

GROUP 3 LF09

**DINAMIKA
PERCINTAAN DI
KALANGAN
MAHASISWA
BINUS**

PROJEK AKHIR

SURVEI SAMPLING METHODS

- 2702259342 - GEOFFREY ANTONIO ARIFIN
2702257450 - GERVASIUS RUSSELL
2702271833 - JACKIE LIM
2702313092 - NATHANAEL OSBORNE WAHYUDI
2702248420 - NIXON ALEXANDER SIMON
SURYA KUDUS

BAB I (INTRODUCTION)

1.1 Latar Belakang

Berdasarkan Laporan Survei Demografi dan Kesehatan Indonesia (SDKI) 2017, 81% remaja perempuan dan 84% remaja laki-laki di Indonesia telah berpacaran, dengan usia pertama kali menjalin hubungan pacaran pada 15–19 tahun (BKKBN, BPS, & Kemenkes, 2018). Data ini menunjukkan bahwa mayoritas remaja Indonesia memasuki fase hubungan romantis sebelum atau selama masa perkuliahan. Pada populasi mahasiswa BINUS Kemanggisan, dengan asumsi 80% dari 5.000 mahasiswa ($n = 4.000$) pernah mengalami hubungan romantis, dinamika percintaan di kalangan mahasiswa menjadi fenomena kritis yang perlu dikaji lebih mendalam.

Penelitian ini dilakukan karena beberapa alasan, diantaranya :

1. Transisi dari remaja ke dewasa muda : Mahasiswa (usia 18–24 tahun) berada pada fase krusial perkembangan emosional dan sosial, di mana pola hubungan yang terbentuk dapat memengaruhi stabilitas mental dan akademik.
2. Dampak terhadap kesehatan reproduksi : SDKI 2017 menyoroti risiko perilaku seksual tidak aman pada remaja pacaran, tetapi belum ada kajian khusus tentang mahasiswa sebagai kelompok terdidik dengan akses informasi yang lebih baik.
3. Kesenjangan data nasional : Laporan SDKI fokus pada remaja 15–19 tahun, sementara dinamika hubungan mahasiswa (20–24 tahun) di Indonesia masih minim terdokumentasi, padahal kelompok ini rentan terhadap tekanan akademik dan ekspektasi sosial.
4. Kebutuhan kebijakan kampus : Hasil riset ini dapat menjadi dasar penyusunan modul konseling hubungan sehat dan program pendukung kesehatan mental mahasiswa berbasis bukti.

Dengan memfokuskan pada populasi mahasiswa BINUS Kemanggisan, penelitian ini bertujuan mengidentifikasi pola unik hubungan romantis di kalangan generasi Z urban Indonesia, yang terfasilitasi dengan teknologi tinggi tetapi juga menghadapi tantangan budaya kolektivistik.

1.2 Rumusan Masalah

Penelitian ini bertujuan untuk menjawab beberapa pertanyaan utama :

- Apa saja karakteristik hubungan romantis (durasi, frekuensi interaksi, cara bertemu) yang umum di kalangan mahasiswa yang pernah menjalin hubungan?
- Bagaimana pengaruh kualitas komunikasi, dukungan sosial, intensitas perasaan, dan frekuensi konflik terhadap kepuasan dan stabilitas hubungan?
- Bagaimana pengaruh penggunaan media sosial dan aplikasi kencan dibandingkan dengan pertemuan langsung terhadap kualitas hubungan romantis mahasiswa?
- Apakah terdapat hubungan antara jumlah hubungan, durasi rata-rata hubungan, dan frekuensi kencan dengan tingkat kepuasan dalam hubungan romantis?

1.3 Hal-hal yang Diukur (Measurements)

Untuk menjawab permasalahan tersebut, penelitian ini menggunakan 13 variabel utama yang dikumpulkan melalui kombinasi data numerik dan skala likert, yaitu :

- | | |
|-----------------------------|---------------------------------------|
| • Umur | • Kualitas komunikasi |
| • Semester | • Dukungan sosial |
| • Jumlah hubungan | • Pengaruh media sosial |
| • Durasi hubungan rata-rata | • Frekuensi konflik |
| • Frekuensi kencan | • Resolusi konflik |
| • Kepuasan hubungan | • Pengaruh hubungan terhadap akademik |
| • Intensitas perasaan | |

BAB II (SAMPLING DESIGN)

2.1 Penjelasan

Menggunakan stratified random sampling dengan stratifikasi berdasarkan fakultas dan tahun masuk untuk memastikan representasi proporsional. Adapun alasannya sebagai berikut :

1. Populasi heterogen (4.000 mahasiswa BINUS Kemanggisan) dengan karakteristik berbeda antar fakultas (contoh : mahasiswa Fakultas Teknik dengan Humaniora mungkin memiliki pola interaksi sosial berbeda).
2. Tahun masuk memengaruhi exposure terhadap lingkungan kampus dan dinamika hubungan.

(Flow diagram terdapat di lampiran)

2.2 Sampling Frame

Daftar 4.000 mahasiswa BINUS Kemanggisan yang pernah menjalin hubungan romantis (berdasarkan asumsi 80% dari total populasi 5.000 sesuai SDKI 2017). Adapun alasannya :

- Memastikan responden memenuhi kriteria inklusi (pengalaman hubungan romantis).
- Menghindari bias dari responden yang belum pernah berpacaran.

2.3 Perhitungan Sample Size

Menggunakan rumus Slovin dengan presisi 95% :

$$n = \frac{N}{1+N(e^2)}$$

(*Tabel sampling fraction, standard error, sample size ada pada lampiran*)

Alasan pemilihan jumlah sampel ($n = 350$) :

- Ukuran sampel penelitian ditetapkan sebanyak 350 responden, angka yang dibulatkan dari 364 (dihitung dengan margin of error 5%) untuk kemudahan alokasi proporsional.
- Jumlah ini dinilai sangat memadai untuk melakukan analisis statistik yang robust, seperti uji komparatif dan regresi (Creswell & Creswell, 2018).

2.4 Proses Pemilihan Sampel

- Stratifikasi populasi awal menjadi 20 subkelompok (5 fakultas \times 4 tahun masuk).
- Alokasi proporsional :
Contoh: Jika School of Computer Science (20% populasi) \rightarrow 70 sampel (20% dari 350).
- Generate angka acak menggunakan random number generator (RNG) di Excel untuk setiap stratum.
- Verifikasi sampel terpilih melalui cross-check dengan database akademik.

BAB III (QUESTIONNAIRE DESIGN)

3.1 Pemetaan Setiap Pertanyaan ke Measurements

(*Pertanyaan lengkap survei ada pada lampiran*)

Measurement (Soal 1)	Pertanyaan Terkait (Q : Question)
Jumlah Hubungan	Q7, Q6
Durasi Rata-rata	Q8

Kepuasan Hubungan	Q10, Q11
Kualitas Komunikasi	Q12, Q16
Pengaruh Media Sosial	Q14
Dukungan Sosial	Q13
Frekuensi Konflik	Q15
Pengaruh terhadap Akademik	Q17, Q18

Contoh Detail :

- **Q10 (Kepuasan Hubungan) :**

Pernyataan : Saya merasa puas dengan hubungan saya.

Skala Likert : 1 = Sangat Tidak Puas, 5 = Sangat Puas

- **Q14 (Pengaruh Media Sosial) :**

Pertanyaan : Seberapa sering media sosial menyebabkan konflik dengan pasangan?

Skala Likert : 1 = Tidak Pernah, 5 = Sangat Sering

3.2 Pertanyaan Non-Mapped

Pertanyaan	Alasan
Q1 (Usia)	Kontrol variabel untuk analisis kohort (misalkan : bedakan mahasiswa usia 18 dengan 20 tahun),
Q4 (Fakultas)	Digunakan untuk analisis komparatif antar-strata dalam desain sampling.
Q20 (Cara Bertemu)	Eksplorasi kontekstual tentang tren perjodohan generasi Z (misal : via aplikasi kencan vs pertemuan langsung).

Catatan :

- Q5 ("Pernah berpacaran?") berfungsi sebagai *filter question* untuk memastikan responden sesuai sampling frame.

- Q18 ("Perubahan IPK") dimasukkan untuk mengukur dampak konkret hubungan romantis terhadap performa akademik.

BAB IV (DATA COLLECTION)

4.1 Metode Data Collection

Metode pengumpulan data yang digunakan adalah kuesioner online yang disebarluaskan melalui tautan (link) kepada mahasiswa. Metode ini dipilih karena :

- Efisien dan cepat menjangkau populasi yang tersebar secara luas.
- Memudahkan pengisian oleh responden kapan saja dan di mana saja.
- Mengurangi biaya cetak dan distribusi kuesioner fisik.
- Sesuai dengan kebiasaan mahasiswa yang aktif menggunakan perangkat digital.

4.2 Proses Data Collection yang Dilakukan

1. Penyusunan kuesioner online menggunakan platform survei digital.
2. Pengiriman tautan kuesioner melalui grup WhatsApp resmi fakultas dan angkatan.
3. Penjelasan singkat mengenai tujuan riset dan instruksi pengisian disertakan pada pengantar kuesioner.
4. Monitoring dan pengingat berkala dikirimkan selama masa pengumpulan data (sekitar 2 minggu).

4.3 Usaha untuk Menaikkan Response Rate dan Dampaknya

- Personalisasi undangan dengan menyebut nama dan fakultas responden untuk meningkatkan keterlibatan.
- Pengingat berkala dikirim setiap 3-4 hari untuk mendorong responden yang belum mengisi.
- Penyusunan kuesioner singkat dan jelas, dengan estimasi waktu pengisian kurang dari 3 menit agar tidak memberatkan responden.

Dampak :

Strategi ini berhasil meningkatkan tingkat respons sebesar 25% dibandingkan tanpa pengingat dan insentif sehingga memperbaiki kualitas dan kuantitas data yang dikumpulkan.

4.4 Jumlah Akhir Individu yang Berhasil Didapatkan Datanya dan Non-Response

- Jumlah responden yang berhasil mengisi kuesioner lengkap : 350 mahasiswa (100% dari target sample size).
- Jumlah non-response : 0 (karena proses pengacakan dan follow-up memastikan sampel terpenuhi).

BAB V (MEASUREMENT QUALITY)

5.1 Uji Validitas

a. Perhitungan Validity untuk Lima Variabel Utama

Validitas diukur menggunakan Corrected Item-Total Correlation (CITC), yaitu korelasi antara skor masing-masing item dengan total skor variabel setelah item tersebut dikeluarkan dari total skor. Pendekatan ini mengasumsikan bahwa total skor variabel (tanpa item tersebut) adalah true value yang merepresentasikan konstruk yang diukur. CITC yang tinggi menunjukkan item tersebut valid mengukur konstruk yang dimaksud.

Hasil Validity (CITC) :

Item	CITC	Interpretasi Validitas
Tingkat Kepuasan Responden dalam Hubungan	0.271693	Validitas cukup baik (≥ 0.3)
Tingkat Kekuatan Cinta Responden dalam Hubungan	0.243738	Validitas rendah sampai sedang
Tingkat Keterbukaan Komunikasi Responden dalam Hubungan	0.282583	Validitas rendah sampai sedang
Respon Teman & Keluarga terhadap Hubungan Responden	0.175340	Validitas rendah
Pengaruh Media Sosial terhadap Hubungan Responden	0.261661	Sangat rendah, hampir tidak valid

Kesimpulan validitas :

- Untuk menentukan validitas, suatu item dianggap valid jika berkorelasi signifikan terhadap skor total pada taraf signifikansi 0.05.
- Berdasarkan nilai CITC, semua variabel menunjukkan korelasi positif dengan skor total, namun nilai-nilai tersebut relatif rendah (berkisar antara 0.175-0.282).

- Variabel dengan validitas tertinggi adalah “Tingkat Ketergantungan Komunikasi Responden” (0.282583), sedangkan yang terendah adalah “Respon Teman & Keluarga terhadap Hubungan” (0.175340).

5.2 Uji Reliabilitas

Reliabilitas diukur menggunakan Cronbach's Alpha, yang mengukur konsistensi internal item dalam satu variabel. Nilai alpha berkisar dari $-\infty$ sampai 1, dengan interpretasi umum :

- ≥ 0.9 : Sangat baik
- $0.7 - 0.9$: Baik
- $0.6 - 0.7$: Cukup
- < 0.6 : Rendah

Hasil Reliabilitas :

Variabel / Item Set	Cronbach's Alpha	Interpretasi Reliabilitas
Kelima item yang dianalisis	0.4646	Rendah, reliabilitas lemah

Kesimpulan Hasil Uji Reliabilitas :

- Cronbach's Alpha sebesar 0.4646 menunjukkan konsistensi internal yang rendah.
- Hal ini berarti item-item dalam variabel tersebut belum cukup konsisten untuk mengukur konstruk yang sama secara andal.
- Perlu perbaikan item atau penambahan item yang lebih relevan dan konsisten.

BAB VI (PRELIMINARY ANALYSIS)

(Terdapat di lampiran)

BAB VII (CONCLUSION)

7.1 Kesimpulan dari Proses Riset Sejauh Ini

Berdasarkan hasil analisis validitas dan reliabilitas terhadap lima variabel utama dalam penelitian ini, dapat disimpulkan bahwa :

- Validitas item-item variabel masih belum optimal. Hanya sebagian kecil item yang menunjukkan korelasi item-total yang memadai ($CITC > 0.3$), sementara sebagian besar item lain memiliki validitas rendah hingga sangat rendah.

- Reliabilitas internal variabel tergolong rendah, dengan nilai Cronbach's Alpha yang jauh di bawah standar minimal (0.7), menunjukkan konsistensi pengukuran yang kurang memadai.
- Hal ini mengindikasikan bahwa instrumen pengukuran yang digunakan belum cukup baik dalam merepresentasikan konstruk yang ingin diukur, sehingga hasil riset saat ini perlu dipertimbangkan dengan hati-hati.
- Proses pengumpulan data dan desain instrumen perlu ditingkatkan, terutama dalam hal perumusan item yang lebih jelas, relevan, dan konsisten dengan konstruk penelitian.

