

Pengaruh Faktor Demografis terhadap Kondisi Depresi dan Kecemasan



Kelompok 3

Filda Zamardah	(22031554047)
Alivia Nayla Wibisono	(22031554041)
Ilham Warmandev	(220315540)

UNIVERSITAS NEGERI SURABAYA
2025

1. Latar Belakang

Kesehatan mental merupakan aspek penting dalam kualitas hidup individu, dan dua kondisi yang paling sering dialami oleh masyarakat global adalah depresi dan kecemasan (*anxiety*). Kedua kondisi ini tidak hanya berdampak pada kesejahteraan psikologis, tetapi juga dapat mempengaruhi produktivitas, hubungan sosial, serta kesehatan fisik secara keseluruhan. Organisasi Kesehatan Dunia (WHO) mencatat bahwa prevalensi depresi dan kecemasan terus meningkat setiap tahun, sehingga analisis faktor yang mempengaruhi kondisi ini menjadi semakin penting. Dalam kajian kesehatan mental, faktor demografis seperti usia, gender dan status pekerjaan sering ditemukan memiliki kontribusi terhadap variasi tingkat depresi dan kecemasan. Beberapa penelitian menunjukkan bahwa perempuan cenderung memiliki tingkat kecemasan dan depresi yang lebih tinggi dibandingkan laki-laki, meskipun penyebabnya dapat bervariasi, mulai dari faktor biologis hingga tekanan sosial. Sementara itu, pengaruh usia terhadap kondisi mental menunjukkan hasil yang beragam, beberapa kelompok usia tertentu lebih rentan mengalami gejala depresi atau kecemasan tergantung pada tahap perkembangan kehidupan, kondisi sosial, serta tekanan lingkungan.

2. Manfaat

- Mendukung perencanaan program kesehatan mental
- Membantu identifikasi kelompok risiko
- Mendorong peningkatan kesadaran masyarakat
- Sebagai dasar pengembangan analisis lanjutan

3. Tujuan

- Menganalisis hubungan demografis seperti age, gender, status pekerjaan tingkat depresi serta kecemasan
- Mengevaluasi perbedaan tingkat depresi dan kecemasan berdasarkan gender
- Mengidentifikasi pola atau kecenderungan umum (usia, gender dan status pekerjaan)
- Menjadi dasar untuk analisis lanjutan

4. Sumber Data

- **Kaggle** : Mental Health in Tech Survey
(<https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey>)
- **Mendeley Data** : Dataset Kesehatan Mental berdasarkan DASS-21
(<https://data.mendeley.com/datasets/br82d4xkj7/1>)
- **Figshare** : Mental health data of 776 Mexican medical students (PHQ-9, GAD-7 and Epworth Sleepiness Scale Scores)
(https://figshare.com/articles/dataset/PHQ_9_GAD_7_and_Epworth_Sleepiness_Scale_Scores_of_783_Mexican_Medical_Students/1305817?file=12990629)

5. Referensi

	<p>Array</p> <p>Volume 25, March 2025, 100375</p>	
<p>Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI</p> <p>https://www.sciencedirect.com/science/article/pii/S2590005625000025</p>		
	<p>Journal of Affective Disorders Reports</p> <p>Volume 12, April 2023, 100512</p>	
<p>Research Paper</p> <p>Exploring the correlates of mental healthcare approaches and employment status among persons with depression and anxiety: A secondary data analysis of a global survey</p> <p>https://www.sciencedirect.com/science/article/pii/S2666915323000501</p>		

6. Proses Wrangling

Pada proyek ini, data diambil dari tiga platform berbeda : Kaggle, Mendeley Data, dan Figshare. Setiap sumber memiliki format, variabel, serta instrumen pengukuran yang berbeda (misalnya PHQ-9, GAD-7, DASS-21).

- a. Mengakses dataset dari tiga platform Kaggle, Mendeley Data, dan Figshare melalui tautan.

- Mengunduh file dalam format csv.
- Melakukan pembacaan data menggunakan Python (pandas).

