols)

```

Gambar 4.16 Penanganan outlier

Dapat dilihat pada visualisasi boxplot di bawah ini setelah outlier pada atribut usia dan pendapatan bulanan di tangani.



Gambar 4.17 Visualisasi data setelah penanganan outlier

4.3.3 Transformasi Data

1. Mengubah data kategorikal menjadi data numerik dengan label encoding.

Sebelum dilakukan pemodelan, data harus diproses terlebih dahulu. Salah satu tahap pra-pemrosesan yang dilakukan adalah mengubah data kategorikal menjadi numerik menggunakan label encoding agar data dapat digunakan dalam analisis k-means clustering.

```

# Mengidentifikasi kolom kategorikal
categorical_cols = df.select_dtypes(include=['object']).columns.tolist()

# Melakukan label encoding pada kolom kategorikal
label_encoders = {}
for col in categorical_cols:
    le = LabelEncoder()
    df[col] = le.fit_transform(df[col])
    label_encoders[col] = le # Menyimpan encoder jika diperlukan untuk decoding di masa mendatang

# Menampilkan informasi setelah label encoding
print("\nInformasi Dataset Setelah Label Encoding:")
print(df.info())

# Menampilkan beberapa baris pertama setelah label encoding
print("\nBeberapa Baris Pertama Setelah Label Encoding:")
print(df.head())

```

Gambar 4.18 Code label encoding

Kode pada gambar di atas bertujuan untuk mempersiapkan data kategorikal agar dapat digunakan dalam analisis lebih lanjut atau dalam model machine learning dengan mengubahnya menjadi nilai numerik melalui proses label encoding, serta memeriksa hasil perubahan tersebut.

Usia	Jenis Kelamin	Pendidikan	Pekerjaan	Pendapatan	Frekuensi	Belanja Tahunan	Jenis Barang	Jumlah Barang	Pengeluaran	Pemicu Belanja	Tempat / Alasan	Metode Pembayaran	Masalah Finansial
24	1	4	10	1000000	3	1	46	4	750000	12	4	4	4
22	0	4	9	1000000	1	3	26	3	750000	12	4	4	5
22	1	4	10	1000000	3	3	13	2	750000	12	4	4	5
22	1	1	15	4125000	1	0	42	3	750000	14	4	4	5
23	1	4	9	2250000	1	1	58	3	500000	12	4	4	5
20.5	1	4	7	2250000	3	1	42	3	500000	5	4	0	2
20.5	1	4	4	2250000	0	0	58	2	500000	6	0	0	6
22	0	1	10	1000000	2	2	13	2	500000	14	1	0	6
21	1	4	9	4125000	4	4	58	2	750000	14	4	0	6
23	1	4	14	200	0	4	58	2	200	14	4	0	0
24.5	1	1	7	4125000	3	1	58	1	500000	6	4	4	5
24.5	0	4	7	4125000	2	4	56	1	1500000	12	1	5	5

Gambar 4.19 Sampel dataset setelah label encoding

Gambar di atas menunjukkan sampel dataset setelah proses label encoding diterapkan pada kolom-kolom kategorikal. Setelah label encoding, setiap kolom yang sebelumnya bertipe object (seperti string atau teks) sekarang menjadi numerik. Misalnya, kolom Jenis Kelamin yang awalnya berisi nilai "Pria" dan "Wanita" menjadi berisi nilai 0 dan 1, yang memudahkan model dalam menghitung dan mengolah data.

2. Memilih dan menyiapkan fitur yang akan digunakan dalam model.

Proses ini sering disebut juga sebagai feature selection dan feature engineering.

```
# Pilih atribut yang relevan
columns = [
    'Usia',
    'Jenis Kelamin',
    'Frekuensi Belanja Tidak Terencana',
    'Belanja Tanpa Kebutuhan',
    'Pengeluaran Bulanan',
    'Pemicu Belanja Impulsif',
    'Tempat / Platform Belanja',
    'Masalah Finansial'
]
```

Gambar 4.20 Pemilihan atribut

Feature selection (pemilihan fitur) bertujuan untuk memilih fitur yang memiliki pengaruh besar terhadap hasil yang ingin dicapai oleh pemodelan. Dalam dataset yang digunakan, terdapat beberapa kolom atau atribut yang memiliki potensi untuk memberikan informasi yang penting mengenai perilaku yang dianalisis, dalam hal ini terkait dengan doom spending pada generasi Z. Kolom-kolom yang dipilih sebagai fitur dalam model adalah:

- a. Usia : menjadi fitur yang relevan karena perilaku belanja impulsif atau doom spending bisa sangat dipengaruhi oleh kelompok usia. Kelompok usia muda, khususnya generasi Z, sering kali menunjukkan kecenderungan perilaku belanja tertentu yang perlu dianalisis.
- b. Jenis Kelamin : dapat mempengaruhi pola belanja, karena studi menunjukkan bahwa perilaku konsumsi antara pria dan wanita sering kali berbeda. Dalam konteks ini, jenis kelamin digunakan

sebagai variabel untuk melihat apakah terdapat perbedaan perilaku belanja berdasarkan jenis kelamin.

- c. Frekuensi Belanja Tidak Terencana : adalah indikator penting dalam menganalisis perilaku *doom spending*. Belanja *impulsif* sering kali tidak terencana, dan fitur ini akan memberikan wawasan terkait kebiasaan mahasiswa dalam pengelolaan pengeluaran mereka.
- d. Belanja Tanpa Kebutuhan : merupakan ciri khas dari perilaku *doom spending*, di mana seseorang membeli barang-barang yang tidak diperlukan. Fitur ini relevan untuk mengukur kecenderungan konsumsi yang berlebihan tanpa alasan yang jelas.
- e. Pengeluaran Bulanan : adalah fitur yang penting karena menggambarkan seberapa besar pengeluaran seseorang dalam kehidupan sehari-hari. Ini terkait langsung dengan *doom spending*, yang sering kali melibatkan pengeluaran yang lebih besar dari pendapatan.
- f. Pemicu Belanja *Impulsif* : berfungsi untuk mengidentifikasi faktor-faktor atau situasi yang mendorong individu untuk membeli barang tanpa perencanaan. Fitur ini dapat memberikan pemahaman lebih dalam tentang perilaku konsumsi yang tidak terkendali.
- g. Tempat / Platform Belanja : mengacu pada lokasi atau platform tempat mahasiswa melakukan pembelian, seperti belanja online atau offline. Platform atau tempat yang digunakan dapat mempengaruhi jenis dan frekuensi pembelian, serta kecenderungan belanja impulsif.