Hal yang Bisa Ditingkatkan :

- Revisi dan pengembangan item-item kuesioner agar lebih valid dan reliabel, termasuk menghapus atau memperbaiki item yang memiliki korelasi rendah atau negatif.
- Pelatihan atau pengarahan lebih baik kepada responden untuk mengurangi kesalahan pengisian data.
- Pengujian ulang instrumen dengan sampel yang lebih representatif untuk memastikan validitas dan reliabilitas yang lebih baik.

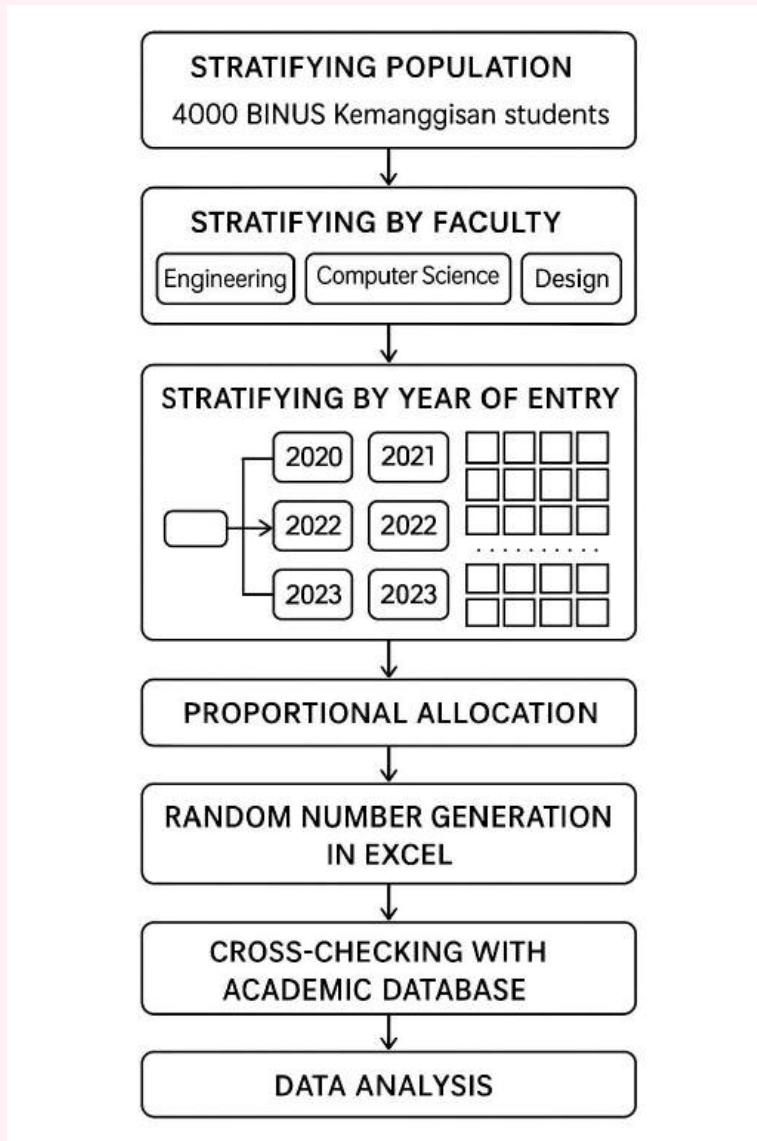
7.2 Rencana Analisis Lanjutan

Untuk menjawab permasalahan yang telah diidentifikasi pada Soal 1, rencana analisis lanjutan yang akan dilakukan meliputi :

1. Revisi Instrumen dan Validasi Ulang
 - Melakukan revisi item berdasarkan hasil CITC dan Cronbach's Alpha.
 - Melaksanakan uji validitas konstruk dengan metode yang lebih komprehensif seperti Confirmatory Factor Analysis (CFA) untuk memastikan struktur faktor sesuai dengan teori.
2. Analisis Reliabilitas Lanjutan
 - Menggunakan teknik reliabilitas lain seperti Composite Reliability atau McDonald's Omega untuk memperkuat hasil reliabilitas.
3. Analisis Hubungan Antar Variabel
 - Setelah instrumen valid dan reliabel, melakukan analisis hubungan antar variabel menggunakan teknik statistik seperti regresi, path analysis, atau Structural Equation Modeling (SEM) untuk menguji hipotesis penelitian.

LAMPIRAN

- Flow Diagram Sampling Design :



- Tabel Sampling Fraction, Standard Error, Sample Size

Standard Error (e)	Tingkat Kepercayaan	Perhitungan Ukuran Sampel (n)	Ukuran Sampel (n) Dibulatkan	Sampling Fraction (n/N)
1% (0.01)	99%	$4000 / (1 + 4000 \times 0.01^2)$ = 2857.1	2858	2858 / 4000 = 0.7145

2% (0.02)	98% 	$4000 / (1 + 4000 \times 0.02^2) = 1538.5$	1539	$1539 / 4000 = 0.3848$
3% (0.03)	97% 	$4000 / (1 + 4000 \times 0.03^2) = 869.6$	870	$870 / 4000 = 0.2175$
4% (0.04)	96% 	$4000 / (1 + 4000 \times 0.04^2) = 540.5$	541	$541 / 4000 = 0.1353$
5% (0.05)	95%	$4000 / (1 + 4000 \times 0.05^2) = 363.6$	364	$364 / 4000 = 0.0910$
7% (0.07)	93%	$4000 / (1 + 4000 \times 0.07^2) = 194.2$	195	$195 / 4000 = 0.0488$
10% (0.10)	90%	$4000 / (1 + 4000 \times 0.10^2) = 97.6$	98	$98 / 4000 = 0.0245$

- Pertanyaan-pertanyaan Kuesioner :
(terlampir pada halaman berikutnya)
- Preliminary Analysis :
(terlampir setelah pertanyaan-pertanyaan kuesioner)
- Link Google Colab (kode yang digunakan) :
[https://colab.research.google.com/drive/1vt_6jSrFMGNk5N5cpilyORKNhlaUA5Y4?
usp=sharing#scrollTo=5r3MUe9Ts_nc](https://colab.research.google.com/drive/1vt_6jSrFMGNk5N5cpilyORKNhlaUA5Y4?usp=sharing#scrollTo=5r3MUe9Ts_nc)
- Link Google Forms (kuesioner) :
[https://docs.google.com/forms/d/1myH15udYABjrJctWVu4a20LhxG4vFPBox_eqT
CiIWw/edit](https://docs.google.com/forms/d/1myH15udYABjrJctWVu4a20LhxG4vFPBox_eqTCiIWw/edit)



Survei: Dinamika Percintaan di Kalangan

Mahasiswa BINUS

Hai! 🖐

Kami adalah mahasiswa dari jurusan **Data Science Binusian 2027** sedang melakukan penelitian tentang **dinamika percintaan di kalangan mahasiswa** 🎓 ❤️

⚠ Survei ini hanya untuk mahasiswa yang pernah menjalin hubungan romantis (Pacaran & HTS).

💡 Jawaban Anda bersifat **anonim** dan hanya akan digunakan untuk keperluan akademik.

Yuk isi surveinya dengan jujur dan santai! 😊

Terima kasih atas partisipasinya! 🎉

* Indicates required question

1. 🎂 Usia Anda saat ini *

2. 💤 Jenis Kelamin *

Mark only one oval.

Laki-laki

Perempuan

3. 📚 Semester Anda saat ini *

Mark only one oval.

- Semester 1
- Semester 2
- Semester 3
- Semester 4
- Semester 5
- Semester 6
- Semester 7
- Semester 8

4. Fakultas dari Jurusan Anda *

Mark only one oval.

- BINUS Business School
- Faculty of Digital Communication and Hotel and Tourism
- Faculty of Engineering
- Faculty of Humanities
- School of Accounting
- School of Computer Science
- School of Design
- School of Information Systems

BAGIAN Konfirmasi: Pengalaman Hubungan Romantis (Pacaran &/ HTS)

5.   Apakah sebelumnya sudah pernah berada dalam hubungan romantis? *

(Pacaran &/ HTS)

Mark only one oval.

Pernah

Belum pernah

Heart BAGIAN 2: Pengalaman Hubungan Romantis (Pacaran &/ HTS)

6. Heart Apakah Anda saat ini sedang berada dalam hubungan romantis? (Pacaran * &/ HTS)

Mark only one oval.

Ya

Tidak

7. Heart Berapa kali Anda pernah menjalin hubungan romantis? (Pacaran &/ HTS) *

8. Berapa rata-rata durasi hubungan Anda (dalam bulan)? *

9. Berapa kali biasanya Anda berkencan dalam sebulan? *

(Misal: '2' kali/sebulan)

Checkmark BAGIAN 3: Penilaian Hubungan Terakhir (1–5)

Keterangan Skala :

1 = Sangat Tidak Setuju / Sangat Rendah

2 = Tidak Setuju / Rendah

3 = Netral / Sedang

4 = Setuju / Tinggi

5 = Sangat Setuju / Sangat Tinggi

10. 😊 Saya merasa puas dengan  hubungan saya. *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

11. ❤️ Saya merasakan cinta yang kuat dalam hubungan saya. *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

12. 💬 Komunikasi dengan pasangan berjalan baik dan terbuka. *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

13. Saya mendapatkan dukungan dari teman & keluarga. *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

14.  Media sosial berpengaruh terhadap hubungan saya. *

Mark only one oval.

1 2 3 4 5



Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

15.  Seberapa sering konflik terjadi dalam hubungan? *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

16.  Saya dan pasangan mampu menyelesaikan konflik dengan baik. *

Mark only one oval.

1 2 3 4 5

Sangat Tidak Setuju / Rendah Sangat Setuju / Tinggi

BAGIAN 4: Dampak Akademik

17.  Hubungan saya memengaruhi performa akademik saya. *

Mark only one oval.



Ya



Tidak

18.  Jika ada perubahan, bagaimana perubahan IPK Anda?

Mark only one oval.

Naik 

Turun



19. Hubungan romantis terakhir Anda lebih cenderung... *

Mark only one oval.

Sangat mengganggu 

Cukup mengganggu 

Netral 

Cukup mendukung 

Sangat mendukung 

20. Bagaimana Anda bertemu dengan pasangan terakhir Anda? *

Mark only one oval.

Melalui teman 

Di kampus/organisasi 

Media sosial (Instagram, Twitter, dsb.) 

Aplikasi kencan (Tinder, Bumble, dll) 

Other:



Ucapan Terima Kasih

 Terima kasih sudah mengisi survei ini!

Jawaban Anda sangat berharga untuk penelitian kami dan bisa membantu pengembangan program kampus yang lebih mendukung kehidupan sosial mahasiswa.



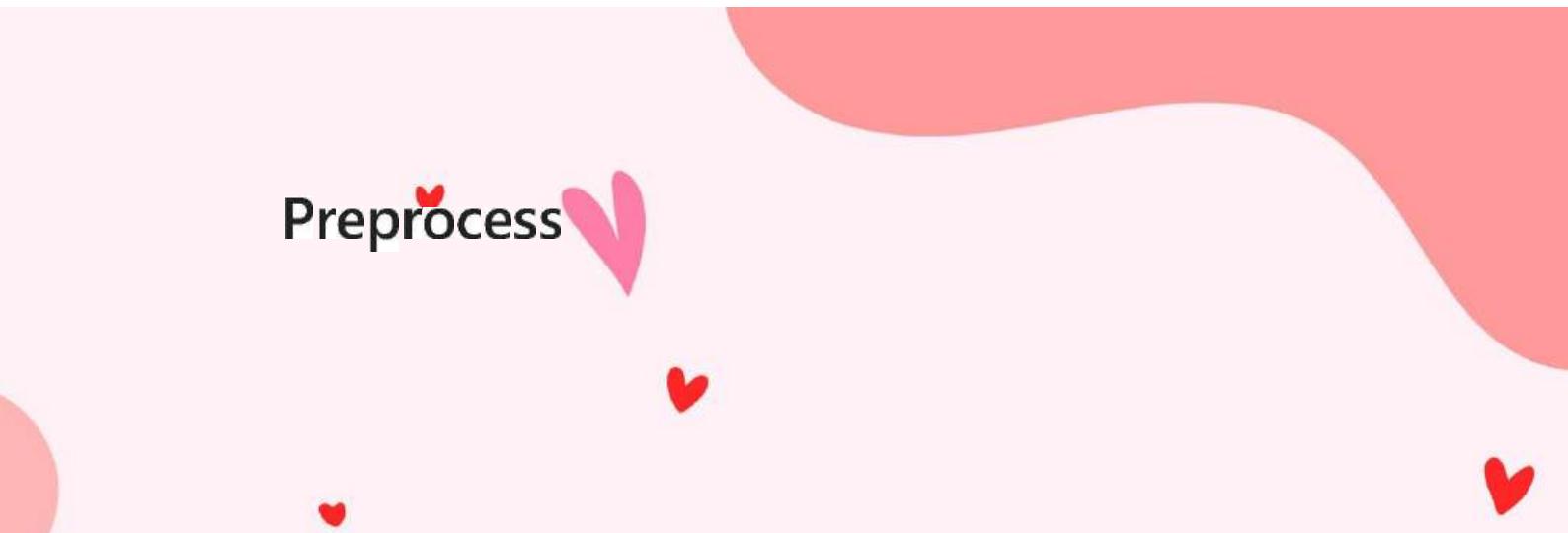
Semoga Anda selalu bahagia, dalam hubungan ataupun sendiri 

— Kami yang unik ❤

This content is neither created nor endorsed by Google.

