

LAPORAN UJIAN TENGAH SEMESTER
ANALISIS VISUALISASI DATA PRAKTIKUM



KELOMPOK 10

ANGGOTA:

[082111633001] | [Fariska Dwi Kartika Sari]

[082111633010] | [Erika Dwi Puspitasari]

[082111633018] | [Perwira Annissa Dyah Permatasari]

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA
SURABAYA

2023

a. Import library

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
```

b. Membaca data

Syntax:

```
data_survei = pd.read_csv("DataSurveiGame.csv")
data = pd.DataFrame(data_survei)
data.head()
```

Output:

A1	B	C	D	E	F	G	H	I	J	K	L	M		
1	Usia	Jenis Kelamin	Pendidikan	1. Apakah Anda	2. Seberapa sering	3. Berapa lama	4. Apa jenis game	5. Apakah Anda	6. Bagaimana	7. Seberapa sering	8. Apa yang	9. Apakah bermain	10. Jika YA, beri 1	
2	21 Laki-laki	Mahasiswa	Ya	Beberapa kali se 1-2 jam	MOBA (Mobile L Tidak	Tinggi	Normal	masuk kelas	Tidak	sebenarnya cuci M	-	-	-	
3	21 Laki-laki	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 n FPS (Valorant, C	Normal	Normal	Realisasikan ekspe	Tidak	Melihat funny vic N	-	-	-	
4	20 Perempuan	Mahasiswa	Ya	Beberapa kali se 30 menit - 1 jam	Sport (8 Ball Pool Netral	Sangat Tinggi	Sering	mata kuliah sem	Tidak	scroll media sos	N	-	-	
5	20 Laki-laki	Mahasiswa	Ya	Beberapa kali se Kurang dari 30 n RPG (Genshin I	Ya	Tinggi	Sering	Tugas kuliah	Tidak	Tidur	N	-	-	
6	21 Perempuan	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 n Arcade (Subway Netral	Tinggi	Sering	tugas kuliah	Tidak	-	-	-	-	
7	19 Perempuan	Mahasiswa	Ya	Jarang	30 menit - 1 jam Arcade (Subway Netral	Tinggi	Sangat sering	Tugas proker	Tidak	Nonton atau cha	M	-	-	
8	20 Perempuan	Mahasiswa	Ya	Setiap hari	30 menit - 1 jam Arcade (Subway Ya	Tinggi	Normal	Tugas Kuliaaaa	Tidak	Scroll X JD	S	-	-	
9	19 Laki-laki	Mahasiswa	Tidak	Jarang	30 menit - 1 jam MOBA (Mobile L Netral	Tinggi	Sering	Kuliah dan aktivi	Tidak	Social media at	M	-	-	
10	17 Laki-laki	Umum / Sudah t	Ya	Setiap hari	Lebih dari 2 jam MOBA (Mobile L Tidak	Normal	Tidak pernah	Nggak punya duit Ya	Tidak	Kepo	S	-	-	
11	18 Perempuan	Mahasiswa	Ya	Beberapa kali se 30 menit - 1 jam Sport (8 Ball Pool Tidak	Sangat Tinggi	Sering	tugas kuliah	Ya	karena bermain	M	-	-	-	
12	20 Perempuan	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 n Arcade (Subway Ya	Tinggi	Sering	TUGAS 🎯🎯	Tidak	Scrolling twiter, Ti	-	-	-	
13	20 Laki-laki	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 n MOBA (Mobile L Netral	Sangat Tinggi	Sangat sering	Banyaknya tuga	Tidak	Berkendara men M	-	-	-	
14	20 Laki-laki	Mahasiswa	Ya	Jarang	30 menit - 1 jam Sport (8 Ball Pool Tidak	Sangat Tinggi	Sangat sering	Hubungan, Orga	Tidak	Menangis	K	-	-	
15	20 Laki-laki	Mahasiswa	Ya	Jarang	Kurang dari 30 n Sport (8 Ball Pool Netral	Tinggi	Sering	Tugas kuliah	Tidak	Streaming YouTi	N	-	-	
16	19 Laki-laki	Mahasiswa	Ya	Jarang	1-2 jam	RPG (Genshin I Tidak	Sangat Tinggi	Sangat sering	Kuliah	Mendengarkan I	M	-	-	
17	20 Laki-laki	Mahasiswa	Ya	Jarang	Kurang dari 30 n MOBA (Mobile L Tidak	Tinggi	Normal	Tugas	Tidak	Dengerin musik	Ti	-	-	
18	20 Laki-laki	Mahasiswa	Ya	Beberapa kali se 30 menit - 1 jam Arcade (Subway Tidak	Tinggi	Normal	kelempahan tuga	Tidak	tidur	N	-	-	-	
19	20 Laki-laki	Mahasiswa	Tidak	Jarang	Kurang dari 30 n Arcade (Subway Tidak	Normal	Normal	Tugas	Tidak	Tugas	Ti	-	-	
20	20 Laki-laki	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 n MOBA (Mobile L Netral	Sangat Tinggi	Normal	Date line	Tidak	Ga suka main g	Ti	-	-	
21	18 Laki-laki	Umum / Sudah t	Ya	Setiap hari	Lebih dari 2 jam	MOBA (Mobile L Netral	Normal	Normal	ganduwe duwek	Ya	Tidur	N	-	-
22	20 Perempuan	Mahasiswa	Ya	Kurang dari 30 n Arcade (Subway Netral	Tinggi	Jarang	Tugas yang ban	Tidak	Tidur	N	-	-	-	
23	21 Perempuan	Mahasiswa	Tidak	Beberapa kali se 30 menit - 1 jam Sport (8 Ball Pool Ya	Sangat Tinggi	Sangat sering	tugasslah, ape	Tidak	nonton film sih, i	Ti	-	-	-	
24	19 Perempuan	Mahasiswa	Ya	Setiap hari	Lebih dari 2 jam RPG (Genshin I Ya	Sangat Tinggi	Sangat sering	Metpen	Ya	Mengejar rewar	S	-	-	
25	20 Laki-laki	Mahasiswa	Ya	Setiap hari	Lebih dari 2 jam Arcade (Subway Ya	Sangat Tinggi	Sangat sering	main game Pou	Tidak	mengaj	S	-	-	
26	19 Laki-laki	Mahasiswa	Ya	Beberapa kali se 1-2 jam	MOBA (Mobile L Netral	Tinggi	Normal	kuliah	Tidak	tidur	M-	-	-	
27	21 Laki-laki	Mahasiswa	Ya	Reberanya kali se 1-2 lam	MOBA (Mobile L Ya	Sangat Tinggi	Tidak pernah	Tidak arta	Tidak	Semakin mend	K	-	-	

c. Mengubah Nama Kolom

Langkah selanjutnya kami melakukan pembaruan pada nama kolom dalam DataFrame 'data' dengan menggunakan metode `.rename()`. Tujuannya adalah untuk memberikan kolom-kolom tersebut nama yang lebih deskriptif dan mudah dimengerti. Contoh paling sederhana adalah mengganti nama kolom '1. Apakah Anda bermain game online?' menjadi 'bermain_game'. Dengan menggunakan nama kolom yang lebih singkat dan deskriptif, kami membuat data lebih mudah diakses dan digunakan dalam analisis. Kami melakukan hal serupa untuk kolom-kolom lain seperti 'durasi_bermain', 'tingkat_stress', 'klasifikasi', dan lainnya, sehingga mewakili pertanyaan atau variabel dengan cara yang lebih jelas. Pembaruan nama kolom ini juga membantu dalam membuat kode lebih bersih dan mempermudah pemahaman dataset, baik untuk kami maupun pembaca lain yang mungkin akan melihat data tersebut. Selain itu, nama kolom yang lebih deskriptif dapat meminimalkan kebingungan dan kesalahan dalam analisis data, sehingga memudahkan dalam interpretasi hasil dan pengambilan keputusan berdasarkan data tersebut.

Syntax:

```
df = data.rename(columns={
    '1. Apakah Anda bermain game online?': 'bermain_game',
    '2. Seberapa sering Anda bermain game online dalam sebulan terakhir?': 'durasi_bermain',
    '3. Berapa lama waktu yang Anda habiskan setiap hari untuk bermain game online? ': 'durasi_waktu_bermain',
    '4. Apa jenis game online yang paling sering Anda mainkan? (Contoh: MOBA, FPS, RPG, dll.)' : 'jenis_game',
    '5. Apakah Anda merasa bermain game online mengganggu waktu belajar atau aktivitas Anda? ' : 'game_mengganggu_aktivitas',
    '6. Bagaimana Anda menilai tingkat rasa stres Anda sebulan terakhir? ' : 'tingkat_stress',
    '7. Seberapa sering Anda merasa cemas, depresi, tertekan, atau memiliki masalah kesehatan mental lainnya dalam sebulan terakhir? ': 'penyebab_utama_stress',
    '8. Apa yang menjadi penyebab utama stres Anda sebulan terakhir? ' : 'penyebab_utama_stress',
    '9. Apakah bermain game online menjadi pilihan pertama Anda ketika Anda merasa stres? atau apakah Anda memiliki metode lain yang efektif? ': 'alasan_pilih',
    '10. Jika YA, berikan alasan mu. Jika TIDAK, aktivitas apa yang menjadi pilihan pertama ketika merasa stres? ' : 'alasan_pilih',
    '11. Apakah Anda merasa bahwa bermain game online telah membantu Anda mengalihkan perhatian dari sumber stres yang sedang Anda alami? ': 'membantu_mengalihkan_stress',
    '12. Apakah Anda pernah mengalami situasi di mana bermain game untuk menghilangkan stres malah menambah stres atau masalah lain? ': 'membantu_mengalihkan_stress',
    '13. Bagaimana pendapat Anda tentang sejauh mana bermain game sebagai cara untuk menghilangkan stres merupakan alternatif yang efektif? ': 'pendapat_stres'
})
```

Output:

	Usia	Jenis Kelamin	Pendidikan	bermain_game	durasi_bermain	durasi_waktu_bermain	jenis_game	game_mengganggu_aktivitas	tingkat_stress
0	21	Laki-laki	Mahasiswa	Ya	Beberapa kali seminggu	1-2 jam	MOBA (Mobile Legends, DOTA, League of Legends)	Tidak	Netral
1	21	Laki-laki	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 menit	FPS (Valorant, CSGO, Point Blank)	Ya	Netral
2	20	Perempuan	Mahasiswa	Ya	Beberapa kali seminggu	30 menit - 1 jam	Sport (8 Ball Pool, FIFA, eFootball)	Netral	Netral
3	20	Laki-laki	Mahasiswa	Ya	Beberapa kali seminggu	Kurang dari 30 menit	RPG (Genshin Impact, Call of Duty, Toram)	Ya	Netral
4	21	Perempuan	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 menit	Arcade (Subway Surf, Temple Run, Pou)	Netral	Netral

d. Mengubah Nama Value Per Kolom Menjadi Data Ordinal

Syntax:

```
df['Jenis Kelamin']=np.where(df['Jenis Kelamin']=='Laki-laki',1,2)
df['bermain_game']=np.where(df['bermain_game']=='Tidak',1,2)
df['Pendidikan'] = np.where(df['Pendidikan'] == 'Pelajar', 1,
                            np.where(df['Pendidikan'] == 'Mahasiswa', 3,
                                    np.where(df['Pendidikan'] == 'Umum / Sudah bekerja', 2, 0)))
df['durasi_bermain'] = np.where(df['durasi_bermain'] == 'Setiap hari', 5,
                                 np.where(df['durasi_bermain'] == 'Beberapa kali seminggu', 4,
                                         np.where(df['durasi_bermain'] == 'Sekali seminggu', 3,
                                                 np.where(df['durasi_bermain'] == 'Jarang', 2,
                                                         np.where(df['durasi_bermain'] == 'Tidak pernah', 1,0))))
df['durasi_waktu_bermain'] = np.where(df['durasi_waktu_bermain'] == 'Lebih dari 2 jam', 4,
                                       np.where(df['durasi_waktu_bermain'] == '1-2 jam', 3,
                                               np.where(df['durasi_waktu_bermain'] == '30 menit - 1 jam', 2,
                                                       np.where(df['durasi_waktu_bermain'] == 'Kurang dari 30 menit', 1,0)))
df['game_mengganggu_aktivitas'] = np.where(df['game_mengganggu_aktivitas'] == 'Tidak', 1,
                                             np.where(df['game_mengganggu_aktivitas'] == 'Ya', 3,
                                                     np.where(df['game_mengganggu_aktivitas'] == 'Netral', 2, 0)))
df['tingkat_stress'] = np.where(df['tingkat_stress'] == 'Sangat Tinggi', 5,
                                np.where(df['tingkat_stress'] == 'Tinggi', 4,
                                        np.where(df['tingkat_stress'] == 'Normal', 3,
                                                np.where(df['tingkat_stress'] == 'Rendah', 2,
                                                        np.where(df['tingkat_stress'] == 'Sangat Rendah', 1,0))))
df['frekuensi_kesehatan_mental'] = np.where(df['frekuensi_kesehatan_mental'] == 'Tidak Pernah', 1,
                                              np.where(df['frekuensi_kesehatan_mental'] == 'Jarang', 2,
                                                      np.where(df['frekuensi_kesehatan_mental'] == 'Normal', 3,
                                                              np.where(df['frekuensi_kesehatan_mental'] == 'Sering', 4,
                                                                      np.where(df['frekuensi_kesehatan_mental'] == 'Sangat Sering', 5)))
df['pilihan_pertama_stress']=np.where(df['pilihan_pertama_stress']=='Tidak',1,2)
df['membantu_mengalihkan_stress'] = np.where(df['membantu_mengalihkan_stress'] == 'Sangat membantu', 5,
                                              np.where(df['membantu_mengalihkan_stress'] == 'Membantu', 4,
                                                      np.where(df['membantu_mengalihkan_stress'] == 'Netral', 3,
                                                              np.where(df['membantu_mengalihkan_stress'] == 'Kurang membantu', 2,
                                                                      np.where(df['membantu_mengalihkan_stress'] == 'Tidak membantu', 1)))
```

Output:

	Usia	Jenis Kelamin	Pendidikan	bermain_game	durasi_bermain	durasi_waktu_bermain	jenis_game	game_mengganggu_aktivitas	tingkat_stress
0	21	Laki-laki	Mahasiswa	Ya	Beberapa kali seminggu	1-2 jam	MOBA (Mobile Legends, DOTA, League of Legends)	Tidak	Netral
1	21	Laki-laki	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 menit	FPS (Valorant, CSGO, Point Blank)	Ya	Netral
2	20	Perempuan	Mahasiswa	Ya	Beberapa kali seminggu	30 menit - 1 jam	Sport (8 Ball Pool, FIFA, eFootball)	Netral	Setengah
3	20	Laki-laki	Mahasiswa	Ya	Beberapa kali seminggu	Kurang dari 30 menit	RPG (Genshin Impact, Call of Duty, Toram)	Ya	Setengah
4	21	Perempuan	Mahasiswa	Tidak	Tidak pernah	Kurang dari 30 menit	Arcade (Subway Surf, Temple Run, Pou)	Netral	Setengah

Setelah merubah nama variabel, langkah selanjutnya yaitu transformasi data kategorikal menjadi data numerik dalam DataFrame 'df'. Tujuan dari langkah-langkah ini adalah untuk memungkinkan analisis statistik dan pemodelan data lebih lanjut dengan variabel-variabel tersebut. Misalnya, kami telah mengganti nilai 'Laki-laki' dan 'Perempuan' pada kolom 'Jenis Kelamin' menjadi 1 dan 2, yang mencerminkan jenis kelamin dalam format numerik untuk analisis lebih lanjut. Demikian pula, kami melakukan penggantian nilai dalam berbagai kolom lainnya seperti 'Pendidikan', 'durasi_bermain', 'tingkat_stress', dan sebagainya, sehingga variabel-variabel ini memiliki representasi numerik yang sesuai untuk analisis data.

Dalam konteks ini, pendekatan penggantian ini membantu dalam menyederhanakan dan mempersiapkan data yang kami punya untuk analisis yang lebih lanjut. Bagaimanapun, sangat penting untuk memastikan bahwa penggantian nilainya mempertahankan makna asli dari data tersebut, dan hasil transformasi sesuai dengan pertanyaan penelitian atau tujuan analisis kami. Selain itu, kami harus memastikan bahwa data yang telah diubah sesuai dengan konteks dan pengetahuan domain yang relevan agar hasil analisis dapat diinterpretasikan secara benar.

Soal UTS:

- Deteksi Mising value + Penanganannya

Syntax:

```
df.isna().sum()
```

Output:

```
Usia                      0
Jenis Kelamin              0
Pendidikan                 0
bermain_game                0
durasi_bermain               0
durasi_waktu_bermain        0
jenis_game                  0
game_mengganggu_aktivitas   0
tingkat_stress                0
frekuensi_kesehatan_mental  0
penyebab_utama_stress       0
pilihan_pertama_stress      0
alasan_pilihan_aktivitas     0
membantu_mengalihkan_stress  0
game_memambah_stress         0
klasifikasi                   0
dtype: int64
```

Setelah perubahan nama variabel, kemudian dilakukan pengecekan missing value. Hasil analisis data dengan menggunakan syntax 'isna().sum()' menunjukkan bahwa tidak ada missing value dalam dataset yang digunakan. Ketiadaan missing value adalah hal yang

positif dalam analisis data, karena memastikan bahwa semua informasi dari setiap variabel telah tercatat dengan lengkap. Dengan dataset yang bebas dari missing value, analisis dan pengolahan data dapat berjalan lebih lancar dan akurat, serta meminimalkan potensi bias yang dapat muncul akibat pengabaian data yang hilang. Hal ini memberikan dasar yang kuat untuk analisis statistik dan pemodelan yang dapat diandalkan dalam mengambil keputusan atau mengambil wawasan dari data tersebut.

Sebelum melakukan visualisasi data, kami melakukan pemisahan tipe data yaitu data ordinal (dengan nama variabel 'num') dan data yang bukan ordinal (dengan nama variabel 'category'). Kami melakukan hal tersebut dengan alasan bahwa data yang berada pada variabel 'category' yang kami miliki merupakan penjelasan alasan dari beberapa pertanyaan kami lontarkan, sehingga tidak dapat dilakukan visualisasi terhadap variabel tersebut. Oleh karena itu, kami memisahkan kedua tipe data tersebut dan melakukan visualisasi dengan variabel 'num' yang berisi data ordinal.

```
category=['jenis_game','penyebab_utama_stress','alasan_pilihan_aktivitas']

category
['jenis_game', 'penyebab_utama_stress', 'alasan_pilihan_aktivitas']

num=[x for x in df.columns if x not in category]

num
['Usia',
 'Jenis Kelamin',
 'Pendidikan',
 'bermain_game',
 'durasi_bermain',
 'durasi_waktu_bermain',
 'game_mengganggu_aktivitas',
 'tingkat_stress',
 'frekuensi_kesehatan_mental',
 'pilihan_pertama_stress',
 'membantu_mengalihkan_stress',
 'game_menambah_stress',
 'klasifikasi']
```

2. Deteksi Outlier + Penanganannya

Syntax:

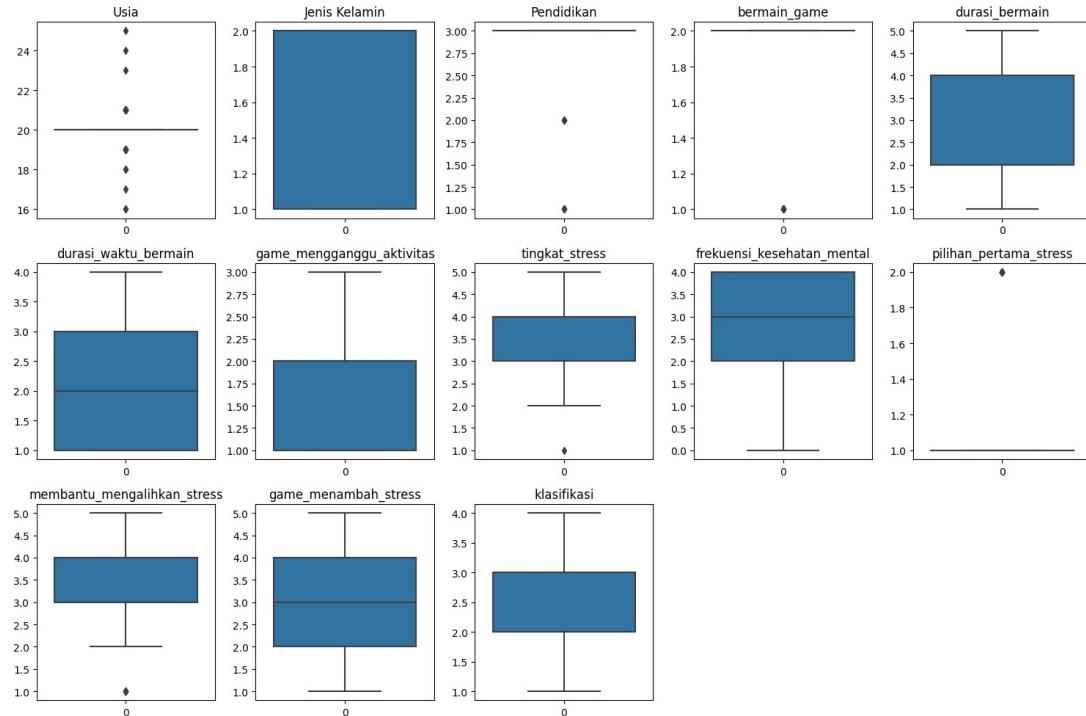
```
import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(15, 10))

for i in range(0, len(df[num].columns)):
    plt.subplot(3, 5, i + 1)
    sns.boxplot(data=df[num][df[num].columns[i]])
    plt.title(df[num].columns[i])
    plt.tight_layout()