- h. Masalah Finansial : merupakan variabel yang menggambarkan tantangan atau kesulitan keuangan yang dihadapi oleh individu. Ini dapat membantu mengidentifikasi hubungan antara masalah keuangan dan kecenderungan untuk terjebak dalam perilaku *doom spending*, karena individu yang memiliki masalah finansial mungkin lebih rentan terhadap pengelolaan keuangan yang buruk

Pemilihan fitur yang relevan merupakan langkah krusial dalam pra-pemrosesan data. Dengan memilih atribut yang tepat, kita dapat memastikan bahwa model yang dibangun akan memiliki basis data yang kuat dan mampu memberikan prediksi yang akurat terkait perilaku *doom spending* pada generasi Z.

3. Standarisasi data

Standarisasi data adalah bagian dari tahapan transformasi/pra-pemrosesan data yang harus dilakukan sebelum tahap modeling. Langkah ini memastikan bahwa data memiliki skala yang seragam, yang sangat penting untuk algoritma machine learning yang sensitif terhadap skala data, seperti *K-Means Clustering*.

```
# Normalisasi data
scaler = StandardScaler()
df_scaled = scaler.fit_transform(df_selected)
```

Gambar4.21 Code normalisasi data

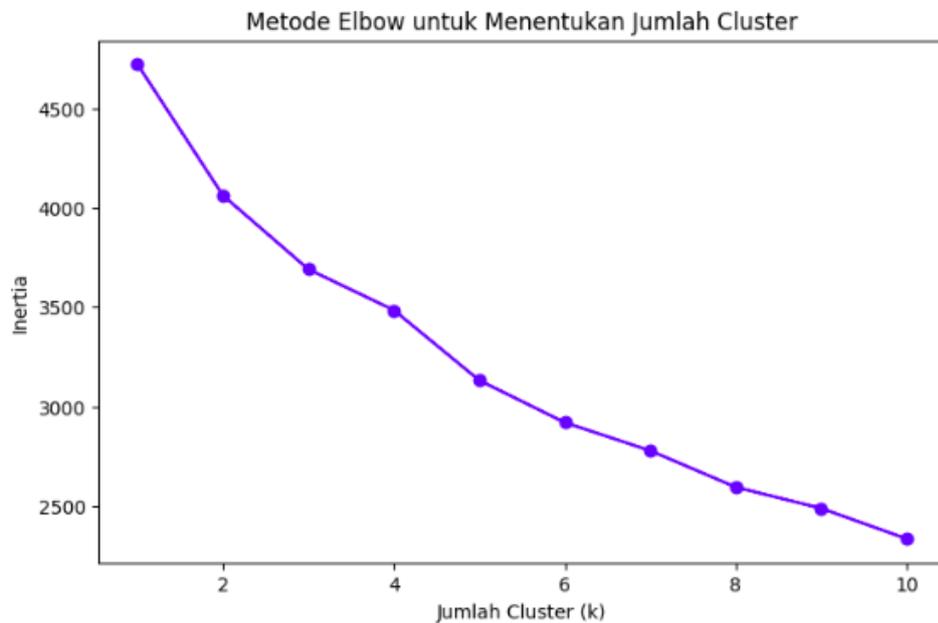
Kode di atas digunakan untuk menormalkan data numerik (yang sudah dipilih sebelumnya). *Standardscaler* adalah salah satu metode yang digunakan untuk normalisasi atau standarisasi data numerik agar setiap fitur berada pada skala yang seragam dengan rata-rata 0 dan standar deviasi 1. Hal ini akan memastikan bahwa algoritma *machine learning* yang digunakan nantinya tidak terpengaruh oleh perbedaan skala antara fitur-fitur yang ada.

4.4 Modeling

4.4.1 Penentuan Jumlah Cluster

Penentuan jumlah cluster menggunakan metode elbow, metode elbow adalah teknik yang digunakan dalam analisis cluster untuk menentukan jumlah kelompok (*cluster*) yang paling tepat untuk suatu dataset. Tujuannya

adalah untuk menemukan titik optimal dimana penambahan cluster tidak lagi memberikan peningkatan signifikan dalam kualitas model.



Gambar 4.22 Visualisasi metode elbow

Analisis metode elbow menunjukkan bahwa jumlah cluster optimal untuk data penelitian ini adalah 3. Hal ini terlihat dari penurunan WCSS (*Within-cluster sum of squares*) yang semakin melambat setelah jumlah cluster mencapai 3. Dengan demikian, data penelitian dibagi menjadi 3 kelompok yang berbeda. Hasil visualnya menunjukkan bahwa terdapat tiga segmen pola perilaku yang berbeda-beda. Pemilihan 3 cluster ini dianggap optimal karena dapat mengidentifikasi pola yang jelas dalam data tanpa mengorbankan interpretabilitas hasil.

4.4.2 Tahapan K-Means Clustering

Teknik *clustering* merupakan suatu pendekatan analisis data yang bertujuan untuk mengelompokkan entitas data secara otomatis berdasarkan kemiripan karakteristiknya, tanpa adanya label kelas yang telah ditentukan sebelumnya. *Clustering*, seringkali diidentifikasi sebagai metode *unsupervised learning*, memberikan kemampuan untuk mengklasifikasikan

data yang belum memiliki label kelas yang jelas atau terdefinisi. Metode ini menjadi esensial dalam konteks analisis data kompleks, di mana informasi tentang struktur kelas tidak diketahui atau sulit ditentukan. Proses clustering melibatkan identifikasi pola-pola intrinsik dalam data, dengan tujuan membentuk kelompok-kelompok homogen yang memiliki tingkat kemiripan tinggi di antara anggotanya.

Adapun pada penelitian ini, saya menggunakan klasterisasi metode k-means. Algoritma ini bekerja dengan menjalankan perintah banyak kelas klaster (k) yang diinginkan. Selain itu, k-means bekerja dengan menunggu jumlah pengulangan eksekusi perintah, yang dikenal dengan istilah iterasi, maksimum dan sebuah nilai konstanta yang digunakan untuk penentuan kapan iterasi dihentikan.

