

**MODEL PEMBELAJARAN DAN LAPORAN AKHIR
PROJECT-BASED LEARNING
MATA KULIAH EXPLONATORY DATA ANALYSIS**



“Persepsi Mahasiswa Administrasi Publik Mengenai Dampak Media Sosial Terhadap Kesehatan Mental dengan Metode Analisis Faktor dan Analisis Data Text”

DISUSUN OLEH KELOMPOK II:

- | | |
|---------------------------|---------------|
| 1. INTAN ADIBA | (22083010007) |
| 2. MUHAMMAD NASHIF FARID | (22083010024) |
| 3. MAULIDYA PRASTITA SYAH | (22083010039) |
| 4. AFFA LELIRA IBRAHIM | (22083010065) |
| 5. WAHYU PUTRA PRATAMA | (22083010092) |

DOSEN PENGAMPU:

AVIOLLA TERZA DAMALIANA, S.Si., M.Stat (NIP. 19940802 202203 2015)

**PROGRAM STUDI SAINS DATA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAWA TIMUR
2023**

DAFTAR ISI

BAB I : PENDAHULUAN	3
1.1. Latar Belakang	3
1.2. Permasalahan	5
1.3. Tujuan	5
1.4. Manfaat	6
BAB II: TINJAUAN PUSTAKA	7
2.1. Teori Penunjang	7
2.2. Penelitian Terkait	12
BAB III : METODOLOGI PENELITIAN	16
3.1 Jenis Penelitian.....	16
3.2 Sumber Data Penelitian.....	16
3.3 Flowchart	16
3.4 Uji Validitas dan Reliabilitas	17
BAB IV : HASIL ANALISIS	20
4.1 Analisis Deskriptif.....	20
4.1.1 Gambaran Responden Yang Mengacu Pada Jenis Kelamin	20
4.1.1 Gambaran Responden Berdasarkan Usia	20
4.1.2 Gambaran Penggunaan Media Sosial dalam Dataset.....	21
4.1.3 Gambaran Perbandingan Penggunaan Aplikasi Sosial Media Terhadap Jenis Kelamin ..	22
4.1.4 Gambaran Seberapa Setuju Bahwa Aktif dalam Bermedia Sosial Berpengaruh Terhadap Kesehatan Mental.....	23
4.2 Analisis Faktor	23
4.2.1 Data Preparation.....	23
4.2.2 Menentukan jumlah faktor	29
4.2.3 Interpretasi Hasil Analisis Faktor	35
4.3 Analisis Data Teks.....	36
4.3.1 Preparation Data.....	36
4.3.2 Removing Stop Words	40
4.3.3 Performing Stemming	44
4.3.4 Creating Word Cloud.....	46
BAB V : KESIMPULAN	48
DAFTAR PUSTAKA.....	49

BAB I : PENDAHULUAN

1.1. Latar Belakang

Pada masa sekarang ini kata globalisasi mulai dikenal oleh masyarakat. Theodore Levitte merupakan orang pertama kali yang menggunakan istilah Globalisasi pada tahun 1985, tetapi istilah tersebut belum terdefinisikan secara lengkap, kecuali sekedar definisi kerja (*working definition*) sehingga bergantung dari sisi mana orang melihatnya. Beberapa melihatnya sebagai suatu proses sosial, proses sejarah atau proses alamiah yang akan membawa seluruh bangsa dan negara di dunia semakin terhubung satu sama lain[1]. Globalisasi merupakan proses dimana hilangnya sekat-sekat atau batas-batas antar negara dalam setiap aspek kehidupan seperti ekonomi, politik, sosial, dan budaya. Maka dapat disimpulkan, globalisasi menjadikan orang melebur menjadi satu dengan semua orang di berbagai belahan dunia (Harahap, 2015).

Tentunya kata globalisasi tidak lepas dari yang namanya teknologi. Teknologi yang kita kenal saat ini telah berkembang begitu pesat dari zaman ke zaman. Salah satu kemajuan dari teknologi yaitu media sosial. Media sosial adalah sebuah platform dimana berbagai bentuk komunikasi dan berbagai jenis informasi di semua tingkatan dapat diinformasikan kepada masyarakat[2]. Media sosial juga memungkinkan penggunaanya dengan mudah berpartisipasi, berbagi, dan membuat konten seperti blog, jejaring sosial, wiki, dsb. Media sosial memberikan banyak fasilitas yang membuat penggunaanya betah berlama-lama berselancar di dunia maya. Pengguna media sosial dapat dengan bebas berkreaitivitas, berbisnis, berkomentar serta dapat berkomunikasi kapanpun dan dimanapun tanpa batasan lokasi dan waktu tanpa memandang batasan umur.

Berdasarkan sumber databoks.katadata, terdapat sebesar 16,1% laki-laki dan 14,2% perempuan berumur 18-24 tahun yang menggunakan media sosial di Indonesia. Kelompok usia tersebut menjadi pengguna dengan posisi ke dua di Indonesia. Dalam rentang umur 18-24 tahun rata-rata umur tersebut merupakan seorang mahasiswa. Artinya mahasiswa tidak lepas kaitannya dengan media sosial karena memang mahasiswa berada pada fase *emerging adulthood* yakni masa transisi dari masa remaja akhir menuju dewasa muda dan sedang mengalami dinamika psikologis (Kandel, 1998:5).

Pada tahun 2018 GlobalWebIndex mengeluarkan hasil laporan bahwasanya umur 16 – 24 tahun rata-rata menghabiskan tiga jam untuk bermain media sosial setiap harinya. Dalam jurnal “JAMA Psychiatry” juga menemukan bahwa remaja atau mahasiswa yang menggunakan media sosial lebih dari tiga jam per hari beresiko tinggi terhadap masalah kesehatan mental terutama masalah internalisasi alias jati diri. Hal ini tentunya menjadi urgensi dalam menghadapi permasalahan dampak media sosial terhadap kesehatan mental[5]

Walaupun demikian, pemanfaatan media sosial sebagai sarana penyebaran informasi, hiburan, bisnis, dsb, membawa banyak dampak hal positif dan negatif bagi kehidupan sehari-hari. Dampak positif jika dikaitkan dengan pendidikan di lingkungan perkuliahan, media sosial menjadi sumber informasi utama dalam mahasiswa untuk mencari apapun yang mereka butuhkan seperti webinar, artikel, berita, *bootcamp*, dsb. Dampak negatif dari media sosial juga semakin maraknya muncul berita hoax. Media sosial dipandang sebagai media yang mudah diakses sehingga penyebaran informasi melalui media sosial dapat dengan cepat menyebar dari satu pengguna ke pengguna lain. Maraknya penyebaran berita hoax bertujuan untuk membuat atau menggiring opini publik dan membentuk persepsi yang negatif ditambah mayoritas pengguna media sosial adalah para mahasiswa yang berada pada fase *emerging adulthood*.

Persepsi mahasiswa terhadap media sosial telah diteliti oleh penelitian Kosbi Alfi Sahid(2019) dengan sasaran penelitian pada remaja dengan judul “Persepsi dan Upaya Remaja Dalam Memanfaatkan Media Sosial Sebagai Sumber Informasi Keagamaan”. Penelitian ini menggunakan pendekatan kualitatif dengan metode deskriptif. Hasil penelitian ini menunjukkan (1) remaja mempersepsikan media sosial sebagai sumber informasi tercepat dan terbaru dalam memperoleh informasi, (2) terdapat perbedaan persepsi dalam mengakses jenis informasi dan preferensi terhadap jenis informasi keagamaan antara informan Fakultas Dakwah dengan Fakultas Umum. Dari hasil penelitian ini, persepsi mahasiswa terhadap media sosial bermakna positif.

Jika dikaji secara luas, persepsi mahasiswa terhadap media sosial juga dapat berdampak pada kesehatan mental penggunanya. Hal ini sejalan dengan hasil dari penelitian oleh Nada Bikriyah (2020) dengan judul “Pengaruh Media Sosial terhadap Kesehatan Mental Peserta Didik di SMPN 166 Jakarta”. Dari hasil penelitian tersebut dengan menggunakan metode kuantitatif (analisis regresi linear) menunjukkan bahwa

Pengaruh penggunaan media sosial terhadap kesehatan mental peserta didik di SMPN 166 Jakarta secara keseluruhan memberikan pengaruh pada tingkatan rendah/lemah yang diperoleh nilai koefisien determinasinya sebanyak 17,8%. Perolehan ini menunjukkan 82,2% kesehatan mental dipengaruhi oleh faktor lain selain dari media sosial instagram. Faktor lain yang menyebabkan kesehatan mental baik ataupun buruk, yaitu faktor internal peserta didik seperti kondisi bawaan peserta didik; faktor eksternal peserta didik seperti media sosial selain instagram, keluarga, orang tua, teman, maupun lainnya.

Dilihat dari latar belakang diatas, ada dorongan untuk melakukan penelitian lebih lanjut dengan sasaran penelitian yaitu mahasiswa administrasi publik terkait bagaimana persepsi mahasiswa mengenai dampak media sosial terhadap kesehatan mental menggunakan metode Analisis Faktor untuk mencari faktor – faktor mendorong persepsi mahasiswa terkait dampak media sosial terhadap kesehatan mental dan metode analisis data teks untuk melihat persepsi dari mahasiswa mengenai permasalahan tersebut. Maka dari itu, dibuatlah penelitian dengan judul “Persepsi Mahasiswa Administrasi Publik Mengenai Dampak Media Sosial Terhadap Kesehatan Mental dengan Metode Analisis Faktor dan Analisis Data Teks”.

1.2. Permasalahan

1. Apa saja informasi yang didapat dari hasil survey terkait dampak media sosial terhadap kesehatan mental Mahasiswa Administrasi Publik UPN “Veteran” Jawa Timur dengan metode Analisis Deskriptif?
2. Apa saja faktor-faktor yang mendorong persepsi mahasiswa administrasi publik terkait dampak media sosial terhadap kesehatan mental Mahasiswa Administrasi Publik UPN “Veteran” Jawa Timur dengan metode Analisis Faktor?
3. Bagaimana persepsi mahasiswa administrasi publik mengenai dampak media sosial terhadap kesehatan mental dengan metode Analisis Data Text?

1.3. Tujuan

4. Untuk mengetahui informasi yang didapat dari hasil survey terkait dampak media sosial terhadap kesehatan mental Mahasiswa Administrasi Publik UPN “Veteran” Jawa Timur dengan menggunakan metode Analisis Deskriptif.

1. Untuk mengetahui faktor-faktor yang mendorong persepsi mahasiswa mengenai dampak dari media sosial terhadap kesehatan mental Mahasiswa Administrasi Publik UPN “Veteran” Jawa Timur dengan menggunakan metode Analisis Faktor.
2. Untuk mengetahui proporsi nilai persepsi mahasiswa mengenai media sosial terhadap kesehatan mental secara positif dan negatif bagi mahasiswa Administrasi Publik UPN “Veteran” Jawa Timur dengan menggunakan metode Analisis Data Text.

1.4. Manfaat

1. Secara Akademis, penelitian ini diharapkan dapat memberikan kajian teori terkait faktor dan persepsi dampak media sosial terhadap kesehatan mental bagi Mahasiswa Administrasi Publik.
2. Secara Praktis, penelitian ini diharapkan dapat memberikan manfaat melalui hasil faktor dan persepsi dampak media sosial terhadap kesehatan mental bagi Mahasiswa Administrasi Publik.

BAB II: TINJAUAN PUSTAKA

2.1. Teori Penunjang

2.1.1 Sosial Media

berdasarkan jurnal penelitian berjudul “Peran Media Sosial Dalam Mempererat Interaksi Antar Keluarga Di Desa Esandom Kecamatan Tombatu Timur Kabupaten Minahasa Tenggara” yang ditulis oleh Tongkotow Liedfray pada tahun 2022, Media sosial adalah sebuah media online, dengan para penggunanya bisa dengan mudah berpartisipasi, berbagi, dan menciptakan isi meliputi blog, jejaring sosial, wiki, forum dan dunia virtual. Blog, jejaring sosial dan wiki merupakan bentuk media sosial yang paling umum digunakan oleh masyarakat di seluruh dunia. Pengertian media sosial menurut ahli, adalah seperangkat alat komunikasi dan kolaborasi baru yang memungkinkan terjadinya berbagai jenis interaksi yang sebelumnya tidak tersedia bagi orang awam (Chris Brogan, 2010).

Terdapat berbagai jenis sosial media yang dapat digunakan sesuai dengan kebutuhan kita, Beberapa jenis sosial media tersebut yaitu aplikasi media sosial berbagi video (Video Sharing) seperti media sosial *YouTube*, *Vimeo* dan *DailyMotion* lalu ada jenis sosial media mikroblog seperti *Twitter* dan *Tumblr*, Aplikasi media sosial berbagi jaringan sosial seperti *Facebook*, *Google Plus*, serta *Path*, Aplikasi berbagi jaringan profesional Sejumlah aplikasi jaringan profesional yang cukup populer di Indonesia antara lain *LinkedIn*, *Scribd* dan *Slideshare* dan aplikasi sosial media berbagi foto seperti *Pinterest*.