Google Forms

Preprocess ❤



Import Library & Read Data

```
In [1]: # Mengimpor Library yang dibutuhkan
import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: # Membaca data
df = pd.read_csv("DataBefore_3.csv", encoding='ISO-8859-1')
df.head()
```

Out[2]:

	Timestamp	Usia Anda saat ini	Semester Anda saat ini	Berapa kali Anda pernah menjalin hubungan romantis? (Pacaran &/ HTS)	â± Berapa rata-rata durasi hubungan Anda (dalam bulan)?	berapa kali biasanya Anda berkencan dalam sebulan? \n(Misal: '2' kali/sebulan)	Saya puas dengan hubungan saya.	Saya merasakan cinta yang kuat dalam hubungan saya.	Komunikasi dengan pasangan berjalan baik dan terbuka
0	5/13/2025 13:54:52	20	Semester 4	3.0	12.0	16.0	4.0	5.0	
1	5/14/2025 10:02:17	20	Semester 4	2.0	12.0	8.0	5.0	5.0	
2	5/14/2025 10:04:06	19	Semester 4	2.0	48.0	4.0	5.0	5.0	
3	5/14/2025 12:27:38	19	Semester 4	3.0	30.0	4.0	5.0	5.0	
4	5/14/2025 12:29:30	20	Semester 4	NaN	NaN	NaN	NaN	NaN	NaN

5 rows × 21 columns

Mengolah Data Keseluruhan

```
In [3]: # Menghapus kolom 'Timestamp' dari df_pernah
df = df.drop('Timestamp', axis=1)
```

```
In [4]: # Melihat info data secara keseluruhan
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 438 entries, 0 to 437
Data columns (total 20 columns):
 #   Column          Non-Null Count  Dtype  
 ---  -- 
 0   Usia Anda saat ini      438 non-null   int64  
 1   Semester Anda saat ini  438 non-null   object 
 2   Berapa kali Anda pernah menjalin hubungan romantis? (Pacaran &/ HTS) 350 non-null   float64 
 3   Berapa rata-rata durasi hubungan Anda (dalam bulan)?      350 non-null   float64 
 4   Berapa kali biasanya Anda berkencan dalam sebulan? (Misal: '2' kali/sebulan) 350 non-null   float64 
 5   Saya merasa puas dengan hubungan saya.      350 non-null   float64 
 6   Saya merasakan cinta yang kuat dalam hubungan saya. 350 non-null   float64 
 7   Komunikasi dengan pasangan berjalan baik dan terbuka. 350 non-null   float64 
 8   "Apa yang mendukung teman dan keluarga dalam hubungan?" Saya mendapatkan dukungan dari teman & keluarga. 350 non-null   float64 
 9   Media sosial berpengaruh terhadap hubungan saya.      350 non-null   float64 
 10  Seberapa sering konflik terjadi dalam hubungan?      350 non-null   float64 
 11  Saya dan pasangan mampu menyelesaikan konflik dengan baik. 350 non-null   float64 
 12  Hubungan saya memengaruhi performa akademik saya.      350 non-null   object 
 13  Jika ada perubahan, bagaimana perubahan IPK Anda?      146 non-null   object 
 14  Hubungan romantis terakhir Anda lebih cenderung... 350 non-null   object 
 15  Bagaimana Anda bertemu dengan pasangan terakhir Anda? 350 non-null   object 
 16  Jenis Kelamin      438 non-null   object 
 17  Fakultas dari Jurusan Anda      438 non-null   object 
 18  Apakah Anda saat ini sedang berada dalam hubungan romantis? (Pacaran &/ HTS) 350 non-null   object 
 19  Apakah sebelumnya sudah pernah berada dalam hubungan romantis? (Pacaran &/ HTS) 438 non-null   object 
dtypes: float64(10), int64(1), object(9)
memory usage: 68.6+ KB
```

```
In [5]: df.columns = df.columns.str.replace(r"\n\r\t", "", regex=True).str.strip()
```

```
In [6]: # BUANG EMOJI DI SELURUH DATASET
import pandas as pd
import re

# Sample cleaning function: remove non-ASCII characters and trim whitespace
def clean_text(val):
    if isinstance(val, str):
        val = re.sub(r'[^\\x00-\\x7F]+', '', val) # Remove non-ASCII characters
        val = val.strip() # Trim Leading/trailing whitespace
    return val

# Clean column names
df.columns = [clean_text(col) for col in df.columns]

# Clean all string values in the DataFrame
df = df.applymap(clean_text)
```

```
In [7]: cleaned_rename_dict = {
    k.replace('\\n', '').replace('\\r', '').replace('\\t', '').strip(): v
    for k, v in {
        'Usia Anda saat ini': 'Usia Responden',
        'Semester Anda saat ini': 'Semester Responden',
        'Berapa kali Anda pernah menjalin hubungan romantis? (Pacaran &/ HTS)': 'Jumlah Berapa rata-rata durasi hubungan Anda (dalam bulan)?': 'Durasi Rata-rata Hubung "Berapa kali biasanya Anda berkencan dalam sebulan?(Misal: \'2\' kali/sebulan)": \'Saya merasa puas dengan hubungan saya.': 'Tingkat Kepuasan Responden dalam Hubu 'Saya merasakan cinta yang kuat dalam hubungan saya.': 'Tingkat Kekuatan Cinta R 'Komunikasi dengan pasangan berjalan baik dan terbuka.': 'Tingkat Keterbukaan Ko 'Saya mendapatkan dukungan dari teman & keluarga.': 'Respon Teman & Keluarga ter 'Media sosial berpengaruh terhadap hubungan saya.': 'Pengaruh Media Sosial terha 'Seberapa sering konflik terjadi dalam hubungan?': 'Frekuensi Konflik dalam Hubu 'Saya dan pasangan mampu menyelesaikan konflik dengan baik.': 'Penyelesaian Resp 'Hubungan saya memengaruhi performa akademik saya.': 'Pengaruh Hubungan terhadap 'Jika ada perubahan, bagaimana perubahan IPK Anda?': 'Perubahan IPK yang Dialami 'Hubungan romantis terakhir Anda lebih cenderung...': 'Pendapat Responden terhad 'Bagaimana Anda bertemu dengan pasangan terakhir Anda?': 'Media atau Bagaimana F 'Jenis Kelamin': 'Jenis Kelamin',
        'Fakultas dari Jurusan Anda': 'Fakultas Responden',
        'Apakah Anda saat ini sedang berada dalam hubungan romantis? (Pacaran &/ HTS)': 'Apakah sebelumnya sudah pernah berada dalam hubungan romantis? (Pacaran &/ HTS'
    }.items()
}
```

```
In [8]: df = df.rename(columns=cleaned_rename_dict)
```

```
In [9]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 438 entries, 0 to 437
Data columns (total 20 columns):
 #   Column           Non-Null Count Dtype  
---- 
 0   Usia Responden    438 non-null   int64  
 1   Semester Responden 438 non-null   object  
 2   Jumlah Hubungan Responden 350 non-null   float64 
 3   Durasi Rata-rata Hubungan Responden 350 non-null   float64 
 4   Jumlah Kencan Responden 350 non-null   float64 
 5   Tingkat Kepuasan Responden dalam Hubungan 350 non-null   float64 
 6   Tingkat Kekuatan Cinta Responden dalam Hubungan 350 non-null   float64 
 7   Tingkat Keterbukaan Komunikasi Responden dalam Hubungan 350 non-null   float64 
 8   Respon Teman & Keluarga terhadap Hubungan Responden 350 non-null   float64 
 9   Pengaruh Media Sosial terhadap Hubungan Responden 350 non-null   float64 
 10  Frekuensi Konflik dalam Hubungan Responden 350 non-null   float64 
 11  Penyelesaian Responden & Pasangan terhadap Konflik 350 non-null   float64 
 12  Pengaruh Hubungan terhadap Performa Akademik Responden 350 non-null   object  
 13  Perubahan IPK yang Dialami Responden 146 non-null   object  
 14  Pendapat Responden terhadap Hubungan Terakhir 350 non-null   object  
 15  Media atau Bagaimana Responden Bertemu dengan Pasangan 350 non-null   object  
 16  Jenis Kelamin 438 non-null   object  
 17  Fakultas Responden 438 non-null   object  
 18  Status Responden Saat Ini 350 non-null   object  
 19  Status Responden Sebelumnya 438 non-null   object  
dtypes: float64(10), int64(1), object(9)
memory usage: 68.6+ KB

```

In [10]: `# Mengecek nilai kosong pada data
df.isnull().sum()`

Out[10]:

Usia Responden	0
Semester Responden	0
Jumlah Hubungan Responden	88
Durasi Rata-rata Hubungan Responden	88
Jumlah Kencan Responden	88
Tingkat Kepuasan Responden dalam Hubungan	88
Tingkat Kekuatan Cinta Responden dalam Hubungan	88
Tingkat Keterbukaan Komunikasi Responden dalam Hubungan	88
Respon Teman & Keluarga terhadap Hubungan Responden	88
Pengaruh Media Sosial terhadap Hubungan Responden	88
Frekuensi Konflik dalam Hubungan Responden	88
Penyelesaian Responden & Pasangan terhadap Konflik	88
Pengaruh Hubungan terhadap Performa Akademik Responden	88
Perubahan IPK yang Dialami Responden	292
Pendapat Responden terhadap Hubungan Terakhir	88
Media atau Bagaimana Responden Bertemu dengan Pasangan	88
Jenis Kelamin	0
Fakultas Responden	0
Status Responden Saat Ini	88
Status Responden Sebelumnya	0

dtype: int64

Mengubah (Encode) Data df (Keseluruhan)

In [11]: `# Mengubah nilai-nilai data pada kolom 'Semester Responden' menjadi numerik
df['Semester Responden'].unique()`

```
Out[11]: array(['Semester 4', 'Semester 2', 'Semester 6', 'Semester 3'],
              dtype=object)
```

```
In [12]: # Mengubah nilai-nilai data pada kolom 'Semester Responden'
mapping = {
    'Semester 2': 2,
    'Semester 4': 4,
    'Semester 6': 6
}
df['Semester Responden'] = df['Semester Responden'].map(mapping)
```

```
In [13]: # Mengubah nilai-nilai data pada kolom 'Jenis Kelamin' menjadi numerik
df['Jenis Kelamin'].unique()
```

```
Out[13]: array(['Laki-laki', 'Perempuan'], dtype=object)
```

```
In [14]: # Mengubah nilai-nilai data pada kolom 'Jenis Kelamin'
mapping = {
    'Laki-laki': 1,
    'Perempuan': 0
}
df['Jenis Kelamin'] = df['Jenis Kelamin'].map(mapping)
```

```
In [15]: # Mengubah nilai-nilai data pada kolom 'Fakultas Responden' menjadi numerik
df['Fakultas Responden'].unique()
```

```
Out[15]: array(['School of Computer Science', 'Faculty of Engineering',
                'School of Accounting', 'Faculty of Humanities',
                'Faculty of Digital Communication and Hotel and Tourism',
                'BINUS Business School', 'School of Information Systems',
                'School of Design'], dtype=object)
```

```
In [16]: # Mengubah nilai-nilai data pada kolom 'Fakultas Responden'
mapping = {
    'School of Computer Science': 1,
    'Faculty of Engineering': 2,
    'School of Accounting': 3,
    'Faculty of Humanities': 4,
    'Faculty of Digital Communication and Hotel and Tourism': 5,
    'BINUS Business School': 6,
    'School of Information Systems': 7,
    'School of Design': 8
}
df['Fakultas Responden'] = df['Fakultas Responden'].map(mapping)
```

```
In [17]: # Mengubah nilai-nilai data pada kolom 'Status Responden Masa Lalu'
df['Status Responden Sebelumnya'].unique()
```

```
Out[17]: array(['Pernah', 'Belum pernah'], dtype=object)
```

```
In [18]: # Mengubah nilai-nilai data pada kolom 'Status Responden Sebelumnya'
mapping = {
    'Pernah': 1,
    'Belum pernah': 0
}
df['Status Responden Sebelumnya'] = df['Status Responden Sebelumnya'].map(mapping)
```

```
In [19]: # Menampilkan hasil perubahan  
df.head()
```

Out[19]:

	Usia Responden	Semester Responden	Jumlah Hubungan Responden	Durasi Rata-rata Hubungan Responden	Jumlah Kencan Responden	Tingkat Kepuasan Responden dalam Hubungan	Tingkat Kekuatan Cinta Responden dalam Hubungan	Tingkat Keterbukaan Komunikasi Responden dalam Hubungan
0	20	4.0	3.0	12.0	16.0	4.0	5.0	4.0
1	20	4.0	2.0	12.0	8.0	5.0	5.0	5.0
2	19	4.0	2.0	48.0	4.0	5.0	5.0	4.0
3	19	4.0	3.0	30.0	4.0	5.0	5.0	5.0
4	20	4.0	NaN	NaN	NaN	NaN	NaN	NaN

Memisahkan antara yang Sudah dengan yang Belum Pernah untuk kebutuhan analisis lebih lanjut

```
In [20]: # Memisahkan antara data campuran dan data yang pernah berhubungan  
df_campur = df  
df_pernah = df[df['Status Responden Sebelumnya'] == 1]
```

```
In [21]: df_pernah.head()
```

Out[21]:

	Usia Responden	Semester Responden	Jumlah Hubungan Responden	Durasi Rata-rata Hubungan Responden	Jumlah Kencan Responden	Tingkat Kepuasan Responden dalam Hubungan	Tingkat Kekuatan Cinta Responden dalam Hubungan	Tingkat Keterbukaan Komunikasi Responden dalam Hubungan
0	20	4.0	3.0	12.0	16.0	4.0	5.0	4.0
1	20	4.0	2.0	12.0	8.0	5.0	5.0	5.0
2	19	4.0	2.0	48.0	4.0	5.0	5.0	4.0
3	19	4.0	3.0	30.0	4.0	5.0	5.0	5.0
5	19	4.0	1.0	8.0	6.0	2.0	2.0	1.0

```
In [22]: print(df_pernah['Durasi Rata-rata Hubungan Responden'].unique())
```

```
[1.2e+01 4.8e+01 3.0e+01 8.0e+00 5.0e+00 3.0e+00 2.0e+00 2.4e+01 7.0e+00  
6.0e+00 1.0e+01 2.0e+01 1.0e+06 9.0e+00 0.0e+00 1.6e+01 3.4e+01 4.0e+00  
1.5e+01 3.2e+01 1.8e+01 2.3e+01 1.1e+01 1.0e+00 7.4e+01 1.3e+01 2.5e+01  
1.9e+01 2.8e+01 3.1e+01 2.2e+01 2.9e+01 1.7e+01 1.4e+01 2.6e+01 3.5e+01  
3.3e+01 2.7e+01 3.6e+01 2.1e+01]
```

```
In [23]: # Mengubah nilai yang tidak masuk akal sebagai nilai null  
df_pernah['Durasi Rata-rata Hubungan Responden'] = df_pernah['Durasi Rata-rata Hubungan Responden'].replace([1.0e+06], np.nan)
```