Setelah jumlah cluster ditentukan, langkah selanjutnya adalah menerapkan algoritma K-Means clustering pada data yang sudah dinormalisasi. K-Means clustering bekerja dengan cara mengelompokkan data ke dalam k cluster berdasarkan kedekatannya dengan pusat cluster (centroid). Setiap data akan dikelompokkan ke dalam cluster yang memiliki jarak terdekat dengan centroid tersebut. . Berikut penggunaan algoritma k-means pada penelitian ini dengan nilai K= 3 seperti yang ditampilkan pada gambar 4.23 dan gambar 4.24

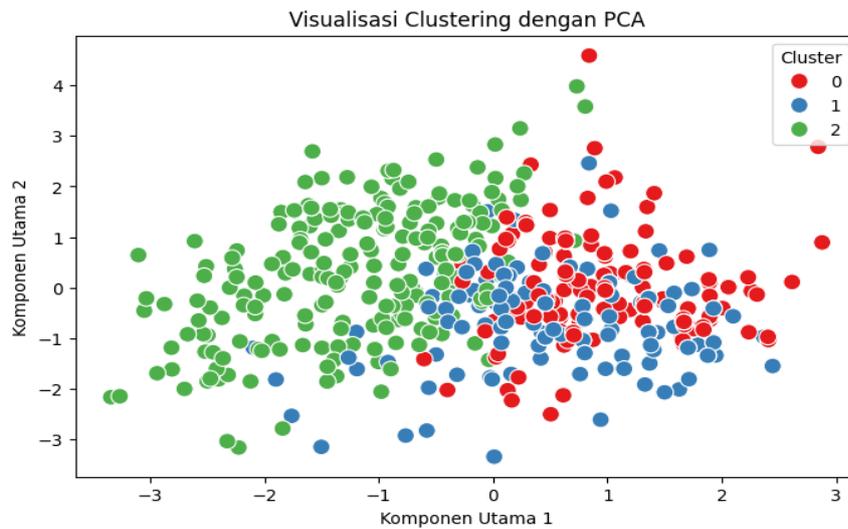
```
# Clustering menggunakan K-Means (misalnya, 3 cluster)
kmeans = KMeans(n_clusters=3, random_state=42)
clusters = kmeans.fit_predict(df_scaled)

# Menambahkan hasil clustering ke data asli
df['cluster'] = clusters

# Analisis hasil clustering
# Visualisasi PCA untuk 2D plot
pca = PCA(n_components=2)
df_pca = pca.fit_transform(df_scaled)

plt.figure(figsize=(8, 5))
sns.scatterplot(x=df_pca[:, 0], y=df_pca[:, 1], hue=df['cluster'], palette='Set1', s=100)
plt.title('Visualisasi Clustering dengan PCA')
plt.xlabel('Komponen Utama 1')
plt.ylabel('Komponen Utama 2')
plt.legend(title='Cluster')
plt.show()
```

Gambar 4.23 Code metode k-means clustering



Gambar 4.24 Visualisasi clustering dengan PCA

Dapat dilihat pada visualisasi clustering dengan PCA tersebut terdapat 3 cluster optimal yang telah ditentukan nilai K sebelumnya menggunakan metode elbow, dapat dilihat bahwa cluster 0 persebarannya ditandai dengan warna merah, cluster 1 berwarna biru dan cluster 2 berwarna hijau.

Setelah dilakukan pengolahan data dan penerapan K-Means Clustering, berikut adalah persentase per cluster yang dihasilkan dari analisis:

Tabel 4.2 Persentase per cluster

Cluster	Jumlah Data	Persentase
0	201	34.067797
1	151	25.593220
2	238	40.338983

Tabel di atas menunjukkan jumlah data dan persentase data yang termasuk dalam masing-masing kluster setelah dilakukan analisis K-Means Clustering dengan jumlah kluster (K) = 3.

Cluster 0: Memiliki 201 data, yang berarti sekitar 34.07% dari total data berada dalam kluster ini.

Cluster 1: Memiliki 151 data, yang berarti sekitar 25.59% dari total data berada dalam kluster ini.

Cluster 2: Memiliki 238 data, yang berarti sekitar 40.34% dari total data berada dalam kluster ini.

Persentase ini memberikan gambaran mengenai distribusi data dalam setiap kluster, yang membantu dalam memahami seberapa besar proporsi masing-masing kluster dalam keseluruhan dataset.

4.5 Hasil Analisis Clustering dengan Nilai K = 3

Analisis Cluster Berdasarkan hasil analisis menggunakan metode K-Means clustering, data generasi Z berhasil dikelompokkan ke dalam tiga cluster dengan karakteristik perilaku yang berbeda-beda terkait dengan doom spending. Setiap cluster menggambarkan kelompok dengan pola perilaku yang berbeda dalam pengelolaan keuangan, frekuensi belanja impulsif, serta tingkat pengeluaran.

Tabel 4.3 Hasil rata-rata analisis cluster

Cluster	Usia	Jenis Kelamin	Pendapatan bulanan	Frekuensi Belanja Tidak Terencana	Belanja Tanpa Kebutuhan	Pengeluaran Bulanan	Pemicu Belanja Impulsif	Tempat / Platform Belanja	Masalah Finansial
0	22.920398	0.492537	2.349502e+06	2.044776	2.009950	1.547264e+06	13.014925	2.333333	2.094527
1	22.807947	0.430464	2.294702e+06	1.993377	1.854305	1.389073e+06	3.304636	3.311258	2.125828
2	21.974790	0.668067	1.645485e+06	1.701681	1.710084	6.882367e+05	11.159664	4.857143	1.886555

1. Cluster 0: Kelompok dengan Doom Spending yang Tinggi (High risk doom spending)

- Cluster 0 mencerminkan kelompok yang menunjukkan tingkat doom spending yang cukup tinggi. Dalam kelompok ini, frekuensi belanja impulsif, pengeluaran bulanan, dan belanja tanpa kebutuhan tercatat

tinggi. Meskipun penghasilan bulanan mereka terbilang relatif besar, dengan rata-rata pendapatan sekitar Rp1.749.466,31, perilaku belanja mereka menunjukkan kecenderungan untuk mengeluarkan uang secara impulsif dan tidak terencana.

- Rata-rata frekuensi belanja impulsif di cluster ini mencapai 37,8%, yang menunjukkan bahwa mereka cenderung melakukan belanja impulsif lebih sering dibandingkan dengan kelompok lainnya. Selain itu, mereka juga lebih sering terlibat dalam belanja tanpa kebutuhan, yang tercermin dari nilai belanja tanpa kebutuhan yang mencapai 43,6%.
- Meskipun pendapatan mereka tinggi, kelompok ini lebih rentan terhadap masalah finansial dengan rata-rata 52,2% yang menunjukkan bahwa meskipun pengeluaran mereka tinggi, mereka tidak dapat mengelola keuangan mereka secara optimal. Dalam jangka panjang, pola belanja seperti ini dapat menimbulkan masalah finansial yang serius jika tidak dikelola dengan bijak.

2. Cluster 1: Kelompok dengan Pengelolaan Keuangan Moderat (Moderate risk doom spending)

- Cluster 1 menggambarkan kelompok yang lebih terkontrol dalam perilaku belanja, namun masih memiliki potensi untuk melakukan doom spending. Kelompok ini memiliki rata-rata pendapatan bulanan yang lebih tinggi dibandingkan dengan Cluster 2, yakni sekitar Rp2.417.229,73. Meskipun pengeluarannya lebih moderat, kelompok ini masih menunjukkan kecenderungan untuk melakukan belanja impulsif dengan rata-rata frekuensi belanja impulsif sebesar 41,2%.