2.1.2 Kesehatan Mental

Secara singkat dapat dikatakan ilmu kesehatan mental adalah ilmu yang memperhatikan perawatan mental atau jiwa. Semium, (2006) mengatakan bahwa Ilmu kesehatan mental mempunyai objek khusus untuk diteliti dan objek tersebut adalah manusia. Ilmu kesehatan mental merupakan terjemahan dari istilah *mental hygiene*. *Mental* (dari kata Latin: *mens*, *mentis*) berarti jiwa, nyawa roh, sedangkan *hygiene* (dari kata Yunani: *hugine*) berarti ilmu tentang kesehatan, dapat diartikan bahwa ilmu kesehatan mental itu adalah ilmu yang membicarakan kehidupan mental manusia dengan memandang manusia sebagai totalitas psikofisik yang kompleks.

Artinya kesehatan mental adalah sebagai wujud karena adanya penyesuaian diri yang berhasil atau tidak adanya psikopatologi dan sebagai keadaan dimana seseorang

digambarkan tidak memiliki gangguan pada bidang psikologis, emosional, perilaku, dan sosial. 2 Sehingga dapat disimpulkan bahwa orang berada dalam keadaan sakit atau sehat psikisnya. Sehat jika tidak terdapat sedikitpun gangguan psikisnya dan jika ada gangguan psikis

2.1.3 Persepsi

Berdasarkan penelitian berjudul “Persepsi Mahasiswa Terhadap Pelayanan Perpustakaan Universitas Trunojoyo Madura”, Pengertian persepsi ialah salah satu aspek psikologis yang penting bagi manusia dalam merespon kehadiran berbagai aspek dan gejala di sekitarnya. Persepsi mengandung pengertian yang sangat luas, menyangkut intern dan ekstern. Menurut Kamus Besar Bahasa Indonesia, persepsi adalah tanggapan (penerimaan) langsung dari sesuatu. Menurut salah satu ahli, Bimo Walgito (2008: 70), persepsi merupakan suatu proses pengorganisasian, penginterpretasian terhadap stimulus yang diterima oleh organisme atau individu sehingga menjadi sesuatu yang berarti, dan merupakan aktivitas yang integrated dalam diri individu.

Persepsi terbagi menjadi dua, yaitu: persepsi terhadap objek (lingkungan fisik) dan persepsi terhadap manusia atau sosial. Persepsi terhadap manusia lebih sulit dan kompleks karena manusia bersifat dinamis. Kedua jenis persepsi ini memiliki perbedaan yaitu:

- a. Persepsi terhadap objek melalui lambang-lambang fisik, sedangkan terhadap manusia melalui lambang-lambang verbal dan nonverbal. Manusia lebih efektif daripada kebanyakan objek dan lebih sulit diramalkan.
- b. Persepsi terhadap objek menanggapi sifat-sifat luar, sedangkan terhadap manusia menanggapi sifat-sifat luar dan dalam (perasaan, motif, harapan, dan sebagainya).
- c. Objek tidak bereaksi, sedangkan manusia bereaksi. Dengan kata lain, objek bersifat statis, sedangkan manusia bersifat dinamis. Oleh karena itu, persepsi terhadap manusia dapat berubah dari waktu ke waktu, lebih cepat daripada persepsi terhadap objek.

2.1.4 Analisis Faktor

Berdasarkan penelitian berjudul “Analisis Faktor dan Penerapannya dalam Mengidentifikasi Faktor – Faktor yang Mempengaruhi Kepuasan Konsumen Terhadap Penjualan Media Pembelajaran” yang ditulis oleh Wiratmanto. Analisis faktor adalah

kajian tentang saling ketergantungan antara variabel – variabel, dengan tujuan untuk menemukan himpunan variabel-variabel baru, yang lebih sedikit jumlahnya dari pada variabel semula, dan yang menunjukkan yang mana di antara variabel-variabel semula itu yang merupakan faktor-faktor persekutuan (Suryanto, 1988:234).

Dalam tulisan Supranto, dikatakan bahwa analisis faktor digunakan untuk mereduksi data/variabel. Analisis faktor dipergunakan dalam kondisi sebagai berikut :

- Mengenali atau mengidentifikasi dimensi yang mendasari (underlying dimensions) atau faktor, yang menjelaskan korelasi antara suatu set variabel.
- Mengenali atau mengidentifikasi suatu set variabel baru yang tidak berkorelasi (independent) yang lebih sedikit jumlahnya untuk
- Menggantikan suatu set variabel asli yang saling berkorelasi di dalam analisis multivariat selanjutnya.
- Mengenali atau mengidentifikasi suatu set variabel yang penting dari suatu set variabel yang lebih banyak jumlahnya untuk dipergunakan dalam analisis multivariat selanjutnya.

Jika vector acak (random vector) $X = X_1, X_2, X_3, \dots X_p$ mempunyai vector rata-rata μ dan matriks ragam peragam Σ , secara linear bergantung pada sejumlah factor yang tidak teramati $F_1, F_2, F_3, \dots F_m$ yang disebut factor umum (*common factor*) dan $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots \varepsilon_p$ yang disebut factor khusus (*specific factors*). Maka model dari analisis factor adalah

$$X_1 - \mu_1 = l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + \varepsilon_1$$

$$X_2 - \mu_2 = l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + \varepsilon_2$$

$$X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \varepsilon_p$$

Jika dituliskan kedalam matriks, bentuknya akan menjadi

$$X_{p \times 1} - \mu = L_{p \times m} F_{m \times 1} + \varepsilon_p$$

Keterangan :

x = vektor variabel asal

μ = vector rata-rata variabel asal

L = matrik loading factor

F = vector factor Bersama

ε = vektor factor spesifik

Varians dari analisis factor sebagai berikut :

$$\sigma_{ii} = l_{i1}^2 + l_{i2}^2 + l_{i3}^2 \dots + l_{im}^2 + \psi_i = h_i^2 + \psi_i$$

Keterangan :

l_{ij} = loading factor

l_i^2 = Communalitiy ke-i

ψ_i = keragaman spesifik ke-i

Nilai loading menunjukkan korelasi antara factor umum yang terbentuk dengan variabel asal, semakin besar nilai loading maka semakin erat hubungan diantara keduanya. Hair (1998) dalam tulisannya menyatakan bahwa nilai minimal loading yang digunakan adalah lebih besar dari $\pm 0,30$. Loading $\pm 0,40$ dianggap penting, dan loading $\pm 0,50$ atau lebih besar dinyatakan signifikan.

Menurut artikel yang ditulis oleh Purwanto berjudul “Analisis Faktor: Konsep, Prosedur Uji Dan Interpretasi”. Penelitian tersebut menjelaskan analisis faktor adalah salah satu analisis multivariat yang dirancang untuk meneliti sifat hubungan antara variabel-variabel dalam satu perangkat tertentu yang pada dasarnya menunjukkan pola hubungan tertentu. Sebelum analisis dilakukan beberapa asumsi harus terpenuhi, sehingga secara keseluruhan langkah-langkah uji analisis faktor adalah:

1. Menguji kelayakan analisis

Kriteria untuk mengetahui apakah suatu data dapat dianalisis faktor ditentukan oleh dua hal yaitu harga koefisien Kaiser-Meyer-Olkin (KMO) sampling adequacy dan Bartlett's sphericity test. Uji KMO sampling adequacy diperlukan untuk melihat kecukupan sampel yang dianalisis dan uji Bartlett untuk melihat normalitas data yang akan dianalisis.

2. Menyajikan matriks korelasi

Matriks korelasi menyajikan interkorelasi antarbutir. Matriks diperlukan untuk mengetahui butir-butir yang saling berkorelasi tinggi dan rendah. Butir yang saling berkorelasi tinggi berarti mengukur dimensi yang sama dan sebaliknya.

3. Melakukan ekstraksi

Proses ekstraksi dilakukan untuk mendapatkan lebih sedikit 162 faktor (eigenvalues factor) dari sejumlah banyak butir / variabel dan sumbangan faktor

terhadap keseluruhan butir (total variance explained). Terdapat beberapa metode untuk melakukan ekstraksi.

4. Melakukan rotasi

Rotasi adalah proses memutar sumbu mendekati koordinat titiktitik butir / variabel. Proses ekstraksi hanya menentukan jumlah faktor yang meringkas keseluruhan butir, namun belum menentukan distribusi butir-butir ke dalam faktor-faktor yang meringkasnya. Rotasi dapat dilakukan dengan satu dari beberapa metode : varimax, quartimax, equamax dan oblimin

5. Memberikan penamaan faktor.

Langkah terakhir adalah menamai (labeling) faktor yang terbentuk dari proses ekstraksi dan rotasi. Nama diberikan berdasarkan kesamaan ciri butir yang menjadi muatan faktor.

6. Interpretasi Hasil Uji

Interpretasi mengacu kepada hasil analisis dari prosedur uji AF, yaitu : (1) KaiserMeyer-Olkin (KMO) sampling adequacy, (2) uji normalitas Bartlett (Bartlett's test of sphericity), (3) matriks korelasi, (4) eigenvalues, (5) communalities, (6) total variance explained, dan (7) factor loadings.

2.1.5 Analisis Data Text

Analisis Data Text adalah proses menganalisis potongan data yang tidak tersruktur untuk menemukan informasi dan insight yang sebelumnya belum ditemukan yang dapat dimanfaatkan untuk pengambilan keputusan dan proses lainnya[4].

2.1.6 Analisis Deskriptif

Analisis deskriptif merupakan suatu metode analisis statistik yang bertujuan untuk memberikan deskripsi atau gambaran mengenai subjek penelitian berdasarkan data variabel yang diperoleh dari kelompok subjek tertentu. Analisis deskriptif dapat ditampilkan dalam bentuk tabel distribusi frekuensi, tabel histogram, nilai mean (rata – rata), nilai standar deviasi dan lainnya. Manfaat yang diperoleh dari penggunaan analisis

deskriptif adalah mendapatkan gambaran lengkap dari data baik dalam bentuk verbal atau numerik yang berhubungan dengan data yang kita teliti. Pengertian analisis deskriptif menurut Sugiyono (2017:147) salah satu ahli yaitu analisis deskriptif adalah statistik yang digunakan untuk menganalisis data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi.

2.2. Penelitian Terkait

Penelitian dilakukan tentu tidak lepas dari penelitian terkait. Penelitian terkait atau terdahulu bertujuan untuk mendapatkan bahan kajian, perbandingan dan acuan. Selain itu, untuk menghindari anggapan kesamaan dengan penelitian ini. Maka dalam tinjauan pustaka ini peneliti mencantumkan hasil-hasil penelitian terdahulu sebagai berikut:

2.2.1 Hasil penelitian Kosbi Alfi Sahid (2019)

Penelitian Kosbi Alfi Sahid (2019), berjudul *“Persepsi dan Upaya Remaja Dalam Memanfaatkan Media Sosial Sebagai Sumber Informasi Keagamaan”*. Penelitian ini menggunakan pendekatan kualitatif dengan metode deskriptif. Tujuan dari penelitian ini adalah: (1) menganalisis persepsi dan upaya remaja dalam memanfaatkan media sosial sebagai sumber informasi keagamaan, (2) mengetahui perbedaan persepsi dan upaya remaja dalam memanfaatkan media sosial sebagai sumber informasi keagamaan.

Berdasarkan hasil penelitian dapat disimpulkan bahwa (1) remaja mempersepsikan media sosial sebagai sumber informasi tercepat dan terbaru dalam memperoleh informasi, (2) terdapat perbedaan persepsi dalam mengakses jenis informasi dan preferensi terhadap jenis informasi keagamaan antara informan Fakultas Dakwah dengan Fakultas Umum.

2.2.2 Hasil Penelitian Nada Bikriyah (2020)

Penelitian ini berjudul *“Pengaruh Media Sosial terhadap Kesehatan Mental Peserta Didik di SMPN 166 Jakarta”*. Tujuan dari penelitian ini adalah untuk mendeskripsikan pengaruh media sosial terhadap kesehatan mental di SMPN 166 Jakarta. Penelitian ini menggunakan metode kuantitatif dengan teknik analisis regresi linear.

Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa kesehatan mental peserta didik dapat dipengaruhi oleh penggunaan media sosial *Instagram*, hal tersebut dapat dibuktikan dengan hasil uji t yakni $t_{hitung} > t_{tabel}$ sebesar $2,945 > 2,021$ dengan persamaan $\hat{y} = 76,178 + 0,231X$ serta hasil koefisien determinasi sebesar 17,8%.