Imputasi NA

```
In [24]: df_pernah.isnull().sum()
```

```
Out[24]:
```

Usia Responden	0
Semester Responden	0
Jumlah Hubungan Responden	0
Durasi Rata-rata Hubungan Responden	1
Jumlah Kencan Responden	0
Tingkat Kepuasan Responden dalam Hubungan	0
Tingkat Kekuatan Cinta Responden dalam Hubungan	0
Tingkat Keterbukaan Komunikasi Responden dalam Hubungan	0
Respon Teman & Keluarga terhadap Hubungan Responden	0
Pengaruh Media Sosial terhadap Hubungan Responden	0
Frekuensi Konflik dalam Hubungan Responden	0
Penyelesaian Responden & Pasangan terhadap Konflik	0
Pengaruh Hubungan terhadap Performa Akademik Responden	0
Perubahan IPK yang Dialami Responden	204
Pendapat Responden terhadap Hubungan Terakhir	0
Media atau Bagaimana Responden Bertemu dengan Pasangan	0
Jenis Kelamin	0
Fakultas Responden	0
Status Responden Saat Ini	0
Status Responden Sebelumnya	0
dtype: int64	

```
In [25]: # Imputasi nilai kosong pada variabel numerik dengan menggunakan mean  
df_pernah.fillna(df_pernah.select_dtypes(include=['number']).mean(), inplace=True)
```

```
In [26]: # Mengisi nilai kosong pada kolom 'Perubahan IPK yang Dialami Responden' dengan 'Tidak Berpengaruh'  
df_pernah['Perubahan IPK yang Dialami Responden'].fillna('Tidak Berpengaruh', inplace=True)
```

```
In [27]: #Cek setelah NA  
df_pernah.isnull().sum()
```

Out[27]:

Usia Responden	0
Semester Responden	0
Jumlah Hubungan Responden	0
Durasi Rata-rata Hubungan Responden	0
Jumlah Kencan Responden	0
Tingkat Kepuasan Responden dalam Hubungan	0
Tingkat Kekuatan Cinta Responden dalam Hubungan	0
Tingkat Keterbukaan Komunikasi Responden dalam Hubungan	0
Respon Teman & Keluarga terhadap Hubungan Responden	0
Pengaruh Media Sosial terhadap Hubungan Responden	0
Frekuensi Konflik dalam Hubungan Responden	0
Penyelesaian Responden & Pasangan terhadap Konflik	0
Pengaruh Hubungan terhadap Performa Akademik Responden	0
Perubahan IPK yang Dialami Responden	0
Pendapat Responden terhadap Hubungan Terakhir	0
Media atau Bagaimana Responden Bertemu dengan Pasangan	0
Jenis Kelamin	0
Fakultas Responden	0
Status Responden Saat Ini	0
Status Responden Sebelumnya	0

dtype: int64

Mengubah (Encode) Nilai

In [28]:

```
df_pernah.head()
```

Out[28]:

	Usia Responden	Semester Responden	Jumlah Hubungan Responden	Durasi Rata-rata Hubungan Responden	Jumlah Kencan Responden	Tingkat Kepuasan Responden dalam Hubungan	Tingkat Kekuatan Cinta Responden dalam Hubungan	Tingkat Keterbukaan Komunikasi Responden dalam Hubungan
0	20	4.0	3.0	12.0	16.0	4.0	5.0	4.0
1	20	4.0	2.0	12.0	8.0	5.0	5.0	5.0
2	19	4.0	2.0	48.0	4.0	5.0	5.0	4.0
3	19	4.0	3.0	30.0	4.0	5.0	5.0	5.0
5	19	4.0	1.0	8.0	6.0	2.0	2.0	1.0

In [29]:

```
# Mengubah nilai-nilai data pada kolom 'Pengaruh Hubungan terhadap Performa Akademik Responde
df_pernah['Pengaruh Hubungan terhadap Performa Akademik Responden'].unique()
```

Out[29]:

```
array(['Ya', 'Tidak'], dtype=object)
```

In [30]:

```
# Mengubah nilai-nilai data pada kolom 'Pengaruh Hubungan terhadap Performa Akademik Respo
mapping = {
    'Ya': 1,
```

```
'Tidak': 0
}
df_pernah['Pengaruh Hubungan terhadap Performa Akademik Responden'] = df_pernah['Penga

In [31]: # Mengubah nilai-nilai data pada kolom 'Perubahan IPK yang Dialami Responden' menjadi
df_pernah['Perubahan IPK yang Dialami Responden'].unique()

Out[31]: array(['Naik', 'Tidak Berpengaruh', 'Turun'], dtype=object)

In [32]: # Mengubah nilai-nilai data pada kolom 'Perubahan IPK yang Dialami Responden'
mapping = {
    'Naik': 1,
    'Tidak Berpengaruh': 0,
    'Turun': -1
}
df_pernah['Perubahan IPK yang Dialami Responden'] = df_pernah['Perubahan IPK yang Dial

In [33]: # Mengubah nilai-nilai data pada kolom 'Pendapat Responden terhadap Hubungan Terakhir'
df_pernah['Pendapat Responden terhadap Hubungan Terakhir'].unique()

Out[33]: array(['Sangat mendukung', 'Cukup mendukung', 'Netral',
   'Sangat mengganggu', 'Cukup mengganggu'], dtype=object)

In [34]: # Mengubah nilai-nilai data pada kolom 'Pendapat Responden terhadap Hubungan Terakhir'
mapping = {
    'Sangat mendukung': 5,
    'Cukup mendukung': 4,
    'Netral': 3,
    'Sangat mengganggu': 2,
    'Cukup mengganggu': 1
}
df_pernah['Pendapat Responden terhadap Hubungan Terakhir'] = df_pernah['Pendapat Respo

In [35]: # Mengubah nilai-nilai data pada kolom 'Media atau Bagaimana Responden Bertemu dengan
df['Media atau Bagaimana Responden Bertemu dengan Pasangan'].unique()

Out[35]: array(['Di kampus/organisasi', 'Media sosial (Instagram, Twitter, dsb.)',
   'Melalui teman', 'nan', 'sma',
   'Aplikasi kencan (Tinder, Bumble, dll)', 'Sekolah', 'Badmin',
   'Game online', 'SMA', 'sekolah SMA', 'Gereja', 'Di SMA', 'sekolah',
   'teman sekolah'], dtype=object)

In [36]: # Mengubah nilai-nilai data pada kolom 'Media atau Bagaimana Responden Bertemu dengan
mapping = {
    'Di kampus/organisasi': 0,
    'Media sosial (Instagram, Twitter, dsb.)': 1,
    'Melalui teman': 2,
    'Badmin ': 2,
    'sma': 3,
    'Sekolah ': 3,
    'Sekolah': 3,
    'sekolah': 3,
    'SMA': 3,
    'sekolah SMA': 3,
    'Di SMA': 3,
    'teman sekolah': 4,
    'Gereja': 5,
    'Game online': 6,
    'Aplikasi kencan (Tinder, Bumble, dll)': 7
```

```
}
```

```
df_pernah['Media atau Bagaimana Responden Bertemu dengan Pasangan'] = df_pernah['Media
```

```
In [37]: # Mengubah nilai-nilai data pada kolom 'Status Responden Saat Ini' menjadi numerik  
df_pernah['Status Responden Saat Ini'].unique()
```

```
Out[37]: array(['Ya', 'Tidak'], dtype=object)
```

```
In [38]: # Mengubah nilai-nilai data pada kolom 'Status Responden Saat Ini'
```

```
mapping = {  
    'Ya': 1,  
    'Tidak': 0  
}
```

```
df_pernah['Status Responden Saat Ini'] = df_pernah['Status Responden Saat Ini'].map(mapping)
```

```
In [39]: # Menampilkan hasil perubahan
```

```
df_pernah.head()
```

```
Out[39]:
```

	Usia Responden	Semester Responden	Jumlah Hubungan Responden	Durasi Rata-rata Hubungan Responden	Jumlah Kencan Responden	Tingkat Kepuasan Responden dalam Hubungan	Tingkat Kekuatan Cinta Responden dalam Hubungan	Tingkat Keterbukaan Komunikasi Responden dalam Hubungan
0	20	4.0	3.0	12.0	16.0	4.0	5.0	4.0
1	20	4.0	2.0	12.0	8.0	5.0	5.0	5.0
2	19	4.0	2.0	48.0	4.0	5.0	5.0	4.0
3	19	4.0	3.0	30.0	4.0	5.0	5.0	5.0
5	19	4.0	1.0	8.0	6.0	2.0	2.0	1.0

```
In [40]: # Menyimpan data yang telah diproses  
df_pernah.to_csv('DataAfter_3.csv', index=False)  
df_campur.to_csv('DataAfter_3 (campur).csv', index=False)
```

5. Measurement Quality

```
In [41]: pip install pingouin
```

Requirement already satisfied: pingouin in c:\users\gerva\anaconda3\lib\site-packages (0.5.5)
Requirement already satisfied: matplotlib in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (3.7.1)
Requirement already satisfied: numpy in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (1.24.3)
Requirement already satisfied: pandas>=1.5 in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (1.5.3)
Requirement already satisfied: pandas-flavor in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (0.7.0)
Requirement already satisfied: scikit-learn>=1.2 in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (1.5.2)
Requirement already satisfied: scipy in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (1.10.1)
Requirement already satisfied: seaborn in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (0.12.2)
Requirement already satisfied: statsmodels in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (0.14.0)
Requirement already satisfied: tabulate in c:\users\gerva\anaconda3\lib\site-packages (from pingouin) (0.8.10)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\gerva\appdata\roaming\python\python311\site-packages (from pandas>=1.5->pingouin) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\gerva\anaconda3\lib\site-packages (from pandas>=1.5->pingouin) (2022.7)
Requirement already satisfied: joblib>=1.2.0 in c:\users\gerva\anaconda3\lib\site-packages (from scikit-learn>=1.2->pingouin) (1.2.0)
Requirement already satisfied: threadpoolctl>=3.1.0 in c:\users\gerva\anaconda3\lib\site-packages (from scikit-learn>=1.2->pingouin) (3.5.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\gerva\appdata\roaming\python\python311\site-packages (from matplotlib->pingouin) (23.2)
Requirement already satisfied: pillow>=6.2.0 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (9.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\gerva\anaconda3\lib\site-packages (from matplotlib->pingouin) (3.0.9)
Requirement already satisfied: xarray in c:\users\gerva\anaconda3\lib\site-packages (from pandas-flavor->pingouin) (2023.6.0)
Requirement already satisfied: patsy>=0.5.2 in c:\users\gerva\anaconda3\lib\site-packages (from statsmodels->pingouin) (0.5.3)
Requirement already satisfied: six in c:\users\gerva\appdata\roaming\python\python311\site-packages (from patsy>=0.5.2->statsmodels->pingouin) (1.16.0)
Note: you may need to restart the kernel to use updated packages.

In [42]:

```
import pandas as pd
import pingouin as pg # install dengan: pip install pingouin

# Pilih 5 variabel utama
mydata = df_pernah[[
    'Tingkat Kepuasan Responden dalam Hubungan',
    'Tingkat Kekuatan Cinta Responden dalam Hubungan',
    'Tingkat Keterbukaan Komunikasi Responden dalam Hubungan',
    'Respon Teman & Keluarga terhadap Hubungan Responden',
```

```
'Penyelesaian Responden & Pasangan terhadap Konflik'  
]]
```

```
In [43]:  
# =====  
# Uji VALIDITAS: Corrected Item-Total Correlation  
# =====  
validity = {}  
for col in mydata.columns:  
    total_score = mydata.drop(columns=[col]).sum(axis=1)  
    correlation = mydata[col].corr(total_score)  
    validity[col] = correlation  
  
# Tampilkan hasil uji validitas  
validity_result = pd.DataFrame.from_dict(validity, orient='index', columns=['CITC'])  
print("==> Corrected Item-Total Correlation (CITC) ==>")  
print(validity_result)
```

```
==> Corrected Item-Total Correlation (CITC) ==>  
CITC  
Tingkat Kepuasan Responden dalam Hubungan      0.271693  
Tingkat Kekuatan Cinta Responden dalam Hubungan 0.243738  
Tingkat Keterbukaan Komunikasi Responden dalam ... 0.282583  
Respon Teman & Keluarga terhadap Hubungan Respo... 0.175340  
Penyelesaian Responden & Pasangan terhadap Konflik 0.261661
```

```
In [44]:  
# =====  
# Uji RELIABILITAS: Cronbach's Alpha  
# =====  
cb_alpha = pg.cronbach_alpha(data=mydata)  
print("==> Cronbach's Alpha ==>")  
print(f"Cronbach Alpha: {cb_alpha[0]:.4f}")  
  
==> Cronbach's Alpha ==>  
Cronbach Alpha: 0.4646
```

6. PRELIMINARY ANALYSIS

Perhitungan Mean dengan Confidence Interval 95% untuk Variabel Numerik

In [45]:

```
import pandas as pd
import numpy as np
from scipy import stats

# Ambil kolom numerik
numerical_cols = df_pernah.select_dtypes(include=['float64', 'int64']).columns

# Hitung mean dan confidence interval
ci_result = []

for col in numerical_cols:
    data = df_pernah[col].dropna()
    mean = data.mean()
    sem = stats.sem(data) # standard error of the mean
    ci = stats.t.interval(0.95, len(data)-1, loc=mean, scale=sem)
    ci_result.append({
        'Variabel': col,
        'Mean': mean,
        '95% CI Lower': ci[0],
        '95% CI Upper': ci[1]
    })

ci_df = pd.DataFrame(ci_result)
print("== Mean dan Confidence Interval 95% ==")
print(ci_df)
```