plt.show()
```

Output:



Selanjutnya adalah pengecekan outlier, yang dapat dilakukan dengan menggunakan visualisasi boxplot. Outlier adalah nilai yang jauh dari rata-rata atau nilai-nilai lain dalam kumpulan data. Penjelasan Visualisasi Boxplot adalah sebagai berikut:

- Kotak Utama (Box): Kotak pada diagram boxplot menggambarkan kuartil pertama (Q1) hingga kuartil ketiga (Q3) dan mencakup 50% data. Semakin besar kotaknya, semakin besar sebaran dalam kuartil ini.
- Garis di Tengah Kotak (Median): Garis di tengah kotak adalah median, yang merupakan nilai tengah dalam data. Ini membagi data menjadi dua bagian yang sama besar, yang berarti 50% data ada di atas dan 50% di bawah garis ini.
- Whiskers (Garis Atas dan Bawah): Garis atas dan bawah pada diagram boxplot menggambarkan batas atas dan batas bawah. Data yang berada di luar batas ini dianggap sebagai potensial outlier.
- Titik-titik di Luar Whiskers (Outlier): Titik-titik atau simbol-simbol di luar whiskers adalah outlier. Mereka adalah data yang jauh dari pusat data dan dapat menunjukkan data yang ekstrem atau tidak biasa.

Titik-titik di luar Whiskers tersebutlah yang akan kami simpulkan sebagai outlier dari sebuah variabel. Akan tetapi, karena data kami sebagian besar merupakan data ordinal dimana hanya ada beberapa keunikan value saja, tidak didapatkan banyak outlier di dalamnya. Pada variabel Usia juga, terdapat sebuah garis pada boxplot dengan arti bahwa sebagian besar nilai Usia responden adalah usia 20 tahun dan sisanya merupakan responden dengan nilai Usia lainnya. Namun hal tersebut tidak dimaksudkan untuk dilakukannya penanganan terhadap outlier seperti dengan cara dihapus, karena sejak awal kami sudah menentukan keinginan agar usia responden kami memiliki rentang umur yang tidak begitu jauh, yaitu sekitar 15-25 tahun. Dengan begitu, kami tidak melakukan perubahan atau penanganan terhadap outlier pada variabel Usia kami agar tidak kehilangan data yang penting di dalamnya.

3. Buatlah korelasi antar variable output dan input dengan scatter plot dan Heatmap

Pengecekan korelasi dalam pemodelan adalah proses menilai hubungan antara dua variabel atau lebih dalam suatu dataset. Korelasi adalah ukuran statistik yang menggambarkan sejauh mana dua variabel berhubungan satu sama lain. Dalam pemodelan, pemeriksaan korelasi penting karena membantu mengidentifikasi variabel mana yang terkait dengan variabel hasil/klasifikasi, dan seberapa kuat atau lemah hubungannya. Kami harus mencari fitur yang memiliki korelasi positif, karena fitur bersifat independen (harus berdiri sendiri) sehingga harus memilih 1 fitur dari beberapa fitur yang memiliki korelasi positif.

- Korelasi dengan scatter plot

Syntax:

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

# Variabel dependent yang ingin Anda visualisasikan
dependent_variables = ["membantu_mengalihkan_stress", "pilihan_pertama_stress", "durasi_bermain", "durasi_waktu_bermain", "bermain_game"]

# Jumlah variabel dependent
num_dependent = len(dependent_variables)

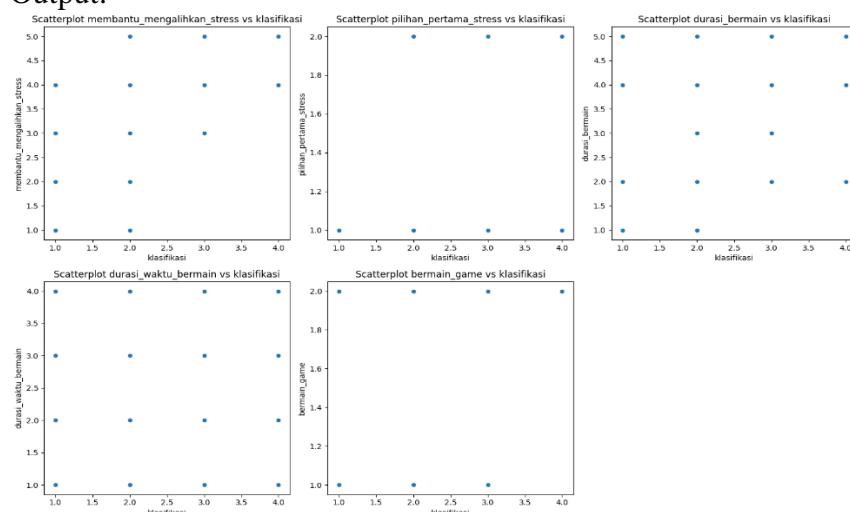
# Jumlah baris dan kolom dalam grid subplot
num_rows = 2
num_cols = 3

plt.figure(figsize=(15, 10))

for i, dependent_var in enumerate(dependent_variables):
    plt.subplot(num_rows, num_cols, i + 1)
    sns.scatterplot(data=df, x="klasifikasi", y=dependent_var)
    plt.title(f"Scatterplot {dependent_var} vs klasifikasi")

plt.tight_layout()
plt.show()
```

Output:



Di bawah ini adalah analisis yang mungkin Anda peroleh berdasarkan output tersebut:

- a. Membantu Mengalihkan Stress vs Klasifikasi:
Scatterplot menunjukkan bahwa variabel "Membantu Mengalihkan Stress" memiliki beberapa titik data yang tersebar di sepanjang variabel "Klasifikasi." Ini menunjukkan bahwa ada variasi dalam "Membantu Mengalihkan Stress" untuk berbagai nilai "Klasifikasi." Namun, tidak ada tren linier yang jelas yang dapat disimpulkan dari scatterplot ini.
- b. Pilihan Pertama Stress vs Klasifikasi:
Scatterplot menunjukkan bahwa variabel "Pilihan Pertama Stress" juga memiliki beberapa variasi dalam nilai untuk berbagai nilai "Klasifikasi." Ada beberapa titik yang tersebar, tetapi tidak ada tren linier yang jelas antara keduanya.
- c. Durasi Bermain vs Klasifikasi:
Scatterplot untuk "Durasi Bermain" dan "Klasifikasi" menunjukkan variasi dalam nilai "Durasi Bermain" untuk berbagai kategori "Klasifikasi." Terlihat beberapa titik data yang menunjukkan tren yang mungkin, tetapi hubungannya tidak terlihat sangat kuat.
- d. Durasi Waktu Bermain vs Klasifikasi:
Scatterplot menunjukkan bahwa "Durasi Waktu Bermain" juga memiliki variasi dalam nilai untuk berbagai nilai "Klasifikasi." Namun, seperti sebelumnya, tidak ada tren linier yang jelas.
- e. Bermain Game vs Klasifikasi:
Scatterplot menunjukkan variasi dalam nilai "Bermain Game" untuk berbagai nilai "Klasifikasi." Namun, seperti variabel lainnya, tidak ada tren linier yang kuat yang dapat disimpulkan.

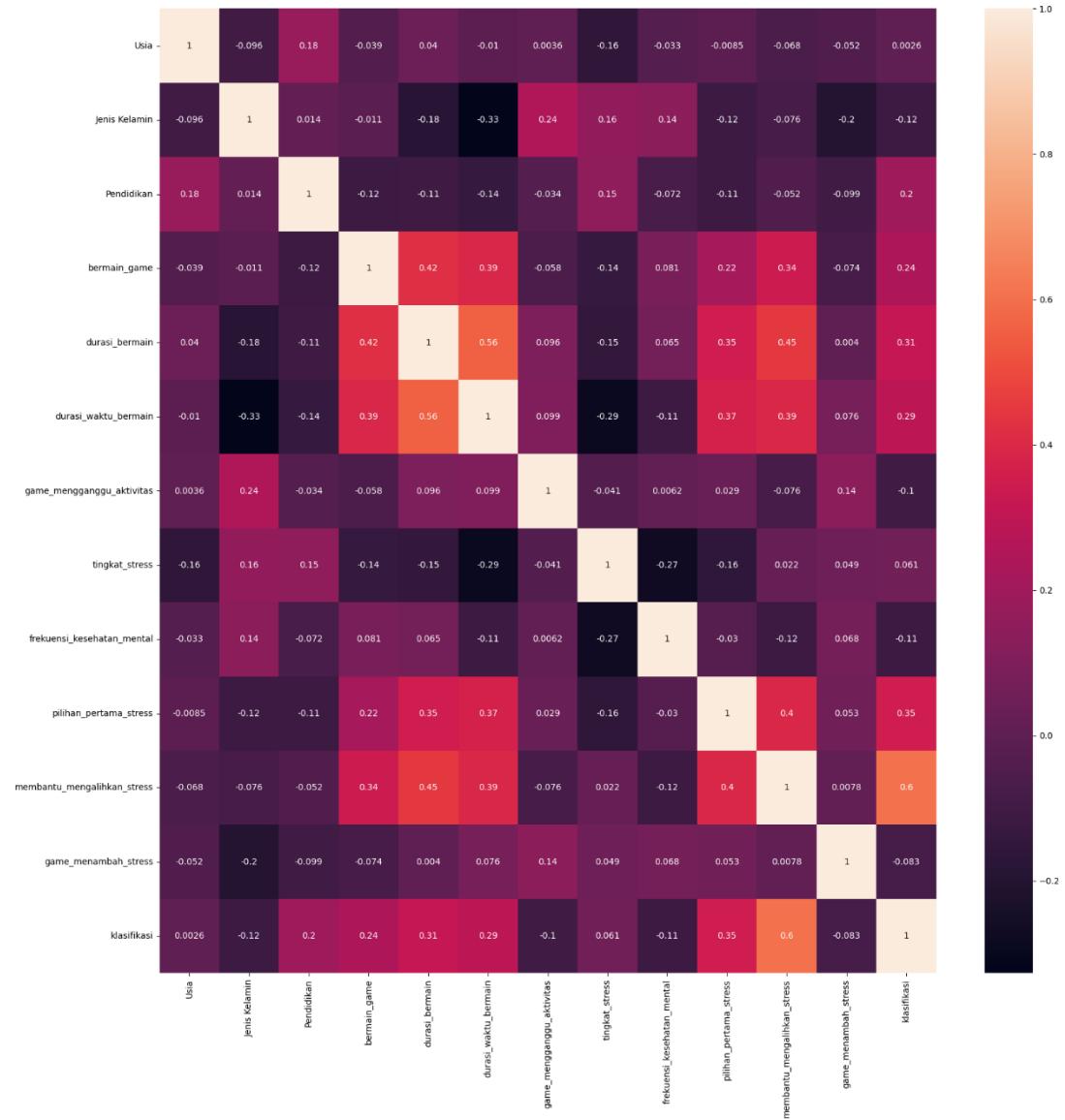
Analisis ini didasarkan pada distribusi titik data dalam scatterplot. Jika Anda ingin mengukur korelasi yang lebih kuat antara variabel-variabel ini, Anda dapat menggunakan metrik statistik seperti korelasi Pearson atau metode lain untuk mengukur hubungan antara variabel-variabel tersebut.

- Korelasi dengan Heatmap

Syntax:

```
plt.figure(figsize=(20,20))
# sns.heatmap = for looking the correlation
sns.heatmap(df[num].corr(),annot=True)
# annot = to result the number
```

Output:



Berdasarkan visualisasi heatmap di atas, semua variabel input saling berkorelasi positif dengan variabel output kecuali beberapa variabel dengan korelasi negatif yaitu "jenis_kelamin", "game_menganggu_aktivitas", dan "frekuensi_kesehatan_mental". Ada beberapa variabel yang memiliki korelasi kuat dan ada juga yang memiliki korelasi rendah. Variabel yang memiliki korelasi kuat akan digunakan untuk visualisasi lebih lanjut menggunakan visualisasi 2 dimensi dan 3 dimensi dari diagram batang, line, pie chart, serta visualisasi scatterplot dan area. Variabel yang memiliki korelasi yang kuat dengan variabel output diantaranya adalah variabel "membantu_menghilangkan_stress", "pilihan_pertama_stress", "durasi_bermain", "durasi_waktu_bermain, dan "bermain_game". Sedangkan variabel yang memiliki korelasi rendah dengan variabel output diantaranya variabel "Usia", "jenis_kelamin", "Pendidikan", "game_menganggu_aktivitas", "tingkat_stress", "frekuensi_kesehatan_mental", dan "game_menambah_stress".

Kemudian, variabel yang memiliki korelasi rendah ini akan dihapus dan tidak digunakan untuk memvisualisasikan bentuk 2 dimensi dan 3 dimensi dari diagram batang, line, pie, chart serta visualisasi scatter plot.

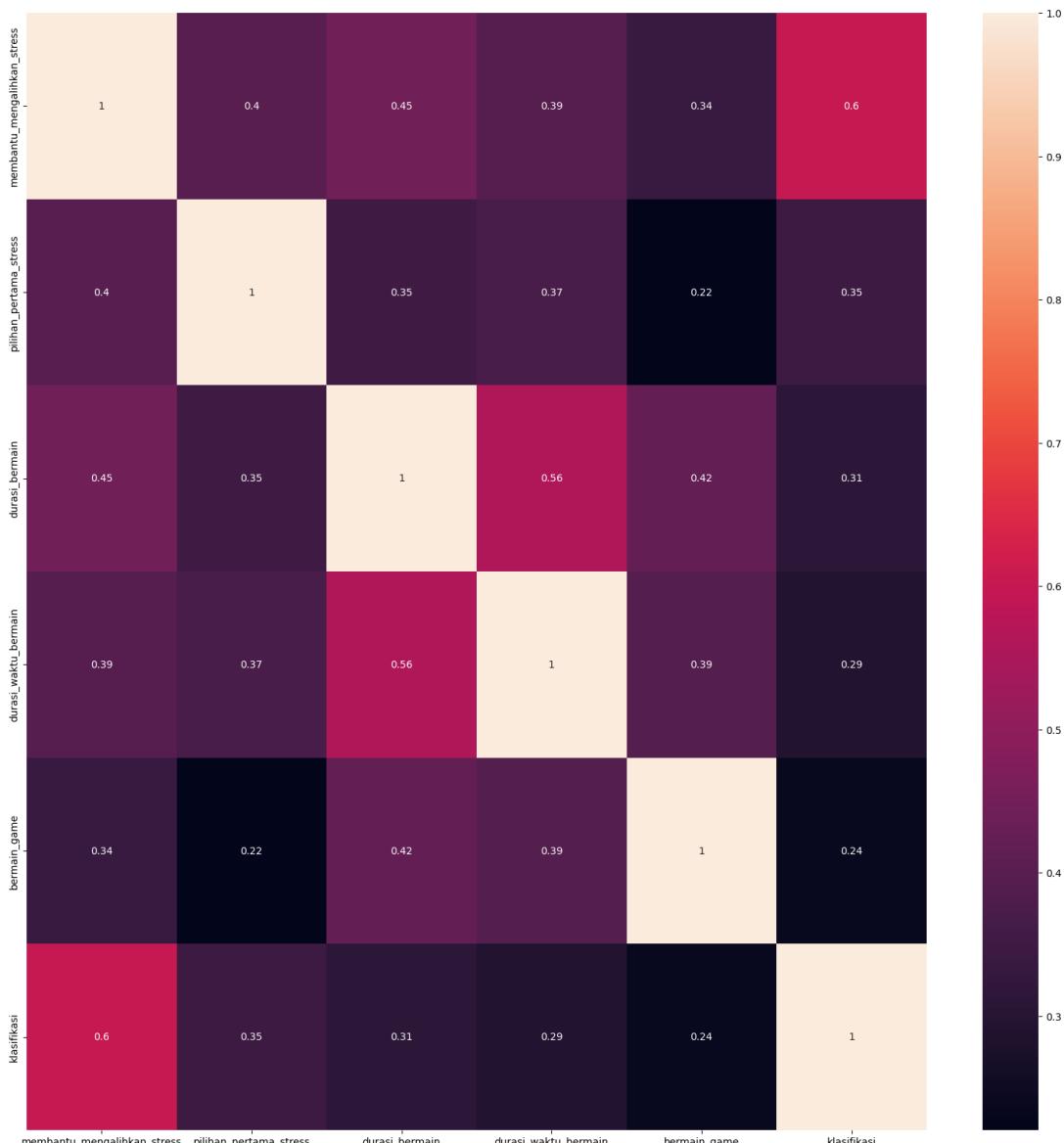
4. Analisis variabel yang mempunyai korelasi dengan output, jika variable tersebut tidak mempunyai korelasi maka bisa dihapus

Langkah selanjutnya kami membuat sebuah daftar bernama 'num' yang tadinya sudah berisi variabel khusus numerik (ordinal), dan menggantinya dengan nama-nama kolom dalam dataset kami yang memiliki korelasi tinggi antara variabel input dengan variabel output. Dengan membuat daftar 'num', kami memberikan panduan yang jelas untuk jenis analisis yang dapat diterapkan pada variabel-variabel ini, serta memudahkan pengelompokan dan pemrosesan data yang relevan dalam analisis yang lebih mendalam.

```
num = ['membantu_mengalihkan_stress','pilihan_pertama_stress','durasi_bermain','durasi_waktu_bermain','bermain_game','klasifikasi']  
  
df[num]  
  
membantu_mengalihkan_stress pilihan_pertama_stress durasi_bermain durasi_waktu_bermain bermain_game klasifikasi  
0 4 1 4 3 2 3  
1 3 1 1 1 1 1  
2 3 1 4 2 2 2  
3 3 1 4 1 2 2  
4 3 1 1 1 1 2  
... ... ... ... ... ... ...  
104 3 1 3 3 2 2  
105 4 1 2 1 2 3  
106 3 1 1 1 1 2  
107 3 1 2 4 2 2  
108 4 2 4 3 2 3  
109 rows x 6 columns
```

Dan membuat heatmap untuk visualisasi korelasi variable-variabel yang telah kami pilih dengan syntax berikut:

```
plt.figure(figsize=(20,20))  
# sns.heatmap = for looking the correlation  
sns.heatmap(df[num].corr(), annot=True)  
# annot = to result the number
```



Heatmap di atas merupakan tampilan heatmap dari variabel input yang memiliki korelasi yang kuat dengan variabel output. Setelah dilakukan analisis, variabel yang pasti akan kami gunakan untuk visualisasi lebih lanjut ada lima variabel. Dengan adanya heatmap ini sangat membantu untuk melihat tingkat korelasi antara variabel input dengan variabel output yang telah dipilih.

- Buatlah visualisasi dua dimensi menggunakan Diagram Batang, Garis, Pie, dan Area dari variabel yang terpilih

- Diagram Batang

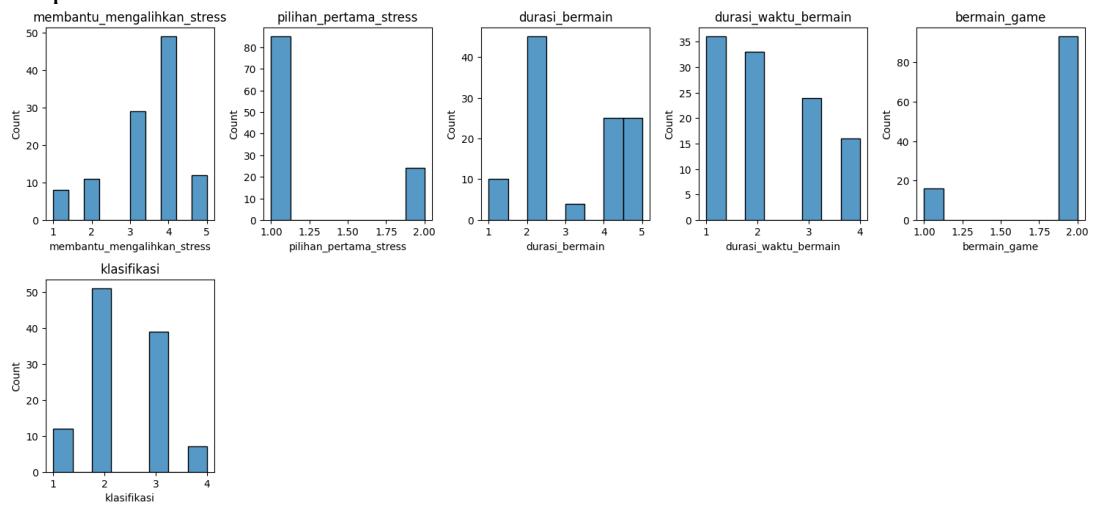
Syntax:

```
plt.figure(figsize=(15, 10))

for i in range(0, len(df[num].columns)):
    plt.subplot(3, 5, i + 1)
    sns.histplot(data=df[num][df[num].columns[i]])
    plt.title(df[num].columns[i])
    plt.tight_layout()

plt.show()
```

Output:



Tampilan histogram yang telah kami buat memberikan wawasan tentang distribusi data untuk setiap variabel. Berikut adalah analisis singkat untuk setiap variabel dari lima variabel yang Anda visualisasikan:

a. Variabel membantu_mengalihkan_stress

Histogram pertama menunjukkan distribusi data yang memiliki puncak yang terletak di sekitar nilai kategori 4. Distribusi ini mencerminkan karakteristik unik yang dapat memberikan wawasan yang berharga dalam analisis lebih lanjut. Pada kategori tersebut, nilai keempat memiliki tingkat frekuensi yang paling tinggi dalam distribusi. Ini mengindikasikan bahwa sebagian besar data cenderung berkumpul di sekitar nilai atau rentang tertentu yang sesuai dengan kategori keempat.

b. Variabel pilihan_pertama_stress

Dalam analisis histogram, terdapat satu variabel di antara dua variabel yang menonjol dengan distribusi yang memiliki kategori tertinggi pada kategori pertama. Distribusi ini menunjukkan bahwa sebagian besar data terkonsentrasi pada kategori pertama dan mungkin memiliki sedikit variasi dalam nilai-nilai lainnya. Fokus yang tinggi pada kategori pertama mengindikasikan bahwa mayoritas observasi dalam dataset tersebut terletak di rentang nilai yang sesuai dengan kategori tersebut.

c. Variabel durasi_bermain

Dari hasil histogram di atas, kami menemukan satu variabel dengan distribusi yang menonjol, di mana kategori kedua memiliki tingkat frekuensi tertinggi. Ini mengindikasikan bahwa sebagian besar data terkonsentrasi di sekitar kategori kedua, yang mungkin mencerminkan suatu pola atau karakteristik dalam dataset tersebut. Selain itu, menariknya, kategori keempat dan kelima memiliki tingkat frekuensi yang sama, yang dapat menunjukkan persamaan atau kesamaan dalam jumlah data di kedua kategori tersebut.

d. Variabel durasi_Waktu_bermain

Berdasarkan hasil histogram di atas, kita mengidentifikasi satu variabel dengan distribusi yang menonjol, di mana kategori pertama memiliki tingkat frekuensi tertinggi. Distribusi ini menggambarkan bahwa sebagian besar data terkonsentrasi

pada kategori pertama, menciptakan puncak tertinggi dalam histogram. Selanjutnya, kategori-kategori berikutnya menurun secara bertahap dalam frekuensi, menunjukkan bahwa jumlah data secara signifikan lebih rendah dibandingkan dengan kategori pertama. Analisis distribusi ini dapat memberikan wawasan tentang pola data yang mendominasi kategori pertama dan kemungkinan adanya outlier dalam data yang mungkin berada di kategori-kategori berikutnya. Data pada kategori pertama mungkin memiliki karakteristik yang sangat berbeda dari data di kategori-kategori selanjutnya, yang dapat menjadi fokus penelitian lebih lanjut.

e. Variabel bermain_game

Histogram kelima menunjukkan analisis bahwa terdapat satu variabel yang menunjukkan kami mengidentifikasi satu variabel dengan distribusi yang menonjol, di mana kategori kedua memiliki tingkat frekuensi tertinggi. Sedangkan tidak ada observasi atau data yang terdistribusi dalam rentang antara kategori pertama dan kategori kedua. Hal ini mengindikasikan adanya kesenjangan atau perbedaan yang signifikan dalam data, dengan sebagian besar observasi terfokus pada kategori kedua dan tidak ada yang terletak di antara kategori tersebut.

f. Variabel klasifikasi

Berdasarkan visualisasi histogram di atas, ditemukan variabel dengan distribusi yang menunjukkan bahwa kategori kedua memiliki tingkat frekuensi tertinggi, sementara kategori pertama memiliki frekuensi yang paling rendah. Ini menggambarkan pola distribusi yang cukup kontras antara kategori pertama dan kedua. Pola ini bisa menunjukkan adanya kelompok data yang lebih kecil dengan karakteristik unik dalam kategori pertama, sementara mayoritas data terkonsentrasi dalam kategori kedua.