	Q1_1	Q1_2	Q1_3	Q1_4	Q1_5	Q1_6	Q3_1_S1	Q3_2_S2	Q3_3_S3	Q3_4_S4	...	Anxiety_Level	Q3_15_D1	Q3_16_D2	Q3_17_D3	Q3_18_D4	Q3_19_D5	Q3_20_D6	Q3_21_D7	Depression_Score	Depression_Level
0	21	2	0	4	4	1	3	3	3	3	...	3	0	3	3	2	3	3	3	17	5
1	18	2	0	4	4	1	3	3	3	3	...	2	0	2	3	0	0	2	0	7	3
2	40	2	1	2	1	1	3	3	1	2	...	5	2	1	1	2	2	1	2	11	4
3	24	1	0	4	2	0	3	3	1	2	...	5	2	1	1	2	2	1	2	11	4
4	50	1	1	2	2	0	3	3	3	2	...	2	1	0	0	1	2	0	0	4	1

5 rows × 33 columns

Dataset 1 (Mendeley Data) : Dataset Kesehatan Mental berdasarkan DASS-21

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	have	mental_health_consequence	phys_health_consequence	concerns	supervisor	mental_health_interview	phys_health_interview	mental_vs_physical	edu_consequence	comment
0	2014-08-27 11:29:31	37	Female	United States	IL	NaN	No	Yes	Often	6-25	...	Somewhat easy	No	No	Some of them	No	No	Maybe	Yes	No	NaN
1	2014-08-27 11:29:37	44	M	United States	IN	NaN	No	No	Rarely	More than 1000	...	Don't know	Maybe	No	No	No	No	Don't know	No	No	NaN
2	2014-08-27 11:29:44	32	Male	Canada	NaN	NaN	No	No	Rarely	6-25	...	Somewhat difficult	No	No	Yes	Yes	Yes	No	No	No	NaN
3	2014-08-27 11:29:48	31	Male	United Kingdom	NaN	NaN	Yes	Yes	Often	25-100	...	Somewhat difficult	Yes	Yes	Some of them	No	Maybe	Maybe	No	Yes	NaN
4	2014-08-27 11:30:22	31	Male	United States	TX	NaN	No	No	Never	100-500	...	Don't know	No	No	Some of them	Yes	Yes	Yes	Don't know	No	NaN
...
1254	2015-09-12 11:17:21	26	male	United Kingdom	NaN	No	No	Yes	NaN	25-100	...	Somewhat easy	No	No	Some of them	Some of them	No	No	Don't know	No	NaN
1255	2015-09-28 01:07:35	32	Male	United States	IL	No	Yes	Yes	Other	25-100	...	Somewhat difficult	No	No	Some of them	Yes	No	No	Yes	No	NaN
1256	2015-11-07 12:36:58	34	male	United States	CA	No	Yes	Yes	Sometimes	More than 1000	...	Somewhat difficult	Yes	Yes	No	No	No	No	No	No	NaN
1257	2015-11-30 21:23:08	48	f	United States	NC	No	No	No	NaN	100-500	...	Don't know	Yes	No	No	No	No	No	No	No	NaN
1258	2016-02-01 23:04:31	25	Male	United States	IL	No	No	Yes	Sometimes	25-100	...	Don't know	Maybe	No	Some of them	No	No	No	Don't know	No	NaN
1259 rows ~ 27 columns																					