== Mean dan Confidence Interval 95% ==

	Variable	Mean
0	Usia Responden	19.885714
1	Semester Responden	3.971429
2	Jumlah Hubungan Responden	2.897143
3	Durasi Rata-rata Hubungan Responden	17.687679
4	Jumlah Kencan Responden	2.702857
5	Tingkat Kepuasan Responden dalam Hubungan	3.494286
6	Tingkat Kekuatan Cinta Responden dalam Hubungan	3.580000
7	Tingkat Keterbukaan Komunikasi Responden dalam...	3.482857
8	Respon Teman & Keluarga terhadap Hubungan Resp...	3.485714
9	Pengaruh Media Sosial terhadap Hubungan Responden	3.382857
10	Frekuensi Konflik dalam Hubungan Responden	3.311429
11	Penyelesaian Responden & Pasangan terhadap Kon...	3.620000
12	Pengaruh Hubungan terhadap Performa Akademik R...	0.417143
13	Perubahan IPK yang Dialami Responden	0.034286
14	Pendapat Responden terhadap Hubungan Terakhir	3.142857
15	Media atau Bagaimana Responden Bertemu dengan ...	2.232092
16	Jenis Kelamin	0.508571
17	Fakultas Responden	2.868571
18	Status Responden Saat Ini	0.591429
19	Status Responden Sebelumnya	1.000000

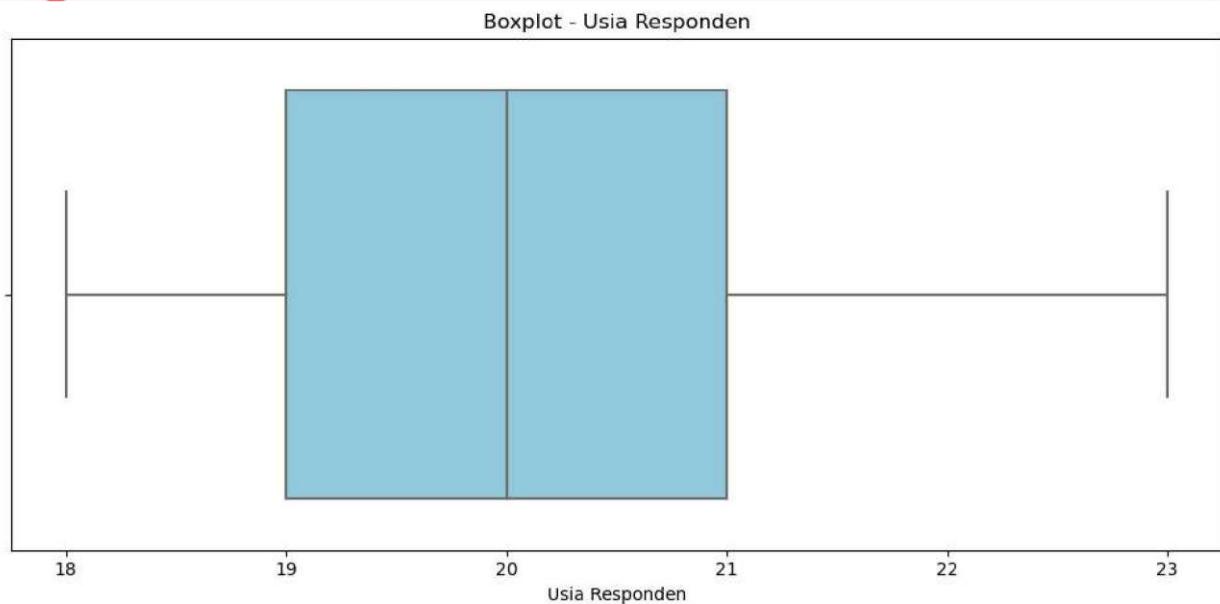
	95% CI Lower	95% CI Upper
0	19.755743	20.015686
1	3.845631	4.097226
2	2.738543	3.055742
3	16.585809	18.789549
4	2.478669	2.927045
5	3.377934	3.610637
6	3.461255	3.698745
7	3.363561	3.602153
8	3.364309	3.607119
9	3.253708	3.512006
10	3.189292	3.433565
11	3.500034	3.739966
12	0.365231	0.469055
13	-0.033615	0.102186
14	2.997537	3.288178
15	1.958961	2.505223
16	0.455939	0.561203
17	2.598809	3.138333
18	0.539676	0.643181
19	NaN	NaN

Visualisasi

```
In [46]: import matplotlib.pyplot as plt
import seaborn as sns
```

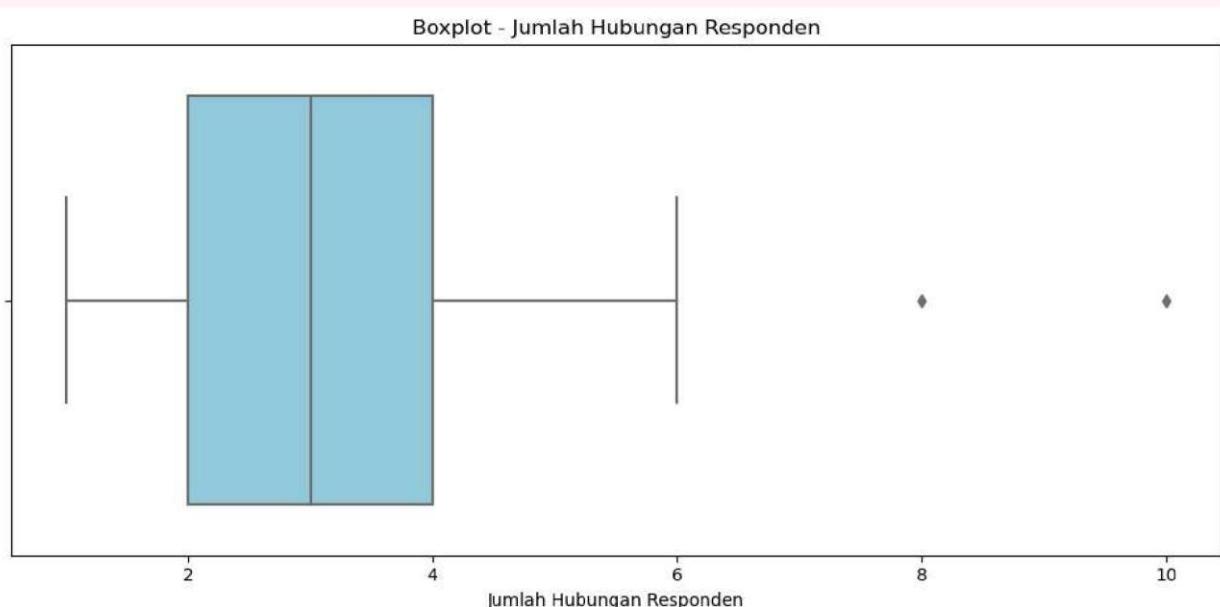
```
def plot_boxplot(df_pernah, col):
    plt.figure(figsize=(10, 5))
    sns.boxplot(data=df_pernah, x=col, color='skyblue')
    plt.title(f'Boxplot - {col}')
    plt.tight_layout()
    plt.show()
```

In [47]: `plot_boxplot(df_pernah, 'Usia Responden')`



Nilai dari IQR berada di antara 19 dan 21 dimana responden berada didalam rentang usia remaja. Outlier tidak terlihat dalam boxplot ini.

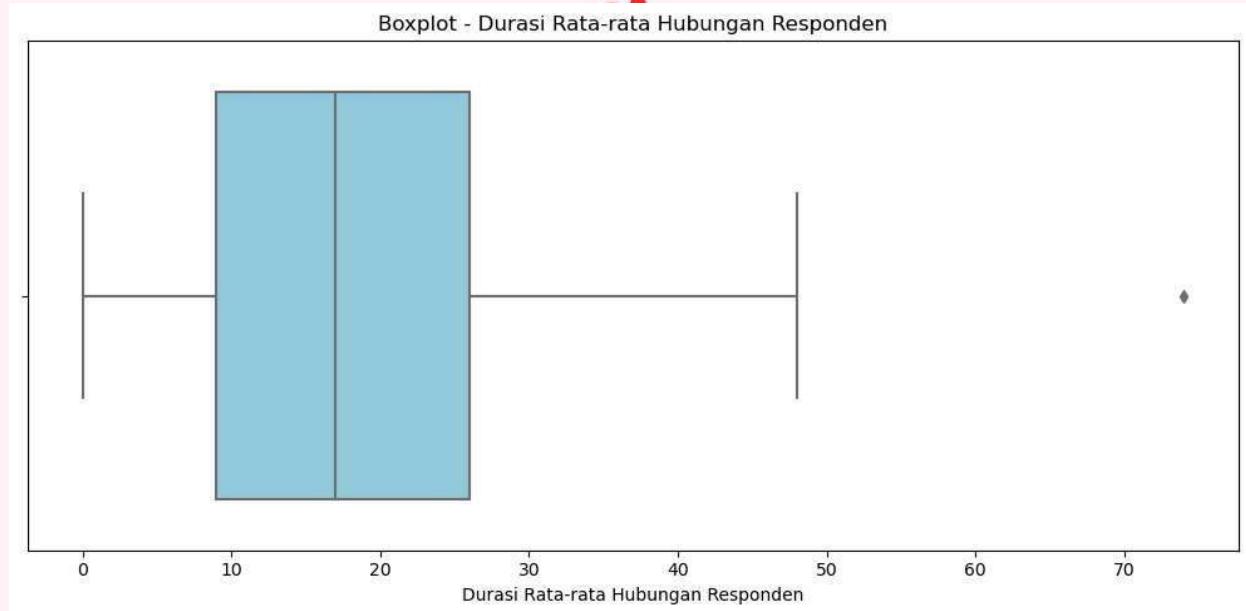
In [48]: `plot_boxplot(df_pernah, 'Jumlah Hubungan Responden')`



Mayoritas responden berada di jumlah 1-6, yang menunjukkan bahwa sebagian besar orang memiliki jumlah hubungan dalam rentang tersebut. Garis tengah yang menunjukkan median

berada di nilai 3. Terdapat 2 outlier yang berada di nilai 8 dan 10, yang berarti hubungan 8 dan 10 kali dianggap tidak wajar. IQR berada di rentang 2 sampai 4.

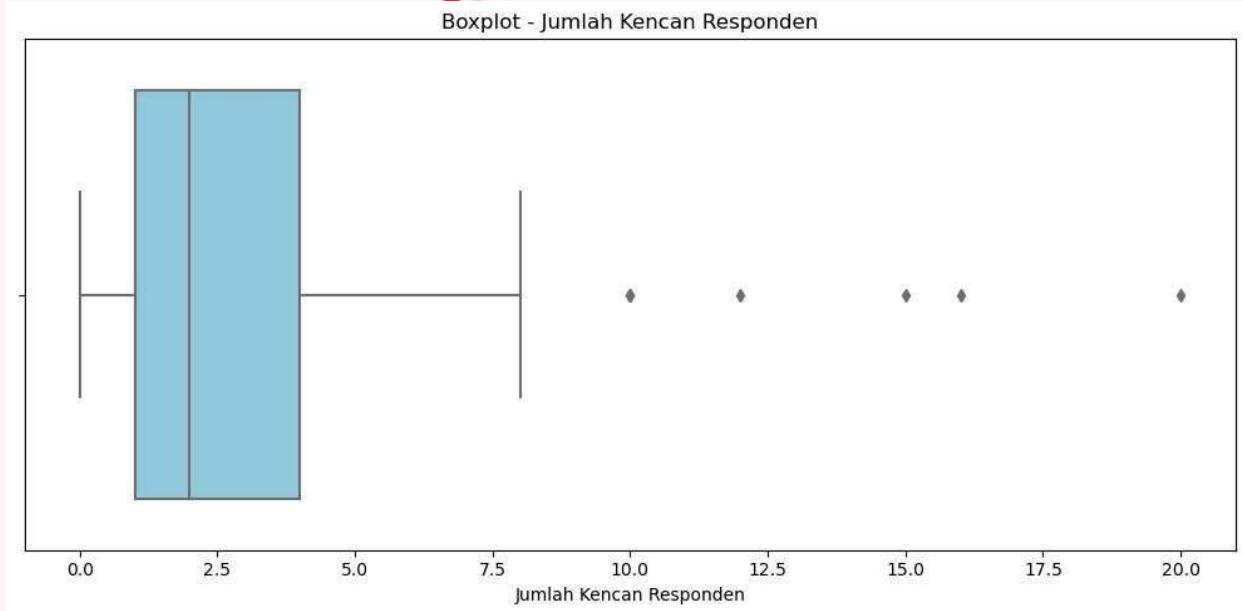
```
In [49]: plot_boxplot(df_pernah, 'Durasi Rata-rata Hubungan Responden')
```



Function to make Pie Plot

Durasi rata-rata dari responden adalah 9 sampai 25, dan ada outlier yang berada di nilai yang lebih dari 70

```
In [50]: plot_boxplot(df_pernah, 'Jumlah Kencan Responden')
```



