

# AI IMPACT

ANALISA AI IMPACT  
TERHADAP JOB TRENDS  
MENGUNAKAN MACHINE  
LEARNING

kelompok 1



# MEET THE TEAM



**ASFAH AF KARINAH**

2012210006



**ENI TRISNIA**

2012210008



**SABRINA NASWA A**

2012210706



# OUTLINE PEMBAHASAN



**STUDY CASE**



**PENGOLAHAN**



**KESIMPULAN**



# STUDY CASE

Seiring dengan berkembangnya teknologi kecerdasan buatan (AI), berbagai sektor industri mulai merasakan dampaknya terhadap struktur tenaga kerja. Beberapa jenis pekerjaan mengalami pertumbuhan pesat karena didukung oleh AI, sementara yang lain justru menurun drastis akibat otomatisasi dan efisiensi teknologi. Studi ini bertujuan untuk memahami dinamika tersebut dengan mengevaluasi data proyeksi pasar kerja global tahun 2024 hingga 2030.



# TUJUAN



**Memahami Faktor-faktor yang Berhubungan dengan Status Pekerjaan**



**Membangun Model Prediktif**



**Mengidentifikasi Fitur Penting**



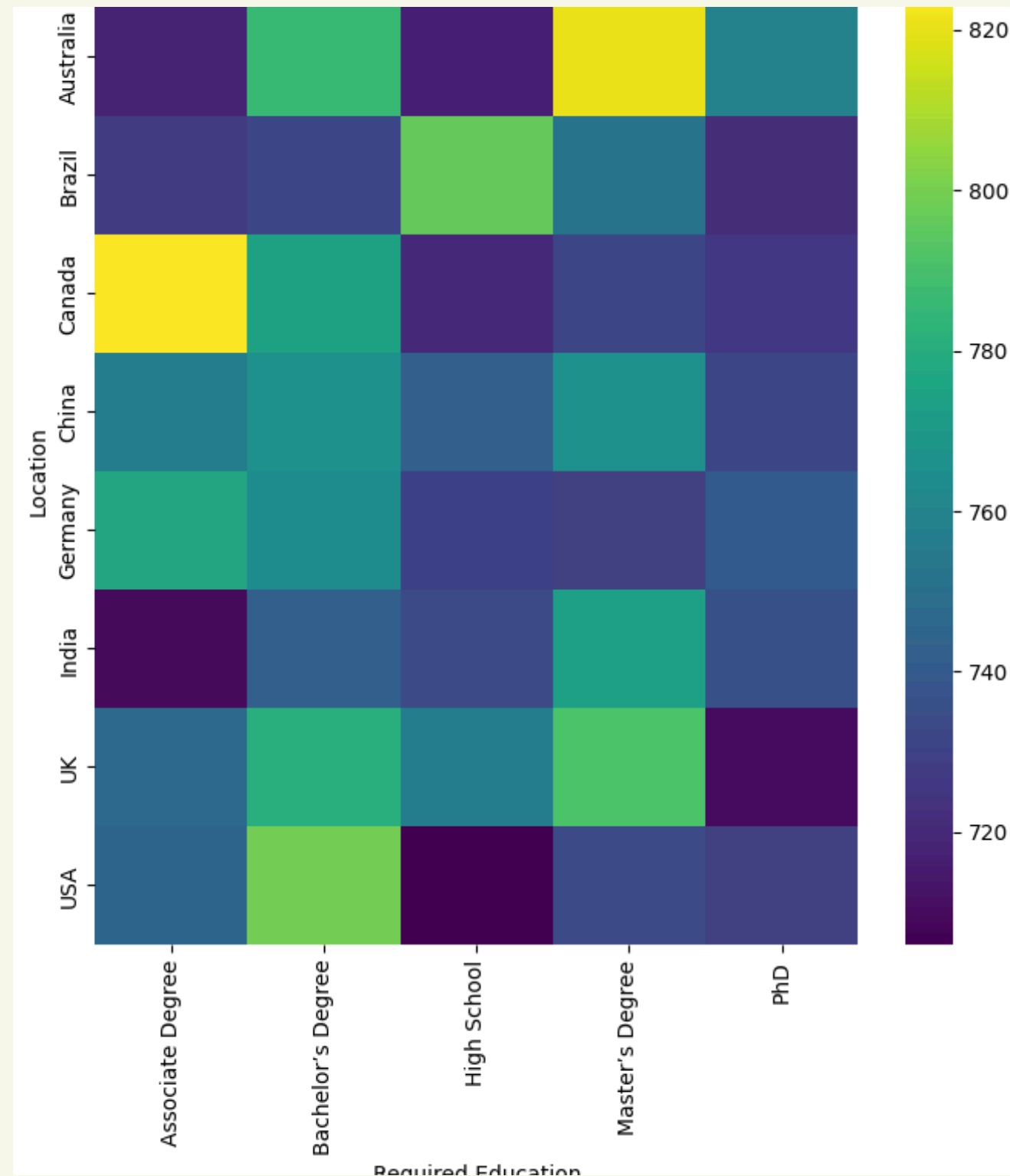
# ISI DATA

- **JOB TITLE**
- **INDUSTRY**
- **JOB STATUS**
- **AI IMPACT LEVEL**
- **MEDIAN SALARY (USD)**
- **REUIRED EDUCATION**
- **EXPERINECE REQUIRED (YEARS)**
- **JOB OPENINGS (2024)**
- **PROJECTED OPENINGS (2030)**
- **REMOTE WORK RATIO (%)**
- **AUTOMATION RISK(%)**
- **LOCATION**
- **GENDER DELIVERY**



# EXPLORATORY DATA ANALYSIS (EDA)

- REQUIRED EDUCATION VS LOCATION



Sumbu X (horizontal): Menunjukkan tingkat pendidikan yang dibutuhkan untuk suatu pekerjaan.

Sumbu Y (vertikal): Location Menunjukkan lokasi atau negara tempat pekerjaan tersedia.

Warna (Color Intensity): Menunjukkan jumlah atau konsentrasi pekerjaan (semakin terang, semakin tinggi jumlahnya).

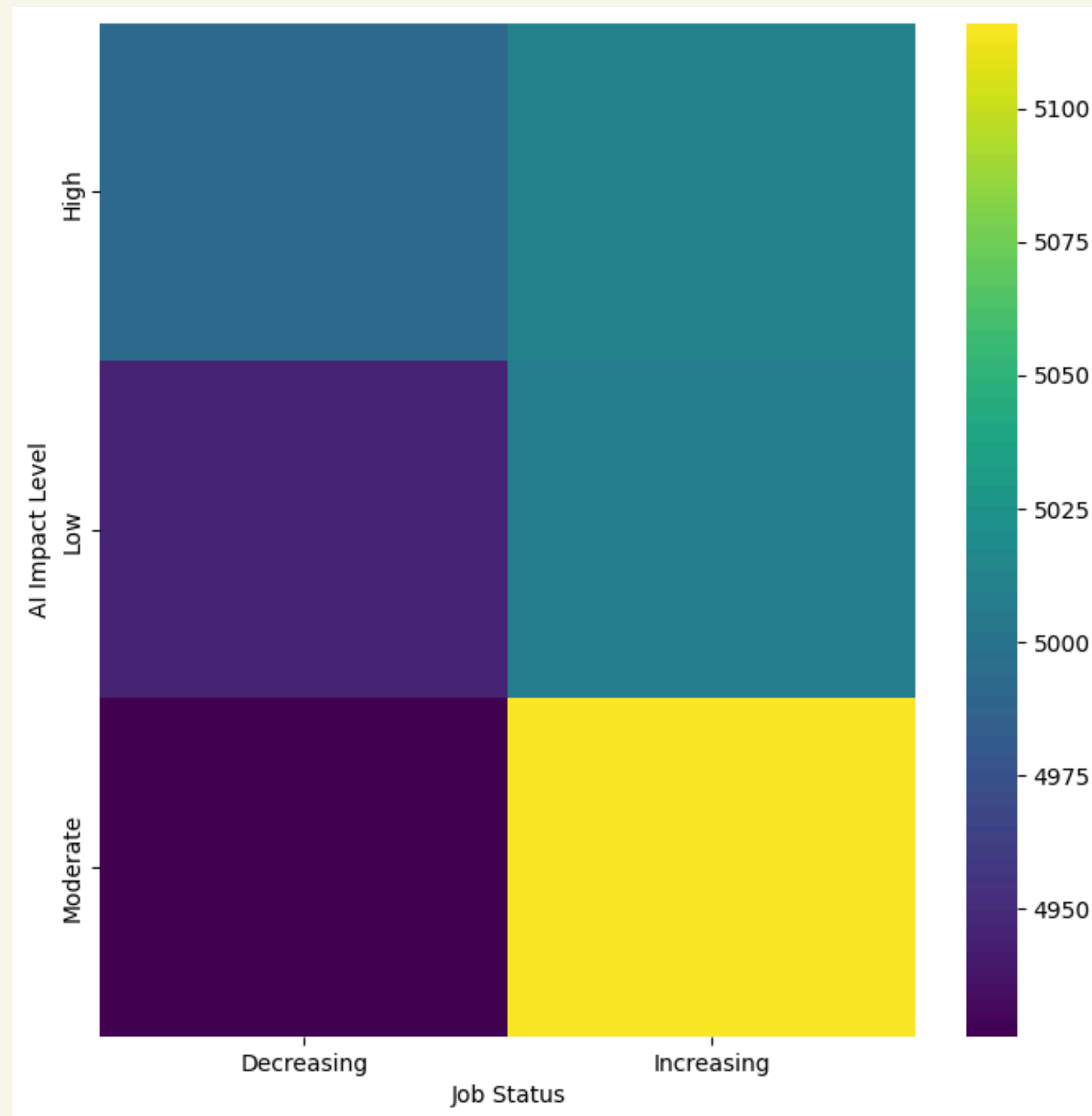
## Kesimpulannya :