2. Diagram Garis

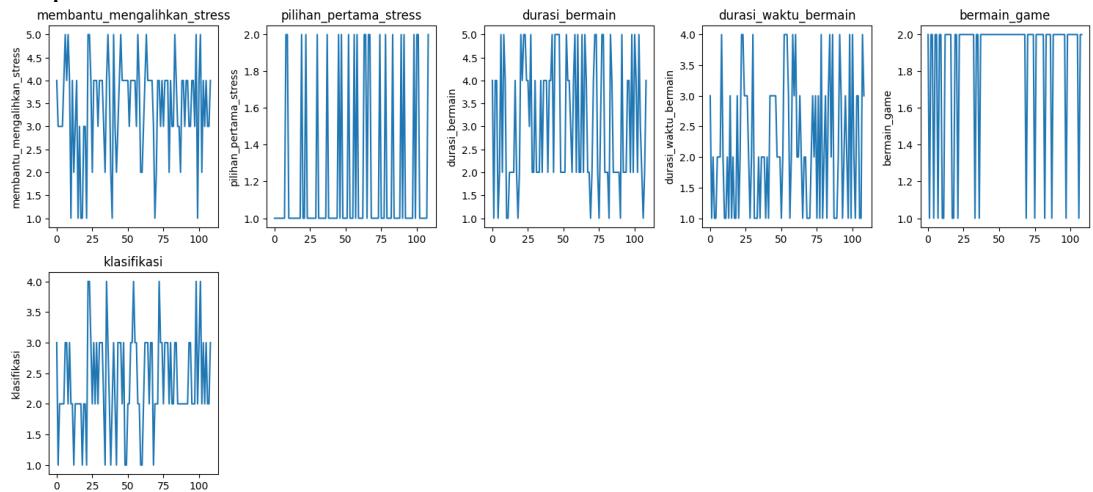
Syntax:

```
plt.figure(figsize=(15, 10))

for i in range(0, len(df[num].columns)):
    plt.subplot(3, 5, i + 1)
    sns.lineplot(data=df[num][df[num].columns[i]])
    plt.title(df[num].columns[i])
    plt.tight_layout()

plt.show()
```

Output:



Analisis dari line graph ini memberikan wawasan yang lebih mendalam tentang variabel-variabel yang terlibat dalam penanganan stres dan aktivitas bermain game dalam berbagai konteks. Dalam grafik tersebut, kita dapat mengidentifikasi pola dan tren yang bermanfaat sebagai berikut:

a. Tren Variabel "Membantu Mengalihkan Stress"

Line graph untuk variabel "Membantu Mengalihkan Stress" memungkinkan kita untuk melihat bagaimana tingkat aktivitas yang membantu mengatasi stres berkembang sepanjang waktu atau dalam berbagai rentang nilai. Tren ini dapat memberikan pemahaman tentang apakah ada peningkatan atau penurunan dalam cara individu mengatasi stres dari waktu ke waktu atau berdasarkan nilai tertentu dalam variabel lain.

b. Perubahan dalam "Pilihan Pertama Stress"

Line graph untuk variabel "Pilihan Pertama Stress" memungkinkan kita untuk melihat perubahan dalam pilihan pertama yang diambil individu ketika merasa stres. Tren ini dapat memberikan wawasan tentang apakah ada pergeseran dalam strategi yang digunakan individu untuk mengatasi stres.

c. Perubahan dalam Variabel Durasi Bermain dan Durasi Waktu Bermain

Line graph untuk variabel "Durasi Bermain" dan "Durasi Waktu Bermain" memungkinkan kita untuk melihat perubahan dalam durasi bermain game sepanjang waktu atau berdasarkan nilai kategori lain. Ini dapat memberikan pemahaman tentang bagaimana perilaku bermain game berubah seiring waktu atau dalam situasi tertentu.

d. Perubahan dalam "Bermain Game" dan "Klasifikasi"

Line graph untuk variabel "Bermain Game" dan "Klasifikasi" dapat memberikan wawasan tentang bagaimana preferensi bermain game dan distribusi klasifikasi berubah seiring waktu atau dalam situasi tertentu.

Secara keseluruhan, line graph memberikan pemahaman tentang perkembangan dan perubahan dalam variabel-variabel ini sepanjang waktu atau dalam situasi yang berbeda. Informasi ini dapat digunakan untuk pengambilan keputusan yang lebih baik,

perencanaan strategi yang lebih efektif, dan pemahaman yang lebih baik tentang bagaimana faktor-faktor tertentu berinteraksi dalam konteks yang relevan.

3. Pie Chart

Syntax:

```
# Memastikan df[num] berisi hanya kolom-kolom numerik
# Jika tidak semua kolom adalah numerik, pastikan untuk memilih yang sesuai
numeric_columns = df[num]

# Mengatur ukuran gambar
plt.figure(figsize=(15, 10))

# Menyusun subplot untuk setiap kolom numerik
for i, column in enumerate(numeric_columns.columns):
    plt.subplot(3, 5, i + 1)

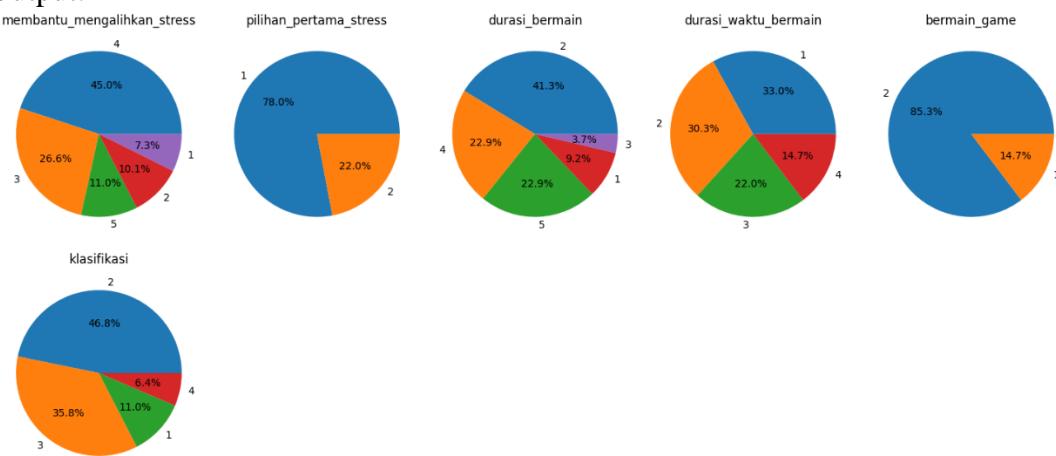
    # Menghitung jumlah masing-masing nilai dalam kolom
    value_counts = df[num][column].value_counts()

    # Membuat diagram pie
    plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%')
    plt.title(column)

    # Mengatur tata letak subplot
    plt.tight_layout()

# Menampilkan gambar
plt.show()
```

Output:



Kode yang Anda berikan mencoba untuk membuat diagram pie untuk masing-masing variabel numerik dalam DataFrame `df[num]`, serta menghitung proporsi masing-masing nilai dalam variabel tersebut. Ini dapat memberikan wawasan visual tentang distribusi nilai di setiap variabel numerik. Berikut adalah analisis per variabel:

a. Variabel 1: membantu_mengalihkan_tress

- Diagram pie menunjukkan distribusi nilai variabel ini.
- Terdapat nilai-nilai kategori 1, 2, 3, 4, dan 5 dengan proporsi tertentu masing masing. Keterangan:
 - Angka 1: Tidak membantu
 - Angka 2: Kurang membantu
 - Angka 3: Netral
 - Angka 4: Membantu

- Angka 5: Sangat membantu
 - Dari diagram pie di atas, kategori yang paling banyak dari variabel "membantu_mengalihkan_stress" adalah kategori 4 dengan jumlah 45% dari total data yang berarti bahwa responden banyak yang merasa terbantu dengan adanya game online dapat mengalihkan dari sumber stress yang telah dialami.
- b. Variabel 2: pilihan_pertama_stress
- Diagram pie menggambarkan distribusi nilai dalam variabel kedua.
 - Kategori 1 dan 2 memiliki sebagian besar proporsi dengan memiliki keterangan masing masing.
 - Angka 1: Ya
 - Angka 2: Tidak
 - Jika diamati kategori yang paling mendominasi dengan sekitar 78% dari total data yaitu kategori 1 yang berarti bahwa responden merasa jika bermain game online menjadi pilihan pertama responden ketika merasa stres.
- c. Variabel 3: durasi_bermain
- Diagram pie memperlihatkan cara nilai-nilai variabel ketiga terdistribusi.
 - Terdapat variasi di antara kategori 1, 2, 3, 4, dan 5 dengan memiliki keterangan masing masing.
 - Angka 1: Tidak pernah
 - Angka 2: Jarang
 - Angka 3: Sekali seminggu
 - Angka 4: Beberapa kali seminggu
 - Angka 5: Setiap hari
 - Berdasarkan hasil diagram pie di atas, kategori 2(Jarang) memiliki proporsi sekitar 41,3% dari total data yang berarti bahwa banyak responden yang jarang bermain game online dalam waktu sebulan terakhir.
- d. Variabel 4: durasi_waktu_bermain
- Diagram pie memvisualisasikan cara nilai-nilai variabel ketiga terdistribusi.
 - Terdapat variasi di antara kategori 1 dan 2 dengan memiliki keterangan masing masing.
 - Angka 1: Kurang dari 30 menit
 - Angka 2: 30 menit - 1 jam
 - Angka 3: 1 - 2 jam
 - Angka 4: Lebih dari 2 jam
 - Dapat dianalisa berdasarkan diagram pie di atas, kategori 1(Kuarang dari 30 menit) memiliki proporsi sekitar 33% dari total data yang berarti bahwa banyak responden yang menghabiskan waktunya untuk bermain game selama kurang dari 30 menit dalam setiap harinya.
- e. Variabel 5: bermain_game
- Diagram pie menunjukkan distribusi nilai variabel ini.
 - Terdapat nilai-nilai kategori 1 dan 2 dengan proporsi tertentu masing-masing.
Keterangan:
 - Angka 1: Tidak
 - Angka 2: Ya

- Dari diagram pie di atas, kategori yang paling banyak dari variabel "bermain_game" adalah kategori 2(Ya) dengan jumlah 85,3% dari total data yang berarti bahwa ternyata responden banyak yang bermain game online.

f. Variabel 6: klasifikasi

- Diagram pie menggambarkan distribusi nilai dalam variabel kedua.
- Pada variabel ini terdapat beberapa kategori yaitu kategori 1, 2, 3, dan 4 yang memiliki sebagian besar proporsi dengan keterangan masing-masing.
 - Angka 1: Tidak setuju
 - Angka 2: Netral
 - Angka 3: Setuju
 - Angka 4: Sangat setuju
- Jika diamati kategori yang paling mendominasi dengan sekitar 46,8% dari total data yaitu kategori 2 yang berarti bahwa responden merasa bermain game sebagai cara untuk menghilangkan stres merupakan alternatif yang sehat dan efektif.

4. Diagram Area

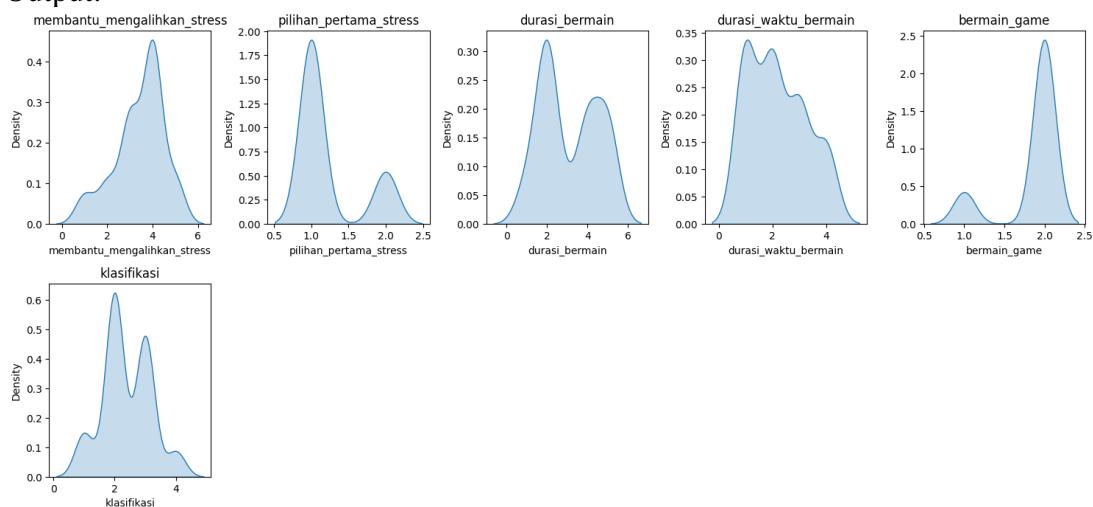
Syntax:

```
plt.figure(figsize=(15, 10))

for i in range(0, len(df[num].columns)):
    plt.subplot(3, 5, i + 1)
    sns.kdeplot(data=df[num][df[num].columns[i]], fill=True) # Menggunakan kdeplot dengan fill=True
    plt.title(df[num].columns[i])
    plt.tight_layout()

plt.show()
```

Output:



Keterangan:

Hasil diagram menunjukkan bahwa kolom variable pilihan_pertama_stress, durasi_bermain, dan durasi_waktu_bermain memiliki distribusi positif (positively skewed) yang artinya bahwa distribusi dengan nilai terbanyak condong berada di sebelah kiri. Sedangkan kolom variable membantu_mengalihkan_stress, dan bermain_game memiliki distribusi negatif (negatively skewed) yang artinya bahwa distribusi dengan nilai terbanyak condong berada di sebelah kanan.

Dan kolom variabel klasifikasi memiliki distribusi simetris yang artinya bahwa distribusi dengan nilai terbanyak berada di bagian tengah.

6. Analisis diagram yang telah dibuat pada nomer 5

Pada visualisasi ini, kami ingin dapat melihat korelasi antar variabel input (dependent) yang sudah kami pilih dengan variabel output (independent) dengan menggunakan diagram batang

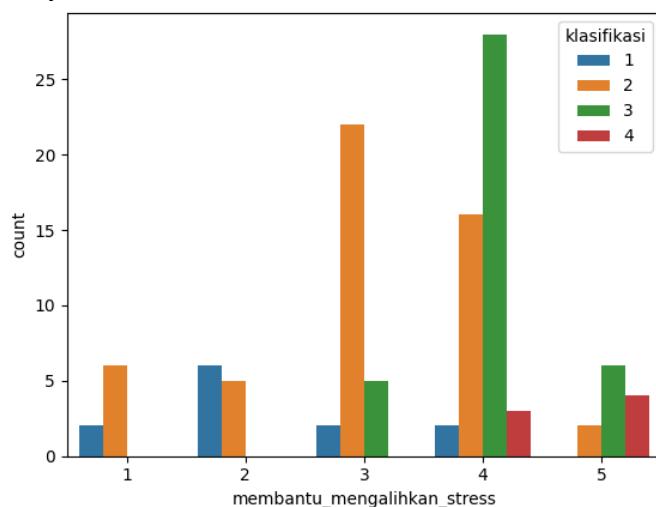
1. Korelasi Dua Variabel dengan Diagram Batang

a. Korelasi Variabel “membantu_mengalihkan_stress” dengan “klasifikasi”

Syntax:

```
#Korelasi antara variabel membantu_mengalihkan_stress dan klasifikasi  
a=df.groupby(['membantu_mengalihkan_stress','klasifikasi'])['membantu_mengalihkan_stress'].size().reset_index(name='count')  
sns.barplot(data=a,x=a['membantu_mengalihkan_stress'],y=a['count'],hue='klasifikasi')
```

Output:



- Penjelasan nilai:

Variabel membantu_mengalihkan_stress:

- 1 = Tidak membantu
- 2 = Kurang membantu
- 3 = Netral
- 4 = Membantu
- 5 = Sangat membantu

- Klasifikasi

- 1 = Tidak setuju
- 2 = Netral
- 3 = Setuju
- 4 = Sangat setuju

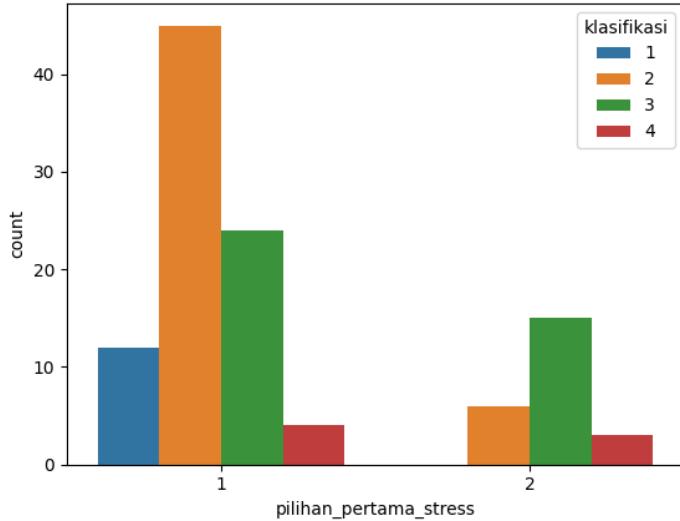
Berdasarkan diagram tersebut, dapat disimpulkan bahwa responden dengan nilai 4 (Membantu) pada variabel membantu_mengalihkan_stress dan nilai 3 (Setuju) pada variabel klasifikasi memiliki jumlah terbanyak. Hal tersebut dapat diartikan bahwa responden setuju bahwa bermain game dapat membantu mengalihkan stress yang sedang dihadapi, akan tetapi tidak begitu menyetujui atau bahkan tidak menyetujui bahwa bermain game merupakan cara alternatif untuk menghilangkan stres yang sehat dan efektif

b. Korelasi Variabel “pilihan_pertama_stress” dengan “klasifikasi”

Syntax:

```
#Korelasi antara variabel pilihan_pertama_stress dan klasifikasi
b=df.groupby(['pilihan_pertama_stress','klasifikasi'])['pilihan_pertama_stress'].size().reset_index(name='count')
sns.barplot(data=b,x=b['pilihan_pertama_stress'],y=b['count'],hue='klasifikasi')
```

Output:



- Penjelasan nilai:

Variabel pilihan_pertama_stress:

1 = Tidak

2 = Ya

- Klasifikasi

1 = Tidak setuju

2 = Netral

3 = Setuju

4 = Sangat setuju

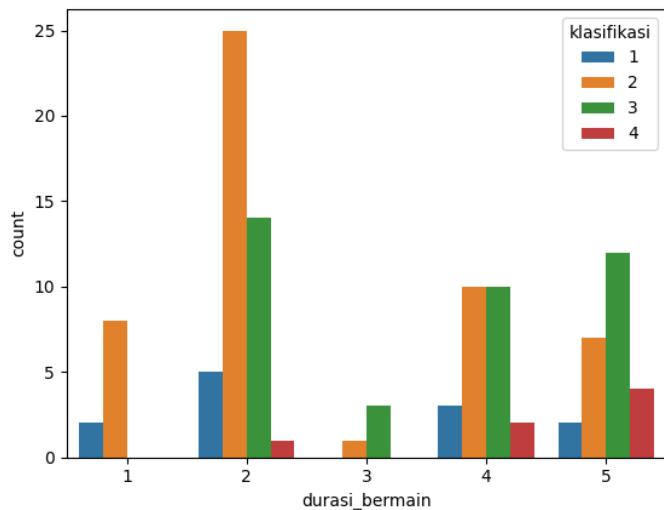
Berdasarkan diagram tersebut, dapat disimpulkan bahwa responden dengan nilai 1 (Tidak) pada variabel pilihan_pertama_stress dan nilai 2 (Netral) pada variabel klasifikasi memiliki jumlah terbanyak. Hal tersebut dapat diartikan bahwa responden tidak menjadikan bermain game online menjadi pilihan pertamanya ketika sedang stress, akan tetapi netral bahwa bermain game merupakan cara alternatif untuk menghilangkan stres yang sehat dan efektif.