1259 rows × 27 columns

Dataset 2 (Kaggle) : Mental Health in Tech Survey

	id	school_year	semester	age	gender	height	weight	phq1	phq2	phq3	...	epw6	epw7	epw8	bed_time	wake_up_time	reported_sleep_hours	times_week_nap	nap_duration	weekly_study_hours	grades
0	A1	1	1	19.0	m	1.50	75.0	2.0	1.0	1.0	—	0.0	1.0	1.0	01:00	06:30	02:00	1.0	02:00	12:00	8.0
1	B2	1	1	18.0	m	1.68	56.0	2.0	0.0	1.0	—	0.0	2.0	2.0	00:00	08:15	07:00	2.0	02:00	05:00	6.5
2	C3	1	1	19.0	m	1.74	76.0	1.0	1.0	0.0	—	0.0	1.0	1.0	02:00	06:30	04:00	1.0	00:30	42:00	7.0
3	D4	1	1	18.0	f	1.68	67.0	2.0	2.0	3.0	—	0.0	3.0	0.0	00:30	07:00	05:00	4.0	02:00	12:30	6.0
4	E5	1	1	18.0	m	1.80	83.0	0.0	1.0	1.0	—	0.0	0.0	0.0	23:00	07:30	08:00	2.0	01:00	15:00	NaN
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
771	R772	4	8	24.0	f	1.57	52.0	3.0	0.0	1.0	—	0.0	1.0	0.0	00:30	06:00	06:00	4.0	02:00	01:00	6.1
772	S773	4	8	22.0	m	1.80	82.0	1.0	0.0	0.0	—	0.0	0.0	0.0	23:30	06:20	07:00	2.0	00:30	01:00	8.0
773	T774	4	8	22.0	m	1.73	68.0	1.0	0.0	0.0	—	0.0	1.0	0.0	23:00	05:30	06:00	1.0	01:00	03:00	9.0
774	U775	4	8	22.0	f	1.58	57.5	2.0	2.0	3.0	—	0.0	1.0	3.0	23:00	05:15	06:00	4.0	02:00	NaN	8.0
775	V776	4	8	22.0	m	1.76	70.0	1.0	0.0	0.0	—	0.0	0.0	0.0	00:00	05:50	06:00	4.0	01:00	02:00	8.0

776 rows × 43 columns

Dataset 3 (Figshare) : Mental Health data off 776 Mexican students (PHQ-9, GAD-7, and Epworth Sleepiness Scale Scores

- Memeriksa jumlah baris dan kolom
- Melakukan rename pada dataset 1 (Mendeley Data) sesuai metadata

age	gender	marital_status	educational_status	occupational_status	sleeping_problem	stress_item1	stress_item2	stress_item3	stress_item4	...	anxiety_level_category	depression_item1	depression_item2	depression_item3	depression_item4	depression_item5	depression_item6	depression_item7	depression	depression_level_category	
0	21	2	0	4	4	1	3	3	3	3	...	3	0	3	3	2	3	3	17	5	
1	18	2	0	4	4	1	3	3	3	3	...	2	0	2	3	0	2	0	7	3	
2	40	2	1	2	1	1	3	3	1	2	...	5	2	1	1	2	2	1	2	11	4
3	24	1	0	4	2	0	3	3	1	2	...	5	2	1	1	2	2	1	2	11	4
4	50	1	1	2	2	0	3	3	3	2	...	2	1	0	0	1	2	0	0	4	1
...
1807	52	2	1	2	1	0	0	0	0	0	...	5	2	3	0	1	2	0	0	8	3
1808	33	2	1	2	1	0	0	0	1	2	...	3	0	1	0	2	1	1	1	6	2
1809	24	2	1	2	1	0	0	0	1	1	...	1	0	0	0	1	0	0	0	1	1
1810	50	2	1	1	1	0	0	0	0	0	...	2	0	1	1	1	1	2	0	6	2
1811	28	2	1	1	1	0	0	0	0	1	...	1	0	0	0	1	1	2	1	5	2
1812 rows x 21 columns																					

1812 rows × 33 columns

- Memeriksa missing values
- Memeriksa tipe data
- Mengambil data demografis seperti umur, gender, status pekerjaan, kesehatan mental (kecemasan), status pekerjaan
- Melakukan penyeragaman variabel gender, status pekerjaan, depresi dan anxiety
- Menentukan kolom yang akan digunakan yaitu age, gender, status pekerjaan, depresi, dan anxiety dan melakukan penyatuan kolom dari berbagai sumber.

	age	gender	depression	anxiety
0	19.0	male	9.0	11.0
1	18.0	male	8.0	5.0
2	19.0	male	8.0	6.0
3	18.0	female	19.0	15.0
4	18.0	male	6.0	14.0
...
3842	52.0	female	8.0	10.0
3843	33.0	female	6.0	6.0
3844	24.0	female	1.0	2.0
3845	50.0	female	6.0	5.0
3846	29.0	female	5.0	3.0

3846 rows × 4 columns

	depression	anxiety	occupation_category
0	9.0	11.0	Student
1	8.0	5.0	Student
2	8.0	6.0	Student
3	19.0	15.0	Student
4	6.0	14.0	Student
...
3842	8.0	10.0	Other
3843	6.0	6.0	Other
3844	1.0	2.0	Other
3845	6.0	5.0	Other
3846	5.0	3.0	Other