- Pengeluaran bulanan kelompok ini relatif lebih tinggi, dengan rata-rata Rp1.576.013,51, namun masih dapat dikatakan berada dalam kisaran yang lebih terkontrol jika dibandingkan dengan Cluster 0. Belanja tanpa kebutuhan pada Cluster 1 tercatat pada 39,4%, yang menunjukkan

bahwa mereka melakukan belanja tanpa kebutuhan dengan frekuensi yang lebih rendah dibandingkan Cluster 0.

Namun demikian,

- kelompok ini tetap memiliki risiko terhadap doom spending jika perilaku impulsif dan pengeluaran yang lebih besar ini tidak dikelola dengan lebih baik di masa depan. Rata-rata masalah finansial pada kelompok ini berada pada angka 41,2%, yang menunjukkan adanya potensi masalah finansial meskipun mereka lebih terkontrol dalam pengelolaan keuangan dibandingkan dengan Cluster 0.

3. Cluster 2: Kelompok dengan Pengelolaan Keuangan Baik (Low risk doom spending)

- Cluster 2 menunjukkan kelompok dengan perilaku yang lebih bijak dalam hal pengelolaan keuangan dan belanja. Kelompok ini memiliki pendapatan bulanan yang lebih rendah dibandingkan dengan Cluster 1 dan Cluster 0, dengan rata-rata Rp1.727.572,06, tetapi mereka lebih mampu mengelola pengeluaran mereka dengan baik.
- Frekuensi belanja impulsif di Cluster 2 terbilang rendah dengan 28,4%, serta mereka memiliki tingkat belanja tanpa kebutuhan yang jauh lebih rendah dibandingkan dengan cluster lainnya, yakni 18,6%. Hal ini menunjukkan bahwa kelompok ini cenderung lebih selektif dan rasional dalam melakukan pembelian, yang berkontribusi pada pengelolaan keuangan yang lebih baik.
- Selain itu, kelompok ini juga menunjukkan masalah finansial yang lebih rendah dengan rata-rata 17%, yang menunjukkan bahwa mereka lebih stabil secara finansial dan lebih kecil kemungkinannya untuk terjebak dalam pola doom spending dalam jangka panjang.

4.5.1 Perbandingan Rata-rata Cluster

Berdasarkan hasil analisis clustering, diperoleh rata-rata persentase untuk masing-masing cluster sebagai berikut:

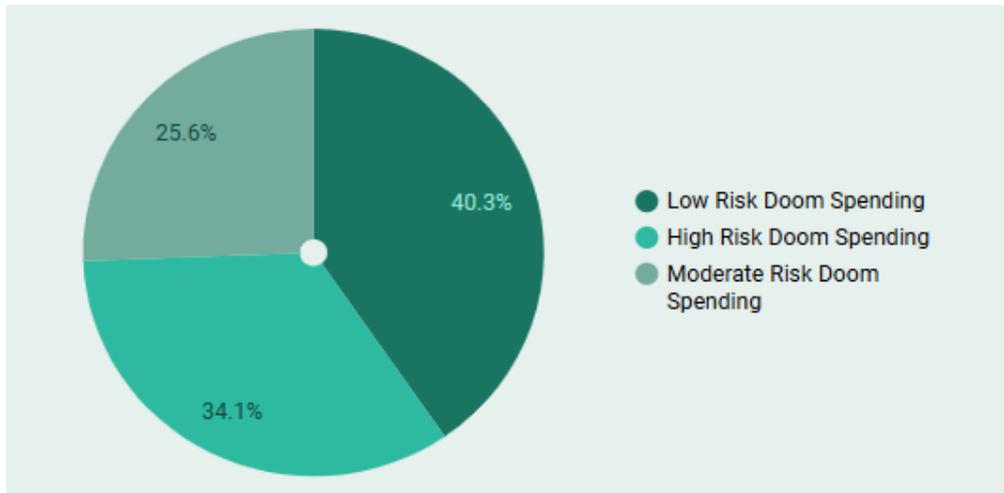
```
Cluster
2      40.338983
0      34.067797
1      25.593220
Name: count, dtype: float64
```

- Cluster 0: 34.07 % - menunjukkan kelompok dengan perilaku belanja impulsif yang lebih tinggi, pengeluaran lebih besar, dan cenderung menghadapi masalah finansial.
- Cluster 1: 25.59 % - kelompok yang lebih moderat dalam belanja, namun masih memiliki potensi untuk terjebak dalam doom spending dengan pengeluaran yang lebih tinggi dan risiko masalah finansial yang lebih besar.
- Cluster 2: 40.34 % - kelompok yang memiliki pengelolaan keuangan yang lebih baik, dengan pengeluaran lebih rendah dan frekuensi belanja impulsif yang lebih jarang, sehingga risiko terhadap doom spending sangat rendah.

4.6 Hasil Analisis Visualisasi

Tahap visualisasi dashboard peneliti menggunakan tools looker studio untuk membuat dan menampilkan data penelitian berdasarkan visualisasi agar mempermudah melihat temuan analisis dari penelitian.

1. Visualisasi Data Berdasarkan Cluster



Gambar 4.25 Visualisasi data cluster

Diagram lingkaran ini memberikan informasi penting mengenai proporsi responden yang termasuk dalam masing-masing kategori perilaku doom spending. Ketiga kategori tersebut adalah :

a. Hight Risk Doom Spending

Kelompok ini memiliki proporsi terbesar kedua dalam diagram, yaitu 34.1%. Ini menunjukkan bahwa sekitar sepertiga responden memiliki risiko tinggi terhadap perilaku doom spending, mereka sangat rentan terhadap pembelian impulsif .

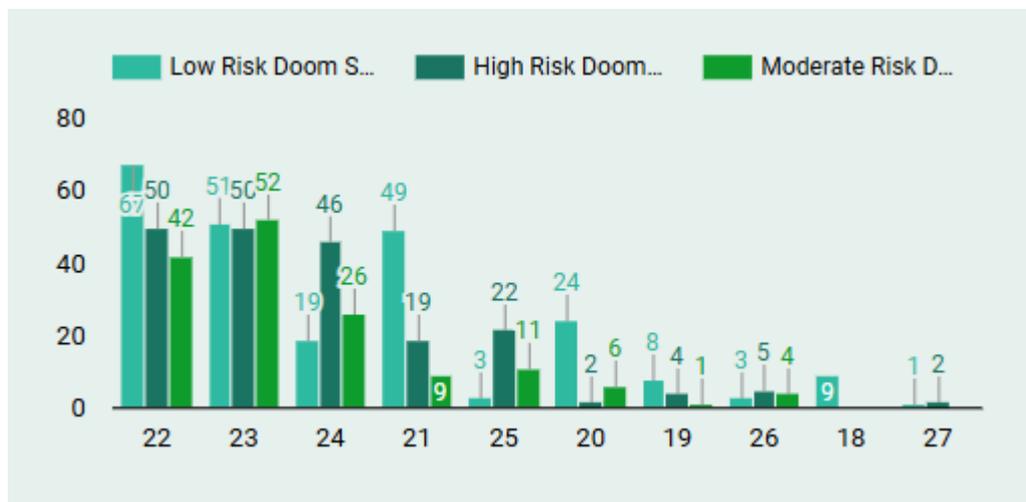
b. Moderat Doom Spending

Kelompok ini memiliki proporsi terkecil dalam diagram, yaitu 25.6%. Ini menunjukkan bahwa sekitar seperempat responden berada pada kategori risiko sedang, mereka memiliki potensi perilaku doom spending tetapi tidak separah kelompok risiko tinggi.