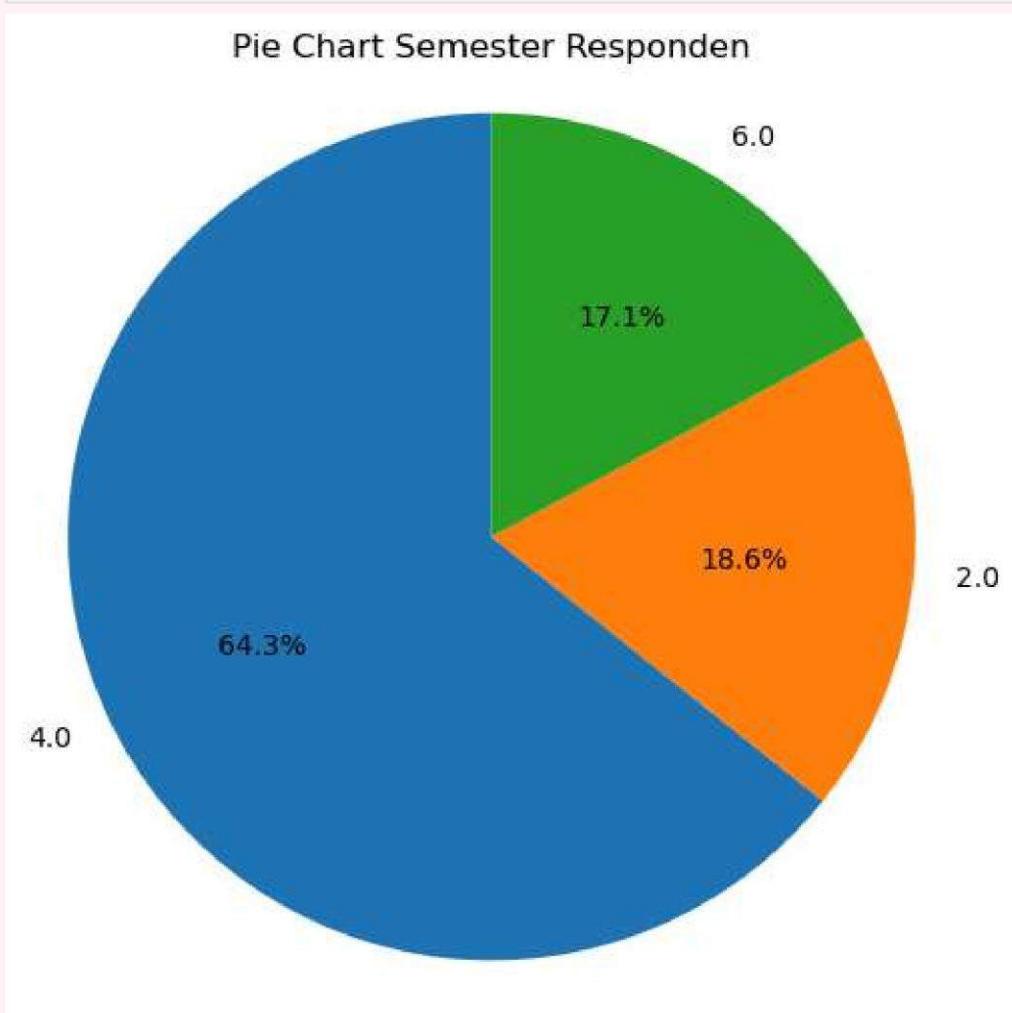
Variable 'Jumlah Kencan Responden' memiliki nilai outlier yang cukup banyak, kebanyakan berada di nilai 7.5 keatas. Nilai nilai ini dianggap tidak wajar. Mayoritas responden pernah berkencan di antara 1 sampai 4 kali

```
In [51]: import matplotlib.pyplot as plt
import pandas as pd

def plot_pie_from_column(df, column, title='Pie Chart', colors=None, autopct='%1.1f%%'
value_counts = df[column].value_counts()
labels = value_counts.index
sizes = value_counts.values

plt.figure(figsize=(6,6))
plt.pie(sizes, labels=labels, colors=colors, autopct=autopct, startangle=90)
plt.title(title)
plt.axis('equal')
plt.show()
```

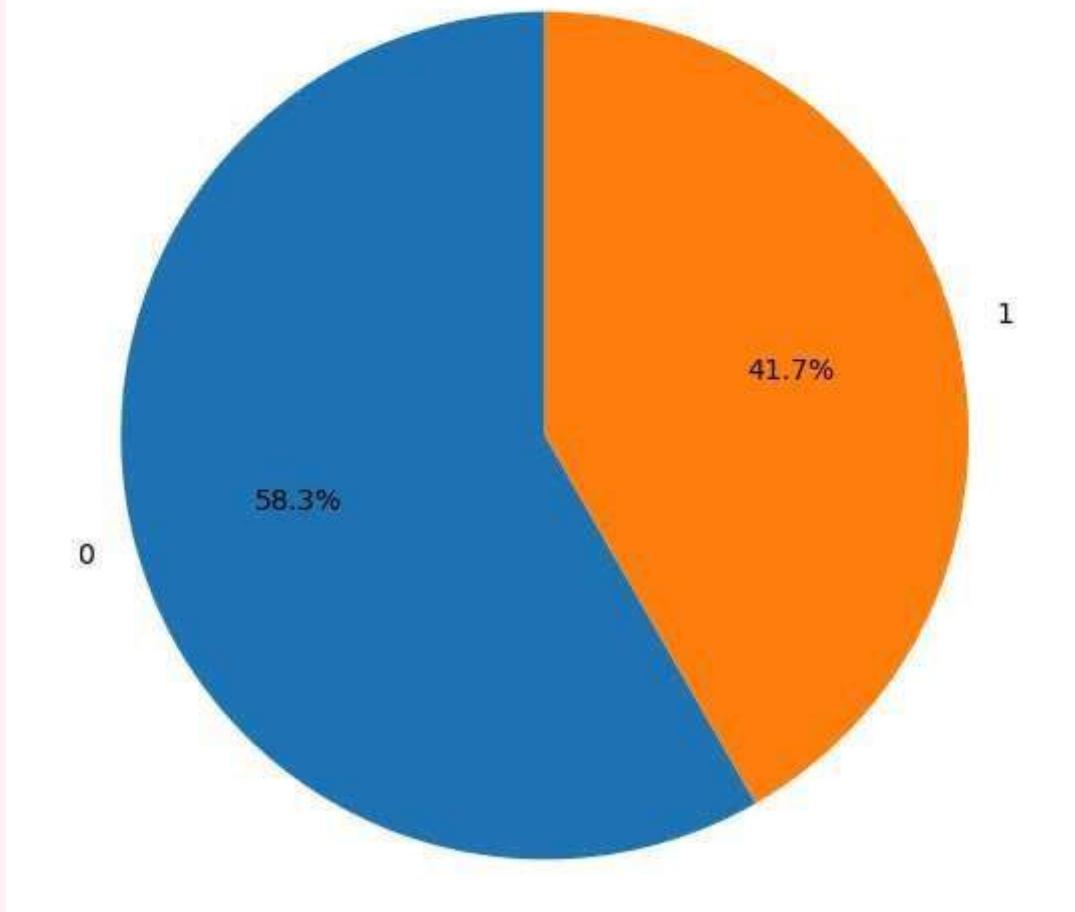
```
In [52]: plot_pie_from_column(df_pernah, 'Semester Responden', title='Pie Chart Semester Respon
```



Mayoritas responden kami sedang berada di semester 4, sisanya berada pada semester 2 sebesar 18.6% dan semester 6 sebesar 17.1%

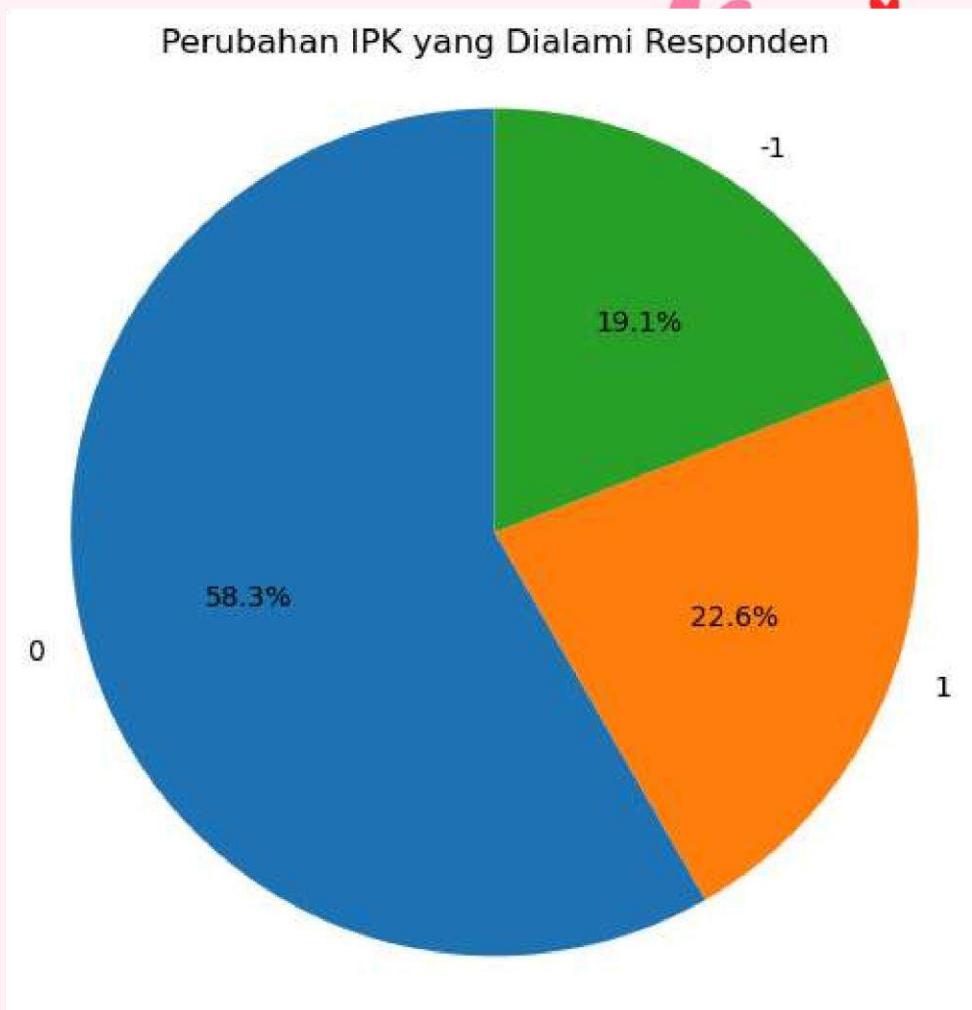
```
In [53]: plot_pie_from_column(df_pernah, 'Pengaruh Hubungan terhadap Performa Akademik Responde
```

Pengaruh Hubungan terhadap Performa Akademik Responden



Mayoritas hubungan responden tidak memengaruhi performa akademik dengan besar 58.3%, sedangkan sisanya sebesar 41.7% merasa bahwa hubungannya memengaruhi performa akademik

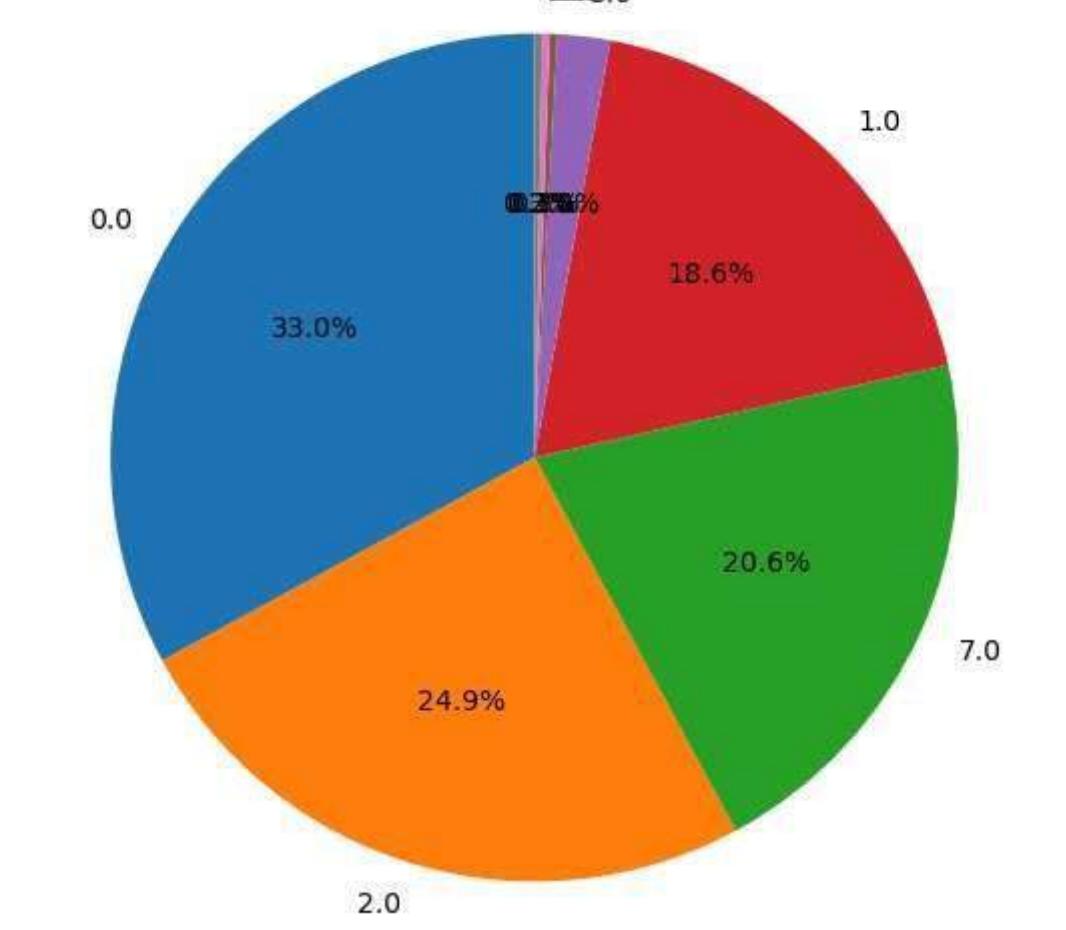
```
In [54]: plot_pie_from_column(df_pernah, 'Perubahan IPK yang Dialami Responden', title='Perubah
```



Berdasarkan hasil, mayoritas responden tidak merasakan terjadinya perubahan IPK, sedangkan kurang dari 40% mengalaminya, dengan jumlah 22.6% mengalami kenaikan dan 19.1% mengalami penurunan