2.2.3 Hasil Penelitian Dwi Sinta Rizki Ariani, Miftakhul Ulfa, Waifti Amalia(2022)

Penelitian ini berjudul “*Hubungan Persepsi Mahasiswa dengan Stress Akademik dalam Pembelajaran Online di Masa Pandemi Covid-19*”. Tujuan dari penelitian ini adalah untuk mengetahui hubungan persepsi mahasiswa dengan stres akademik dalam penerapan pembelajaran *online* di masa Pandemi Covid-19 pada mahasiswa semester II STIKes Widyagama Husada Malang. Penelitian ini menggunakan metode *deskriptif korelasi* dengan pendekatan *retrospektif*.

Berdasarkan hasil penelitian tersebut terdapat hubungan persepsi mahasiswa dengan stres akademik dalam penerapan pembelajaran *online* di masa Pandemi Covid-19 pada mahasiswa semester II STIKes Widyagama Husada Malang.

Tabel 1. Penelitian Terkait

Nama Peneliti	Judul	Metode Penelitian	Hasil Penelitian	Perbedaan
Kosbi Alfi Sahid (2019)	Persepsi dan Upaya Remaja Dalam Memanfaatkan Media Sosial Sebagai Sumber Informasi Keagamaan	Deskriptif Kualitatif	Berdasarkan hasil penelitian dapat disimpulkan bahwa (1) remaja mempersepsikan media sosial sebagai sumber informasi tercepat dan terbaru dalam memperoleh informasi, (2) terdapat perbedaan persepsi dalam mengakses jenis informasi dan preferensi terhadap jenis informasi keagamaan antara	-Fokus permasalahan berbeda (Media Sosial sebagai sumber informasi keagamaan dengan dampak media sosial terhadap kesehatan mental) -Metode penelitian berbeda (deskriptif kualitatif dengan analisis data text)

			informan Fakultas Dakwah dengan Fakultas Umum.	
Nada Bikriyah (2020)	Pengaruh Media Sosial terhadap Kesehatan Mental Peserta Didik di SMPN 166 Jakarta	Kuantitatif (Analisis Regresi Linear)	Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa kesehatan mental peserta didik dapat dipengaruhi oleh penggunaan media sosial <i>Instagram</i> , hal tersebut dapat dibuktikan dengan hasil uji t yakni $t_{hitung} > t_{tabel}$ sebesar $2,945 > 2,021$ dengan persamaan $\hat{y} = 76,178 + 0,231X$ serta hasil koefisien determinasi sebesar 17,8%.	-Sasaran target berbeda (Peserta Didik SMPN 166 Jakarta dengan Mahasiswa Administrasi Publik) -Metode penelitian berbeda (Analisis regresi Linear dengan Analisis Faktor)

Dwi Sinta Rizki Ariani, Miftakhul Ulfa, Waifti Amalia(2022)	Hubungan Persepsi Mahasiswa dengan Stress Akademik dalam Pembelajaran Online di Masa Pandemi Covid-19	Deskriptif Korelasi dengan pendekatan Retrospektif.	Terdapat hubungan persepsi mahasiswa dengan stres akademik dalam penerapan pembelajaran <i>online</i> di masa Pandemi Covid-19 pada mahasiswa semester II STIKes Widyagama Husada Malang.	-Fokus permasalahan yang terperinci (tentang stress akademik) -Metode penelitian berbeda (Deskriptif Korelasi pendekatan Retrospektif dengan Analisis Faktor)
---	---	---	---	--

BAB III : METODOLOGI PENELITIAN

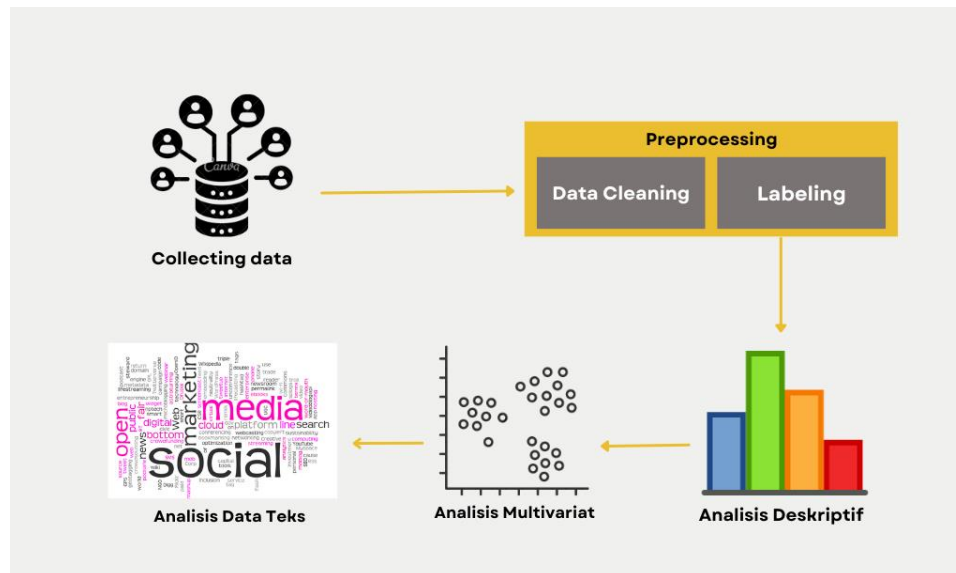
3.1 Jenis Penelitian

Jenis penelitian ini merupakan penelitian analitik multivariat dengan mengaplikasikan pendekatan kuantitatif guna mencari faktor-faktor yang mendorong persepsi mahasiswa administrasi publik mengenai dampak media sosial terhadap kesehatan mental.

3.2 Sumber Data Penelitian

Penelitian ini menggunakan sumber data primer. Sumber data primer dikumpulkan melalui kuesioner *google form* dengan ukuran sampel sebanyak 66 mahasiswa/i program studi Administrasi Publik UPN “Veteran” Jawa Timur.

3.3 Flowchart



Hal yang pertama dilakukan yaitu collecting data. Untuk mengcollect data ini kita harus mengetahui metode apa yang ingin dipakai saat melakukan collecting data dan mengidentifikasi sumber data. Sesuai dengan judul yang ingin kami teliti ini yaitu “Persepsi Mahasiswa Administrasi Publik Mengenai Dampak Media Sosial Terhadap Kesehatan Mental dengan Metode Analisis Faktor dan Analisis Data Teks” maka sumber data kami berasal dari mahasiswa administrasi publik UPN Veteran Jawa Timur yang dilakukan dengan menggunakan pengisian kuisisioner.

Data yang sudah didapatkan selanjutnya dilakukan preprocessing data yaitu data cleaning dan labeling. Pada proses data cleaning kita membersihkan data dari noise dan menyiapkan data untuk analisis lebih lanjut dengan melakukan identifikasi dan menangani missing value, deteksi outlier, normalisasi data dan mengeliminasi data yang tidak valid.

Setelah dilakukan preprocessing data selanjutnya dilakukan analisis deskriptif yaitu menganalisis karakteristik data berupa menghitung statistik deskriptif seperti mean, median, modus, dan visualisasi data menggunakan grafik seperti histogram atau diagram batang. Lalu dilakukan analisis multivariat pada dataset.

Analisis multivariat yang digunakan adalah analisis faktor. Analisis faktor digunakan untuk mengelompokkan variabel-variabel kedalam kelompok lebih kecil agar memudahkan proses menganalisis dan melihat apa saja variabel yang berpengaruh. Setelah itu kita melakukan analisis data teks. Analisis data teks digunakan untuk menganalisis kompleks antar variabel dalam data. Pada analisis data teks ini kita dapat membuat word cloud dari dataset yang dibedakan antara negatif dan positif.

3.4 Uji Validitas dan Reliabilitas

Uji validitas berfungsi menakumulasi data yang akan dihitung dan menentukan ketepatan alat yang digunakan. Pengujian validitas dalam studi ini memanfaatkan skor korelasi elemen total yang disesuaikan (r), tetapi dianggap valid jika nilai korelasi elemen total yang disesuaikan $> r$ (0.2387) $=5\%$ dan $df=n-2$ dan nilai elemen yang dikoreksi dinyatakan tidak valid jika total korelasi (r) $< r$ tabel (Hidayat, 2021)

Selama uji reliabilitas, metode cronbach digunakan untuk menentukan konsistensi internal antar variabel instrumen. Dengan kata lain, reliabilitas pemeriksaan akan meningkat, terlepas dari apakah teknik yang digunakan dalam penelitian ini dapat diterima dan relevan. Jika nilai alfa cronbach mendekati 1, maka skala yang digunakan sangat efektif, bahkan ketika diberikan kepada peserta yang memiliki banyak perbedaan. Selain itu, jika nilai alpha elemen kurang dari 0,6, maka elemen tersebut tidak dapat dipercaya atau tidak baik (Hidayat, 2021).

Pada penelitian yang telah dilakukan peneliti untuk mengenali validitas dan realibilitas jadi mediasi akumulasi pengukuran karena itu hasil 65 koresponden setelah duji didapatkan hasil sebagai berikut:

Tabel 3. 1 Hasil Pengujian Validitas dan Realibilitas pada data

Item	Korelasi	Hubungan dengan $r_{tabel} = 0.2387$	Validitas Item
q1	0.334645	>0.2387	Valid
q2	0.363803	>0.2387	Valid
q3	0.408606	>0.2387	Valid
q4	0.189798	<0.2387	Tidak Valid
q5	0.516921	>0.2387	Valid
q6	0.503751	>0.2387	Valid
q7	0.395956	>0.2387	Valid
q8	0.344679	>0.2387	Valid
q9	0.296065	>0.2387	Valid
q10	0.383046	>0.2387	Valid
q11	0.335032	>0.2387	Valid
q12	0.614389	>0.2387	Valid
q13	0.573260	>0.2387	Valid
q14	0.676583	>0.2387	Valid
q15	0.553309	>0.2387	Valid
q16	0.633159	>0.2387	Valid
q17	0.611341	>0.2387	Valid
q18	0.615052	>0.2387	Valid
q19	0.227302	<0.2387	Tidak Valid
q20	0.284775	>0.2387	Valid
q21	0.510668	>0.2387	Valid
Nilai Alpha	0.812137107741128		Reliable

Berdasarkan tabel 3.1 diatas ditemukan jika nilai Corrected item_Total correlation (r) tidak keseluruhan item dinyatakan valid karena pada q4 dan q19 memiliki nilai kurang dari rtabel 0.2387 yang artinya q4 dan q19 tidak mampu mengukur apa yang ingin diukur dan item lainnya mampu mengukur apa yang ingin diukur. Nilai alpha cronbach dinyatakan reliable karena didapatkan nilai sebesar 0.812137107741128 melebihi 0.6 dan mendekati 1, artinya media pengukuran yang digunakan mempunyai konsistensi apabila penakumulasi nilai dilakukan secara berulang.

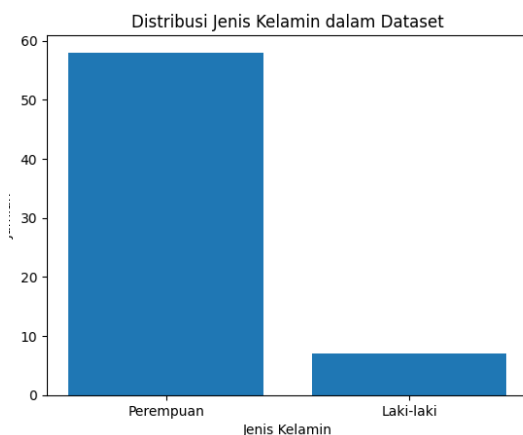
BAB IV : HASIL ANALISIS

4.1 Analisis Deskriptif

Daftar Responden yang dipilih oleh kelompok kami berada pada angka 66 orang. Pemilihan dari responden tersebut berdasarkan penyebaran Google Form yang telah kami sebar kepada mahasiswa Administrasi Publik. Gambar plot dibawah ini merupakan bentuk visualisasi atau penjelasan subjek yang kelompok kami teliti.