c. Low Risk Doom Spending

Kelompok ini memiliki proporsi terbesar dalam diagram, yaitu 40.3%. Ini menunjukkan bahwa mayoritas responden dalam penelitian ini tergolong memiliki risiko rendah terhadap perilaku doom spending. Mereka cenderung lebih terkendali dalam berbelanja dan tidak terlalu di pengaruhi oleh emosi negative dalam keputusan pembelian.

2. Visualisasi Data Usia Berdasarkan Masing - Masing Cluster



Gambar 4.26 Visualisasi data usia

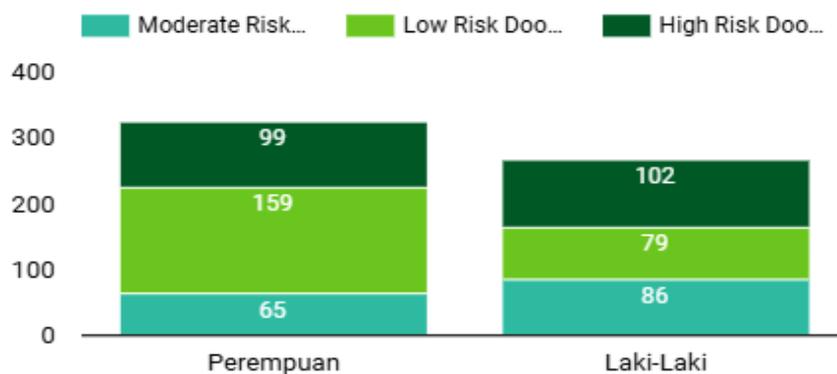
Diagram batang ini menyajikan informasi mengenai distribusi usia responden dalam penelitian, beserta jumlah responden untuk setiap kelompok usia berdasarkan cluster atau kategori kelompok doom spending.

- Pada kelompok usia 22 dan 23 tahun, jumlah responden dari kategori low risk doom spending dan high doom spending hampir sama . Namun pada kelompok usia 22 tahun, jumlah responden dari kategori moderate risk doom spending lebih tinggi dibandingkan kelompok usia 23 tahun.
- Pada kelompok usia 24 tahun, jumlah responden dari kategori high risk doom spending lebih tinggi dibandingkan kategori lainnya.

- c. Kelompok usia yang lebih tua (25 tahun ke atas) dan kelompok usia yang lebih muda (20 tahun kebawah), jumlah responden dari kategori low risk doom spending lebih tinggi dibandingkan kategori lainnya.

Diagram ini memberikan gambaran informasi bahwa pada kelompok usia lebih muda (sekitar 22-24 tahun), risiko terhadap perilaku doom spending lebih tinggi , terutama pada kategori high risk doom spending. Pada kelompok usia yang lebih tua dan lebih muda, risiko terhadap perilaku doom spending lebih rendah dan didominasi oleh kategori low risk doom spending.

3. Visualisasi Jumlah Data Kategori Cluster Berdasarkan Jenis kelamin



Gambar 4.27 Visualisasi data jenis kelamin

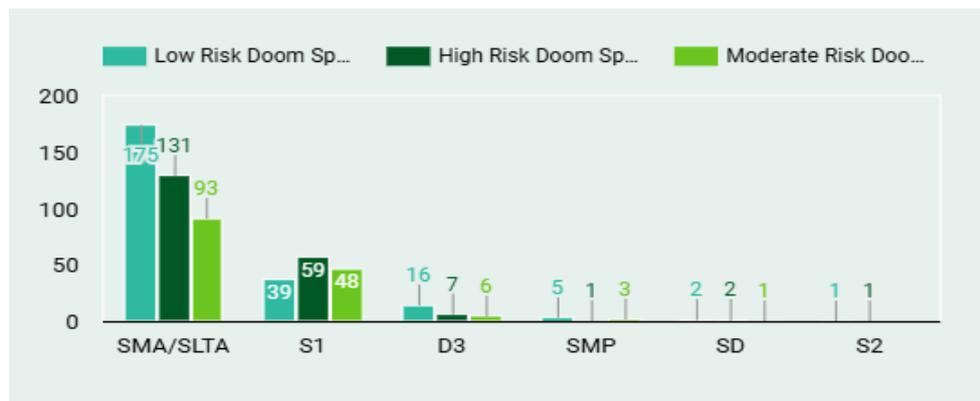
Diagram pada gambar 4.27 memberikan informasi mengenai jumlah responden berdasarkan jenis kelamin dan kategori perilaku doom spending.

- a. Perempuan ; Jumlah data responden keseluruhan berdasarkan jenis kelamin perempuan yaitu terdapat 323 responden. Perempuan dengan kategori high risk doom spending berjumlah 99 orang, dan yang tergolong kategori moderat doom spending berjumlah 65 orang, serta 159 orang termasuk dalam kategori low risk doom spending.

- b. Laki – laki : Jumlah responden keseluruhan laki-laki yaitu 267 responden. Laki-laki dengan kategori high risk doom spending berjumlah 102 orang, dan yang tergolong kategori modert doom spending berjumlah 79 orang, serta 86 orang termasuk dalam kategori low risk doom spending.

Diagram ini menunjukkan bahwa terdapat perbedaan yang signifikan dalam proporsi responden laki-laki dan perempuan dalam setiap kategori perilaku doom spending. Laki-laki cenderung lebih banyak mengalami doom spending, terutama pada kategori risiko tinggi. Sementara itu perempuan lebih banyak yang tergolong kedalam kelompok kategori risiko rendah terhadap doom spending.