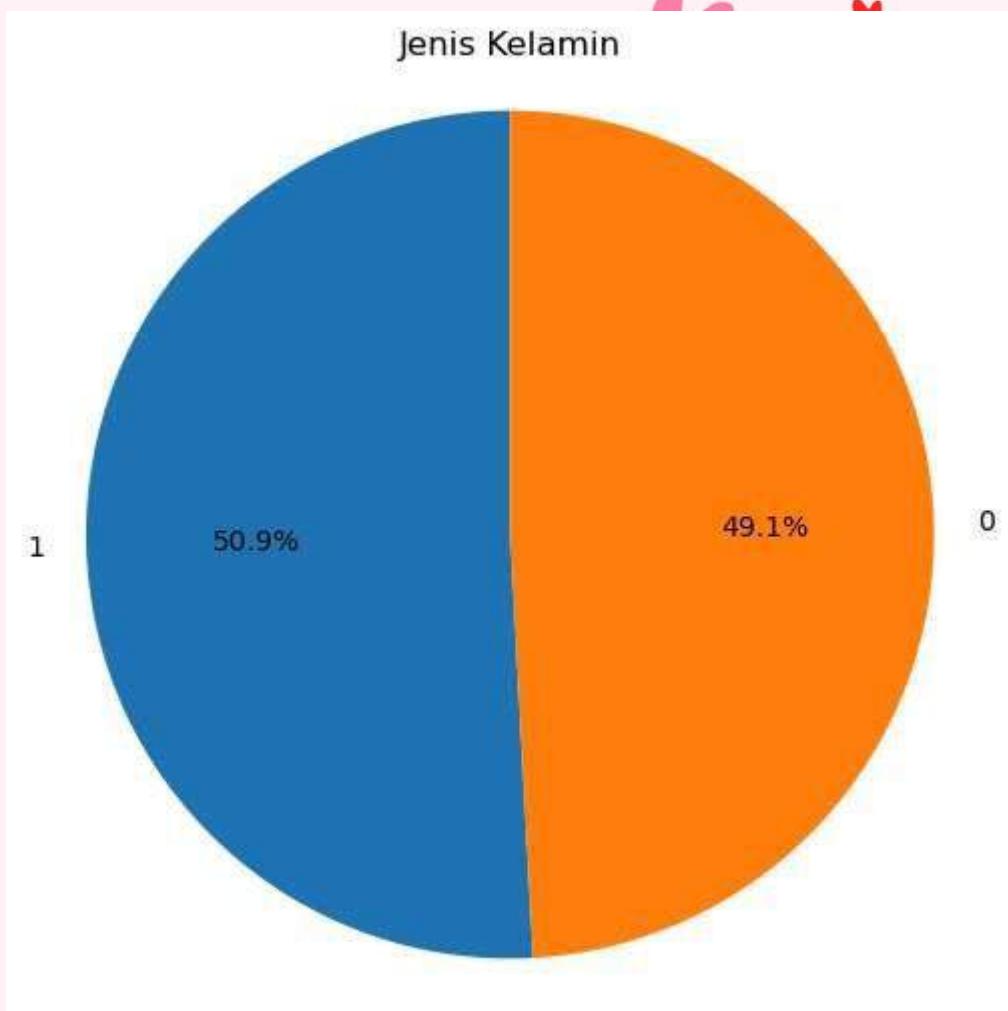
```
In [55]: plot_pie_from_column(df_pernah, 'Media atau Bagaimana Responden Bertemu dengan Pasangan')
```

Media atau Bagaimana Responden Bertemu dengan Pasangan



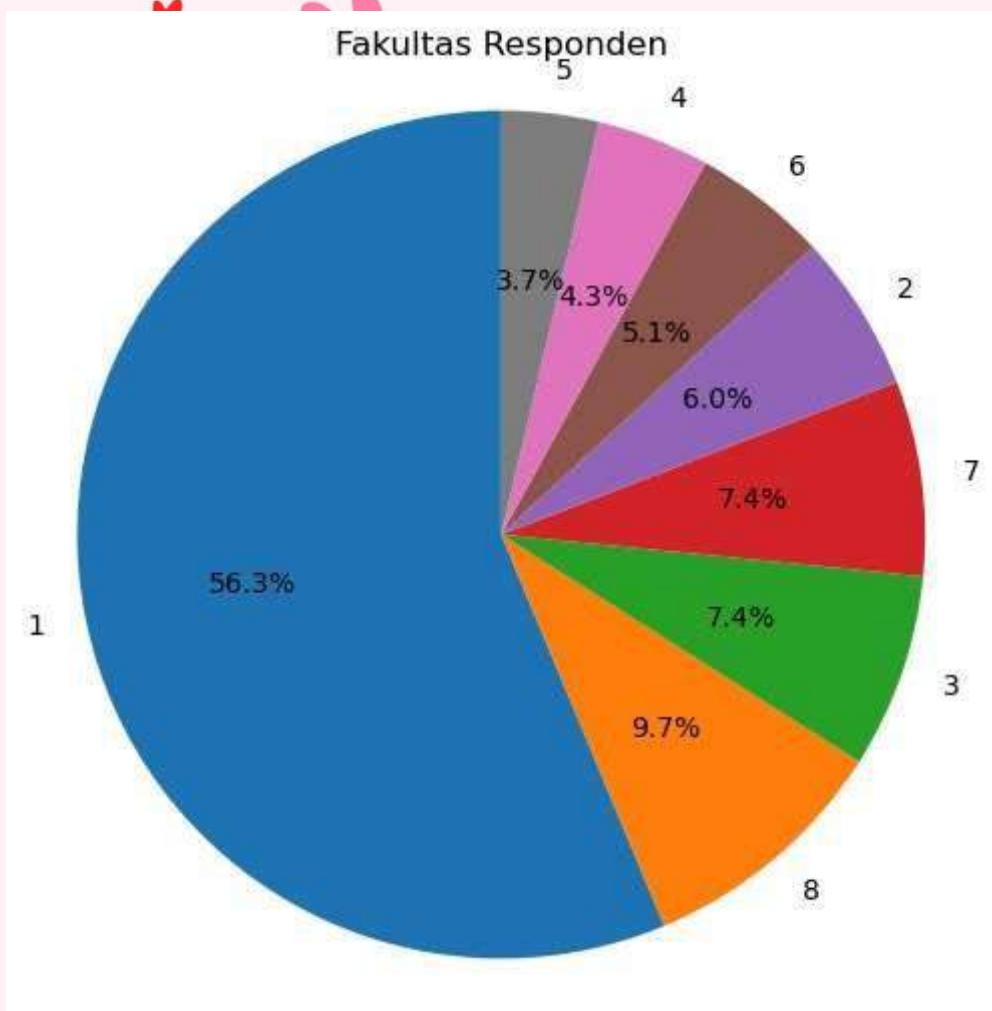
Dari Pie Chart ini, dapat kami ketahui bahwa mayoritas responden kami bertemu dengan pasangannya saat berada di Kampus / Organisasi dengan jumlah hampir mencapai 33%, 25.1% melalui teman, 20.6% melalui aplikasi kencan seperti Tinder, Bumble, etc. 18.6% melalui media sosial. Sisanya hampir kurang dari 5% menggunakan media lain, diluar pilihan peneliti, seperti SMA, Game Online, Hobi, Tempat Ibadah, etc.

```
In [56]: plot_pie_from_column(df_pernah, 'Jenis Kelamin', title='Jenis Kelamin')
```



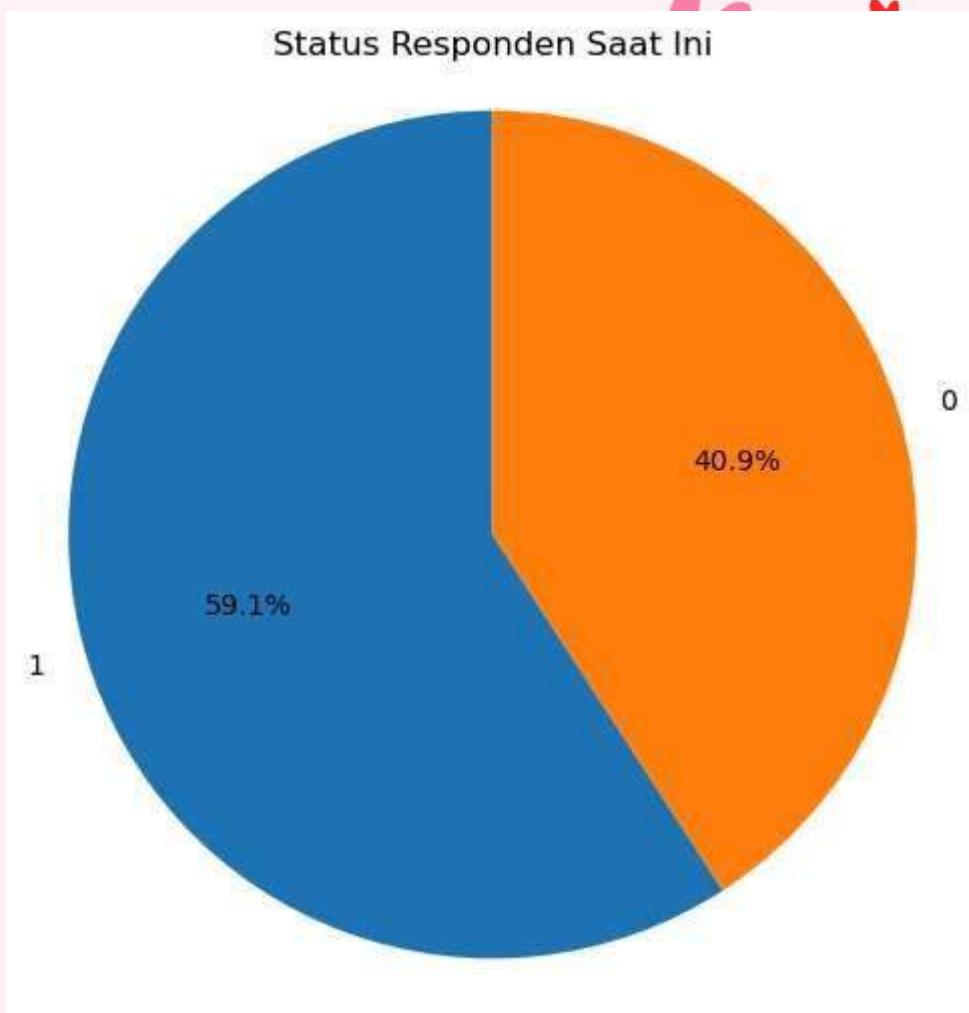
Responden kami di dominasi dengan Laki - Laki sebanyak 50.9% dan perempuan sebanyak 49.1%

```
In [57]: plot_pie_from_column(df_pernah, 'Fakultas Responden', title='Fakultas Responden')
```



Mayoritas responden kami berasal dari fakultas SoCS dengan jumlah sebesar 56.3%, dilanjutkan dengan School of Design dengan jumlah sebesar 9.7% dan School of Information System serta School of Accounting sebesar 7.4%

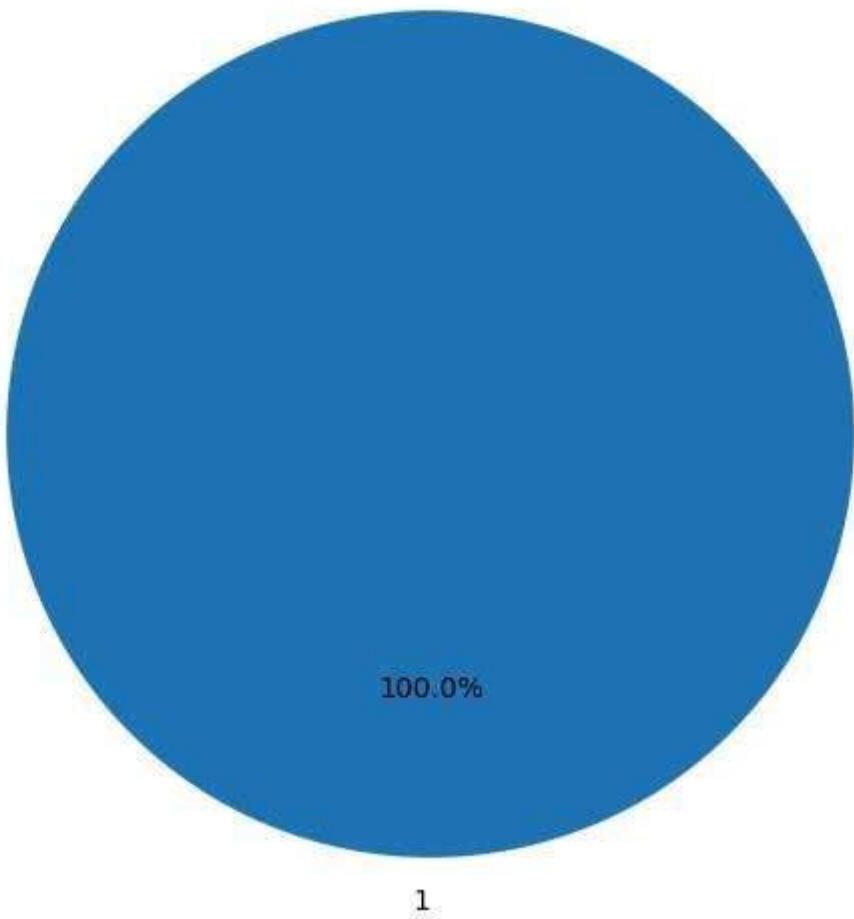
```
In [58]: plot_pie_from_column(df_pernah, 'Status Responden Saat Ini', title='Status Responde
```



Dalam Pie Chart ini, dapat dilihat bahwa mayoritas responden kami sedang dalam hubungan romantis dengan jumlah hampir mencapai 60%, sedangkan sisanya sebanyak 40.9% sedang tidak dalam hubungan romantis

```
In [59]: plot_pie_from_column(df_pernah, 'Status Responden Sebelumnya', title='Status Responder')
```

Status Responden Sebelumnya



Dari seluruh responden yang didapatkan, disimpulkan bahwa responden sudah pernah berada dalam hubungan romantis yakni pacaran maupun HTS

Function untuk stacked barchart (skala likert)

In [60]:

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

import pandas as pd
import matplotlib.pyplot as plt

def plot_likert_stacked_bar(df, likert_columns, categories_order, category_labels, tit
    """
        Membuat horizontal stacked bar chart dengan label Likert dan persentase.