4.1.1 Gambaran Responden Yang Mengacu Pada Jenis Kelamin

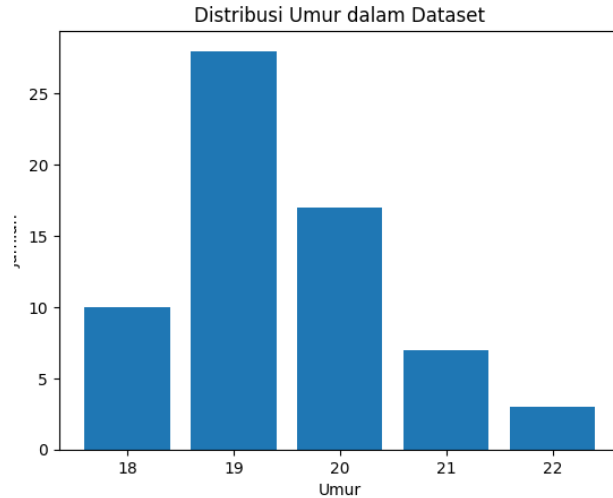
Dalam proses penelitian kami terdapat kriteria yang diambil dari gender 2 jenis, yaitu laki-laki dan perempuan. Maka didapatkan hasil sebagai berikut :



Bisa kita lihat dari diagram diatas berdasarkan hasil penyebaran google form yang telah kita lakukan didapatkan hasil 58 responden berjenis kelamin perempuan dan sedikitnya 7 responden berjenis kelamin laki-laki. Dan kita bisa simpulkan bahwa responden kita lebih banyak perempuan daripada laki-laki.

4.1.1 Gambaran Responden Berdasarkan Usia

Selanjutnya dalam proses penelitian kita juga membagi kelompok berdasarkan usia, berikut adalah bentuk diagram barplot:

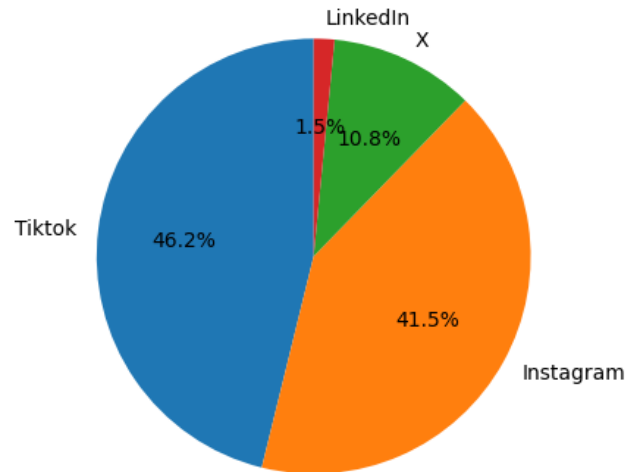


Berdasarkan barplot diatas dapat diklasifikasikan bahwa responden yang berusia 22 tahun sebesar 4,5% atau sebanyak 3 orang responden, responden yang berusia 21 tahun sebesar 11,9% atau sebanyak 8 orang responden, untuk responden yang berumur 20 tahun sebanyak 25,4% atau sebanyak 17 orang responden, responden yang berusia 19 tahun sebanyak 41,8% atau sebanyak 28 orang responden, dan responden yang berumur 18 tahun sebanyak 16,4% atau sebanyak 11 orang responden.

4.1.2 Gambaran Penggunaan Media Sosial dalam Dataset

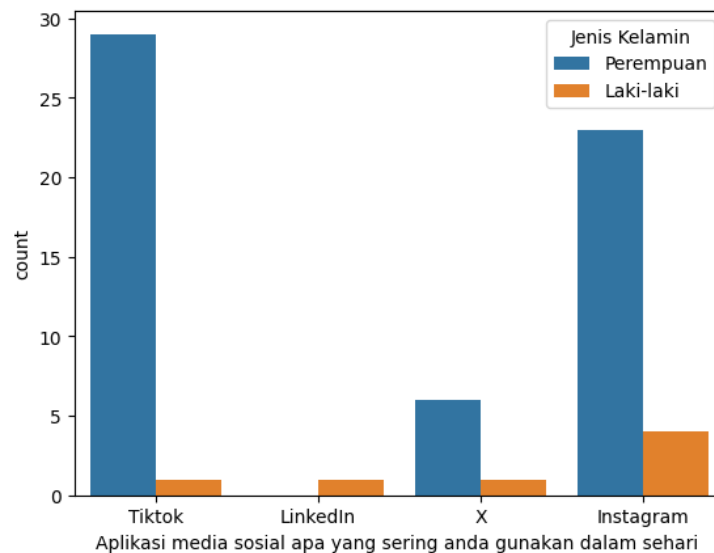
Selain jenis kelamin dan umur dalam proses penelitian kelompok kami juga menggunakan kriteria penggunaan media sosial, yaitu sebagai berikut :

Distribusi Media Sosial dalam Dataset



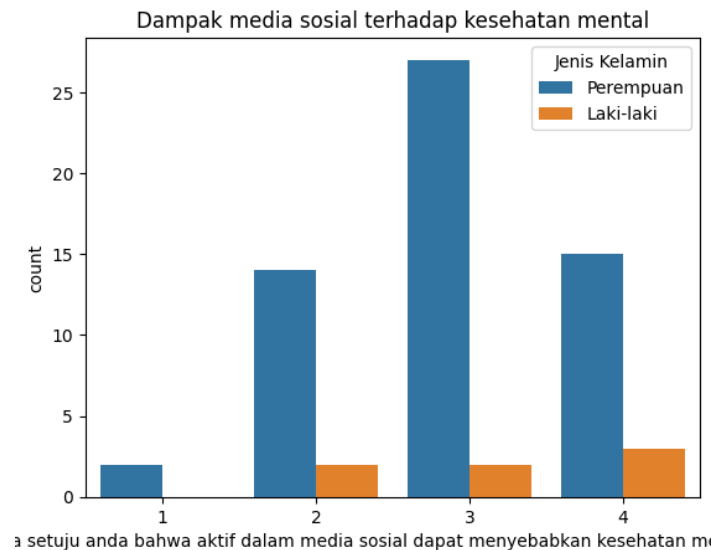
Berdasarkan piechart yang ada di atas bisa kita lihat banyak persen dari beberapa sosial media yang kita gunakan dalam proses penelitian kami. Dengan penggunaan TikTok yang paling besar dengan 46,2% disusul oleh Instagram dengan 41,5% dan yang paling sedikit digunakan oleh mahasiswa Administrasi Publik adalah LinkedIn dengan total presentase 1,5%.

4.1.3 Gambaran Perbandingan Penggunaan Aplikasi Sosial Media Terhadap Jenis Kelamin



Dapat dilihat dari barplot diatas perbandingan penggunaan antara laki laki dan perempuan terhadap penggunaan medai sosial.

4.1.4 Gambaran Seberapa Setuju Bahwa Aktif dalam Bermedia Sosial Berpengaruh Terhadap Kesehatan Mental



Dilihat dari barplot diatas dari skala 1-4 dimana 1 sangat tidak setuju dan 4 sangat setuju. Banyak responden yang memilih setuju dan sedikit mahasiswa memilih sangat tidak setuju untuk dampak media sosial terhadap kesehatan mental.

4.2 Analisis Faktor

4.2.1 Data Preparation

Analisis faktor adalah hasil dari pengolahan data dari responden. Berikut langkah-langkah yang dilakukan:

1. Melakukan import beberapa library yang akan digunakan.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import StandardScaler
from factor_analyzer import FactorAnalyzer
from factor_analyzer.factor_analyzer import calculate_kmo
import numpy as np
```

Mounted at /content/drive

2. Memanggil data set.

[illegible]

3. Memilih Variabel-Variabel yang akan di Analisis

Drop variabel yang tidak diperlukan


```

dataset.drop('Timestamp', axis=1, inplace=True)
dataset.drop('Email', axis=1, inplace=True)
dataset.drop('Nama Lengkap', axis=1, inplace=True)
dataset.drop('Umur ', axis=1, inplace=True)
dataset.drop('Jenis Kelamin', axis=1, inplace=True)
dataset.drop('Jurusan', axis=1, inplace=True)
dataset.drop('NPM', axis=1, inplace=True)
dataset.drop('Aplikasi media sosial apa yang sering anda gunakan dalam sehari ', axis=1, inplace=True)
dataset.drop('Dimana anda biasa menggunakan sosial media', axis=1, inplace=True)
dataset.drop('Dengan apa anda mengakses sosial media', axis=1, inplace=True)
dataset.drop('Apa pendapat anda mengenai kebijakan pemerintah yang telah menutup Tiktok Shop ? Setuju atau tidak setuju mohon diberikan alasannya', axis=1, inplace=True)
dataset.drop('Apa kamu setuju bahwa lingkungan dalam media sosial yang sering anda gunakan saat ini sangat toxic dan dapat mengganggu kesehatan mental anda ? setuju atau tidak setuju mohon diberikan alasannya!', axis=1, inplace=True)
dataset.drop('Informasi atau pengetahuan terkait bagaimana yang anda dapatkan ?', axis=1, inplace=True)
dataset.drop('Apakah dengan informasi terkait jurusan yang anda dapatkan (section 2), dapat merubah gaya hidup (produktivitas) anda ?', axis=1, inplace=True)
dataset.drop('Dengan informasi terkait jurusan yang anda dapatkan (section 2), Apakah Anda pernah merasa tertekan atau stres karena perbandingan dengan kehidupan orang lain di media sosial?', axis=1, inplace=True)
dataset.drop('Dengan informasi terkait jurusan yang anda dapatkan (section 2), Apakah dapat memotivasi diri anda untuk meningkatkan potensi diri dengan berbagai kegiatan positif seperti webinar, seminar dll ?', axis=1, inplace=True)
dataset.drop('Total', axis=1, inplace=True)
dataset.drop('Total', axis=1, inplace=True)
dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena ada hadiah tambahan ', axis=1, inplace=True)
dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena review bagus influencer', axis=1, inplace=True)

```

```

dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena ada diskon yang jauh lebih besar daripada di offline shop', axis=1, inplace=True)
dataset.drop('Seberapa sering menghabiskan sebagian uang untuk membeli barang yang sebenarnya tidak dibutuhkan', axis=1, inplace=True)
dataset.drop('Seberapa sering anda sebelum Tiktok Shop ditutup, anda membeli barang di media sosial tersebut', axis=1, inplace=True)
dataset.drop('Seberapa sering anda mengakses online shop pada media sosial', axis=1, inplace=True)
dataset.drop('Seberapa sering suatu online shop pada media sosial yang membuat anda tertarik untuk membeli produk yang ditawarkan', axis=1, inplace=True)

```

Mendrop variabel yang tidak lulus uji validitas

```
dataset.drop('Seberapa sering anda mendapatkan informasi atau pengetahuan terkait jurusan anda di media sosial'ingat pertanyaan ini', axis=1, inplace=True)
```

```
dataset.drop('Seberapa setuju anda bahwa kita sebaiknya istirahat ber media sosial setidaknya 1 hari dalam satu bulan', axis=1, inplace=True)
```

Drop variabel yang tidak diperlukan

```
dataset.drop('timestamp', axis=1, inplace=True)
dataset.drop('Email', axis=1, inplace=True)
dataset.drop('Nama Lengkap', axis=1, inplace=True)
dataset.drop('Umur', axis=1, inplace=True)
dataset.drop('Jenis Kelamin', axis=1, inplace=True)
dataset.drop('Jurusan', axis=1, inplace=True)
dataset.drop('NPM', axis=1, inplace=True)
dataset.drop('Aplikasi media sosial apa yang sering anda gunakan dalam sehari', axis=1, inplace=True)
dataset.drop('Dinamika anda bisa menggunakan media sosial', axis=1, inplace=True)
dataset.drop('Dengan apa anda mengakses media sosial', axis=1, inplace=True)
dataset.drop('Apa pendapat anda mengenai kebijakan pemerintah yang telah membatasi TikTok Shop? Setuju atau tidak setuju mohon diberikan alasannya', axis=1, inplace=True)
dataset.drop('Apa kamu setuju bahwa lingkungan dalam media sosial yang sering anda gunakan saat ini sangat toxic dan dapat mengganggu kesehatan mental anda? Setuju atau tidak setuju mohon diberikan alasannya', axis=1, inplace=True)
dataset.drop('Informasi atau pengetahuan terkait bagaimana yang anda dapatkan?', axis=1, inplace=True)
dataset.drop('Apakah dengan informasi terkait jurusan yang anda dapatkan (section 2), dapat merubah gaya hidup (produktivitas) anda?', axis=1, inplace=True)
dataset.drop('Dengan informasi terkait jurusan yang anda dapatkan (section 2), Apakah Anda pernah merasa tertekan atau stres karena perbandingan dengan kehidupan orang lain di media sosial?', axis=1, inplace=True)
dataset.drop('Apakah dengan informasi terkait jurusan yang anda dapatkan (section 2), Apakah dapat memotivasi diri anda untuk meningkatkan potensi diri dengan berbagai kegiatan positif seperti webinar, seminar dll?', axis=1, inplace=True)
dataset.drop('Total', axis=1, inplace=True)
dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena ada hadiah tambahan', axis=1, inplace=True)
dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena review bagus influencer', axis=1, inplace=True)
dataset.drop('Seberapa sering membeli produk yang ditawarkan di media sosial karena ada diskon yang jauh lebih besar daripada di offline shop', axis=1, inplace=True)
dataset.drop('Seberapa sering menghabiskan sebagian uang untuk membeli barang yang sebenarnya tidak dibutuhkan', axis=1, inplace=True)
dataset.drop('Seberapa sering anda sebelum TikTok Shop ditutup, anda membeli barang di media sosial tersebut', axis=1, inplace=True)
dataset.drop('Seberapa sering anda mengikuti online shop pada media sosial', axis=1, inplace=True)
dataset.drop('Seberapa sering suatu online shop pada media sosial yang membuat anda tertarik untuk membeli produk yang ditawarkan', axis=1, inplace=True)
```

Mendrop variabel yang tidak lulus uji validitas

```
[ ] dataset.drop('Seberapa sering anda mendapatkan informasi atau pengetahuan terkait jurusan anda di media sosial'ingat pertanyaan ini', axis=1, inplace=True)
dataset.drop('Seberapa setuju anda bahwa kita sebaiknya istirahat ber media sosial setidaknya 1 hari dalam satu bulan', axis=1, inplace=True)
```

Melakukan drop variabel yang tidak digunakan & tidak lulus uji validitas yang totalnya berjumlah 25

4. Merename setiap kolom atau variabel.