- Korelasi Variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
#Korelasi antara variabel durasi_bermain dan klasifikasi
c=df.groupby(['durasi_bermain','klasifikasi'])['durasi_bermain'].size().reset_index(name='count')
sns.barplot(data=c,x=c['durasi_bermain'],y=c['count'],hue='klasifikasi')
```

Output:



- Penjelasan nilai:
Variabel durasi_bermain:
1 = Tidak pernah
2 = Jarang
3 = Sekali seminggu
4 = Beberapa kali seminggu
5 = Setiap hari
- Klasifikasi
1 = Tidak setuju
2 = Netral
3 = Setuju
4 = Sangat setuju

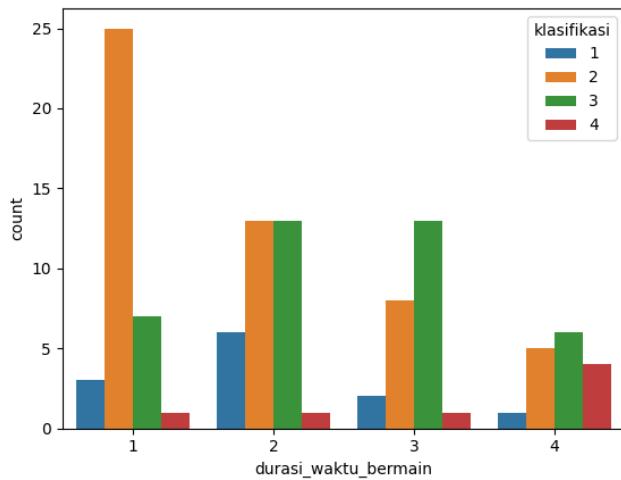
Berdasarkan diagram tersebut, dapat disimpulkan bahwa responden dengan nilai 2 (Jarang) pada variabel durasi_bermain dan nilai 2 (Netral) pada variabel klasifikasi memiliki jumlah terbanyak. Hal tersebut dapat diartikan bahwa sebagian besar responden jarang bermain game namun tetap netral bahwa bermain game merupakan cara alternatif untuk menghilangkan stres yang sehat dan efektif.

- d. Korelasi Variabel “durasi_waktu_bermain” dengan “klasifikasi”

Syntax:

```
#Korelasi antara variabel durasi_waktu_bermain dan klasifikasi
d=df.groupby(['durasi_waktu_bermain','klasifikasi'])['durasi_waktu_bermain'].size().reset_index(name='count')
sns.barplot(data=d,x=d['durasi_waktu_bermain'],y=d['count'],hue='klasifikasi')
```

Output:



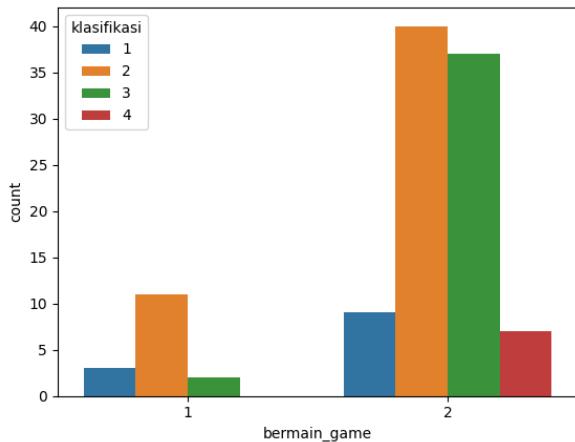
- Penjelasan nilai:
Variabel durasi_waktu_bermain:
1 = Kurang dari 30 menit
2 = 30 menit - 1 jam
3 = 1 - 2 jam
4 = Lebih dari 2 jam
 - Klasifikasi
1 = Tidak setuju
2 = Netral
3 = Setuju
4 = Sangat setuju
- Berdasarkan diagram tersebut, dapat disimpulkan bahwa responden dengan nilai 1 (Kurang dari 30 menit) pada variabel durasi_bermain dan nilai 2 (Netral) pada variabel klasifikasi memiliki jumlah terbanyak. Hal tersebut dapat diartikan bahwa durasi bermain game sebagian besar responden adalah kurang dari 30 menit setiap harinya, namun tetap menjawab netral bahwa bermain game merupakan cara alternatif untuk menghilangkan stres yang sehat dan efektif.

e. Korelasi Variabel “bermain_game” dengan “klasifikasi”

Syntax:

```
#Korelasi antara variabel bermain_game dan klasifikasi
e=df.groupby(['bermain_game','klasifikasi'])['bermain_game'].size().reset_index(name='count')
sns.barplot(data=e,x=e['bermain_game'],y=e['count'],hue='klasifikasi')
```

Output:



- Penjelasan nilai:
Variabel bermain_game:
1 = Tidak
2 = Ya
- Klasifikasi
1 = Tidak setuju
2 = Netral
3 = Setuju
4 = Sangat setuju

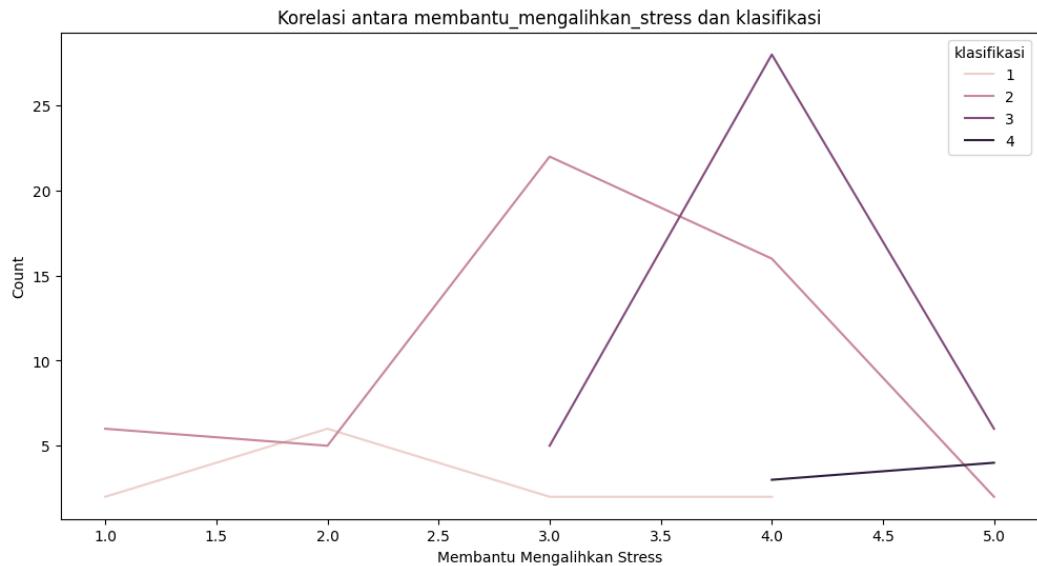
Berdasarkan diagram tersebut, dapat disimpulkan bahwa responden dengan nilai 2 (Ya) pada variabel durasi_bermain dan nilai 2 (Netral) serta nilai 3 (Setuju) pada variabel klasifikasi memiliki jumlah terbanyak. Hal tersebut dapat diartikan bahwa durasi sebagian besar responden bermain game dan menyatakan netral serta setuju bahwa bermain game merupakan cara alternatif untuk menghilangkan stres yang sehat dan efektif

2. Korelasi 2 Variabel dengan Diagram Garis

- a. Korelasi Variabel “membantu_mengalihkan_stress” dengan “klasifikasi”
Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'membantu_mengalihkan_stress' dan 'klasifikasi'  
b = df.groupby(['membantu_mengalihkan_stress', 'klasifikasi'])['membantu_mengalihkan_stress'].size().reset_index(name='count')  
  
plt.figure(figsize=(12, 6))  
sns.lineplot(data=b, x='membantu_mengalihkan_stress', y='count', hue='klasifikasi')  
  
plt.title('Korelasi antara membantu_mengalihkan_stress dan klasifikasi')  
plt.xlabel('Membantu Mengalihkan Stress')  
plt.ylabel('Count')  
  
plt.show()
```

Output:



Keterangan:

Terdapat klasifikasi yang terdiri dari:

1 = Tidak setuju, 2 = Netral, 3 = Setuju, 4= Sangat setuju

Sumbu X dari diagram garis di atas adalah nilai dari variable 'apakah game online membantu mengalihkan dari stress", dimana : 1 = Tidak membantu, 2 = Kurang membantu, 3 = Netral, 4 = Membantu, 5 = Sangat membantu.

Sumbu Y merupakan jumlah banyaknya nilai dari variable.

Dari visualisasi diatas, dapat dilihat bahwa nilai 4 (membantu) memiliki jumlah nilai tertinggi dengan warna yang menunjukkan klasifikasi ke 3 (Setuju), sedangkan nilai 1 (Tidak membantu) memiliki nilai terendah. Hal ini dapat disimpulkan bahwa banyak responden yang merasa jika game online memang dapat membantu mereka dalam mengalihkannya dari sumber stress dan responden setuju bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress. Hanya sedikit sekali yang merasa game online tidak membantu mengalihkan dari stress.

- b. Korelasi Variabel “pilihan_pertama_stress” dengan “klasifikasi”
Syntax:

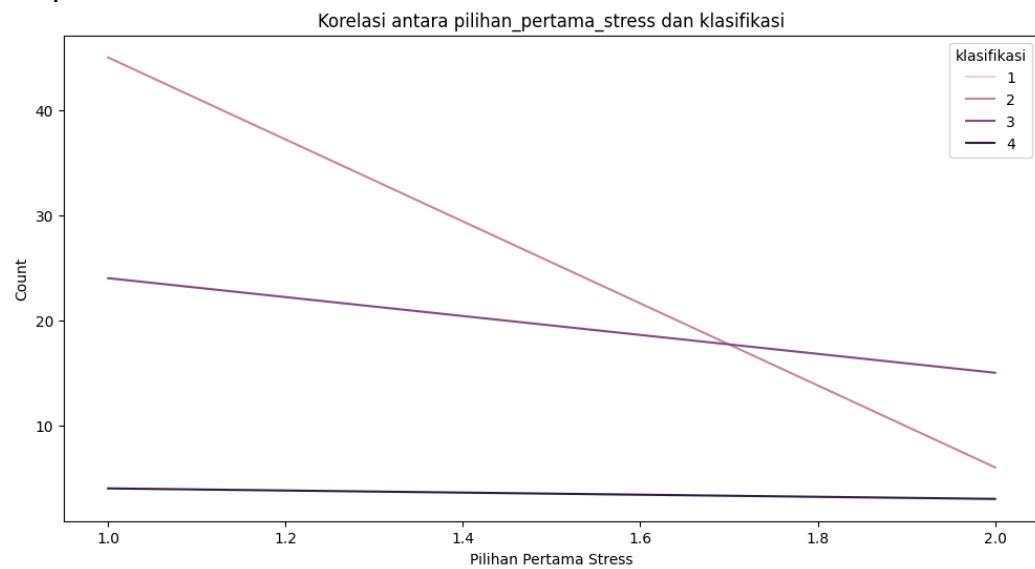
```
# Menghitung frekuensi kemunculan kombinasi 'pilihan_pertama_stress' dan 'klasifikasi'
b = df.groupby(['pilihan_pertama_stress', 'klasifikasi'])['pilihan_pertama_stress'].size().reset_index(name='count')

plt.figure(figsize=(12, 6))
sns.lineplot(data=b, x='pilihan_pertama_stress', y='count', hue='klasifikasi')

plt.title('Korelasi antara pilihan_pertama_stress dan klasifikasi')
plt.xlabel('Pilihan Pertama Stress')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Terdapat klasifikasi yang terdiri dari:

1 = Tidak setuju, 2 = Netral, 3 = Setuju, 4= Sangat setuju

Sumbu X dari diagram garis di atas adalah nilai dari variable 'Pilihan Pertama Stress', dimana:

1 = Ya, 2 = Tidak

Sumbu Y merupakan jumlah banyaknya nilai dari variable.

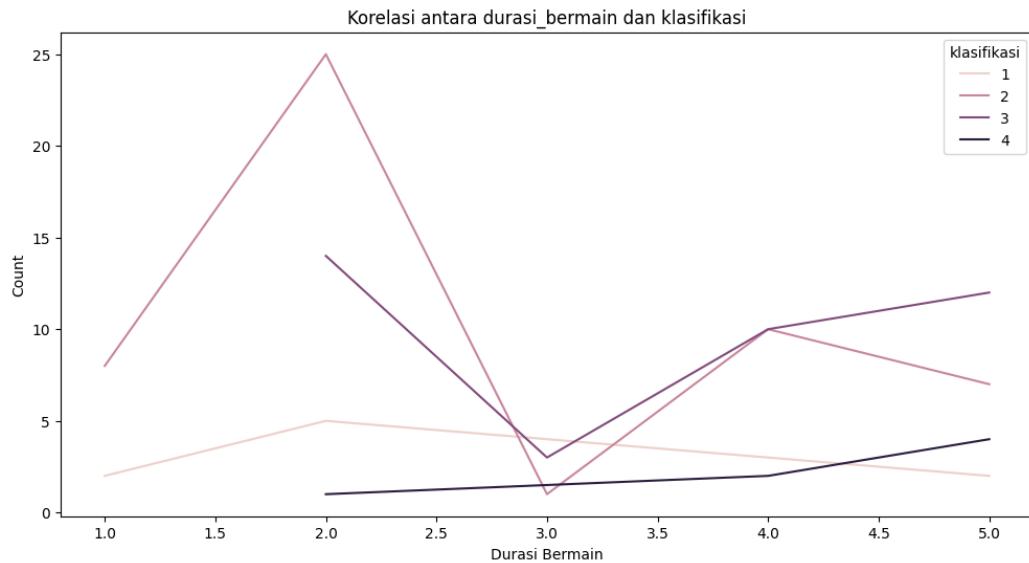
Dari visualisasi diatas, dapat dilihat bahwa nilai 1(Ya) memiliki jumlah nilai tertinggi dengan warna yang menunjukkan klasifikasi ke 2 (Netral) . Hal ini dapat disimpulkan bahwa game online menjadi pilihan pertama yang dilakukan responden ketika dirinya merasa stress. Namun meskipun demikian, responden merasa Netral jika game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress

c. Korelasi Variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'durasi_bermain' dan 'klasifikasi'  
b = df.groupby(['durasi_bermain', 'klasifikasi'])['durasi_bermain'].size().reset_index(name='count')  
  
plt.figure(figsize=(12, 6))  
sns.lineplot(data=b, x='durasi_bermain', y='count', hue='klasifikasi')  
  
plt.title('Korelasi antara durasi_bermain dan klasifikasi')  
plt.xlabel('Durasi Bermain')  
plt.ylabel('Count')  
  
plt.show()
```

Output:



Keterangan:

Terdapat klasifikasi yang terdiri dari:

1 = Tidak setuju, 2 = Netral, 3 = Setuju, 4= Sangat setuju

Sumbu X dari diagram garis di atas adalah nilai dari variable 'Seberapa Sering Durasi Bermain dalam 1 Bulan', dimana:

1 = Tidak pernah, 2 = Jarang, 3 = Sekali seminggu, 4 = Beberapa kali seminggu,
5 = Setiap hari

Sumbu Y merupakan jumlah banyaknya nilai dari variable.

Dari visualisasi diatas, dapat dilihat bahwa nilai 2(jarang bermain) memiliki jumlah nilai tertinggi dengan warna yang menunjukkan klasifikasi ke 2 (Netral) . Hal ini dapat disimpulkan bahwa banyak responden yang jarang bermain game online di sebulan terakhir dan responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- d. Korelasi Variabel “durasi_waktu_bermain” dengan “klasifikasi”

Syntax:

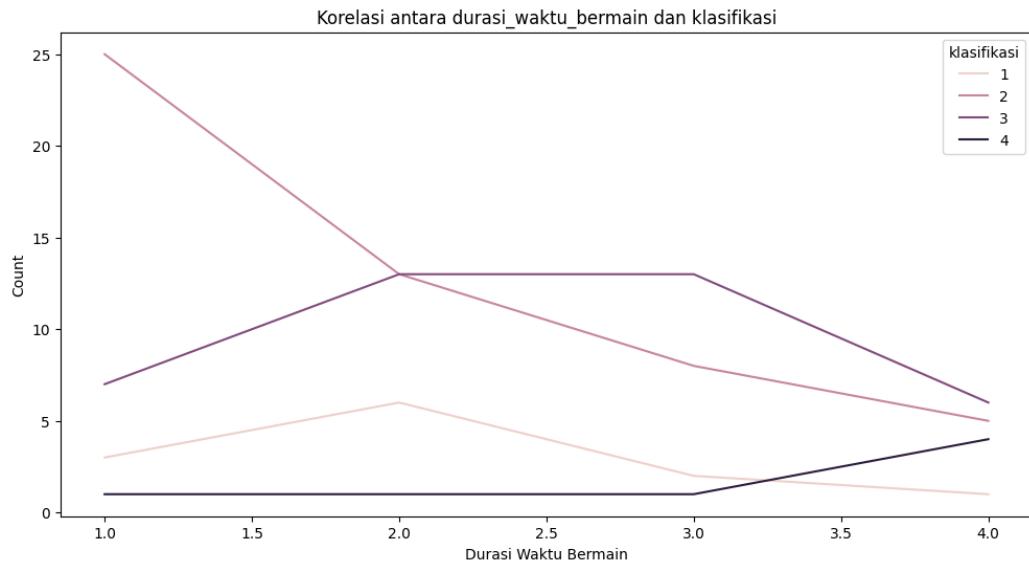
```
# Menghitung frekuensi kemunculan kombinasi 'durasi_waktu_bermain' dan 'klasifikasi'
b = df.groupby(['durasi_waktu_bermain', 'klasifikasi'])['durasi_waktu_bermain'].size().reset_index(name='count')

plt.figure(figsize=(12, 6))
sns.lineplot(data=b, x='durasi_waktu_bermain', y='count', hue='klasifikasi')

plt.title('Korelasi antara durasi_waktu_bermain dan klasifikasi')
plt.xlabel('Durasi Waktu Bermain')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Terdapat klasifikasi yang terdiri dari:

1 = Tidak setuju, 2 = Netral, 3 = Setuju, 4= Sangat setuju

Sumbu X dari diagram garis di atas adalah nilai dari variable 'Seberapa Lama Waktu Bermain Game', dimana:

1 = Kurang dari 30 menit, 2 = 30 menit - 1 jam, 3 = 1-2 jam, 4 = Lebih dari 2 jam
Sumbu Y merupakan jumlah banyaknya nilai dari variable.

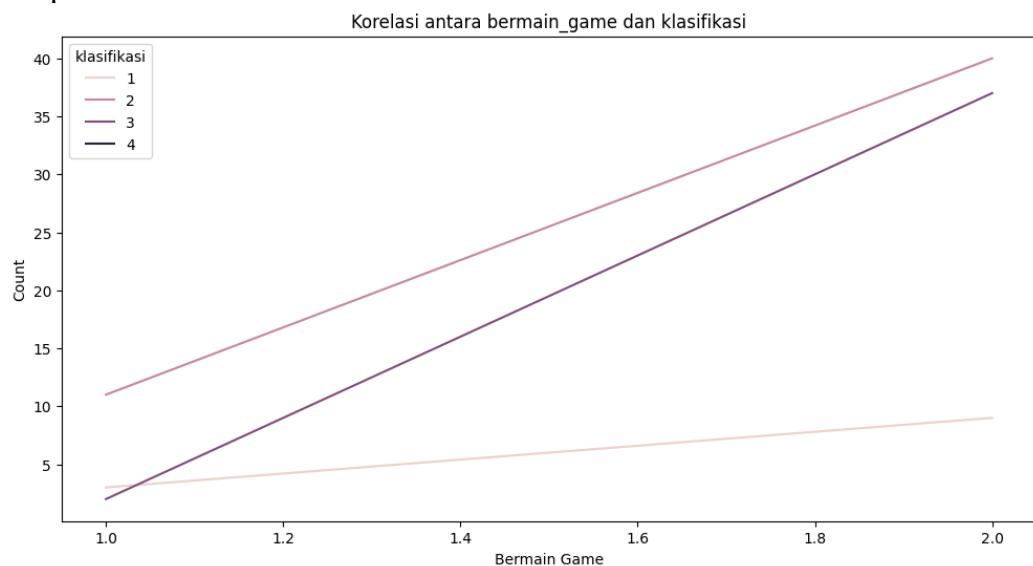
Dari visualisasi diatas, dapat dilihat bahwa nilai 1(Kurang dari 30 menit) memiliki jumlah nilai tertinggi dengan warna yang menunjukkan klasifikasi ke 2 (Netral) . Hal ini dapat disimpulkan bahwa responden biasanya bermain game online dengan durasi lama waktu kurang dari 30 menit dan responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- e. Korelasi Variabel “bermain_game” dengan “klasifikasi”

Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'bermain_game' dan 'klasifikasi'  
b = df.groupby(['bermain_game', 'klasifikasi'])['bermain_game'].size().reset_index(name='count')  
  
plt.figure(figsize=(12, 6))  
sns.lineplot(data=b, x='bermain_game', y='count', hue='klasifikasi')  
  
plt.title('Korelasi antara bermain_game dan klasifikasi')  
plt.xlabel('Bermain Game')  
plt.ylabel('Count')  
  
plt.show()
```

Output:



Keterangan:

Terdapat klasifikasi yang terdiri dari:

1 = Tidak setuju, 2 = Netral, 3 = Setuju, 4= Sangat setuju

Sumbu X dari diagram garis di atas adalah nilai dari variable 'Apakah responden bermain game online', dimana:

1 = Ya, 2 = Tidak

Sumbu Y merupakan jumlah banyaknya nilai dari variable.