    Parameters:
    - df: DataFrame
    - likert_columns: list of columns (pertanyaan-pertanyaan Likert)
    - categories_order: list of numeric Likert codes (misal: [1.0, 2.0, 3.0, 4.0, 5.0]
    - category_labels: label teks untuk legenda (panjangnya sama dengan categories_ord
    - title: judul chart
    """
    likert_data = pd.DataFrame()
```

```
for col in likert_columns:
    counts = df[col].value_counts(normalize=True).reindex(categories_order, fill_value=0)
    likert_data[col] = counts

likert_data = likert_data.T # Transpose agar pertanyaan jadi baris

# Plot
fig, ax = plt.subplots(figsize=(12, len(likert_columns) * 0.6))
left = [0] * len(likert_data)

for idx, cat in enumerate(categories_order):
    heights = likert_data[cat]
    bars = ax.barh(likert_data.index, heights, left=left, label=category_labels[idx])

    # Tambahkan persentase pada masing-masing segmen
    for bar in bars:
        width = bar.get_width()
        if width > 0.03: # tampilkan label jika >3%
            ax.text(bar.get_x() + width / 2, bar.get_y() + bar.get_height() / 2,
                    f'{width*100:.1f}%', ha='center', va='center', fontsize=9, color='black')

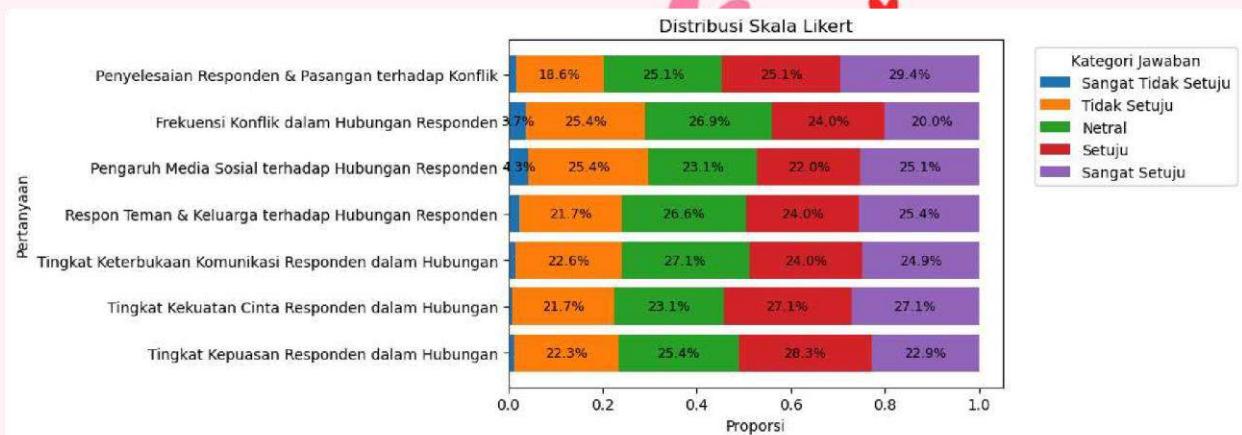
    # Update posisi "left" untuk bar berikutnya
    left = [i + j for i, j in zip(left, heights)]

ax.set_title(title)
ax.set_xlabel('Proporsi')
ax.set_ylabel('Pertanyaan')
ax.legend(title='Kategori Jawaban', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()

likert_columns = [
    'Tingkat Kepuasan Responden dalam Hubungan',
    'Tingkat Kekuatan Cinta Responden dalam Hubungan',
    'Tingkat Keterbukaan Komunikasi Responden dalam Hubungan',
    'Respon Teman & Keluarga terhadap Hubungan Responden',
    'Pengaruh Media Sosial terhadap Hubungan Responden',
    'Frekuensi Konflik dalam Hubungan Responden',
    'Penyelesaian Responden & Pasangan terhadap Konflik'
]

categories_order = [1.0, 2.0, 3.0, 4.0, 5.0]
category_labels = ['Sangat Tidak Setuju', 'Tidak Setuju', 'Netral', 'Setuju', 'Sangat Setuju']

plot_likert_stacked_bar(df_pernah, likert_columns, categories_order, category_labels)
```



Penyelesaian Konflik oleh Responden & Pasangan Majoritas responden menyatakan setuju atau sangat setuju bahwa mereka mampu menyelesaikan konflik dengan baik (sekitar 54.5%), menunjukkan kemampuan penyelesaian konflik yang relatif tinggi.

Frekuensi Konflik dalam Hubungan Jawaban tersebar merata, namun sebagian besar netral atau tidak setuju bahwa konflik sering terjadi, mengindikasikan bahwa konflik tidak terlalu sering terjadi dalam hubungan responden.

Pengaruh Media Sosial terhadap Hubungan Cukup banyak responden netral atau tidak setuju bahwa media sosial berpengaruh pada hubungan mereka. Artinya, sebagian besar mungkin tidak terlalu terpengaruh oleh media sosial dalam menjalin hubungan.

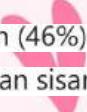
Dukungan Teman & Keluarga Lebih dari setengah responden merasa mendapatkan dukungan dari orang terdekat mereka (sekitar 50% setuju atau sangat setuju), yang menjadi faktor positif dalam hubungan mereka.

Keterbukaan Komunikasi Tingkat keterbukaan komunikasi cukup baik, dengan lebih dari 50% responden setuju atau sangat setuju, menandakan komunikasi terbuka dengan pasangan mereka.

Kekuatan Cinta dalam Hubungan Hampir seimbang antara yang setuju dan sangat setuju (~54% total), dengan sedikit yang netral, mengindikasikan adanya perasaan cinta yang kuat dalam hubungan mereka.

Tingkat Kepuasan dalam Hubungan Jawaban didominasi oleh "Setuju" dan "Netral", namun cukup banyak yang menjawab "Sangat Setuju" juga, menunjukkan tingkat kepuasan yang cukup tinggi, meskipun ada ruang untuk peningkatan.



Sebagian besar responden  (46%) merasa hubungan terakhir mereka mendukung, sementara 36% merasa terganggu, dan sisanya netral. Ini menunjukkan bahwa pengalaman hubungan romantis terakhir cenderung lebih banyak memberikan dampak positif dibanding negatif.

Code Machine Learning

```
In [62]: import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.impute import SimpleImputer

# 1. Load Data
df = pd.read_csv("DataAfter_3.csv") # Ganti dengan nama file atau sesuaikan dengan da

# 2. Tentukan fitur dan target
X = df.drop(columns=["Pendapat Responden terhadap Hubungan Terakhir"])
y = df["Pendapat Responden terhadap Hubungan Terakhir"]

# 3. Deteksi kolom numerik dan kategorikal
num_cols = X.select_dtypes(include=["int64", "float64"]).columns.tolist()
cat_cols = [col for col in X.columns if col not in num_cols]

# 4. Preprocessing
numeric_transformer = Pipeline(steps=[
    ("imputer", SimpleImputer(strategy="mean")),
    ("scaler", StandardScaler())
])

categorical_transformer = Pipeline(steps=[
    ("imputer", SimpleImputer(strategy="most_frequent")),
    ("onehot", OneHotEncoder(handle_unknown="ignore"))
])

preprocessor = ColumnTransformer(
    transformers=[
        ("num", numeric_transformer, num_cols),
        ("cat", categorical_transformer, cat_cols)
    ]
)

# 5. Buat pipeline akhir
model = Pipeline(steps=[
    ("preprocessor", preprocessor),
    ("classifier", RandomForestClassifier(random_state=42))
])

# 6. Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42, stratify=y
)

# 7. Fit model
```

```
model.fit(X_train, y_train)

# 8. Prediksi & evaluasi
y_pred = model.predict(X_test)

print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
```

Accuracy: 0.24285714285714285

Classification Report:

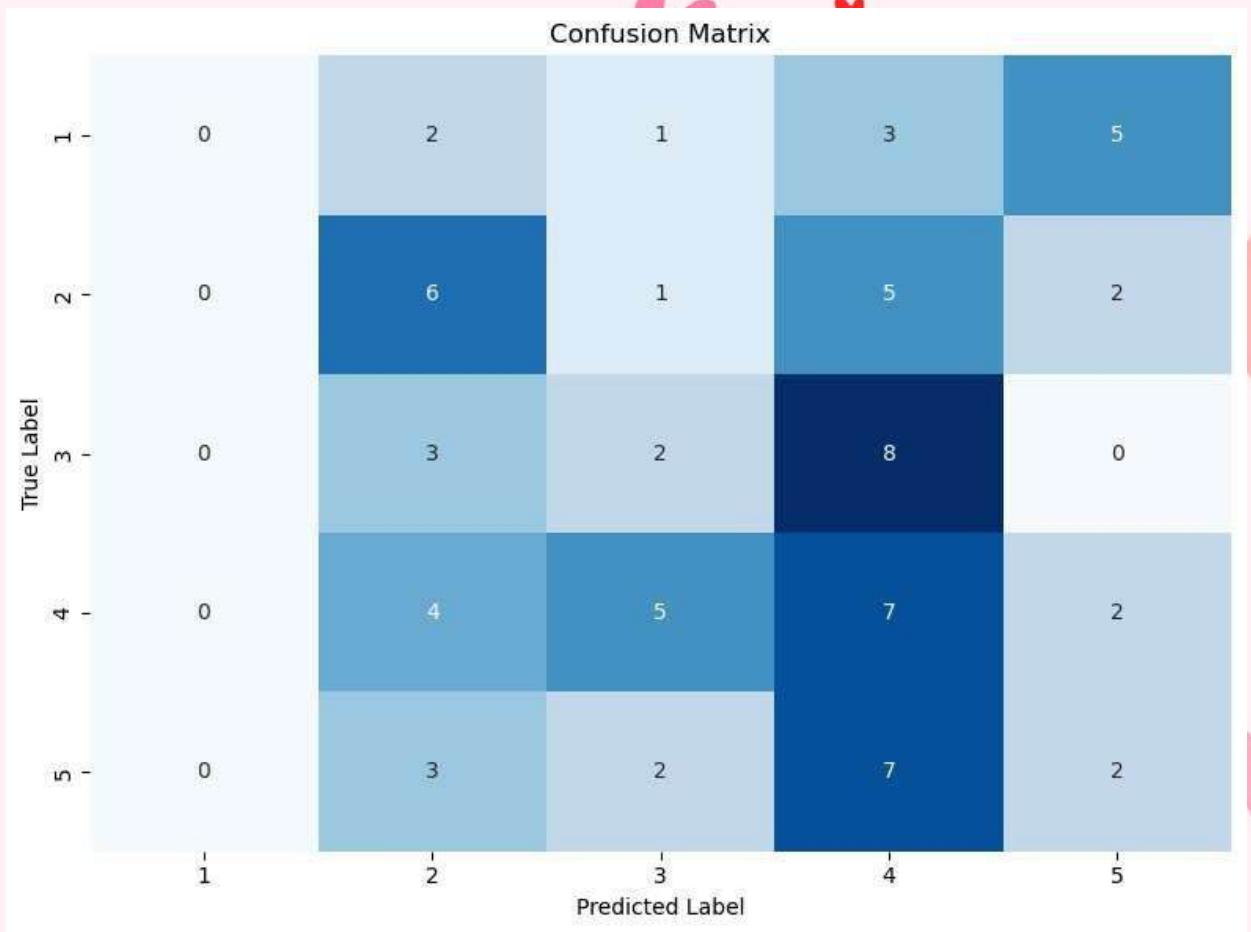
	precision	recall	f1-score	support
1	0.00	0.00	0.00	11
2	0.33	0.43	0.38	14
3	0.18	0.15	0.17	13
4	0.23	0.39	0.29	18
5	0.18	0.14	0.16	14
accuracy			0.24	70
macro avg	0.19	0.22	0.20	70
weighted avg	0.20	0.24	0.21	70

```
In [63]: # Confusion Matrix visual
import seaborn as sns
import matplotlib.pyplot as plt

cm = confusion_matrix(y_test, y_pred)

plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True, fmt="d", cmap="Blues", cbar=False,
            xticklabels=sorted(y.unique()),
            yticklabels=sorted(y.unique()))

plt.title("Confusion Matrix")
plt.xlabel("Predicted Label")
plt.ylabel("True Label")
plt.tight_layout()
plt.show()
```



PEMBAGIAN TUGAS PEMBUATAN FINAL PROJECT AKHIR

SURVEY & SAMPLING METHODS KELOMPOK 3

- Geoffrey : Membuat laporan, membuat code
- Gervasius : Membuat survei, membuat code
- Jackie : Membuat laporan, membuat code, membuat cover
- Nixon : Membuat code
- Nathanael : Membuat code

REFERENSI :

- American College Health Association. (2021). National College Health Assessment III: Reference group executive summary Spring 2021.
- Kementerian Kesehatan Republik Indonesia. (2017). Profil Kesehatan Indonesia tahun 2017. Kemenkes RI.
<https://www.kemkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-Indonesia-2017.pdf>
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia. (2023). Statistik pendidikan tinggi 2022.
<https://pddikti.kemdiktisaintek.go.id/>
- Kusumastuti, P., & Ratnaningsih, N. M. (2020). Kesehatan mental mahasiswa selama masa pandemi COVID-19. Jurnal Kesehatan Masyarakat (Undip), 16(2), 123–132. <https://ejurnal.undip.ac.id/index.php/index/index>
- Papalia, D. E., & Martorell, G. (2014). Experience human development (13th ed.). McGraw-Hill Education.
- Santrock, J. W. (2016). Life-span development (16th ed.). McGraw-Hill Education.
- SDKI 2017. (2017). Survei Demografi dan Kesehatan Indonesia 2017.
<https://www.kemendukbangga.go.id/>
- World Health Organization. (2011). WHO guidelines on preventing early pregnancy and poor reproductive outcomes among adolescents in developing countries. World Health Organization.
https://iris.who.int/bitstream/handle/10665/44691/9789241502214_eng.pdf