4. Visualisasi Jumlah Data Kategori Cluster Berdasarkan Pendidikan Terakhir



Gambar 4.28 Visualisasi data pendidikan

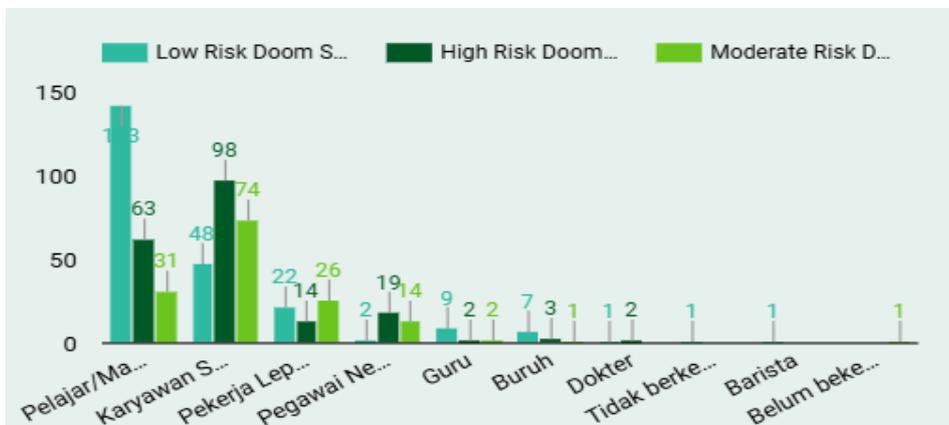
Diagram pada gambar 4.28 memberikan informasi mengenai jumlah responden berdasarkan tingkat pendidikan terakhir yang mereka tempuh, serta kategori risiko doom spending.

- a. SMA/SLTA : Jumlah responden dengan risiko doom spending rendah lebih banyak yaitu terdapat 175 responden, dibandingkan risiko sedang dengan jumlah 93 responden dan risiko tinggi berjumlah 131 responden.

- b. S1: Jumlah responden dengan risiko doom spending tinggi lebih banyak yaitu berjumlah 59 responden dibandingkan dengan risiko doom spending sedang berjumlah 48 responden, dan responden dengan risiko rendah berjumlah 39 responden.
- c. Pada tingkat pendidikan D3,SMP,SD dan S2, jumlah responden dari kategori risiko doom spending rendah lebih banyak dibandingkan kategori risiko lainnya.

Data dalam visualisasi ini menunjukkan bahwa tingkat pendidikan SMA/SLTA memiliki jumlah responden terbanyak dengan risiko doom spending rendah. Namun pada tingkat pendidikan S1 risiko doom spending tinggi lebih banyak.

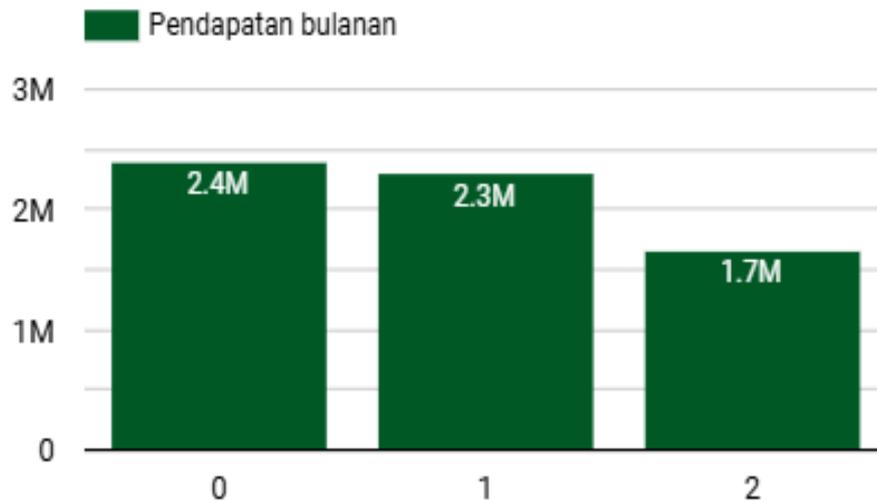
5. Visualisasi Data Pekerjaan Berdasarkan Kategori Cluster Doom Spending



Gambar 4.29 Visualisasi data pekerjaan

Pada Diagram gambar 4.29 tersebut menunjukkan informasi mengenai jumlah responden berdasarkan pekerjaan mereka, serta kategori risiko doom spending. Data tersebut menunjukkan bahwa siswa/ mahasiswa cenderung memiliki risiko doom spending rendah, sedangkan karyawan swasta cenderung memiliki risiko doom spending tinggi.

6. Visualisasi Rata-Rata Pendapatan Bulanan Berdasarkan Cluster



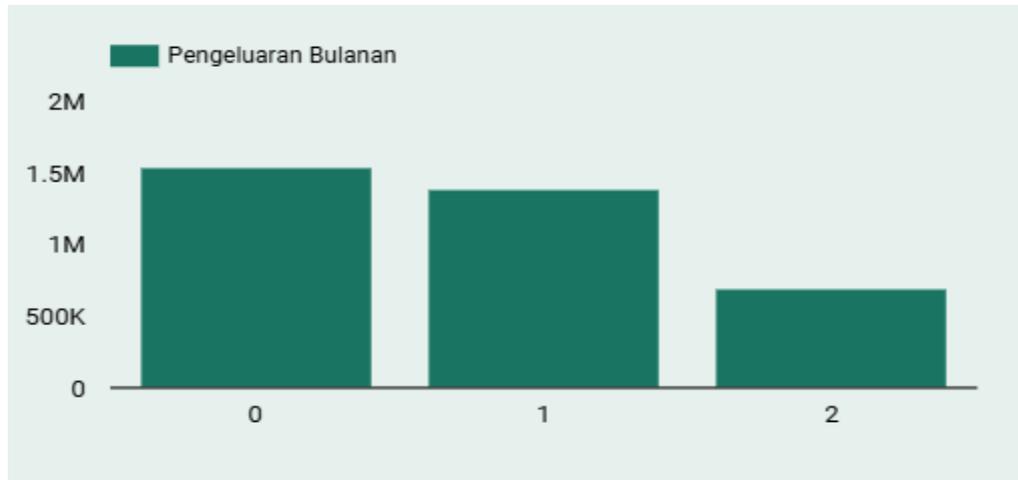
Gambar 4.30 Visualisasi data rata-rata pendapatan bulanan

Tujuan dari visualisasi ini untuk melihat perbandingan rata-rata pendapatan bulanan antar *cluster* .

- Cluster 0 Memiliki rata – rata pendapatan bulanan tertinggi, yaitu sekitar 2,4 juta rupiah
- Cluster 1 Memiliki rata-rata pendapatan bulanan sedang, yaitu sekitar 2,3 juta rupiah.
- Cluster 2 Memiliki rata-rata pendapatan bulanan terendah, yaitu sekitar 1.7 juta rupiah.

Berdasarkan data bahwa terdapat perbedaan signifikan dalam pendapatan bulanan antara ketiga cluster risiko doom spending.