- Semakin terang warnanya, jumlah pekerjaan makin banyak.
- Bachelor's Degree paling banyak dan tersebar luas.
- Master's dan PhD terkonsentrasi di lokasi tertentu.
- High School dan Associate Degree lebih sedikit jumlahnya.



# EXPLORATORY DATA ANALYSIS (EDA)

- **JOB STATUS VS AI IMPACT LEVEL**



Sumbu X (horizontal): Job Status Menunjukkan apakah jumlah pekerjaan meningkat (Increasing) atau menurun (Decreasing).

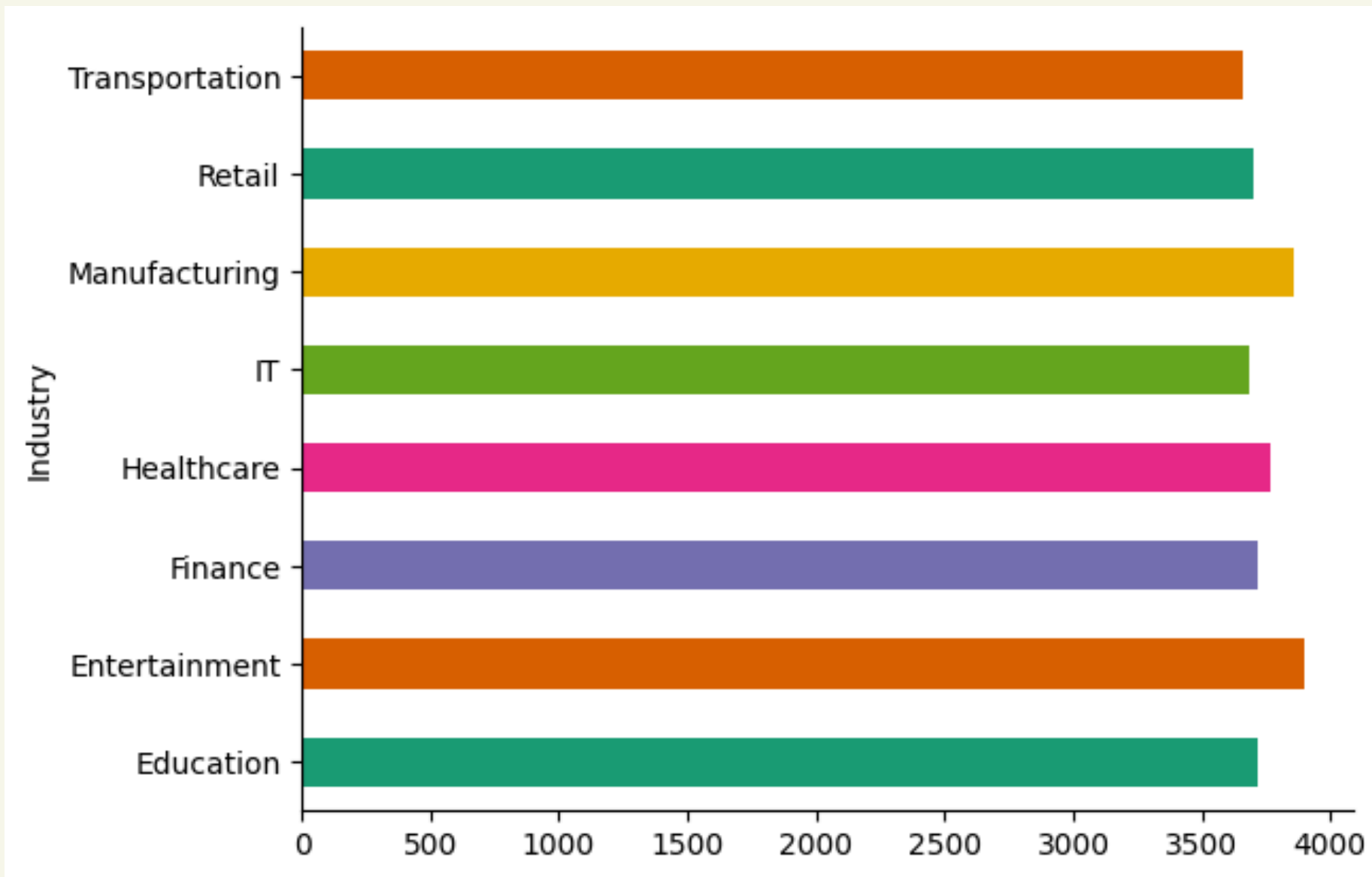
Sumbu Y (vertikal): AI Impact Level Menunjukkan tingkat dampak AI terhadap pekerjaan: Low, Moderate, dan High.

Warna (color intensity): Menunjukkan jumlah pekerjaan. Warna lebih terang berarti jumlah lebih besar. Warna yang hampir merata menunjukkan bahwa dampak AI tidak terlalu menentukan apakah suatu pekerjaan sedang meningkat atau menurun.