```
dataset.columns = ["q1", "q2", "q3", "q4", "q5", "q6", "q7", "q8", "q9", "q10", "q11", "q12"]
dataset.head()
```

	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12
0	4	4	4	4	4	4	4	4	4	4	4	4
1	2	2	3	3	3	3	4	3	3	4	3	3
2	2	2	2	3	3	3	3	3	3	3	3	2
3	1	1	1	4	4	4	4	4	4	4	3	2
4	2	1	1	3	4	4	4	4	4	4	3	4

Setiap kolom diberi nama q1 hingga q12

5. Memeriksa apakah variabel-variabel yang akan dikelompokkan ada missing value atau tidak.

```
dataset.isnull().sum()
```

```

q1      0
q2      0
q3      0
q4      0
q5      0
q6      0
q7      0
q8      0
q9      0
q10     0
q11     0
q12     0
dtype: int64

```

Kode ini menggunakan fungsi `isnull()` dari modul `pandas` untuk memeriksa apakah suatu nilai terdapat missing value. Bisa dilihat pada hasil di atas, tidak terdapat missing value pada data ini, maka kita bisa melanjutkan analisisnya.

6. Melakukan pengecekan multikolinearitas menggunakan metode `corr.` pada library `pandas`

```
dataset.corr()[(dataset.corr().>0.9) & (dataset.corr().<1)]
```

```

q1  q2  q3  q4  q5  q6  q7  q8  q9  q10  q11  q12
q1  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q2  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q3  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q4  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q5  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q6  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q7  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q8  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q9  NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q10 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q11 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN
q12 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN

```

Bagian pertama (`data.corr() > 0.9`) akan menghasilkan matriks boolean yang berisi `True` jika korelasi antara dua variabel lebih besar dari 0.9, dan `False` jika tidak. Bagian kedua (`data.corr().<1`) akan menghasilkan matriks boolean yang berisi `True` jika korelasi antara dua variabel kurang dari 1, dan `False` jika tidak. Karena tidak terdapat multikolinearitas maka analisis dapat dilanjutkan.

7. Melakukan pengujian KMO

```
kmo_all, kmo_model = calculate_kmo(dataset)
kmo_model
```

```
0.8497601022982899
```

Hasil dari KMO menunjukkan 0.84 yang berarti variabel memiliki variansi yang cukup untuk dilakukan analisis. Jika variansi < 0.6 maka data dianggap kurang bervariasi dan tidak akan memberikan hasil yang berguna.

8. Mengimplementasikan Analisis Faktor

```
fa = FactorAnalyzer(n_factors = 3, rotation="varimax")
fa.fit(dataset)
```

```
FactorAnalyzer
FactorAnalyzer(rotation='varimax', rotation_kwargs={})
```

9. Melakukan pengecekan nilai loadings

```
loadings_output = pd.DataFrame(fa.loadings_,
                               index=dataset.columns)
loadings_output
```

	0	1	2
q1	-0.080738	0.486857	0.208562
q2	-0.000265	0.956330	-0.265285
q3	0.040420	0.736587	-0.004742
q4	0.759149	0.167113	-0.003288
q5	0.884846	-0.073723	0.113512
q6	0.830979	-0.058837	0.070173
q7	0.764940	-0.139896	0.304814
q8	0.809187	-0.049262	0.404632
q9	0.711062	0.022078	0.431182
q10	0.900999	-0.155864	-0.046741
q11	0.153126	-0.001730	0.518048
q12	0.368637	0.087992	0.182276

4.2.2 Menentukan jumlah faktor

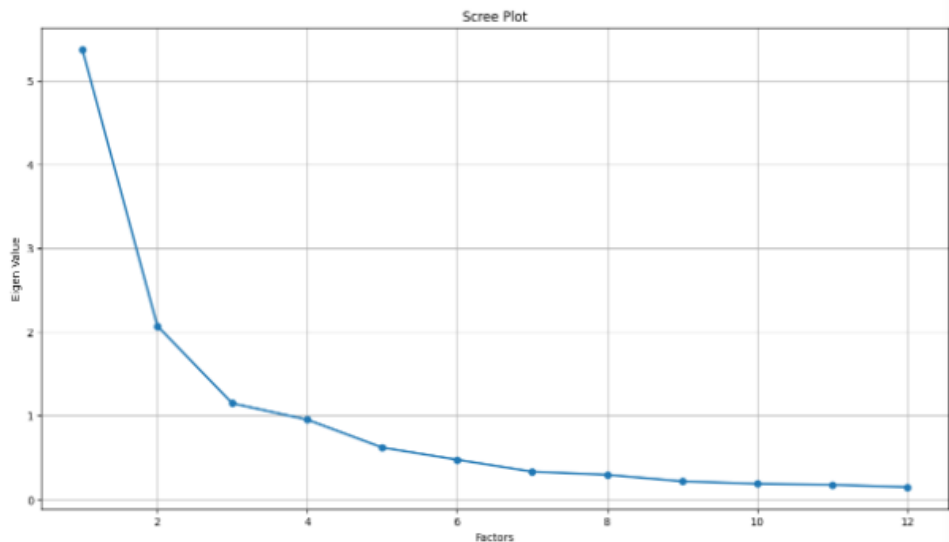
1. Menentukan jumlah faktor

```
fa = FactorAnalyzer(rotation = 'varimax',n_factors=dataset.shape[1])
fa.fit(dataset)
ev,_=fa.get_eigenvalues()

factor_values=np.arange(dataset.shape[1])+1

plt.figure(figsize=(15,8))

plt.plot(factor_values,ev,'o-',linewidth=2)
plt.title("Scree Plot")
plt.xlabel("Factors")
plt.ylabel("Eigen Value")
plt.grid()
```



FactorAnalyzer digunakan untuk menerapkan analisis faktor. rotation='varimax' menunjukkan bahwa rotasi faktor yang digunakan adalah Varimax, suatu metode rotasi yang dapat membantu mempermudah interpretasi faktor. get_eigenvalues() digunakan untuk mendapatkan nilai eigen (eigenvalues) dari faktor-faktor yang dihasilkan oleh analisis faktor. ev akan berisi nilai eigen untuk masing-masing faktor. Lalu dilakukan visualisasi dengan scree plot. Berdasarkan scree plot diatas, n factornya adalah 4, Maka dilakukan lagi penerapan analisis faktor ke kumpulan data menggunakan kelas FactorAnalyzer di factor_analyzer library dengan N_factor = 4

2. Maka dilakukan lagi penerapan analisis faktor ke kumpulan data menggunakan kelas FactorAnalyzer di factor_analyzer library dengan N_factor = 4

```
fa = FactorAnalyzer(n_factors = 4, rotation="varimax")
fa.fit(dataset)
```

```
FactorAnalyzer
FactorAnalyzer(n_factors=4, rotation='varimax', rotation_kwags={})
```

3. Dilakukan pengecekan loadings kembali

```
loadings_output = pd.DataFrame(fa.loadings_,
                               index=dataset.columns)
loadings_output
```

	0	1	2	3
q1	-0.057708	0.520443	-0.163098	0.257630
q2	-0.046087	0.951595	0.113527	-0.264760
q3	0.016651	0.722926	0.073537	-0.011252
q4	0.751277	0.186701	0.079931	-0.019911
q5	0.892596	-0.053225	0.069098	0.079581
q6	0.820426	-0.043392	0.127561	0.043712
q7	0.777635	-0.129731	0.079034	0.260194
q8	0.821981	-0.040048	0.089850	0.356573
q9	0.744265	0.041987	-0.013507	0.420109
q10	0.875016	-0.149422	0.230371	-0.095956
q11	0.142916	-0.021819	0.167976	0.500875
q12	0.241409	0.060378	0.936314	0.240207

Setelah itu Menganalisis kembali faktor yang terbentuk dengan langkah-langkah seperti dibawah berikut :

1. Filter loadings dengan ambang batas tertentu

```
loadings_output.where(abs(loadings_output)>0.5)
```

	0	1	2	3
q1	NaN	0.520443	NaN	NaN
q2	NaN	0.951595	NaN	NaN
q3	NaN	0.722926	NaN	NaN
q4	0.751277	NaN	NaN	NaN
q5	0.892596	NaN	NaN	NaN
q6	0.820426	NaN	NaN	NaN
q7	0.777635	NaN	NaN	NaN
q8	0.821981	NaN	NaN	NaN
q9	0.744265	NaN	NaN	NaN
q10	0.875016	NaN	NaN	NaN
q11	NaN	NaN	NaN	0.500875
q12	NaN	NaN	0.936314	NaN

Dari Hasil diatas menunjukan nilai loadings yang dapat melihat variabel-variabel yang mempunyai pengaruh paling signifikan terhadap faktor-faktor tersebut. Hasil diatas hanya memuat variabel dengan nilai loadings lebih besar dari 0,5.

2. Menghitung Komunalitas

```
pd.DataFrame(fa.get_communalities(),
             index=dataset.columns,
             columns=['Communalities'])
```

	Communalities
q1	0.367165
q2	0.990643
q3	0.528434
q4	0.606060
q5	0.810668
q6	0.693163
q7	0.695495
q8	0.812474
q9	0.732367
q10	0.850259
q11	0.299992
q12	0.996308

Berdasarkan nilai komunalitas didapatkan q1 dan q11 yang memiliki nilai dibawah 0.5 sehingga data tersebut harus dihapus dari data lalu dilakukan pemeriksaan kembali.

```
dataset.drop('q1', axis=1, inplace=True)
dataset.drop('q11', axis=1, inplace=True)
dataset.head()
```

	q2	q3	q4	q5	q6	q7	q8	q9	q10	q12
0	4	4	4	4	4	4	4	4	4	4
1	2	3	3	3	3	4	3	3	4	3
2	2	2	3	3	3	3	3	3	3	2
3	1	1	4	4	4	4	4	4	4	2
4	1	1	3	4	4	4	4	4	4	4

3. Pengecekan kembali multikolinier menggunakan metode corr pada pandas setelah menghapus q1 dan q11

```
dataset.corr()[(dataset.corr())>0.9] & (dataset.corr())<1]
```

	q2	q3	q4	q5	q6	q7	q8	q9	q10	q12
q2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
q12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Bagian pertama ($\text{data.corr()} > 0.9$) akan menghasilkan matriks boolean yang berisi True jika korelasi antara dua variabel lebih besar dari 0.9, dan False jika tidak. Bagian kedua ($\text{data.corr()} < 1$) akan menghasilkan matriks boolean yang berisi True jika korelasi antara dua variabel kurang dari 1, dan False jika tidak. Karena tidak terdapat multikolinearitas maka analisis dilanjutkan.

4. Melakukan pengujian KMO

```
kmo_all,kmo_model=calculate_kmo(dataset)
```



```
kmo_model
```

```
0.8629063347105498
```

Hasil dari nilai KMO yang baru menunjukkan 0.86 yang berarti variabel memiliki variansi yang cukup untuk dilakukan analisis.

5. Mengimplementasikan Analisis Faktor

```
fa = FactorAnalyzer(n_factors = 4, rotation="varimax")  
fa.fit(dataset)
```

```
* FactorAnalyzer  
FactorAnalyzer(n_factors=4, rotation='varimax', rotation_kwargs={})
```

6. Melakukan pengecekan loadings