Dari visualisasi diatas, dapat dilihat bahwa nilai 2(Tidak) memiliki jumlah nilai tertinggi dengan warna yang menunjukkan klasifikasi ke 2 (Netral) . Hal ini dapat disimpulkan bahwa kebanyakan responden tidak bermain game online dan merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress

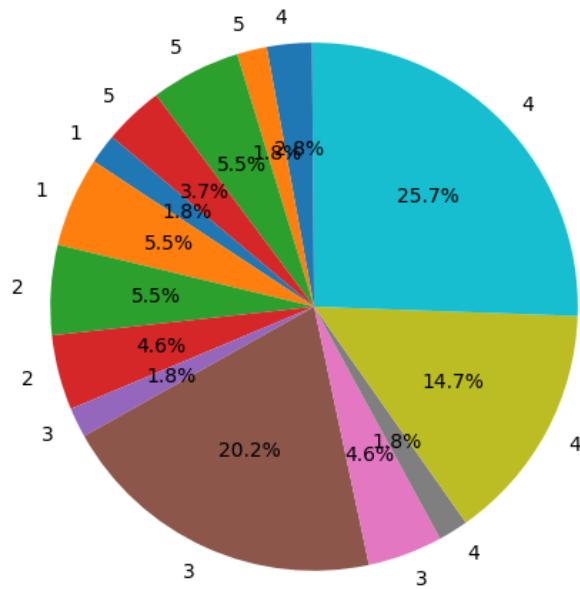
3. Pie Chart

- Korelasi Variabel “membantu_mengalihkan_stress” dengan “klasifikasi”
Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'membantu_mengalihkan_stress' dan 'klasifikasi'  
a = df.groupby(['membantu_mengalihkan_stress', 'klasifikasi']).size().reset_index(name='count')  
  
# Buat diagram pie  
plt.figure(figsize=(6, 6))  
plt.pie(a['count'], labels=a['membantu_mengalihkan_stress'], autopct='%1.1f%%', startangle=140)  
plt.title('Diagram Pie: Korelasi membantu_mengalihkan_stress dan klasifikasi')  
  
plt.show()
```

Output:

Diagram Pie: Korelasi membantu_mengalihkan_stress dan klasifikasi



Keterangan:

- 1 = Tidak membantu,
- 2 = Kurang membantu,
- 3 = Netral,
- 4 = Membantu,
- 5 = Sangat membantu.

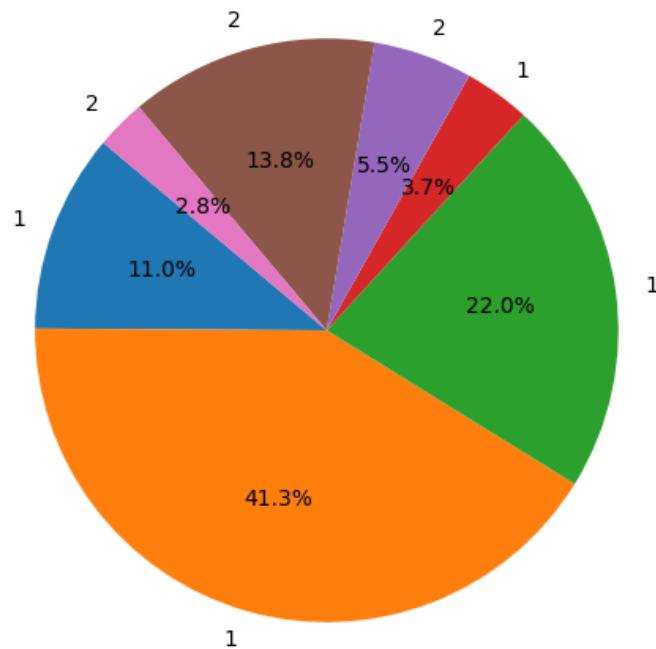
Dari pie chart yang dihasilkan, terlihat bahwa nilai yang mendominasi adalah nilai 4 (Membantu) untuk mengalihkan stress, yaitu sebesar 56.4%.3%, yang artinya bahwa 56.4% responden merasa bahwa game online dapat mengalihkan mereka dari sumber yang membuatnya merasa stress.

- b. Korelasi Variabel “pilihan_pertama_stress” dengan “klasifikasi”
Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'pilihan_pertama_stress' dan 'klasifikasi'  
b = df.groupby(['pilihan_pertama_stress', 'klasifikasi']).size().reset_index(name='count')  
  
# Buat diagram pie  
plt.figure(figsize=(6, 6))  
plt.pie(b['count'], labels=b['pilihan_pertama_stress'], autopct='%1.1f%%', startangle=140)  
plt.title('Diagram Pie: Korelasi pilihan_pertama_stress dan klasifikasi')  
  
plt.show()
```

Output:

Diagram Pie: Korelasi pilihan_pertama_stress dan klasifikasi



Keterangan :

1 = Ya

2 = Tidak

Dari pie chart yang dihasilkan, terlihat bahwa nilai yang mendominasi adalah nilai 1 (YA) untuk pilihan pertama stress, yaitu sebesar 74.3%, yang artinya bahwa 74.3% responden memilih game online sebagai pilihan pertama yang mereka lakukan ketika stress.

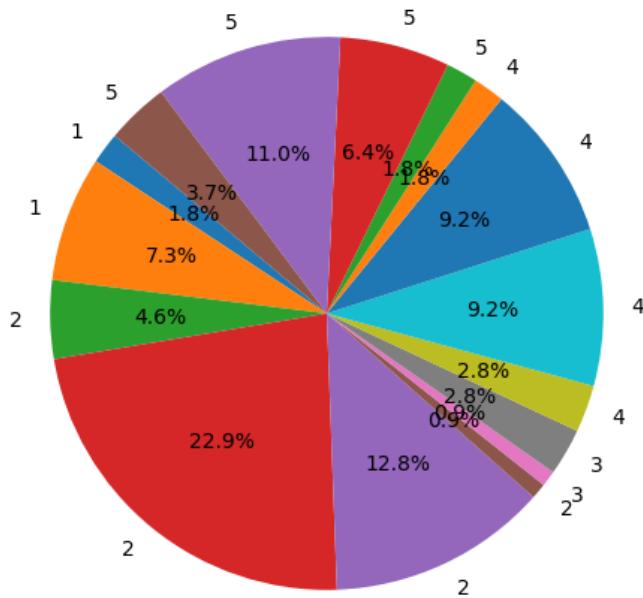
- c. Korelasi Variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'durasi_bermain' dan 'klasifikasi'  
c = df.groupby(['durasi_bermain', 'klasifikasi']).size().reset_index(name='count')  
  
# Buat diagram pie  
plt.figure(figsize=(6, 6))  
plt.pie(c['count'], labels=c['durasi_bermain'], autopct='%.1f%%', startangle=140)  
plt.title('Diagram Pie: Korelasi durasi_bermain dan klasifikasi')  
  
plt.show()
```

Output:

Diagram Pie: Korelasi durasi_bermain dan klasifikasi



Keterangan :

1 = Tidak pernah,

2 = Jarang,

3 = Sekali seminggu,

4 = Beberapa kali seminggu,

5 = Setiap hari

Dari pie chart yang dihasilkan, terlihat bahwa nilai yang mendominasi adalah nilai 2 (Jarang) untuk seberapa sering bermain game online dalam sebulan terakhir, yaitu sebesar 41.2%, yang artinya bahwa 41.2% responden dalam sebulan terakhir jarang bermain game online untuk menghilangkan stress.

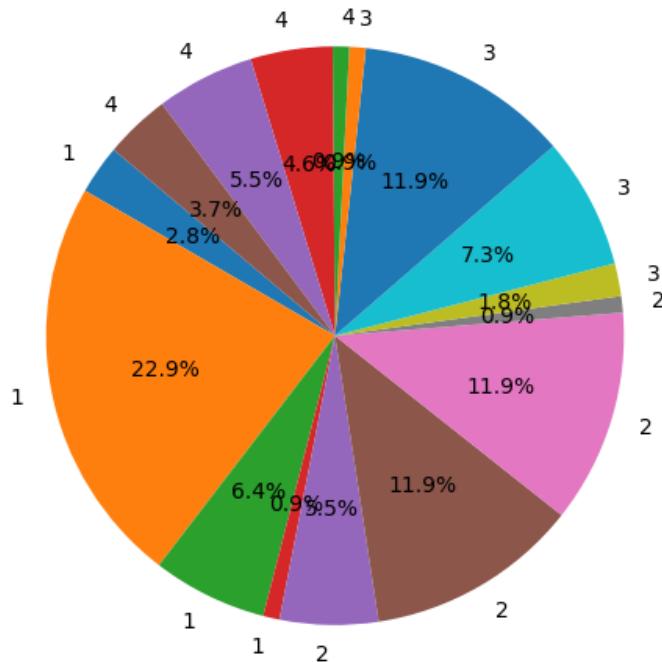
- d. Korelasi Variabel “durasi_waktu_bermain” dengan “klasifikasi”

Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'durasi_waktu_bermain' dan 'klasifikasi'  
d = df.groupby(['durasi_waktu_bermain', 'klasifikasi']).size().reset_index(name='count')  
  
# Buat diagram pie  
plt.figure(figsize=(6, 6))  
plt.pie(d['count'], labels=d['durasi_waktu_bermain'], autopct='%1.1f%%', startangle=140)  
plt.title('Diagram Pie: Korelasi durasi_waktu_bermain dan klasifikasi')  
  
plt.show()
```

Output:

Diagram Pie: Korelasi durasi_waktu_bermain dan klasifikasi



Keterangan :

- 1 = Kurang dari 30 menit,
- 2 = 30 menit - 1 jam,
- 3 = 1-2 jam,
- 4 = Lebih dari 2 jam

Dari pie chart yang dihasilkan, terlihat bahwa nilai yang mendominasi adalah nilai 1 (Kurang dari 30 menit) untuk mberapa lama durasi waktu responden bermain game online, yaitu sebesar 33%, yang artinya bahwa 33% responden biasanya bermain game online hanya sekitar kurang dari 30 menit.

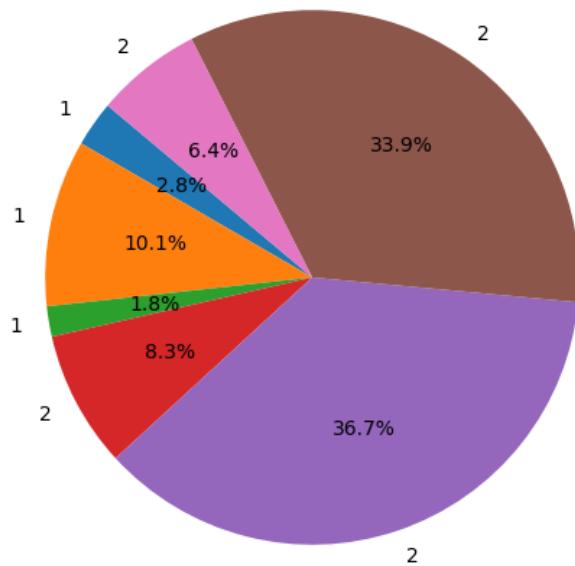
e. Korelasi Variabel “bermain_game” dengan “klasifikasi”

Syntax:

```
# Menghitung frekuensi kemunculan kombinasi 'bermain_game' dan 'klasifikasi'  
e = df.groupby(['bermain_game', 'klasifikasi']).size().reset_index(name='count')  
  
# Buat diagram pie  
plt.figure(figsize=(6, 6))  
plt.pie(e['count'], labels=e['bermain_game'], autopct='%1.1f%%', startangle=140)  
plt.title('Diagram Pie: Korelasi bermain_game dan klasifikasi')  
  
plt.show()
```

Output:

Diagram Pie: Korelasi bermain_game dan klasifikasi



Keterangan :

1 = Ya,

2 = Tidak

Dari pie chart yang dihasilkan, terlihat bahwa nilai yang mendominasi adalah nilai 2 (Tidak) untuk apakah responden bermain game online, yaitu sebesar 85.3%, yang artinya bahwa 85.3% responden tidak bermain game online untuk menghilangkan stress.

4. Area Chart

a. Korelasi Variabel “membantu_mengalihkan_stress” dengan “klasifikasi”

Syntax:

```
a = df.groupby(['klasifikasi', 'membantu_mengalihkan_stress']).size().reset_index(name='count')

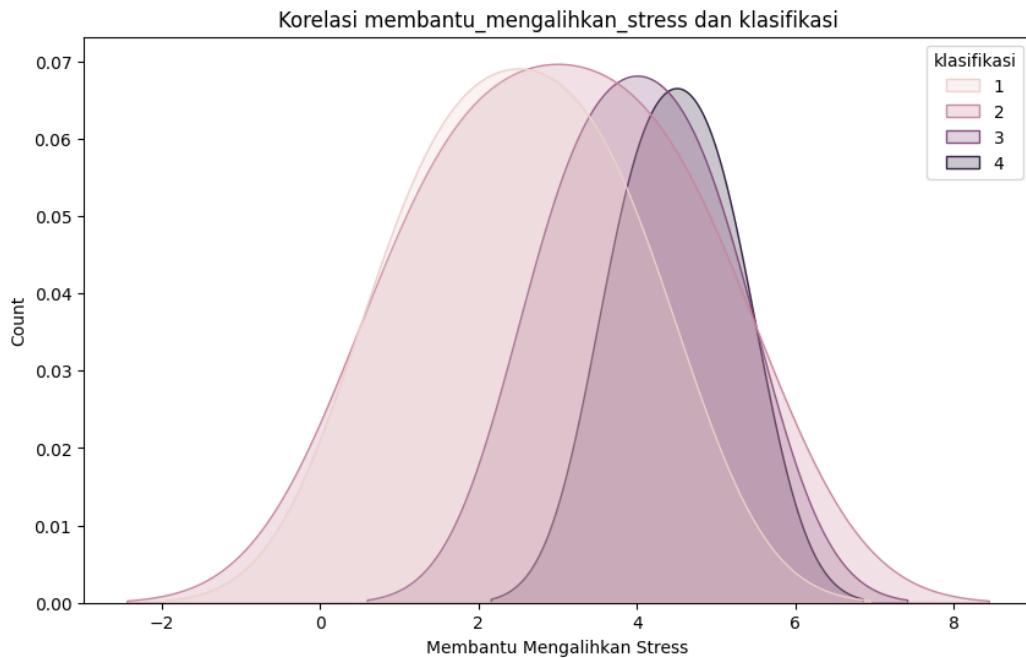
plt.figure(figsize=(10, 6))

# Menggunakan kdeplot untuk menggambarkan Kernel Density Plot
sns.kdeplot(data=a, x='membantu_mengalihkan_stress', hue='klasifikasi', fill=True)

plt.title('Korelasi membantu_mengalihkan_stress dan klasifikasi')
plt.xlabel('Membantu Mengalihkan Stress')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Dari visualisasi diatas, dapat dilihat bahwa varibel memiliki distribusi simetris dimana nilai yang paling tinggi berpusat pada area bagian tengah dengan area paling luas menunjukkan klasifikasi ke 2 yakni responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- b. Korelasi Variabel “pilihan_pertama_stress” dengan “klasifikasi”
Syntax:

```
b = df.groupby(['klasifikasi', 'pilihan_pertama_stress']).size().reset_index(name='count')

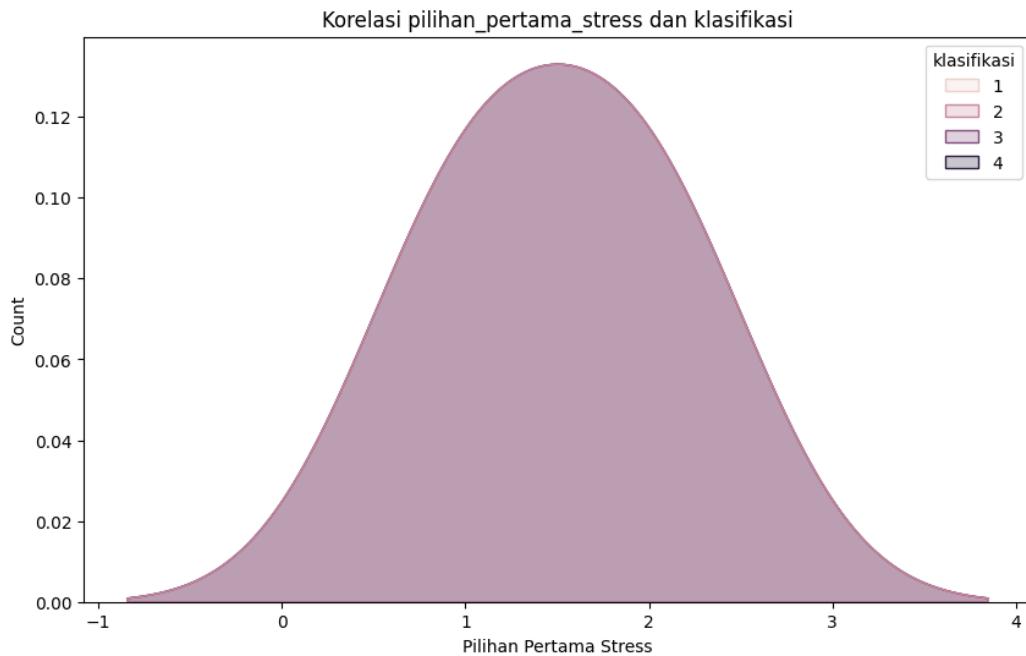
plt.figure(figsize=(10, 6))

# Menggunakan kdeplot untuk menggambarkan Kernel Density Plot
sns.kdeplot(data=b, x='pilihan_pertama_stress', hue='klasifikasi', fill=True)

plt.title('Korelasi pilihan_pertama_stress dan klasifikasi')
plt.xlabel('Pilihan Pertama Stress')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Dari visualisasi diatas, dapat dilihat bahwa varibel memiliki distribusi simetris dimana nilai yang paling tinggi berpusat pada area bagian tengah dengan area paling luas menunjukkan klasifikasi ke 2 yakni responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- c. Korelasi Variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
c = df.groupby(['klasifikasi', 'durasi_bermain']).size().reset_index(name='count')

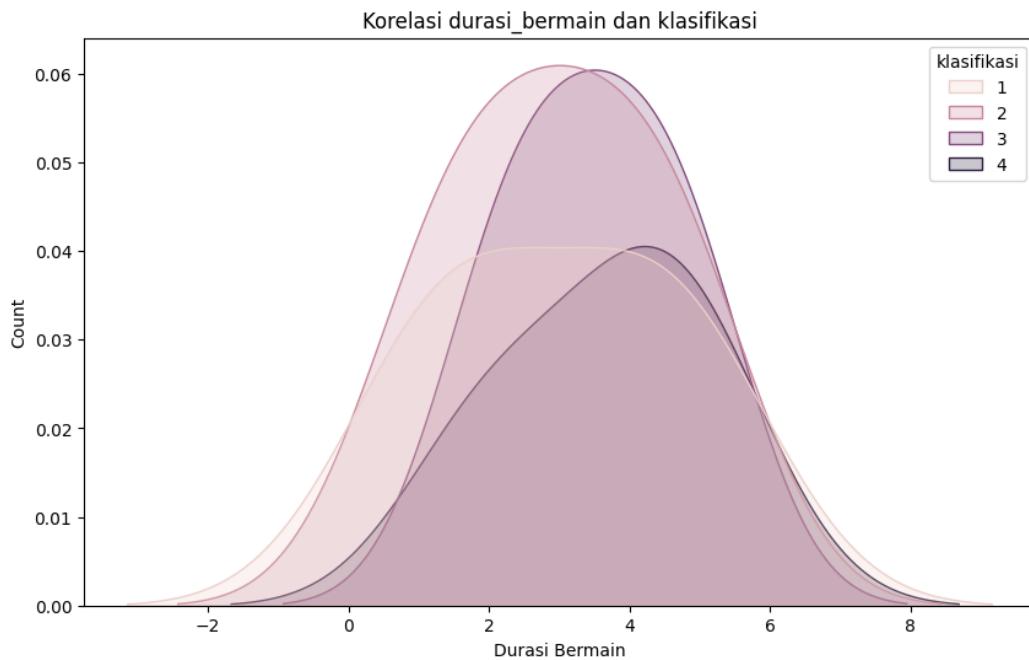
plt.figure(figsize=(10, 6))

# Menggunakan kdplot untuk menggambarkan Kernel Density Plot
sns.kdeplot(data=c, x='durasi_bermain', hue='klasifikasi', fill=True)

plt.title('Korelasi durasi_bermain dan klasifikasi')
plt.xlabel('Durasi Bermain')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Dari visualisasi diatas, dapat dilihat bahwa varibel memiliki distribusi simetris dimana nilai yang paling tinggi berpusat pada area bagian tengah dengan area paling luas menunjukkan klasifikasi ke 2 yakni responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- d. Korelasi Variabel “durasi_waktu_bermain” dengan “klasifikasi”

Syntax:

```
d = df.groupby(['Klasifikasi', 'durasi_waktu_bermain']).size().reset_index(name='count')

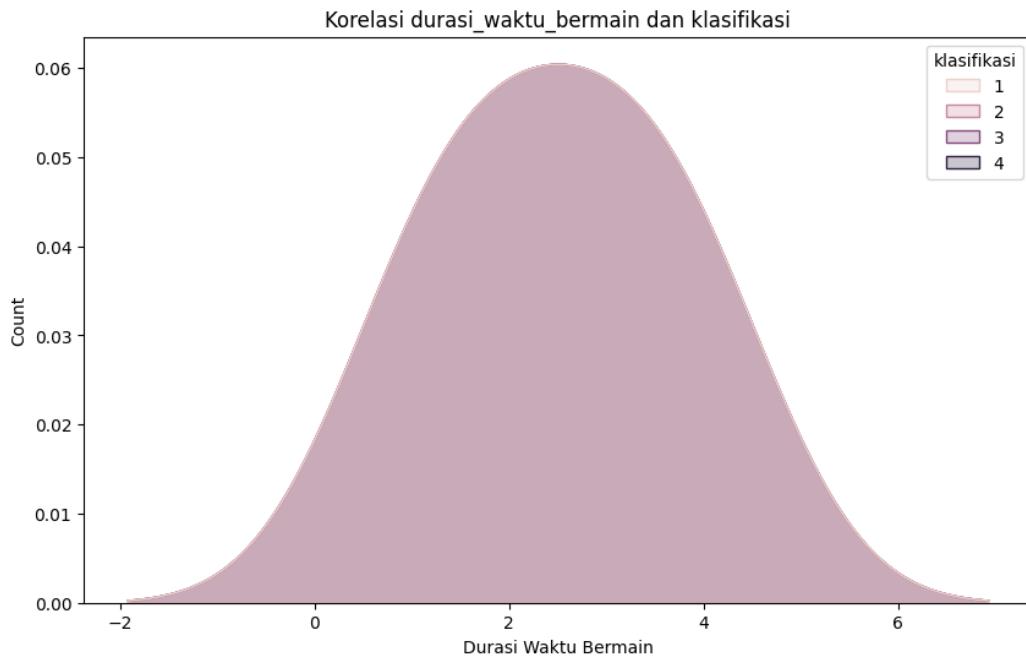
plt.figure(figsize=(10, 6))

# Menggunakan kdeplot untuk menggambarkan Kernel Density Plot
sns.kdeplot(data=d, x='durasi_waktu_bermain', hue='Klasifikasi', fill=True)

plt.title('Korelasi durasi_waktu_bermain dan klasifikasi')
plt.xlabel('Durasi Waktu Bermain')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Dari visualisasi diatas, dapat dilihat bahwa varibel memiliki distribusi simetris dimana nilai yang paling tinggi berpusat pada area bagian tengah dengan area paling luas menunjukkan klasifikasi ke 2 yakni responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

- e. Korelasi Variabel “bermain_game” dengan “klasifikasi”

Syntax:

```
e = df.groupby(['klasifikasi', 'bermain_game']).size().reset_index(name='count')

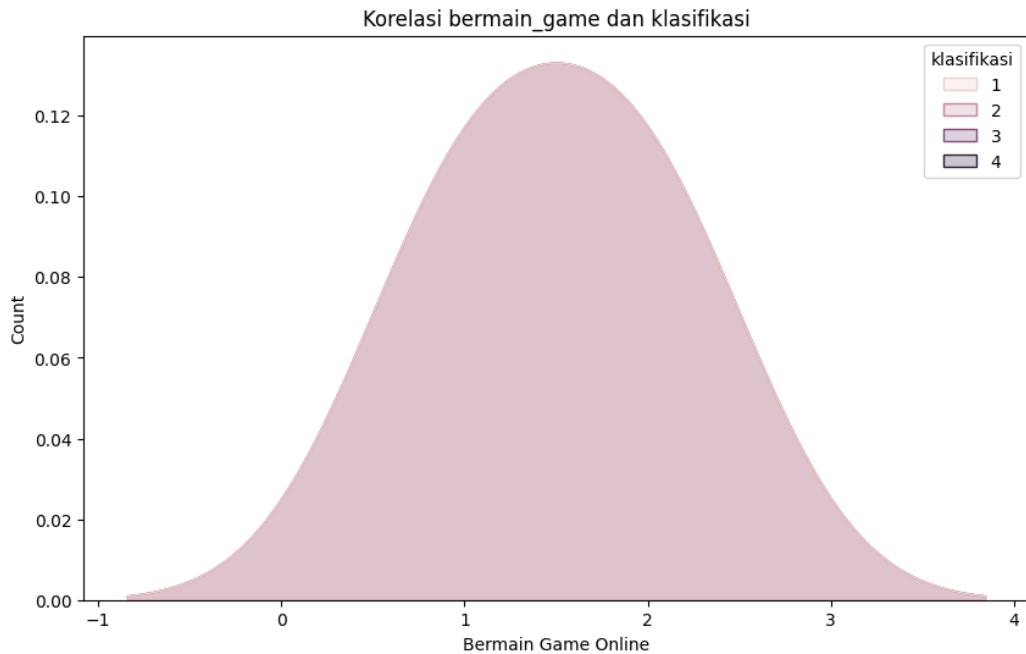
plt.figure(figsize=(10, 6))

# Menggunakan kdeplot untuk menggambarkan Kernel Density Plot
sns.kdeplot(data=e, x='bermain_game', hue='klasifikasi', fill=True)

plt.title('Korelasi bermain_game dan klasifikasi')
plt.xlabel('Bermain Game Online')
plt.ylabel('Count')

plt.show()
```

Output:



Keterangan:

Dari visualisasi diatas, dapat dilihat bahwa varibel memiliki distribusi simetris dimana nilai yang paling tinggi berpusat pada area bagian tengah dengan area paling luas menunjukkan klasifikasi ke 2 yakni responden merasa Netral bahwa game online merupakan cara alternatif yang sehat dan efektif untuk dapat menghilangkan stress.

7. Buatlah visualisasi high dimensional dan multivariate menggunakan Diagram, Batang, Garis, Pie, dan Area dari variable yang terpilih

1. Korelasi 2 Variabel menggunakan diagram batang

- a. Korelasi variabel “membantu_mengalihkan_stress” dengan “klasifikasi”

Syntax:

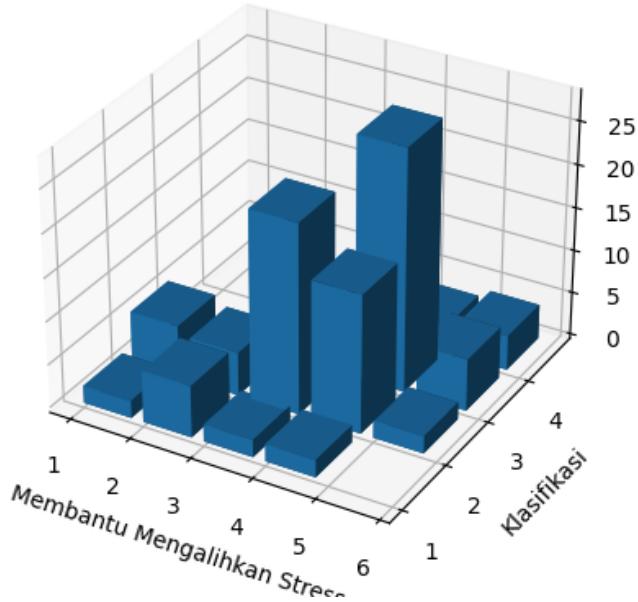
```
#Korelasi high-dimensional diagram batang dari variabel membantu_mengalihkan_stress
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'a' adalah DataFrame Anda dengan kolom 'game_menambah_stress', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram batang
x = a['membantu_mengalihkan_stress']
y = a['klasifikasi']
z = np.zeros(len(a)) # Tinggi batang diatur ke 0, karena tinggi batang akan diwakili oleh 'count'

dx = dy = 0.8 # Lebar diagram batang pada sumbu x dan y
dz = a['count']

ax.bar3d(x, y, z, dx, dy, dz, shade=True)

# Label sumbu x, y, dan z
ax.set_xlabel('Membantu Mengalihkan Stress')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram batang 3D
plt.show()
```

Output:



- b. Korelasi variabel “pilihan_pertama_stress” dengan “klasifikasi
- Syntax:

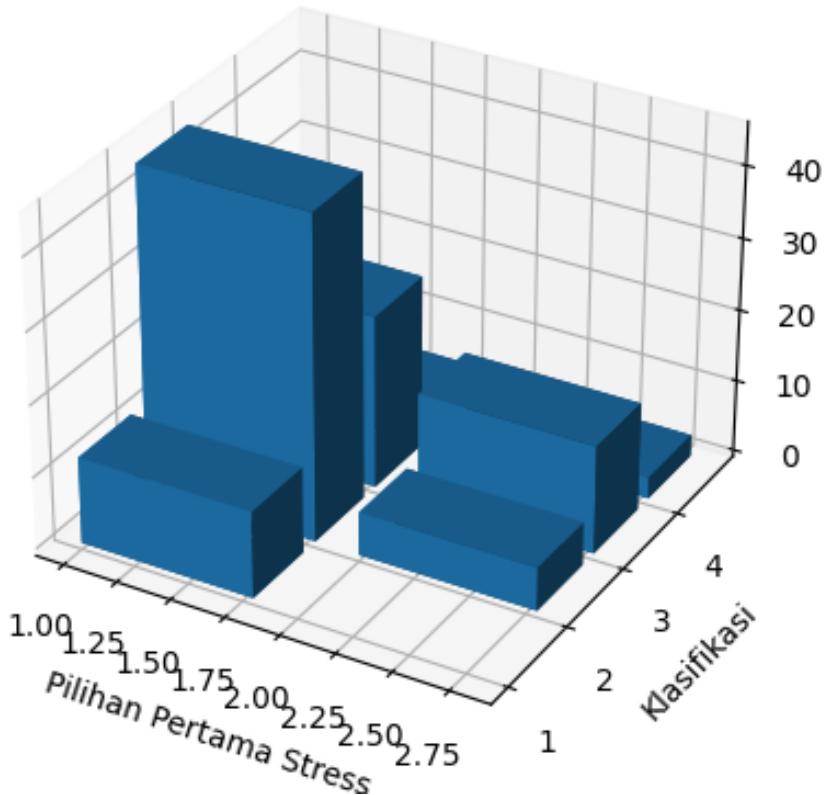
```
#Korelasi high-dimensional diagram batang dari variabel pilihan_pertama_stress dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'b' adalah DataFrame Anda dengan kolom 'pilihan_pertama_stress', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram batang
x = b['pilihan_pertama_stress']
y = b['klasifikasi']
z = np.zeros(len(b)) # Tinggi batang diatur ke 0, karena tinggi batang akan diwakili oleh 'count'

dx = dy = 0.8 # Lebar diagram batang pada sumbu x dan y
dz = b['count']

ax.bar3d(x, y, z, dx, dy, dz, shade=True)

# Label sumbu x, y, dan z
ax.set_xlabel('Pilihan Pertama Stress')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram batang 3D
plt.show()
```

Output:



- c. Korelasi variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

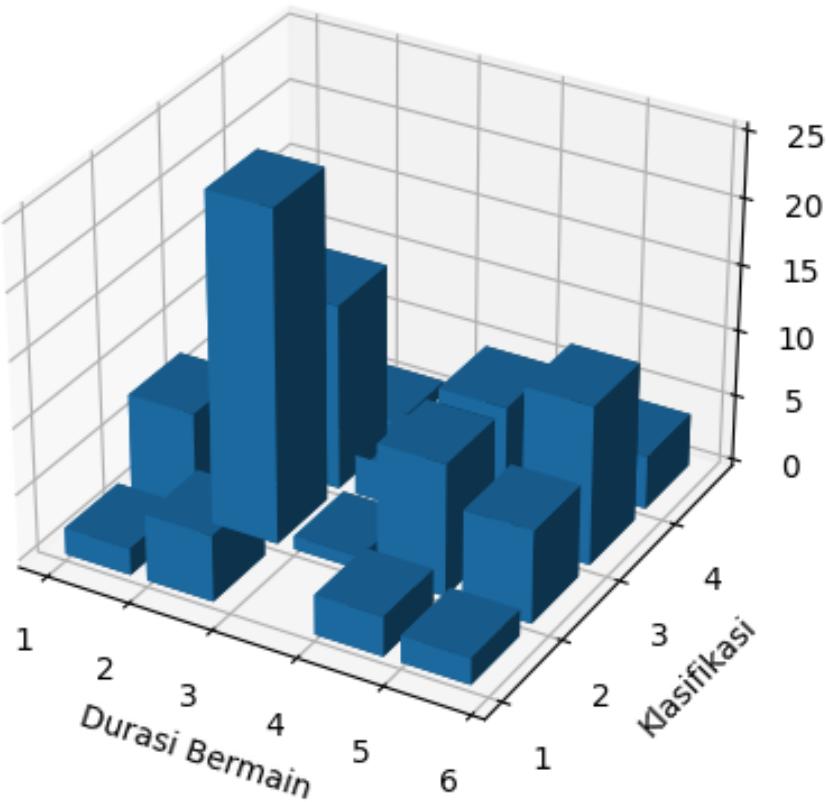
```
#Korelasi high-dimensional diagram batang dari variabel durasi_bermain dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'c' adalah DataFrame Anda dengan kolom 'durasi_bermain', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram batang
x = c['durasi_bermain']
y = c['klasifikasi']
z = np.zeros(len(c)) # Tinggi batang diatur ke 0, karena tinggi batang akan diwakili oleh 'count'

dx = dy = 0.8 # Lebar diagram batang pada sumbu x dan y
dz = c['count']

ax.bar3d(x, y, z, dx, dy, dz, shade=True)

# Label sumbu x, y, dan z
ax.set_xlabel('Durasi Bermain')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram batang 3D
plt.show()
```

Output:



- d. Korelasi variabel “durasi_waktu_bermain” dengan “klasifikasi”
Syntax:

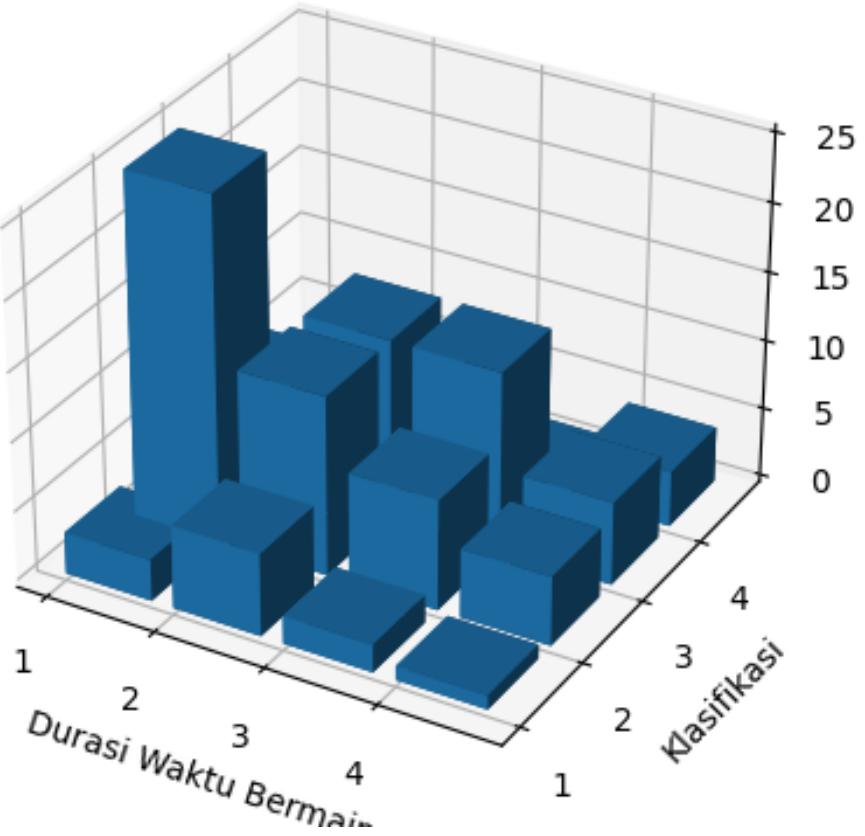
```
#Korelasi high-dimensional diagram batang dari variabel durasi_waktu_bermain dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'a' adalah DataFrame Anda dengan kolom 'durasi_waktu_bermain', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram batang
x = d['durasi_waktu_bermain']
y = d['klasifikasi']
z = np.zeros(len(d)) # Tinggi batang diatur ke 0, karena tinggi batang akan diwakili oleh 'count'

dx = dy = 0.8 # Lebar diagram batang pada sumbu x dan y
dz = d['count']

ax.bar3d(x, y, z, dx, dy, dz, shade=True)

# Label sumbu x, y, dan z
ax.set_xlabel('Durasi Waktu Bermain')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram batang 3D
plt.show()
```

Output:



- e. Korelasi variabel “bermain_game” dengan “klasifikasi”

Syntax:

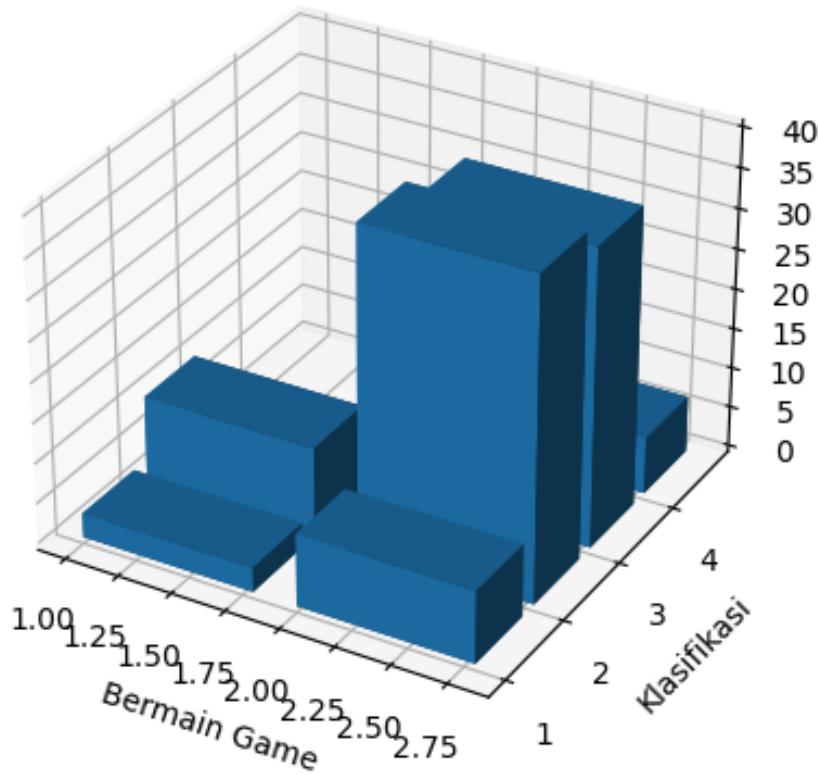
```
#Korelasi high-dimensional diagram batang dari variabel bermain_game dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'e' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram batang
x = e['bermain_game']
y = e['klasifikasi']
z = np.zeros(len(e)) # Tinggi batang diatur ke 0, karena tinggi batang akan diwakili oleh 'count'

dx = dy = 0.8 # Lebar diagram batang pada sumbu x dan y
dz = e['count']

ax.bar3d(x, y, z, dx, dy, dz, shade=True)

# Label sumbu x, y, dan z
ax.set_xlabel('Bermain Game')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram batang 3D
plt.show()
```

Output:

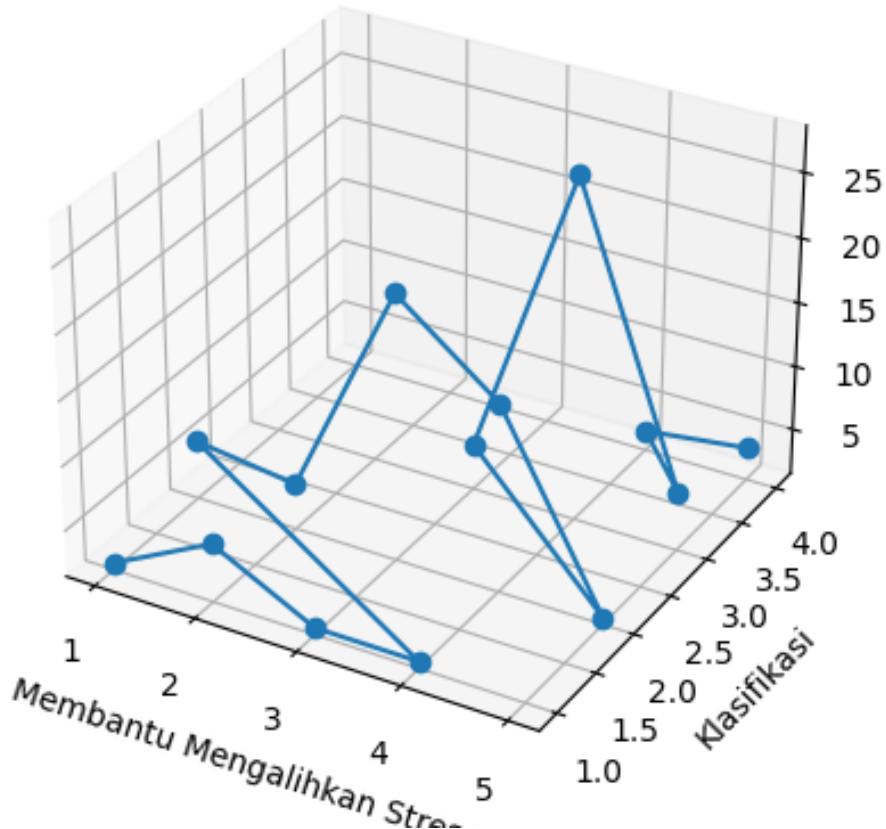


2. Korelasi 2 variabel dengan diagram garis

- a. Korelasi variabel “membantu_mengalihkan_stress” dengan “klasifikasi”
Syntax:

```
#Korelasi high-dimensional diagram garis dari variabel membantu_mengalihkan_stress dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'a' adalah DataFrame Anda dengan kolom 'membantu_mengalihkan_stress', 'klasifikasi', dan 'count'
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram garis
x = a['membantu_mengalihkan_stress']
y = a['klasifikasi']
z = a['count']
# Menggambar diagram garis 3D
ax.plot(x, y, z, marker='o', linestyle='-' )
# Label sumbu x, y, dan z
ax.set_xlabel('Membantu Mengalihkan Stress')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')
# Tampilkan diagram garis 3D
plt.show()
```

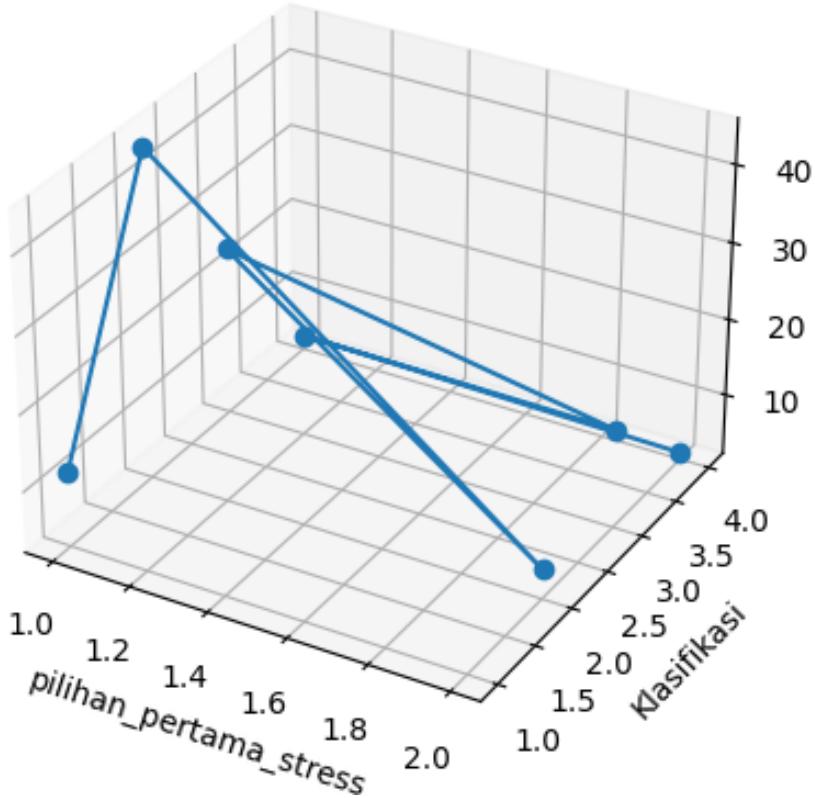
Output:



- b. Korelasi variabel “pilihan_pertama_stress” dengan “klasifikasi”
Syntax:

```
#Korelasi high-dimensional diagram garis dari variabel pilihan_pertama_stress dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'a' adalah DataFrame Anda dengan kolom 'pilihan_pertama_stress', 'klasifikasi', dan 'count'
fig = plt.figure()
bx = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram garis
x = b['pilihan_pertama_stress']
y = b['klasifikasi']
z = b['count']
# Menggambar diagram garis 3D
bx.plot(x, y, z, marker='o', linestyle='-' )
# Label sumbu x, y, dan z
bx.set_xlabel('pilihan_pertama_stress')
bx.set_ylabel('Klasifikasi')
bx.set_zlabel('Count')
# Tampilkan diagram garis 3D
plt.show()
```

Output:

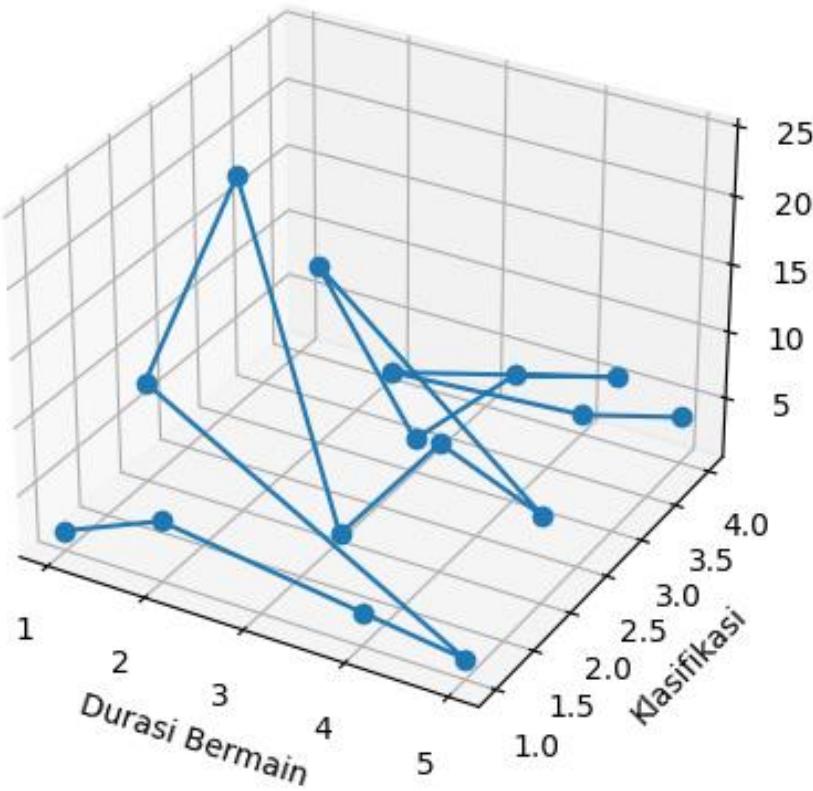


- c. Korelasi variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
#Korelasi high-dimensional diagram garis dari variabel durasi_bermain dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'c' adalah DataFrame Anda dengan kolom 'durasi_bermain', 'klasifikasi', dan 'count'
fig = plt.figure()
cx = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram garis
x = c['durasi_bermain']
y = c['klasifikasi']
z = c['count']
# Menggambar diagram garis 3D
cx.plot(x, y, z, marker='o', linestyle='--')
# Label sumbu x, y, dan z
cx.set_xlabel('Durasi Bermain')
cx.set_ylabel('Klasifikasi')
cx.set_zlabel('Count')
# Tampilkan diagram garis 3D
plt.show()
```

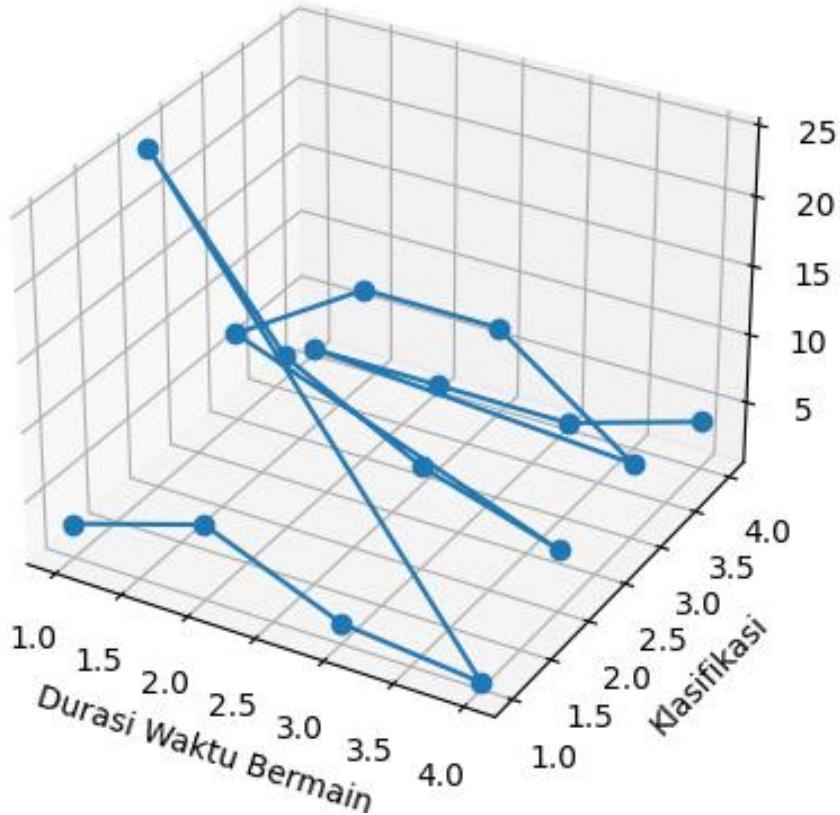
Output:



- d. Korelasi variabel “durasi_waktu_bermain” dengan “klasifikasi”
Syntax:

```
#Korelasi high-dimensional diagram garis dari variabel durasi_waktu_bermain dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'd' adalah DataFrame Anda dengan kolom 'durasi_waktu_bermain', 'klasifikasi', dan 'count'
fig = plt.figure()
dx = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram garis
x = d['durasi_waktu_bermain']
y = d['klasifikasi']
z = d['count']
# Menggambar diagram garis 3D
dx.plot(x, y, z, marker='o', linestyle='-' )
# Label sumbu x, y, dan z
dx.set_xlabel('Durasi Waktu Bermain')
dx.set_ylabel('Klasifikasi')
dx.set_zlabel('Count')
# Tampilkan diagram garis 3D
plt.show()
```

Output:

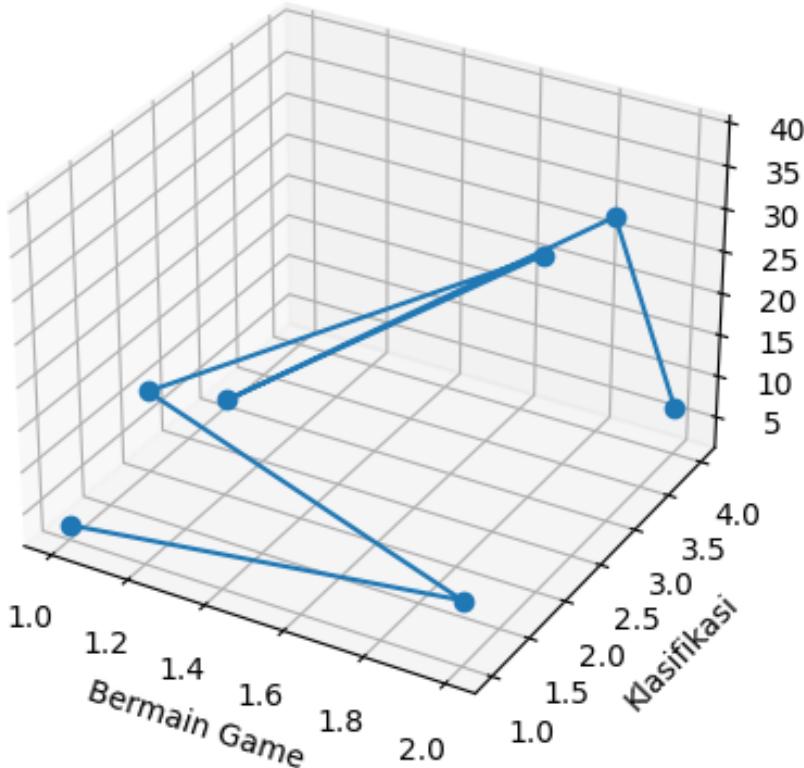


- e. Korelasi variabel “bermain_game” dengan ”klasifikasi”

Syntax:

```
#Korelasi high-dimensional diagram garis dari variabel bermain_game dan klasifikasi
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
# Misalkan 'a' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'
fig = plt.figure()
ex = fig.add_subplot(111, projection='3d')
# Koordinat x, y, dan z untuk diagram garis
x = e['bermain_game']
y = e['klasifikasi']
z = e['count']
# Menggambar diagram garis 3D
ex.plot(x, y, z, marker='o', linestyle='--')
# Label sumbu x, y, dan z
ex.set_xlabel('Bermain Game')
ex.set_ylabel('Klasifikasi')
ex.set_zlabel('Count')
# Tampilkan diagram garis 3D
plt.show()
```

Output:



8. Analisis diagram yang telah dibuat pada nomer 7

1. Analisis korelasi 2 variabel menggunakan diagram batang

a. Korelasi variabel "membantu_mengalihkan_stress" dengan "klasifikasi"

Melalui analisis diagram batang 3D, kita dapat mengamati sebaran data yang mencerminkan hubungan antara variabel 'membantu_mengalihkan_stress' dan 'klasifikasi'. Diagram ini memberikan pandangan visual tentang bagaimana kombinasi nilai dalam kedua variabel ini berdampak pada jumlah pengamatan ('count'). Kita melihat bahwa tinggi batang bervariasi di seluruh diagram, menunjukkan variasi jumlah pengamatan untuk setiap kombinasi 'membantu_mengalihkan_stress' dan 'klasifikasi'. Kita dapat melihat bahwa ada beberapa kombinasi yang memiliki tinggi batang yang jauh lebih tinggi dari yang lain, yang mengindikasikan bahwa kombinasi tersebut muncul lebih sering dalam data. Korelasi antara kedua variabel dapat dianalisa dari hasil visualisasi diagram batang di atas yang menunjukkan bahwa responden merasa dengan bermain game online telah membantu mengalihkan perhatian dari sumber stres yang sedang dialami dan responden juga setuju jika dengan bermain game online dapat menjadi cara yang efektif untuk menghilangkan stress dengan cara yang sehat. Jadi korelasi antara kedua variabel ini yaitu game online menjadi salah satu cara yang efektif dan dapat membantu untuk menghilangkan stess.

b. Korelasi variabel "pilihan_pertama_stress" dengan "klasifikasi"

Melalui analisis diagram batang 3D, kita mendapatkan wawasan tentang bagaimana variabel 'pilihan_pertama_stress' dan 'klasifikasi' berhubungan dengan jumlah pengamatan ('count'). Kita melihat variasi dalam tinggi batang, yang menunjukkan perbedaan jumlah pengamatan untuk setiap kombinasi nilai dalam

kedua variabel. Terdapat kombinasi yang muncul lebih sering daripada yang lain, ditandai dengan tinggi batang yang lebih tinggi. Selanjutnya, pola muncul dalam data. Beberapa kombinasi 'pilihan_pertama_stress' dan 'klasifikasi' memiliki tinggi batang yang berbeda. Ini menyarankan adanya korelasi antara variabel-variabel tersebut. Korelasi antara kedua variabel ini dapat dianalisa dari hasil visualisasi diagram balok di atas yang menunjukkan responden setuju jika bermain game online menjadi pilihan pertama ketika merasa stres dan responden juga memilih setuju jika dengan bermain game dapat menjadi sebuah cara untuk menghilangkan rasa stress yang sehat. Hal ini dibuktikan dengan hasil visualisasi pada diagram balok menunjukkan pada variabel pilihan_pertama_stress balok yang paling tinggi berada pada angka 1 dan pada variabel klasifikasi balok yang paling tinggi berada pada angka 2. Jadi korelasinya yaitu game online dapat menjadi pilihan pertama sebagai langkah untuk menghilangkan stress.

c. Korelasi variabel "durasi_bermain" dengan "klasifikasi"

Berdasarkan hasil diagram 3D di atas, data yang digunakan dalam visualisasi ini yaitu variabel 'durasi_bermain' dan 'klasifikasi' akan digunakan sebagai sumbu x dan y, sementara 'count' akan digunakan untuk menentukan tinggi batang. Diagram di atas memiliki tinggi batang yang sangat variatif. Ada salah satu batang yang paling tinggi yaitu hasil klasifikasi berangka 3. Kemudian yang kedua diikuti oleh tinggi batang dari klasifikasi yang berangka 4. Dari tinggi balok yang variatif menunjukkan bahwa ada beberapa kombinasi yang memiliki tinggi batang yang jauh lebih tinggi dari yang lain, yang mengindikasikan bahwa kombinasi tersebut muncul lebih sering dalam data. Hasilnya adalah diagram batang 3D yang mencerminkan hubungan antara "durasi_bermain" dan "klasifikasi" dengan tinggi batang yang menunjukkan frekuensi atau jumlah pengamatan. Korelasi antara kedua variabel ini dapat dianalisa bahwa responden sering bermain game online hanya beberapa kali seminggu dalam sebulan terakhir dan responden juga setuju bahwa dengan bermain game online dapat menjadi sebuah cara untuk menghilangkan stres yang sehat dan efektif. Pernyataan ini diperkuat dengan hasil visualisasi diagram batang di atas yang menunjukkan pada variabel durasi bermain paling tinggi baloknya berada di angka 2 yang berarti banyak responden yang bermain game online hanya beberapa kali seminggu, sementara pada variabel klasifikasi responden juga banyak yang memilih setuju jika dengan bermain game online dapat menjadi sebuah cara untuk menghilangkan stress. Jadi korelasi dari kedua variabel yaitu dengan bermain game selama beberapa kali seminggu dapat menjadi sebuah langkah untuk membantu menghilangkan rasa stress.

d. Korelasi variabel "durasi_waktu_bermain" dengan "klasifikasi"

Pada hasil diagram batang di atas, Hasilnya adalah diagram batang 3D yang mencerminkan hubungan antara "durasi_waktu_bermain" dan "klasifikasi" dengan tinggi batang yang menunjukkan frekuensi atau jumlah pengamatan. Dengan visualisasi ini, kita dapat dengan mudah melihat pola dan korelasi antara dua variabel ini dan memahami distribusi data dengan lebih baik. Ini memungkinkan untuk menilai apakah ada pola atau hubungan yang signifikan antara "durasi_waktu_bermain" dan "klasifikasi" dalam dataset kita. Pada diagram batang di atas sangat jelas terlihat bahwa durasi waktu bermain game itu sekitar kurang dari 30 menit yang mana ditunjukkan dengan balok yang paling tinggi berada di angka 1 dan hal itu dapat menjadi cara untuk menghilangkan stres secara alternatif yang sehat dan efektif yang mana ditunjukkan dengan hasil klasifikasi di angka 2

yaitu setuju. Jadi korelasi dari kedua variabel yaitu hanya dengan bermain game kurang dari 30 menit dapat menjadi sebuah langkah untuk membantu menghilangkan rasa stress.

e. Korelasi variabel "bermain_game" dengan "klasifikasi"

Diagram batang 3D di atas mencerminkan hubungan antara variabel "bermain_game" dan "klasifikasi" dengan tinggi batang yang menunjukkan frekuensi atau jumlah pengamatan. Dengan visualisasi ini, kita dapat dengan mudah melihat pola dan korelasi antara dua variabel ini dan memahami distribusi data dengan lebih baik. Ini memungkinkan untuk mengevaluasi apakah ada hubungan yang signifikan antara "bermain_game" dan "klasifikasi" dalam dataset kita, dan apakah ada kategori klasifikasi tertentu yang mungkin cenderung bermain game lebih sering atau sebaliknya. Jika diamati dari diagram batang di atas, korelasi antara kedua variabel ini dapat dianalisa bahwa responden banyak yang merasa dengan bermain game online dapat mengganggu waktu belajar dan aktivitas sehari - hari akan tetapi responden juga setuju bahwa bermain game merupakan cara untuk menghilangkan stres yang sehat dan efektif, pernyataan ini dapat didukung dengan melihat hasil visualisasi dari tingginya balok pada variabel bermain_online paling tinggi berada pada angka 2 yang menunjukkan bahwa responden lebih banyak yang memilih netral dan pada variabel klasifikasi paling tinggi juga berada pada angka 2 yang menunjukkan bahwa responden banyak yang memilih setuju.

2. Analisis korelasi 2 variabel dengan diagram garis

a. Korelasi variabel "membantu_mengalihkan_stress" dengan "klasifikasi"

Dari diagram garis 3D yang dihasilkan oleh syntax di atas, kita dapat menganalisis hubungan antara dua variabel, yaitu "Membantu Mengalihkan Stress" dan "Klasifikasi", serta bagaimana hubungannya berkaitan dengan "Count" (kemungkinan frekuensi atau jumlah pengamatan). Dari diagram garis ini, kita dapat melihat bagaimana nilai variabel "Membantu Mengalihkan Stress" berhubungan dengan "Klasifikasi". Jika garis ini menunjukkan tren yang miring ke atas atau ke bawah, ini menunjukkan bahwa ada korelasi antara dua variabel ini. Gambar diagram garis di atas sangat jelas menjelaskan bahwa terdapat korelasi antar kedua variabel karena titik satu dengan yang lain saling terhubung satu sama lain.

b. Korelasi variabel "pilihan_pertama_stress" dengan "klasifikasi"

Dalam diagram garis 3D yang disajikan, kita dapat melihat bagaimana variabel "Pilihan Pertama Stress" dan "Klasifikasi" berkorelasi. Garis yang dihasilkan dalam plot 3D memberikan gambaran tentang hubungan antara keduanya. Jika garis tersebut cenderung naik, ini menunjukkan bahwa ketika skor "Pilihan Pertama Stress" meningkat, kemungkinan "Klasifikasi" juga cenderung meningkat. Dengan kata lain, ada korelasi positif antara kedua variabel ini. Ini bisa mengartikan bahwa, dalam konteks data yang diambil, ketika tingkat "Pilihan Pertama Stress" menjadi lebih tinggi, klasifikasi juga cenderung menjadi lebih tinggi, mengindikasikan bahwa kategori atau label yang lebih tinggi lebih mungkin terkait dengan tingkat stres pertama yang lebih tinggi.

Sebaliknya, jika garis cenderung menurun, itu menandakan korelasi negatif antara kedua variabel tersebut. Artinya, ketika skor "Pilihan Pertama Stress" meningkat, kemungkinan "Klasifikasi" cenderung menurun. Ini bisa berarti bahwa, dalam

konteks data tersebut, kategori yang lebih tinggi atau label yang lebih rendah lebih mungkin berkaitan dengan tingkat stres pertama yang lebih tinggi. Namun, jika garis tersebut mendatar, ini mengindikasikan bahwa tidak ada korelasi yang kuat antara kedua variabel tersebut. Dengan kata lain, perubahan dalam "Pilihan Pertama Stress" tidak memiliki pengaruh yang konsisten terhadap "Klasifikasi".

c. Korelasi variabel "durasi_bermain" dengan "klasifikasi"

Dalam plot garis 3D yang telah dibuat, kita dapat mengamati hubungan antara "Durasi Bermain" dan "Klasifikasi." Garis yang dihasilkan menunjukkan pola hubungan antara kedua variabel tersebut. Apabila garis ini cenderung naik, maka hal ini mengindikasikan adanya korelasi positif antara "Durasi Bermain" dan "Klasifikasi." Dengan kata lain, semakin lama durasi bermain, semakin tinggi kemungkinan seseorang diklasifikasikan dalam kategori atau label yang lebih tinggi. Dalam konteks data yang tersaji, ini bisa diartikan bahwa aktivitas bermain yang lebih lama cenderung berhubungan dengan hasil klasifikasi yang lebih tinggi. Sebaliknya, jika garis tersebut cenderung menurun, hal ini menunjukkan adanya korelasi negatif antara kedua variabel tersebut. Ini berarti bahwa semakin lama durasi bermain, kemungkinan "Klasifikasi" cenderung lebih rendah. Dalam konteks data, ini bisa diartikan sebagai fakta bahwa durasi bermain yang lebih lama cenderung berhubungan dengan klasifikasi yang lebih rendah. Akan tetapi jika garis tersebut mendatar, ini mengindikasikan bahwa tidak ada korelasi yang kuat antara kedua variabel tersebut. Artinya, perubahan dalam "Durasi Bermain" tidak memiliki pengaruh yang konsisten terhadap "Klasifikasi."

d. Korelasi variabel "durasi_waktu_bermain" dengan "klasifikasi"

Pada diagram garis 3D di atas, kita dapat melihat bagaimana variabel "Durasi Waktu Bermain" berkorelasi dengan "Klasifikasi". Garis yang terbentuk dalam plot 3D memberikan gambaran tentang hubungan antara keduanya. Jika garis tersebut cenderung naik, hal ini menandakan adanya korelasi positif antara kedua variabel tersebut. Artinya, semakin lama durasi bermain, semakin tinggi kemungkinan seseorang diklasifikasikan dalam kategori atau label yang lebih tinggi. Sebaliknya, jika garis tersebut cenderung menurun, ini mengindikasikan adanya korelasi negatif. Ini berarti semakin lama durasi bermain, kemungkinan klasifikasi menjadi lebih rendah. Namun, jika garis tersebut mendatar, ini menunjukkan bahwa tidak ada korelasi yang kuat antara kedua variabel tersebut. Dengan kata lain, perubahan dalam "Durasi Waktu Bermain" tidak memiliki pengaruh yang konsisten terhadap "Klasifikasi". Penting untuk diingat bahwa analisis ini merupakan interpretasi awal dari diagram garis 3D.

e. Korelasi variabel "bermain_game" dengan "klasifikasi"

Dari hasil diagram garis 3D di atas, kita dapat mengamati hubungan antara variabel "Bermain Game" dan "Klasifikasi". Garis yang terbentuk dalam plot 3D memberikan gambaran tentang korelasi antara keduanya. Jika garis tersebut cenderung naik, ini mengindikasikan adanya korelasi positif antara "Bermain Game" dan "Klasifikasi". Dengan kata lain, semakin sering atau lama seseorang bermain game, semakin tinggi kemungkinan mereka diklasifikasikan dalam kategori atau label yang lebih tinggi. Sebaliknya, jika garis tersebut cenderung menurun, ini menunjukkan adanya korelasi negatif. Ini berarti semakin banyak bermain game, kemungkinan klasifikasi menjadi lebih rendah. Jika garis tersebut mendatar, ini menunjukkan bahwa tidak ada korelasi yang kuat antara kedua

variabel tersebut. Dalam hal ini, perubahan dalam "Bermain Game" tidak memiliki pengaruh yang konsisten terhadap "Klasifikasi".