3847 rows × 3 columns

Integrasi depresi dan anxiety berdasarkan umur, dan jenis kelamin

Integrasi depresi dan anxiety berdasarkan status pekerjaan

```

Rata-rata depresi per gender:
gender
female    5.873143
male      3.886774
Name: depression, dtype: float64

```

```

Rata-rata anxiety per gender:
gender
female    4.824112
male      3.726177
Name: anxiety, dtype: float64

```

```

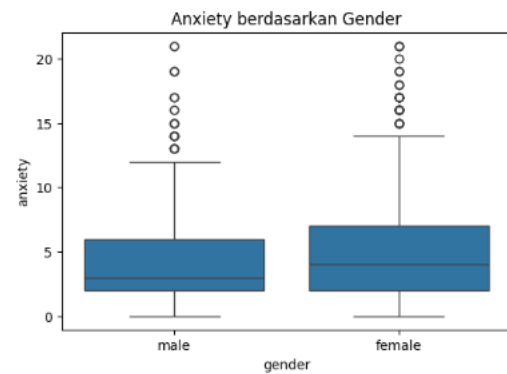
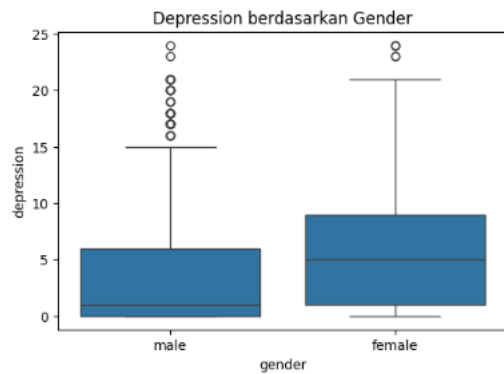
=== Rata-rata Depresi & Anxiety Berdasarkan Status Pekerjaan ===

```

	depression	anxiety
occupation_category		
Other	6.68	4.60
Student	7.12	6.81
Tech_Worker	0.51	1.54

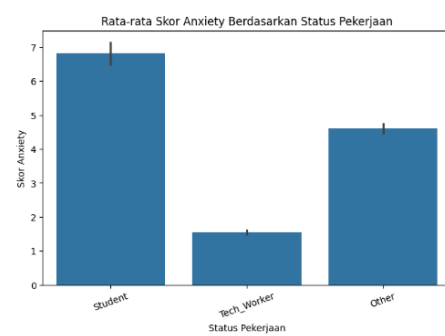
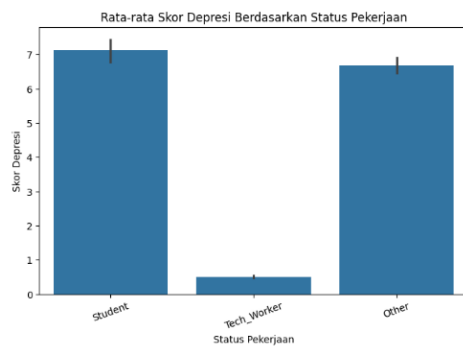
Tingkat depresi dan anxiety yang dialami oleh perempuan lebih tinggi dibandingkan dengan laki-laki

Tingkat depresi dan anxiety pelajar lebih tinggi dibandingkan dengan pekerja teknologi dan lainnya



Visualisasi tingkat depresi antara perempuan dan laki-laki

Visualisasi tingkat anxiety antara perempuan dan laki-laki



Visualisasi tingkat depresi berdasarkan status pekerjaan

Visualisasi tingkat anxiety berdasarkan status pekerjaan

7. Kendala

- Perbedaan struktur dan format dataset
- Ketidaksesuaian definisi variabel
- Perbedaan populasi responden
- Skala skor yang tidak seragam

8. Rencana Lanjutan

- Pembuatan model prediksi
- Segmentasi populasi
- Dashboard interaktif

9. Kontribusi

Filda Zamardah	22031554047	1. Analisis depresi dan anxiety berdasarkan nage dan gender 2. Membuat laporan
Alivia Nayla Wibisono	22031554041	1. Analisis depresi dan anxiety berdasarkan status pekerjaan 2. Membuat PPT
Ilham Warmandev	220315540	1. Mencari dataset 2. Mencari referensi