```
loadings_output = pd.DataFrame(fa.loadings_,  
                                index=dataset.columns)  
loadings_output
```

	0	1	2	3
q2	0.086164	-0.203883	0.842510	0.001635
q3	-0.065915	0.104267	0.878990	0.044027
q4	0.734513	0.296300	0.161818	0.183276
q5	0.696738	0.492035	-0.068771	0.279493
q6	0.710314	0.364626	-0.075640	0.316033
q7	0.316225	0.755895	-0.108295	0.361315
q8	0.418575	0.737175	-0.050252	0.336930
q9	0.455967	0.703975	-0.006882	0.106407
q10	0.579825	0.376701	-0.112845	0.628736
q12	0.195370	0.197621	0.116618	0.343376

Konsep Loading mengungkapkan hubungan antar variabel asli dan faktor yang mendasarinya. Nilai Loadingnya berkisar antara -1 sampai 1. Nilai yang mendekati -1 atau 1 menunjukkan bahwa faktor tersebut mempunyai pengaruh yang signifikan terhadap variabel.

7. Menentukan Jumlah faktor

```
fa = FactorAnalyzer(rotation = 'varimax', n_factors=dataset.shape[1])  
fa.fit(dataset)  
ev,_=fa.get_eigenvalues()
```

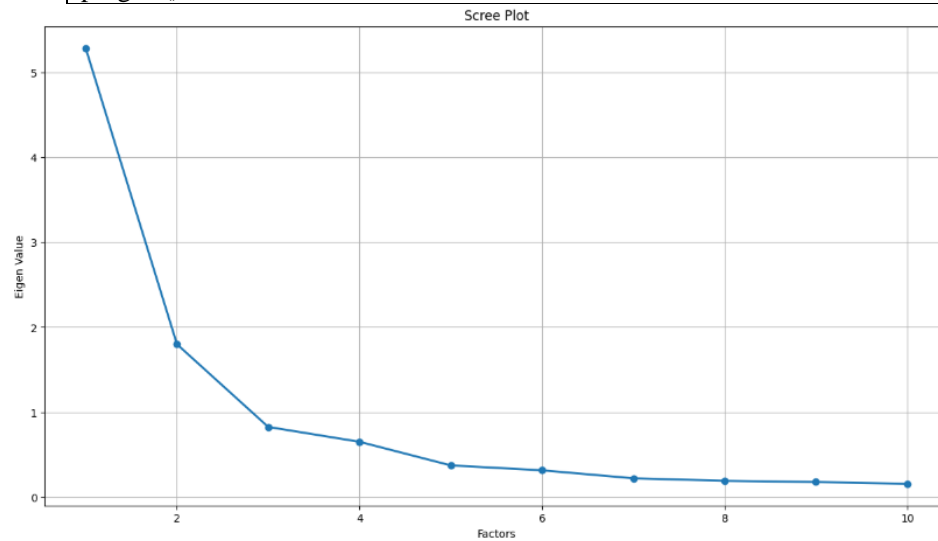
```

factor_values=np.arange(dataset.shape[1])+1

plt.figure(figsize=(15,8))

plt.plot(factor_values,ev,'o-',linewidth=2)
plt.title("Scree Plot")
plt.xlabel("Factors")
plt.ylabel("Eigen Value")
plt.grid()

```



Berdasarkan scree plot diatas, n factornya adalah 4 yang berarti analisis faktor sebelumnya sudah benar dengan $n_factor = 4$

8. Menganalisis faktor yang terbentuk.

a. Nilai loading

```

loadings_output.where(abs(loadings_output)>0.5)

```

	0	1	2	3
q2	NaN	NaN	0.84251	NaN
q3	NaN	NaN	0.87899	NaN
q4	0.734513	NaN	NaN	NaN
q5	0.696738	NaN	NaN	NaN
q6	0.710314	NaN	NaN	NaN
q7	NaN	0.755895	NaN	NaN
q8	NaN	0.737175	NaN	NaN
q9	NaN	0.703975	NaN	NaN
q10	0.579825	NaN	NaN	0.628736
q12	NaN	NaN	NaN	NaN

Dari Hasil diatas menunjukan nilai loadings yang dapat melihat variabel-variabel yang mempunyai pengaruh paling signifikan terhadap faktor-faktor tersebut. Hasil diatas hanya memuat variabel dengan nilai loadings lebih besar dari 0,5.

b. Nilai Komunalitas

```
pd.DataFrame(fa.get_communalities(),
             index=dataset.columns,
             columns=['Communalities'])
```

	Communalities
q2	0.825097
q3	0.803162
q4	0.727375
q5	0.868146
q6	0.832725
q7	0.872532
q8	0.890183
q9	0.765461
q10	0.862485
q12	0.275662

4.2.3 Interpretasi Hasil Analisis Faktor

Berdasarkan nilai loading terakhir, dapat mengumpulkan variabel-variabel dalam satu faktor yang terbentuk. Berikut adalah daftarnya:

Faktor 1 : Perundungan di Media Sosial

- Q4:Seberapa setuju anda bahwa sekarang ini banyak terjadi perundungan di media sosial:
- Q5:Seberapa setuju anda bahwa perundungan di media sosial dapat menyebabkan seseorang yang dibuli depresi atau bahkan bunuh diri:
- Q6:Seberapa setuju anda bahwa perundungan di media sosial kebanyakan karena body shamming:
- Q10:Seberapa setuju anda bahwa banyak kasus pelecehan di media sosial

Faktor 2: Penyebab Perundungan di Media Sosial

- Q7:Seberapa setuju anda bahwa perundungan di media sosial kebanyakan karena meyebarkan kebohongan atau aib orang lain:
- Q8: Seberapa setuju anda bahwa perundungan di media sosial kebanyakan karena saling ejek:
- Q9:Seberapa setuju anda bahwa perundungan di media sosial kebanyakan karena adanya hasutan untuk menebar kebencian terhadap orang lain:

Faktor 3: Aktivitas Posting di Media Sosial

- Q2: Seberapa sering anda update status di media sosial:
- Q3:Seberapa sering anda upload tentang keseharian di media sosial

Faktor 4: Pandangan terhadap Kasus Pelecehan di Media Sosial

- Q10:Seberapa setuju anda bahwa banyak kasus pelecehan di media sosial

4.3 Analisis Data Teks

4.3.1 Preparation Data

1. Import Library

```
pip install contractions
import pandas as pd
import string
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
import contractions
```

Langkah awal yang perlu kita lakukan serta perhatikan yakni mengimport library yang digunakan dalam pemrosesan data teks.

2. Load Dataset

Selanjutnya yakni dengan memanggil dataset hasil kuisioner dengan menggunakan fungsi `pd.read` untuk membaca data serta memanggil variabel data untuk mengetahui gambaran data frame.

```
from google.colab import drive
drive.mount('/content/drive')
dataset = pd.read_excel("drive/MyDrive/Colab Notebooks/KUISSIONER Responses.xlsx")
dataset
```

```
dataset = pd.read_excel("KUISSIONER Kel 2 (Responses).xlsx")
dataset
```

	Nama Lengkap	Lingkungan Sosial Media Toxic	Label1	Insecure (Tertekan/Stress)	Label2
0	Tiffani Valensia Hendrik	Setuju, karena menurut saya banyak orang di so...	P	Iya, karena terkadang saya dapat terhasut deng...	P
1	Noor abika diva ainura	setuju , karena banyaknya orang orang yang men...	P	tidak , karena saya tidak terlalu perduli deng...	N
2	EMIL NAHDIAH	tidak setuju karena media sosial pribadi saya ...	N	tidak karena kemampuan orang berbeda2 jadi tid...	N
3	Veronica Aditya Rahayu Ningtyas		N	tidak pernah	N
4	Dino Dwi	Tidak setuju, karna rata2 teman media sosial s...	N	jarang dan bahkan tidak pernah	N
...
61	Khovifah Nur	Saya setuju karena banyak sekali kasus akibat ...	P	Tidak pernah, karena semua pencapaian ada pors...	N
62	Diana putri wahyudi	Setuju karena dengan kita bermain media sosial...	P	Tidak saya tidak merasa stres karena kehidupan...	N
63	Yossyefa Nabila Choirunnisa Putri	Saya pribadi setuju, karena dalam media sosial...	P	Terikat saya pribadi terkadang merasa stres j...	P
64	Nanda Guritno Nurhaliza	Tidak setuju, penggunaan media sosial bergantu...	N	Tidak pernah, setiap orang memiliki jalan hidu...	N
65	Maulana Yogi Aditya Dwi Putra	Sangat mengganggu karena ada arus informasi ya...	P	Iya karena melihat pencapaian orang lain dapat...	P

66 rows x 5 columns

Proses selanjutnya yakni menghapus variabel yang tidak diperlukan dalam proses analisis data teks dengan fungsi `.drop()` hingga menyisakan 2 variabel yang akan diproses analisis data teks mengenai persepsi mahasiswa terhadap lingkungan sosial media yang toxic dengan memberikan sebuah label pembeda persepsi mahasiswa positif dan persepsi mahasiswa negatif.

```
dataset.drop('Insecure (Tertekan/Stress)', axis=1, inplace=True)
dataset.drop('Labeling kedua', axis=1, inplace=True)
dataset.drop('Nama Lengkap', axis=1, inplace=True)
```

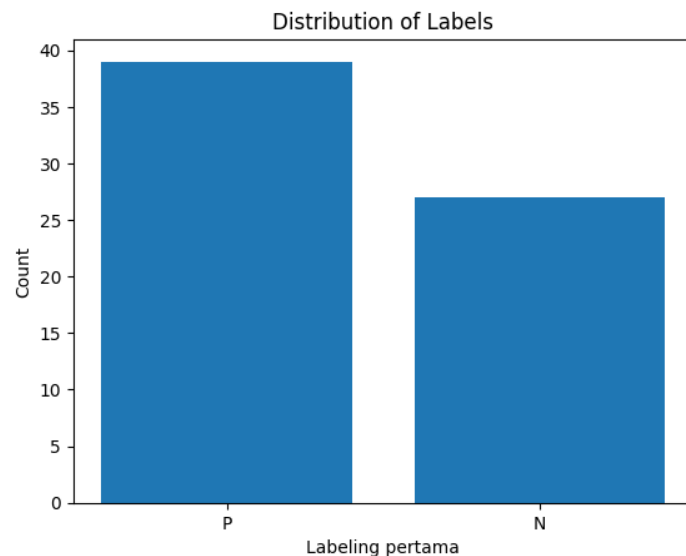
Kemudian disajikan tabel distribusi data persepsi mahasiswa sesuai dengan label yang telah diberikan, dapat diketahui bersama jumlah persepsi mahasiswa positif (pro) lebih banyak ketimbang dengan persepsi mahasiswa negatif (kontra)

```
import matplotlib.pyplot as plt

# Menghitung jumlah setiap label
label_counts = dataset['Labeling pertama'].value_counts()

# Membuat barchart
plt.bar(label_counts.index, label_counts.values)
plt.xlabel('Labeling pertama')
```

```
plt.ylabel('Count')
plt.title('Distribution of Labels')
plt.show()
```



3. Expand Contractions

Proses dalam tahap ini merupakan memperluas kata gabungan atau singkatan dalam data ulasan menggunakan fungsi lambda, bersama dengan metode perbaikan dalam library. Proses ini dibentuk dengan menghilangkan beberapa huruf dan menggantinya dengan apostrof. Kemudian mengkonversi nilai dalam kolom contraction yang diperluas dari list menjadi string

```
dataset['no_contractions'] = dataset['Lingkungan Sosial Media Toxic'].apply(lambda x:
[contractions.fix(word) for word in x.split()])
dataset.head(7)
```

	Lingkungan Sosial Media Toxic	Labeling pertama	no_contractions
0	Setuju, karena menurut saya banyak orang di so...	P	[Setuju,, karena, menurut, saya, banyak, orang...
1	setuju , karena banyaknya orang orang yang men...	P	[setuju, ,, karena, banyaknya, orang, orang, y...
2	tidak setuju karena media sosial pribadi saya ...	N	[tidak, setuju, karena, media, sosial, pribadi...
3	tidak setuju	N	[tidak, setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	[Tidak, setuju,, karna, rata2, teman, media, s...
5	Tidak setuju,karena dengan adanya medsos dunia...	N	[Tidak, setuju,karena, dengan, adanya, medsos...
6	Tidak. Karena lingkungan di media sosial saya ...	N	[Tidak., Karena, lingkungan, di, media, sosial...