7. Visualisasi Rata – Rata Pengeluaran Bulanan Berdasarkan Cluster



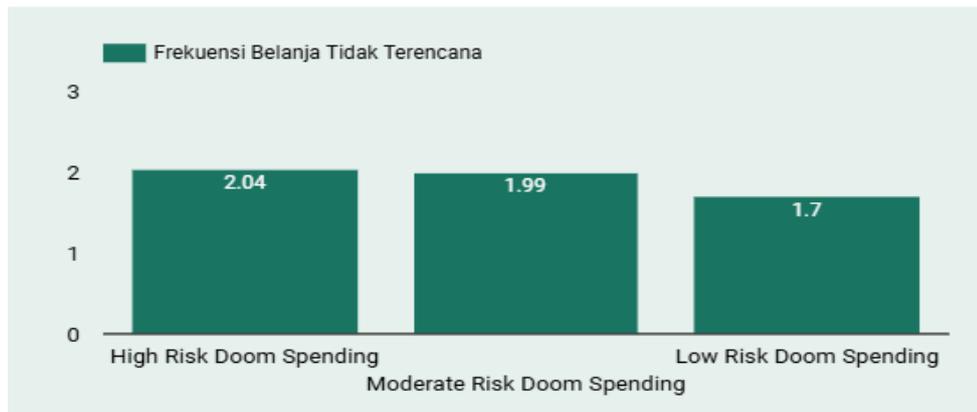
Gambar 4.31 Visualisasi data rata-rata pengeluaran bulanan

Tujuan dari visualisasi ini untuk melihat perbandingan rata-rata pengeluaran bulanan antar cluster .

- d. Cluster 0 Memiliki rata – rata pengeluaran bulanan tertinggi, yaitu sekitar 1,5 juta rupiah
- e. Cluster 1 Memiliki rata-rata pengeluaran bulanan sedang, yaitu sekitar 1,3 juta rupiah.
- f. Cluster 2 Memiliki rata-rata pengeluaran bulanan terendah, yaitu sekitar 600 ribu rupiah.

Berdasarkan data bahwa terdapat perbedaan signifikan dalam pengeluaran bulanan antara ketiga cluster risiko doom spending. Cluster dengan risiko doom tertinggi yaitu cluster 0 memiliki pengeluaran bulanan rata – rata tertinggi, sedangkan cluster dengan risiko doom spending terendah yaitu cluster 2 memiliki pengeluaran bulanan rata-rata terendah, berarti cluster 1 termasuk kedalam kategori risiko doom spending sedang.

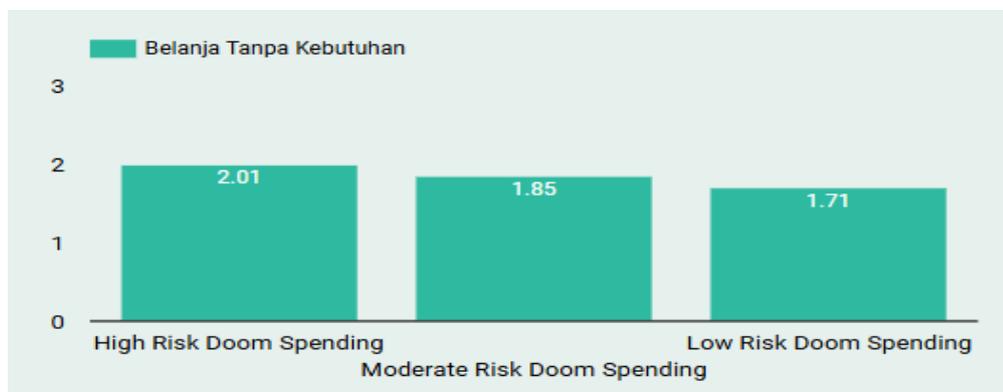
8. Visualisasi Rata-Rata Frekuensi Belanja Tidak Terencana Berdasarkan Kategori Doom Spending



Gambar 4.32 Visualisasi data rata-rata belanja tidak terencana

Berdasarkan data dari visualisasi tersebut memberikan pemahaman bahwa perilaku yang termasuk ke dalam kelompok kategori risiko doom spending tinggi (High Risk Doom Spending), lebih sering melakukan aktifitas belanja tidak terencana dari pada kelompok kategori risiko doom spending yang lain. Hal ini dapat dilihat bahwa ada perbedaan signifikan dalam frekuensi belanja tidak terencana antara ketiga cluster risiko doom spending.

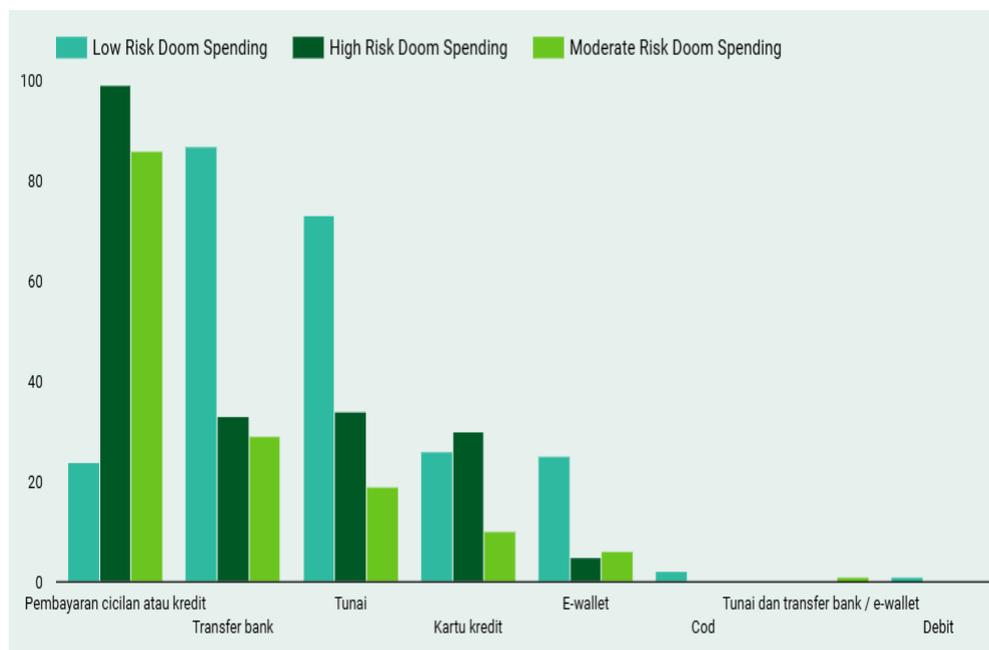
9. Visualisasi Rata-Rata Belanja Tanpa Kebutuhan Berdasarkan Kategori Doom Spending



Gambar 4.33 Visualisasi data rata-rata belanja tanpa kebutuhan

Berdasarkan data visualisasi gambar 4.33 menunjukkan bahwa ada perbedaan signifikan dalam rata-rata belanja tanpa kebutuhan antara ketiga cluster risiko doom spending tertinggi memiliki rata-rata belanja tanpa kebutuhan tertinggi dibandingkan dengan kategori doom spending yang lain.