9. Buatlah visualisasi high dimensional dan multivariate menggunakan Scatter Plot dari variabel yang terpilih

- a. Korelasi Variabel “membantu_mengalihkan_stress” dengan “klasifikasi”

Syntax:

```
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np

# Misalkan 'a' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'

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

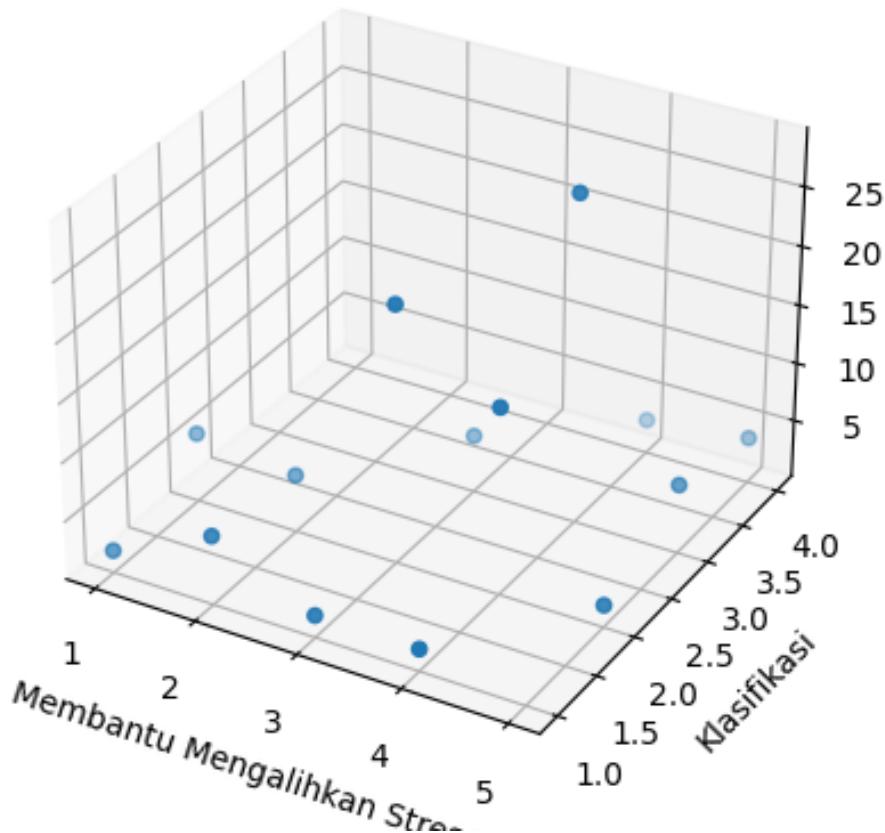
# Koordinat x, y, dan z untuk scatter plot 3D
x = a['membantu_mengalihkan_stress']
y = a['klasifikasi']
z = a['count']

# Menggambar scatter plot 3D
ax.scatter(x, y, z, marker='o')

# Label sumbu x, y, dan z
ax.set_xlabel('Membantu Mengalihkan Stress')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')

# Tampilkan scatter plot 3D
plt.show()
```

Output:



- b. Korelasi Variabel “pilihan_pertama_stress” dengan “klasifikasi”
Syntax:

```
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np

# Misalkan 'a' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'

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

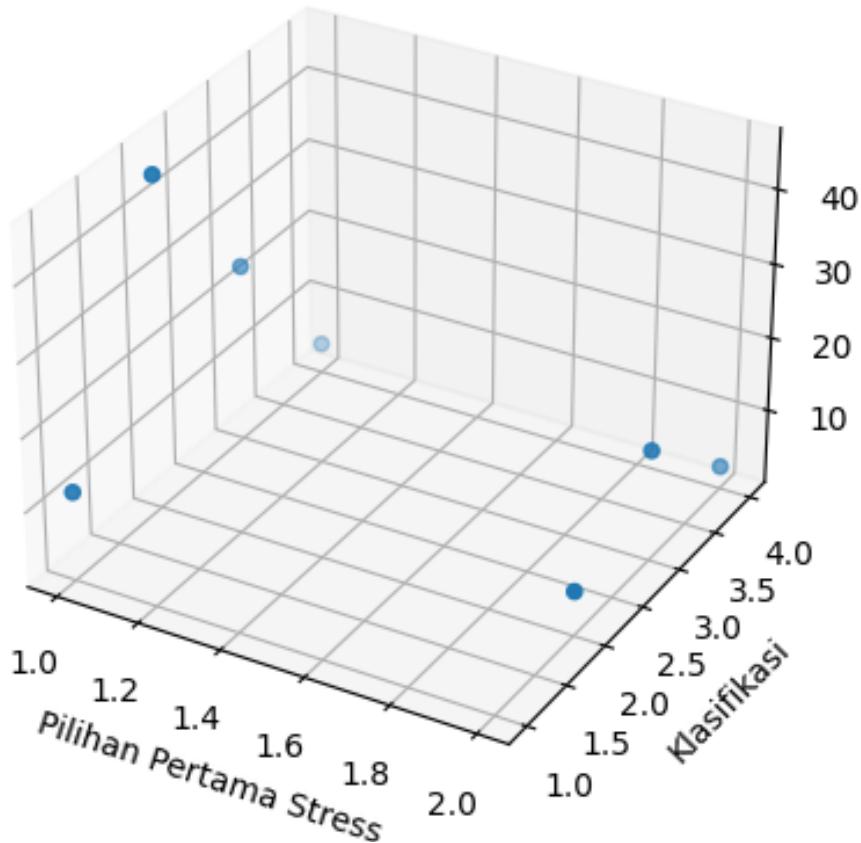
# Koordinat x, y, dan z untuk scatter plot 3D
x = b['pilihan_pertama_stress']
y = b['klasifikasi']
z = b['count']

# Menggambar scatter plot 3D
ax.scatter(x, y, z, marker='o')

# Label sumbu x, y, dan z
ax.set_xlabel('Pilihan Pertama Stress')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')

# Tampilkan scatter plot 3D
plt.show()
```

Output:



- c. Korelasi Variabel “durasi_bermain” dengan “klasifikasi”

Syntax:

```
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np

# Misalkan 'a' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'

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

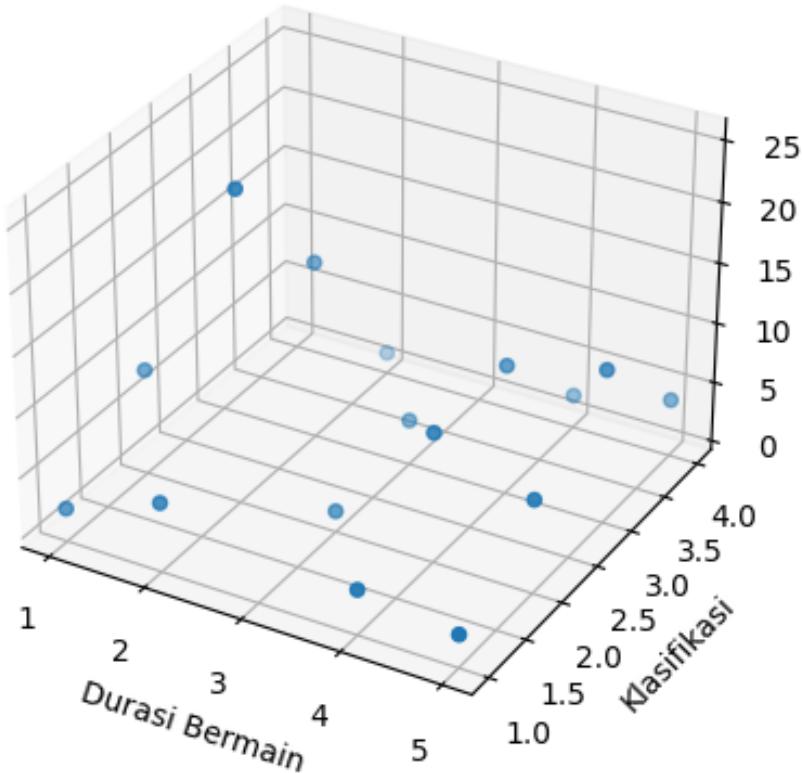
# Koordinat x, y, dan z untuk scatter plot 3D
x = c['durasi_bermain']
y = c['klasifikasi']
z = c['count']

# Menggambar scatter plot 3D
ax.scatter(x, y, z, marker='o')

# Label sumbu x, y, dan z
ax.set_xlabel('Durasi Bermain')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')

# Tampilkan scatter plot 3D
plt.show()
```

Output:



- d. Korelasi Variabel “durasi_waktu_bermain” dengan “klasifikasi”
Syntax:

```
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np

# Misalkan 'a' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'

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

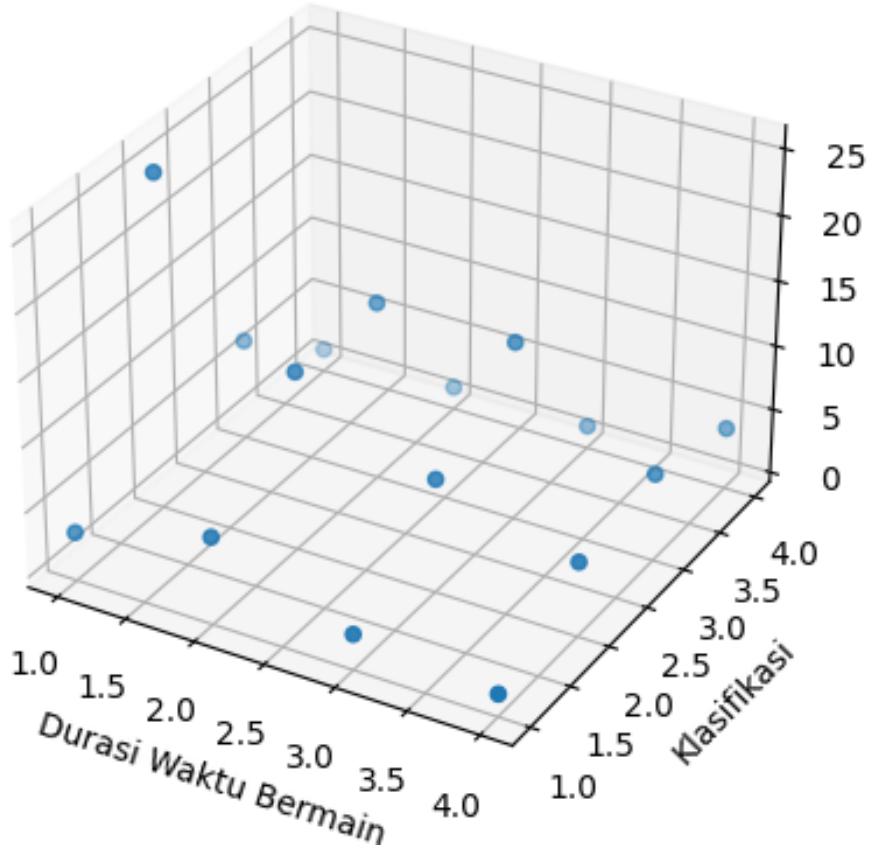
# Koordinat x, y, dan z untuk scatter plot 3D
x = d['durasi_waktu_bermain']
y = d['klasifikasi']
z = d['count']

# Menggambar scatter plot 3D
ax.scatter(x, y, z, marker='o')

# Label sumbu x, y, dan z
ax.set_xlabel('Durasi Waktu Bermain')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')

# Tampilkan scatter plot 3D
plt.show()
```

Output:



- e. Korelasi Variabel “bermain_game” dengan “klasifikasi”

Syntax:

```
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np

# Misalkan 'e' adalah DataFrame Anda dengan kolom 'bermain_game', 'klasifikasi', dan 'count'

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

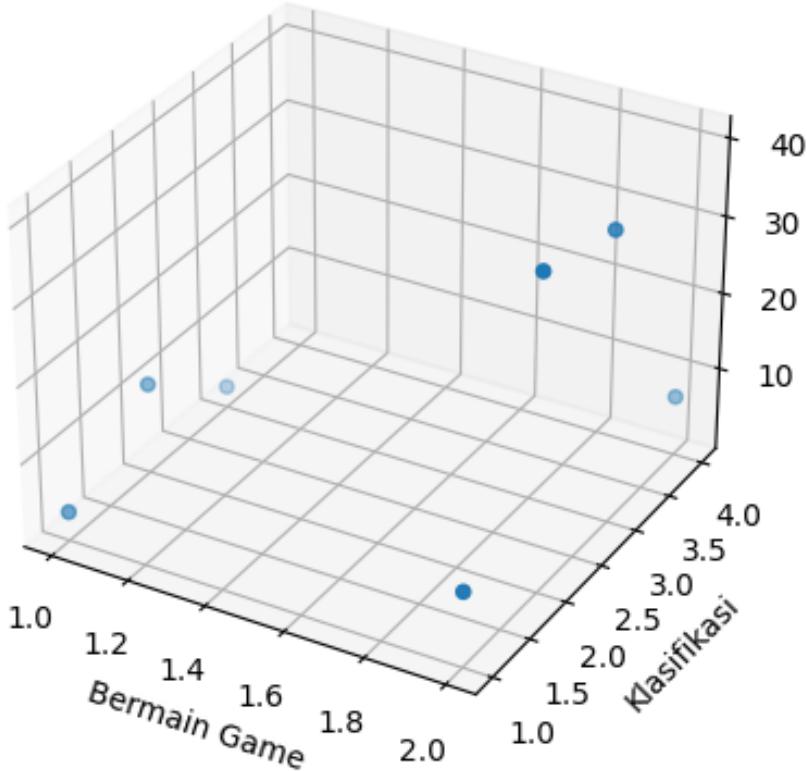
# Koordinat x, y, dan z untuk scatter plot 3D
x = e['bermain_game']
y = e['klasifikasi']
z = e['count']

# Menggambar scatter plot 3D
ax.scatter(x, y, z, marker='o')

# Label sumbu x, y, dan z
ax.set_xlabel('Bermain Game')
ax.set_ylabel('Klasifikasi')
ax.set_zlabel('Count')

# Tampilkan scatter plot 3D
plt.show()
```

Output:



10. Analisis diagram yang telah dibuat pada nomer 9

- a. Korelasi variabel "membantu_mengalihkan_stress" dengan "klasifikasi"

Analisis dari scatterplot 3D ini mengungkap beberapa insight yang berguna. Pertama, dalam hubungan antara "Membantu Mengalihkan Stress" dan "Klasifikasi," terdapat variasi yang ditunjukkan oleh sebaran titik data di sepanjang sumbu x (Klasifikasi). Hal ini mengindikasikan bahwa variasi dalam tingkat "Membantu Mengalihkan Stress" terkait dengan kategori "Klasifikasi." Sebagai contoh, perbedaan dalam "Klasifikasi" mungkin memengaruhi bagaimana seseorang mengalami tingkat stres yang mereka alami.

Selain itu, scatterplot 3D ini juga menggambarkan pengaruh variabel "Count" terhadap hubungan antara "Membantu Mengalihkan Stress" dan "Klasifikasi." Variabel "Count" tercermin dalam tinggi atau besarnya titik data dalam plot pada sumbu z. Ini mengartikan bahwa "Count" berkontribusi terhadap sejauh mana hubungan antara "Membantu Mengalihkan Stress" dan "Klasifikasi" dinyatakan dalam plot. Ketika "Count" memiliki nilai yang tinggi, titik data cenderung lebih tinggi di sumbu z, sementara ketika "Count" rendah, titik data lebih rendah. Ini mengisyaratkan bahwa variabel "Count" memainkan peran penting dalam mempengaruhi hubungan antara tingkat stres dan klasifikasi yang mungkin mencerminkan dampak dari jumlah pengamatan atau frekuensi tertentu dalam dataset. Analisis ini menunjukkan pentingnya mempertimbangkan variabel tambahan seperti "Count" ketika menganalisis hubungan antar variabel yang kompleks.

- b. Korelasi variabel "pilihan_pertama_stress" dengan "klasifikasi"

Analisis dari scatterplot 3D ini membantu memahami hubungan antara variabel dependent "Pilihan Pertama Stress" dan variabel independent "Klasifikasi" dalam

konteks dataset. Scatterplot ini menggambarkan variasi nilai "Pilihan Pertama Stress" sepanjang sumbu x (Klasifikasi), menunjukkan bahwa nilai-nilai "Pilihan Pertama Stress" bervariasi tergantung pada kategori "Klasifikasi." Ini memberikan wawasan bahwa kategori "Klasifikasi" mungkin memengaruhi preferensi pilihan pertama yang digunakan untuk mengatasi stres.

Selain itu, pengaruh variabel "Count" tercermin dalam tinggi atau besarnya titik data dalam plot pada sumbu z. Ini mengisyaratkan bahwa "Count" memiliki peran dalam mengukur sejauh mana hubungan antara "Pilihan Pertama Stress" dan "Klasifikasi" direpresentasikan dalam scatterplot. Ketika "Count" memiliki nilai yang tinggi, titik data cenderung lebih tinggi di sumbu z, sementara ketika "Count" rendah, titik data lebih rendah. Pengamatan ini menunjukkan bahwa variabel "Count" memainkan peran penting dalam mempengaruhi hubungan antara preferensi pilihan pertama dalam menghadapi stres dan klasifikasi, mencerminkan dampak dari jumlah pengamatan atau frekuensi tertentu dalam dataset.

Hasil analisis ini menyediakan wawasan lebih dalam tentang bagaimana preferensi pilihan pertama dalam mengatasi stres (Pilihan Pertama Stress) bervariasi tergantung pada kategori klasifikasi (Klasifikasi) sambil mempertimbangkan peran signifikan dari variabel "Count." Analisis ini relevan dalam memahami faktor-faktor yang mungkin memengaruhi preferensi individu dalam menghadapi stres dalam kelompok yang berbeda dalam dataset tersebut.

c. Korelasi variabel "durasi_bermain" dengan "klasifikasi"

Analisis dari scatterplot 3D ini memberikan wawasan yang berharga terkait hubungan antara variabel dependent "Durasi Bermain" dan variabel independent "Klasifikasi," dengan mempertimbangkan variabel "Count." Scatterplot tersebut menggambarkan variasi nilai "Durasi Bermain" sepanjang sumbu x (Klasifikasi), yang mengindikasikan bahwa preferensi dalam durasi bermain bervariasi tergantung pada kategori "Klasifikasi." Dalam kata lain, bermain dengan durasi tertentu dapat berbeda antara kelompok yang berbeda dalam kategori "Klasifikasi."

Selain itu, scatterplot ini juga menyoroti pengaruh variabel "Count" sebagai sumbu z. Tinggi atau rendahnya titik data dalam plot di sumbu z mencerminkan nilai "Count." Ini menggambarkan bagaimana "Count" berkontribusi dalam sejauh mana hubungan antara "Durasi Bermain" dan "Klasifikasi" direpresentasikan dalam scatterplot. Saat "Count" memiliki nilai yang tinggi, titik data cenderung lebih tinggi di sumbu z, sedangkan "Count" yang rendah membuat titik data lebih rendah. Hal ini menunjukkan bahwa variabel "Count" memainkan peran penting dalam memengaruhi hubungan antara durasi bermain dan klasifikasi, menggambarkan dampak dari jumlah pengamatan atau frekuensi tertentu dalam dataset.

Analisis ini menyediakan pemahaman yang lebih mendalam tentang bagaimana preferensi dalam durasi bermain (Durasi Bermain) berkaitan dengan kategori klasifikasi (Klasifikasi) dalam konteks dataset, sambil menyoroti peran yang signifikan dari variabel "Count." Insight ini relevan dalam pemahaman faktor-faktor yang mungkin memengaruhi keputusan terkait durasi bermain dalam kelompok yang berbeda berdasarkan klasifikasinya.

d. Korelasi variabel "durasi_waktu_bermain" dengan "klasifikasi"

Analisis dari scatterplot 3D memberikan wawasan penting terkait hubungan antara variabel dependent "Durasi Waktu Bermain" dan variabel independent "Klasifikasi," dengan mempertimbangkan peran variabel "Count" dalam konteks visualisasi.

Pertama, scatterplot ini memvisualisasikan bagaimana "Durasi Waktu Bermain" bervariasi tergantung pada kategori "Klasifikasi." Sebaran titik data di sepanjang sumbu x (Klasifikasi) menggambarkan variasi dalam preferensi durasi waktu bermain antara kelompok yang berbeda. Hal ini mengisyaratkan bahwa faktor-faktor yang memengaruhi durasi waktu bermain mungkin berbeda di antara klasifikasi yang berbeda. Selain itu, scatterplot ini memperlihatkan bahwa variabel "Count" memiliki pengaruh yang signifikan pada hubungan antara "Durasi Waktu Bermain" dan "Klasifikasi." Tinggi atau rendahnya titik data di sumbu z mencerminkan nilai "Count," dan ini mencerminkan dampak jumlah pengamatan atau frekuensi tertentu dalam dataset. Dengan demikian, "Count" memainkan peran penting dalam membentuk hubungan antara durasi waktu bermain dan klasifikasi.

Analisis ini memberikan wawasan yang lebih dalam tentang bagaimana preferensi dalam durasi waktu bermain bervariasi antara kelompok yang berbeda, sambil menggarisbawahi peran yang penting dari variabel "Count" dalam memengaruhi hubungan tersebut. Pemahaman ini relevan dalam memahami faktor-faktor yang memengaruhi keputusan tentang durasi waktu bermain dalam berbagai kelompok berdasarkan klasifikasinya.

e. Korelasi variabel “bermain_game” dengan “klasifikasi”

Analisis dari scatterplot 3D memberikan pemahaman yang penting terkait hubungan antara variabel dependent "Bermain Game" dan variabel independent "Klasifikasi," dengan mempertimbangkan pengaruh dari variabel "Count" dalam konteks visualisasi tersebut. Pertama, scatterplot ini menggambarkan variasi nilai "Bermain Game" sepanjang sumbu x (Klasifikasi), menunjukkan bahwa preferensi dalam bermain game dapat bervariasi tergantung pada kategori "Klasifikasi." Hal ini mengindikasikan bahwa mungkin ada faktor-faktor yang memengaruhi pilihan bermain game yang berbeda di antara kelompok yang berbeda dalam kategori "Klasifikasi."