```
dataset['reviews_no_contractions'] = [' '.join(l for l in dataset['no_contractions'])]
dataset.head(7)
```

	Lingkungan Sosial Media Toxic	Labeling pertama	no_contractions	reviews_no_contractions
0	Setuju, karena menurut saya banyak orang di so...	P	[Setuju,, karena, menurut, saya, banyak, orang...	Setuju, karena menurut saya banyak orang di so...
1	setuju , karena banyaknya orang orang yang men...	P	[setuju, ,, karena, banyaknya, orang, orang, y...	setuju , karena banyaknya orang orang yang men...
2	tidak setuju karena media sosial pribadi saya ...	N	[tidak, setuju, karena, media, sosial, pribadi...	tidak setuju karena media sosial pribadi saya ...
3	tidak setuju	N	[tidak, setuju]	tidak setuju
4	Tidak setuju, karna rata2 teman media sosial s...	N	[Tidak, setuju,, karna, rata2, teman, media, s...	Tidak setuju, karna rata2 teman media sosial s...
5	Tidak setuju,karena dengan adanya medsos dunia...	N	[Tidak, setuju,karena, dengan, adanya, medsos,...	Tidak setuju,karena dengan adanya medsos dunia...
6	Tidak. Karena lingkungan di media sosial saya ...	N	[Tidak., Karena, lingkungan, di, media, sosial...	Tidak. Karena lingkungan di media sosial saya ...

4. Performing Tokenization

Tokenisasi data menggunakan fungsi `word_tokenize` di `nlTK`. Tokenisasi melibatkan pemecahan teks menjadi potongan-potongan yang lebih kecil, seperti seperti kata atau frasa. Memecah teks menjadi unit-unit yang lebih kecil membantu kita menganalisisnya dengan lebih efektif.

```
import nltk
nltk.download('punkt')
dataset['reviews_tokenized'] = dataset['reviews_no_contractions'].apply(word_tokenize)
dataset.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	no_contractions	reviews_no_contractions	reviews_tokenized
0	Setuju, karena menurut saya banyak orang di so...	P	[Setuju,, karena, menurut, saya, banyak, orang...	Setuju, karena menurut saya banyak orang di so...	[Setuju, ,, karena, menurut, saya, banyak, ora...
1	setuju , karena banyaknya orang orang yang men...	P	[setuju, ,, karena, banyaknya, orang, orang, y...	setuju , karena banyaknya orang orang yang men...	[setuju, ,, karena, banyaknya, orang, orang, y...
2	tidak setuju karena media sosial pribadi saya ...	N	[tidak, setuju, karena, media, sosial, pribadi...	tidak setuju karena media sosial pribadi saya ...	[tidak, setuju, karena, media, sosial, pribadi...
3	tidak setuju	N	[tidak, setuju]	tidak setuju	[tidak, setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	[Tidak, setuju,, karna, rata2, teman, media, s...	Tidak setuju, karna rata2 teman media sosial s...	[Tidak, setuju, ,, karna, rata2, teman, media,...

5. Converting to Lowercase

Ubah data ulasan menjadi huruf kecil menggunakan metode `lower` di Python. Hal ini dilakukan karena data teks biasanya berupa kombinasi huruf besar dan huruf kecil. Namun, hal ini perlu distandarisasi untuk memudahkan analisis. Oleh karena itu, semua huruf perlu dikonversi ke bentuk huruf kecil sebelum dianalisis.

```
dataset['reviews_lower'] = dataset['reviews_tokenized'].apply(lambda x: [word.lower() for word in x])
dataset.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	no_contractions	reviews_no_contractions	reviews_tokenized	reviews_lower
0	Setuju, karena menurut saya banyak orang di so...	P	[Setuju,, karena, menurut, saya, banyak, orang...	Setuju, karena menurut saya banyak orang di so...	[Setuju,, karena, menurut, saya, banyak, ora...	[setuju,, karena, menurut, saya, banyak, ora...
1	setuju , karena banyaknya orang orang yang men...	P	[setuju,, , karena, banyaknya, orang, orang, y...	setuju , karena banyaknya orang orang yang men...	[setuju,, , karena, banyaknya, orang, orang, y...	[setuju,, , karena, banyaknya, orang, orang, y...
2	tidak setuju karena media sosial pribadi saya ...	N	[tidak, setuju, karena, media, sosial, pribadi...	tidak setuju karena media sosial pribadi saya ...	[tidak, setuju, karena, media, sosial, pribadi...	[tidak, setuju, karena, media, sosial, pribadi...
3	tidak setuju	N	[tidak, setuju]	tidak setuju	[tidak, setuju]	[tidak, setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	[Tidak, setuju,, karna, rata2, teman, media, s...	Tidak setuju, karna rata2 teman media sosial s...	[Tidak, setuju,, karna, rata2, teman, media,...	[tidak, setuju,, karna, rata2, teman, media,...

6. Remove Punctuations

Menghapus tanda baca dari data ulasan menggunakan atribut tanda baca di dalam library string. Tanda baca berguna untuk memisahkan kalimat, klausa, atau frasa, namun, sebagian besar tidak diperlukan untuk analisis teks karena tidak menyampaikan makna yang signifikan. Setelah tanda baca dihapus, selanjutnya perlu untuk mengkonversi nilai dalam kolom no_punctuations yang diperluas dari list menjadi string.

```
punctuations = string.punctuation
dataset['reviews_no_punctuation'] = dataset['reviews_lower'].apply(lambda x: [word for word in x if word not in punctuations])
dataset.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	no_contractions	reviews_no_contractions	reviews_tokenized	reviews_lower	reviews_no_punctuation
0	Setuju, karena menurut saya banyak orang di so...	P	[Setuju,, karena, menurut, saya, banyak, orang...	Setuju, karena menurut saya banyak orang di so...	[Setuju,, karena, menurut, saya, banyak, ora...	[setuju, karena, menurut, saya, banyak, orang...	[setuju, karena, menurut, saya, banyak, orang...
1	setuju , karena banyaknya orang orang yang men...	P	[setuju,, , karena, banyaknya, orang, orang, y...	setuju , karena banyaknya orang orang yang men...	[setuju,, , karena, banyaknya, orang, orang, y...	[setuju, karena, banyaknya, orang, orang, y...	[setuju, karena, banyaknya, orang, orang, y...
2	tidak setuju karena media sosial pribadi saya ...	N	[tidak, setuju, karena, media, sosial, pribadi...	tidak setuju karena media sosial pribadi saya ...	[tidak, setuju, karena, media, sosial, pribadi...	[tidak, setuju, karena, media, sosial, pribadi...	[tidak, setuju, karena, media, sosial, pribadi...
3	tidak setuju	N	[tidak, setuju]	tidak setuju	[tidak, setuju]	[tidak, setuju]	[tidak, setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	[Tidak, setuju,, karna, rata2, teman, media, s...	Tidak setuju, karna rata2 teman media sosial s...	[Tidak, setuju,, karna, rata2, teman, media,...	[tidak, setuju,, karna, rata2, teman, media,...	[tidak, setuju, karna, rata2, teman, media, so...

7. Export Cleaned Data text

Dalam proses ini, data yang sudah dibersihkan dan disiapkan, diekspor menjadi dataset untuk dilanjutkan dalam proses selanjutnya.

```
dataset[['Lingkungan Sosial Media Toxic','Labeling pertama','reviews_cleaned']].to_csv('cleaned_data.csv',index = False)
```

4.3.2 Removing Stop Words

Stop words adalah kata yang sangat sering muncul dalam suatu bahasa. Kata-kata ini umumnya tidak menambah arti yang signifikan makna pada teks. Beberapa kata henti

yang umum termasuk kata ganti, kata depan, kata sambung, dan konjungsi. Dalam bahasa Indonesia, contoh kata henti meliputi di, dengan, karena, hingga, sampai, pada, dsb. Daftar ini dapat bervariasi berdasarkan bahasa dan konteksnya.

Sebelum menganalisis teks, kita harus menghapus kata henti agar kita dapat fokus pada kata-kata yang lebih relevan yang lebih relevan di dalam teks. Kata henti biasanya tidak memiliki informasi yang signifikan dan dapat menyebabkan *noise* dalam kumpulan data kita. Oleh karena itu, menghapusnya akan membantu kita menemukan wawasan dengan mudah dan fokus pada hal yang paling relevan.

Langkah-langkah untuk melakukan stopwords adalah:

1. Import Libraries

Mengimpor pustaka yang relevan yang dibutuhkan dalam proses menghilangkan stopwords

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import string
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
```

2. Load Dataset

```
dataset1 = pd.read_csv("cleaned_data.csv")
dataset1
```

Memuat data ulasan dari file csv yang telah dipersiapkan melalui proses sebelumnya. Kemudian memeriksa lima baris pertama dengan menggunakan metode head() dan mengecek jumlah kolom dan baris menggunakan metode shape

Lingkungan Sosial Media Toxic Labeling pertama			reviews_cleaned
0	Setuju, karena menurut saya banyak orang di so...	P	setuju karena menurut saya banyak orang di sos...
1	setuju , karena banyaknya orang orang yang men...	P	setuju karena banyaknya orang orang yang mengg...
2	tidak setuju karena media sosial pribadi saya ...	N	tidak setuju karena media sosial pribadi saya ...
3	tidak setuju	N	tidak setuju
4	Tidak setuju, karna rata2 teman media sosial s...	N	tidak setuju karna rata2 teman media sosial sa...
...
61	Saya setuju karena banyak sekali kasus akibat ...	P	saya setuju karena banyak sekali kasus akibat ...
62	Setuju karena dengan kita bermain media sosial...	P	setuju karena dengan kita bermain media sosial...
63	Saya pribadi setuju, karena dalam media sosial...	P	saya pribadi setuju karena dalam media sosial ...
64	Tidak setuju, penggunaan media sosial bergantu...	N	tidak setuju penggunaan media sosial bergantun...
65	Sangat mengganggu karena ada arus informasi ya...	P	sangat mengganggu karena ada arus informasi ya...

66 rows × 3 columns

3. Performing Tokenization

Proses ini menggunakan fungsi `word_tokenize` dalam modul `nlTK`, fungsi `word_tokenize` akan membagi string menjadi daftar substring yang berisi kata-kata dan tanda baca.

```
dataset1['reviews_tokenized'] = dataset1['reviews_cleaned'].apply(word_tokenize)
```

4. Performing Top Word

Dalam menunjukkan top word atau kata yang paling sering muncul dalam ulasan, kita perlu membuat fungsi kustom `combine_words` yang menggabungkan semua kata ke dalam list serta membuat fungsi kustom `count_topwords` yang menghitung frekuensi setiap kata dan mengembalikan 20 kata teratas berdasarkan frekuensi

```
def combine_words(word_list):
    all_words = []
    for word in word_list:
        all_words += word
    return all_words

def count_topwords(all_words):
    counts = dict()
    for word in all_words:
        if word in counts:
            counts[word] += 1
        else:
            counts[word] = 1
```

```

word_count = pd.DataFrame([counts])
word_count_transposed = word_count.T.reset_index()
word_count_transposed.columns = ['words','word_count']
word_count_sorted = word_count_transposed.sort_values("word_count",ascending
= False)
word_count_sorted
return word_count_sorted[:20]

```

Setelah membuat fungsi sebelumnya, langkah selanjutnya yakni menerapkan fungsi kustom `combine_words` ke data ulasan yang diberi token untuk mendapatkan list semua kata. Proses ini dilakukan untuk memvalidasi apakah semua ulasan sudah menjadi satu dalam suatu list

```

reviews = dataset1['reviews_tokenized']
reviews_words = combine_words(reviews)
reviews_words[:10]

```

```

['setuju',
 'karena',
 'menurut',
 'saya',
 'banyak',
 'orang',
 'di',
 'sosial',
 'media',
 'yang']

```

Setelah data ulasan menjadi satu dalam list, terapkan fungsi custom `count_topwords` untuk membuat data frame dan melihat Top Words yang ada dalam kuisioner persepsi mahasiswa terhadap dampak media sosial. Beberapa hasil Top Words berupa stopwords yang dapat dan tidak dapat memberikan suatu insight dalam proses analisis teks ini, contohnya seperti kata setuju, sosial, dan media merupakan kata yang dapat memberikan insight. Sementara itu, kata-kata karena, yang, di merupakan kata yang tidak dapat memberikan insight dalam data sehingga stopwords tersebut per dihapuskan dengan proses selanjutnya.

5. Removing Stopwords

Menghapus stopwords dari data ulasan dapat menggunakan metode stopwords dalam modul `nlTK`. Dengan menjalankan perintah code dibawah ini stopwords akan dihapus dan disimpan dalam kolom tersendiri.

```

stop_words = set(stopwords.words('indonesian'))

```

```
dataset1['reviews_no_stopwords'] = dataset1['reviews_tokenized'].apply(lambda x: [word
for word in x if word not in stop_words])
dataset1.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	reviews_cleaned	reviews_tokenized	reviews_no_stopwords
0	Setuju, karena menurut saya banyak orang di so...	P	setuju karena menurut saya banyak orang di sos...	[setuju, karena, menurut, saya, banyak, orang, ...]	[setuju, orang, sosial, media, gampang, terhas...
1	setuju, karena banyaknya orang orang yang men...	P	setuju karena banyaknya orang orang yang mengg...	[setuju, karena, banyaknya, orang, orang, yang, ...]	[setuju, banyaknya, orang, orang, media, sosia...
2	tidak setuju karena media sosial pribadi saya ...	N	tidak setuju karena media sosial pribadi saya ...	[tidak, setuju, karena, media, sosial, pribadi, ...]	[setuju, media, sosial, pribadi, arah, hal2, p...
3	tidak setuju	N	tidak setuju	[tidak, setuju]	[setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	tidak setuju karna rata2 teman media sosial sa...	[tidak, setuju, karna, rata2, teman, media, so...	[setuju, karna, rata2, teman, media, sosial, t...