10. Visualisasi Metode Pembayaran Berdasarkan Kategori Doom Spending

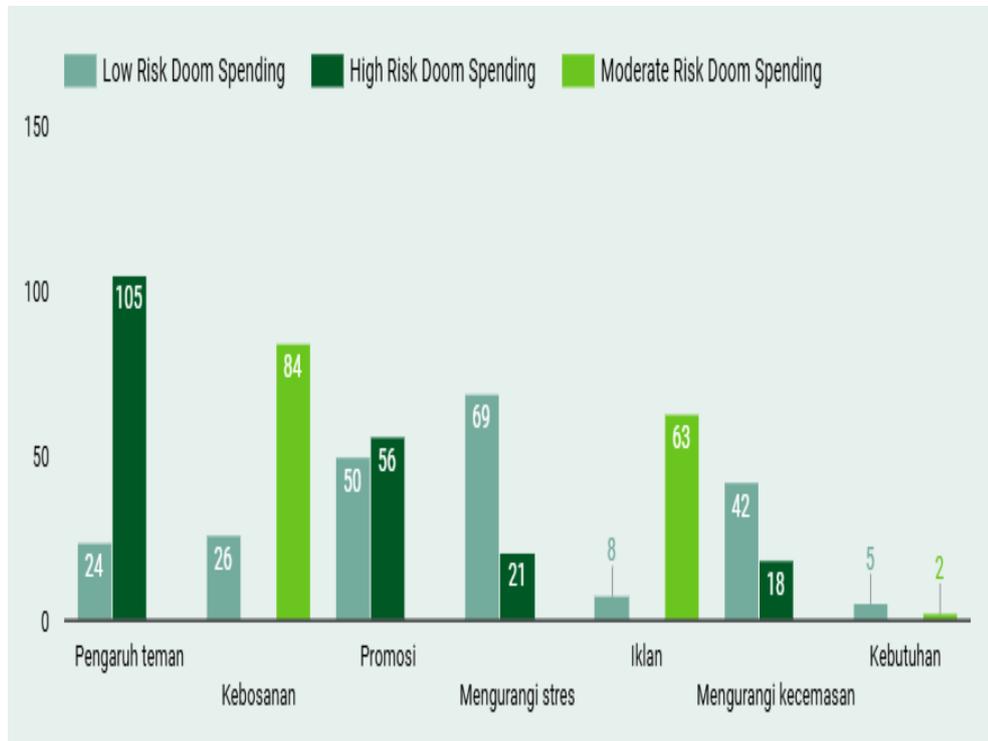


Gambar 4.34 Visualisasi data pembayaran kategori doom spending

Diagram ini menunjukkan bahwa pembayaran cicilan atau kredit (misalnya menggunakan pembayaran Spaylater, Gopaylater, dan pembayaran yang menggunakan cicilan lainnya) merupakan metode pembayaran yang paling banyak digunakan oleh responden dengan risiko doom spending tinggi. Sedangkan transfer bank merupakan metode pembayaran yang paling banyak digunakan oleh responden dengan risiko doom spending rendah.

Data ini dapat membantu memahami bagaimana metode pembayaran yang digunakan seseorang terkait dengan risiko terhadap doom spending.

11. Visualisasi Rata – Rata Pemicu Belanja Impulsif Berdasarkan Kategori Doom Spending



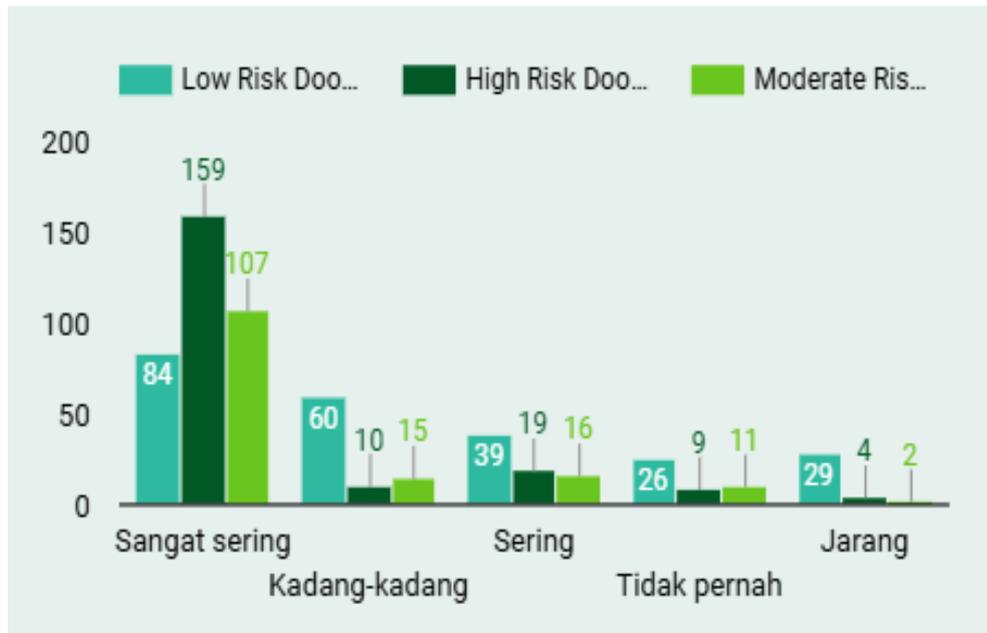
Gambar 4.35 Visualisasi data pemicu doom spending

Visualisasi ini menunjukkan bahwa kelompok dengan risiko doom spending tinggi (High Risk Doom Spending) rata-rata pemicu belanja impulsif nya paling tinggi di dasari oleh pengaruh teman dengan jumlah responden sebanyak 105 orang .

Pada kelompok dengan risiko doom spending sedang (Moderate Risk Doom Spending) rata-rata pemicu belanja impulsifnya yaitu karena kebosanan dengan jumlah responden 84 orang.

Selanjutnya yang terakhir yaitu pada kelompok dengan kategori risiko doom spending rendah(Low Risk Doom Spending) rata-rata pemicu belanja impulsifnya yang paling tinggi dengan jumlah responden 69 orang pemicunya untuk mengurangi stress.

12. Visualisasi Rata-Rata Masalah Finansial Berdasarkan Kategori Doom Spending.



Gambar 4.36 Visualisasi data rata-rata masalah finansial

Visualisasi ini menampilkan data bahwa responden dengan risiko doom spending tinggi (High Risk Doom Spending) lebih sering mengalami masalah finansial dibandingkan responden dengan risiko doom spending sedang (Moderate Risk Doom Spending) dan rendah (Low Risk Doom Spending).