Data ulasan yang telah bersih dari stopwords perlu diubah dari format list menjadi string untuk memudahkan proses selanjutnya.

```
dataset1['reviews_cleaned_stopwords'] = [' '.join(l) for l in
dataset1['reviews_no_stopwords']]
dataset1.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	reviews_cleaned	reviews_tokenized	reviews_no_stopwords	reviews_cleaned_stopwords
0	Setuju, karena menurut saya banyak orang di so...	P	setuju karena menurut saya banyak orang di sos...	[setuju, karena, menurut, saya, banyak, orang, ...]	[setuju, orang, sosial, media, gampang, terhas...	setuju orang sosial media gampang terhasut omo...
1	setuju, karena banyaknya orang orang yang men...	P	setuju karena banyaknya orang orang yang mengg...	[setuju, karena, banyaknya, orang, orang, yang, ...]	[setuju, banyaknya, orang, orang, media, sosia...	setuju banyaknya orang orang media sosial meny...
2	tidak setuju karena media sosial pribadi saya ...	N	tidak setuju karena media sosial pribadi saya ...	[tidak, setuju, karena, media, sosial, pribadi, ...]	[setuju, media, sosial, pribadi, arah, hal2, p...	setuju media sosial pribadi arah hal2 postif ...
3	tidak setuju	N	tidak setuju	[tidak, setuju]	[setuju]	setuju
4	Tidak setuju, karna rata2 teman media sosial s...	N	tidak setuju karna rata2 teman media sosial sa...	[tidak, setuju, karna, rata2, teman, media, so...	[setuju, karna, rata2, teman, media, sosial, t...	setuju karna rata2 teman media sosial teman

Langkah selanjutnya yakni dengan menerapkan fungsi kustom combine_words yang telah dibuat sebelumnya dalam data ulasan yang diberi token untuk mendapatkan list semua kata. Kemudian data dapat diterapkan fungsi kustom count_topwords untuk mengetahui Top Words setelah stopwords dihapus.

```
reviews_no_stopwords = dataset1['reviews_no_stopwords']
reviews_words = combine_words(reviews_no_stopwords)
reviews_words[:10]
```

```
['setuju',
 'orang',
 'sosial',
 'media',
 'gampang',
 'terhasut',
 'omongan',
 'orang',
 'menyinggung',
 'setuju']
```

6. Mengekspor cleaned data yang sudah dihilangkan stop wordsnya.

```
dataset1[['Lingkungan Sosial Media Toxic','Labeling
pertama','reviews_cleaned','reviews_cleaned_stopwords']].to_csv("cleaned_no_stopwords.csv",ind
ex = False)
```

4.3.3 Performing Stemming

Ketika menganalisis data teks, kita biasanya perlu mengurangi kata-kata ke bentuk dasar atau akarnya. Proses ini disebut stemming. Stemming diperlukan karena kata-kata dapat muncul dalam beberapa variasi tergantung pada konteksnya. Stemming memastikan

kata-kata tersebut direduksi menjadi bentuk yang umum. Hal ini membantu meningkatkan akurasi analisis kami karena beberapa variasi dari kata yang sama dapat menyebabkan noise dalam kumpulan data kami.

Langkah - langkah stemming :

1. Mengimport library yang relevan.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import wordnet
from nltk import pos_tag
from nltk.stem import PorterStemmer
from nltk.stem import WordNetLemmatizer
from nltk.corpus import stopwords
```

2. Load Dataset

Langkah awal dalam proses ini yakni dengan memasukkan data cleaned_no_stopwords.csv ke dataframe menggunakan read_csv.

```
dataset2 = pd.read_csv("cleaned_no_stopwords.csv")
dataset2
```

3. Tokenisasi data menggunakan fungsi word_tokenize di library nltk. Fungsi word_tokenize akan membagi string menjadi daftar substring yang berisi kata-kata dan tanda baca.

```
dataset2['reviews_tokenized'] =
dataset2['reviews_cleaned_stopwords'].apply(word_tokenize)
```

4. Lakukan stemming pada data ulasan menggunakan kelas PorterStemmer di nltk.

```
stemmer = nltk.PorterStemmer()
dataset2['reviews_stemmed_data'] = dataset2['reviews_tokenized'].apply(lambda x:
[stemmer.stem(word) for word in x])
dataset2.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	reviews_cleaned	reviews_cleaned_stopwords	reviews_tokenized	reviews_stemmed_data
0	Setuju, karena menurut saya banyak orang di so...	P	setuju karena menurut saya banyak orang di sos...	setuju orang sosial media gampang terhasut ome...	[setuju, orang, sosial, media, gampang, terhas...	[setuju, orang, sosial, media, gampang, terhas...
1	setuju, karena banyaknya orang orang yang men...	P	setuju karena banyaknya orang orang yang mengg...	setuju banyaknya orang orang media sosial meny...	[setuju, banyaknya, orang, orang, media, sosia...	[setuju, banyaknya, orang, orang, media, sosia...
2	tidak setuju karena media sosial pribadi saya ...	N	tidak setuju karena media sosial pribadi saya ...	setuju media sosial pribadi arah hat2 positif...	[setuju, media, sosial, pribadi, arah, hat2, p...	[setuju, media, sosial, pribadi, arah, hat2, p...
3	tidak setuju	N	tidak setuju	setuju	[setuju]	[setuju]
4	Tidak setuju, karna rata2 teman media sosial s...	N	tidak setuju karna rata2 teman media sosial sa...	setuju karna rata2 teman media sosial teman	[setuju, karna, rata2, teman, media, sosial, t...	[setuju, karna, rata2, teman, media, sosial, t...

5. Mengonversi nilai dalam kolom reviews_lemmatized dari daftar menjadi string

```
dataset2['reviews_cleaned_stemmed'] = [' '.join(l for l in dataset2['reviews_stemmed_data'])
dataset2.head()
```

	Lingkungan Sosial Media Toxic	Labeling pertama	reviews_cleaned	reviews_cleaned_stopwords	reviews_tokenized	reviews_stemmed_data	reviews_cleaned_stemmed
0	Setuju, karena menurut saya banyak orang di so...	P	setuju karena menurut saya banyak orang di so...	setuju orang sosial media gampang terhasut omo...	[setuju, orang, sosial, media, gampang, terhas...	[setuju, orang, sosial, media, gampang, terhas...	setuju orang sosial media gampang terhasut omo...
1	setuju , karena banyaknya orang orang yang man...	P	setuju karena banyaknya orang orang yang man...	setuju banyaknya orang orang media sosial many...	[setuju, banyaknya, orang, orang, media, socia...	[setuju, banyaknya, orang, orang, media, socia...	setuju banyaknya orang orang media sosial many...
2	tidak setuju karena media sosial pribadi saya ...	N	tidak setuju karena media sosial pribadi saya ...	setuju media sosial pribadi arah hal2 positif ...	[setuju, media, sosial, pribadi, arah, hal2, p...	[setuju, media, sosial, pribadi, arah, hal2, p...	setuju media sosial pribadi arah hal2 positif ...
3	tidak setuju	N	tidak setuju	setuju	[setuju]	[setuju]	setuju
4	Tidak setuju, karna rata2 teman media sosial sa...	N	Tidak setuju karna rata2 teman media sosial sa...	setuju karna rata2 teman media sosial teman	[setuju, karna, rata2, teman, media, sosial, t...	[setuju, karna, rata2, teman, media, sosial, t...	setuju karna rata2 teman media sosial teman

6. Setelah melalui proses stemming data akan diimport ke sebuah dataset baru

```
dataset2[['Lingkungan Sosial Media Toxic','Labeling
pertama','reviews_cleaned_stemmed','reviews_cleaned']].to_csv("cleaned_reviews_stemmed.
csv",index = False)
```

4.3.4 Creating Word Cloud

Word cloud adalah representasi visual dari kata-kata yang paling umum dalam sebuah teks. Word Cloud akan mengukur frekuensi setiap kata dalam teks dan mewakili frekuensi ini dengan ukuran kata. Kata yang ukurannya lebih besar adalah kata yang muncul lebih sering dalam teks, sedangkan kata yang ukurannya lebih kecil adalah kata yang jarang muncul. Sebuah word cloud memberikan ringkasan yang sangat berguna tentang distribusi kata dalam sebuah teks. Word Cloud juga merupakan cara yang bagus untuk mendapatkan wawasan cepat tentang kata-kata yang menonjol dalam data teks.

1. Membuat visualisasi Word Cloud berdasarkan Label Positif dengan Negatif

```
from wordcloud import WordCloud
import matplotlib.pyplot as plt

# Fungsi untuk membuat WordCloud
def create_wordcloud(text):
    wordcloud = WordCloud(width=800, height=800, background_color='white', stopwords=None,
min_font_size=10).generate(text)
    plt.figure(figsize=(3, 3), facecolor=None)
    plt.imshow(wordcloud)
    plt.axis("off")
    plt.tight_layout(pad=0)
    plt.show()

# Membuat WordCloud untuk setiap label
for label in ['P', 'N']:
    text = ' '.join(dataset3.loc[dataset3['Labeling pertama'] == label, 'Lingkungan Sosial Media
Toxic'])
    print(f"WordCloud for label {label}:")
    create_wordcloud(text)
```

Wordcloud untuk label P:
(positif/setuju bahwa sosmed toxic)

Top 100 Most Common Words



Wordcloud untuk label N (negatif):
(negatif/tidak setuju sosmed toxic)

Top 100 Most Common Words



Berdasarkan hasil wordcloud masing-masing label menunjukkan bahwa persepsi mahasiswa terhadap lingkungan sosial media yang toxic sebagaimana distribusi kata yang terdapat pada visualisasi. Dapat dilihat pada word cloud yang menunjukkan pandangan negatif bahwa tidak setuju terhadap statement lingkungan sosial media toxic dengan dibuktikan kata yang terbanyak yaitu sosial, mental, orang, positif, berguna, bijak, teman dan lain sebagainya yang dapat dilihat pada gambar

BAB V : KESIMPULAN

Berdasarkan hasil analisis yang kami lakukan dengan judul “Persepsi Mahasiswa Administrasi Publik mengenai dampak media sosial terhadap kesehatan mental dengan metode Analisis Faktor dan Analisis Data Teks” yang kami proses dengan melihat gambaran distribusi keadaan mahasiswa dalam menyikapi sosial media. Maka dapat disimpulkan bahwa secara garis besar sosial media memberikan dampak terhadap kesehatan mental mahasiswa, dibuktikan dengan persepsi mahasiswa yang positif atau setuju dengan lingkungan sosial media yang toxic dapat mempengaruhi kesehatan mental lebih banyak dibandingkan dengan persepsi mahasiswa negatif atau kontra dengan statement tersebut.

Dengan analisis faktor kami dapat simpulkan bahwa terdapat 3 faktor yang mempengaruhi sosial media berdampak pada kesehatan mental yakni Perundungan yang terjadi pada sosial media, penyebab atau pendorong untuk melakukan perundungan dan aktivitas yang dilakukan pada sosial media. Oleh karena itu besar harapan kami eksplorasi ini dijadikan bahan evaluasi bagi mahasiswa hingga masyarakat luas dalam berperilaku dan menyikapi lingkungan sosial media saat ini.

DAFTAR PUSTAKA

- [1] Wikipedia. (2023b, December 12). *Globalisasi*. <https://id.wikipedia.org/wiki/Globalisasi>
- [2] Ferlitasari Reni (2018). Pengaruh Media Sosial Instagram Terhadap Perilaku Keagamaan Remaja (Studi Pada Rohis Di SMA Perintis 1 Bandar Lampung). Fakultas Ushuluddin dan Studi Agama Universitas Islam Negeri Raden Intan Lampung.
- [3] says:, N., & says:, S. S. (2021, January 30). *Metode Analisis Faktor Dan Rumus Pengujiannya [lengkap]*. Statmat. <https://www.statmat.net/metode-analisis-faktor/>
- [4] adminlp2m. (2022, April 28). *Analisis Teks : Definisi, Teknik Serta contohnya*. Lembaga Penelitian dan Pengabdian Masyarakat. <https://lp2m.uma.ac.id/2022/04/28/analisis-teks-definisi-teknik-serta-contohnya/>
- [5] Halodoc, R. (2023, July 13). *Pengaruh media Sosial Pada Kesehatan mental remaja*. halodoc. <https://www.halodoc.com/artikel/pengaruh-media-sosial-pada-kesehatan-mental-remaja>