# EXPLORATORY DATA ANALYSIS (EDA)



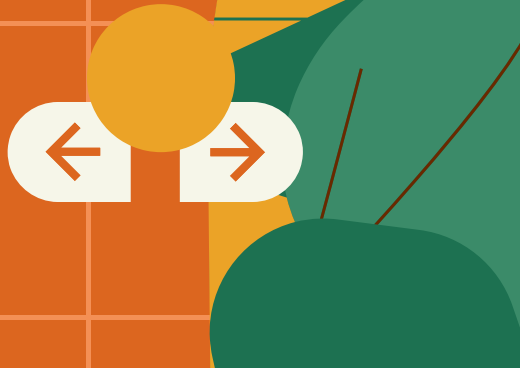
Sumbu Y: Menunjukkan berbagai kategori industri dalam dataset.

Sumbu X: Menunjukkan jumlah pekerjaan atau frekuensi untuk setiap industri.

## Kesimpulannya :

Dari bar plot ini, dapat dengan mudah melihat industri mana yang paling banyak diwakili dalam dataset, dan industri mana yang paling sedikit. Panjang setiap batang menunjukkan jumlah pekerjaan di industri tersebut.

Melihat grafik tersebut, dapat mengidentifikasi industri teratas dengan jumlah pekerjaan terbanyak dan industri terbawah. Ini memberikan gambaran tentang komposisi industri dari data pekerjaan AI ini



# PREPROCESSING DATA

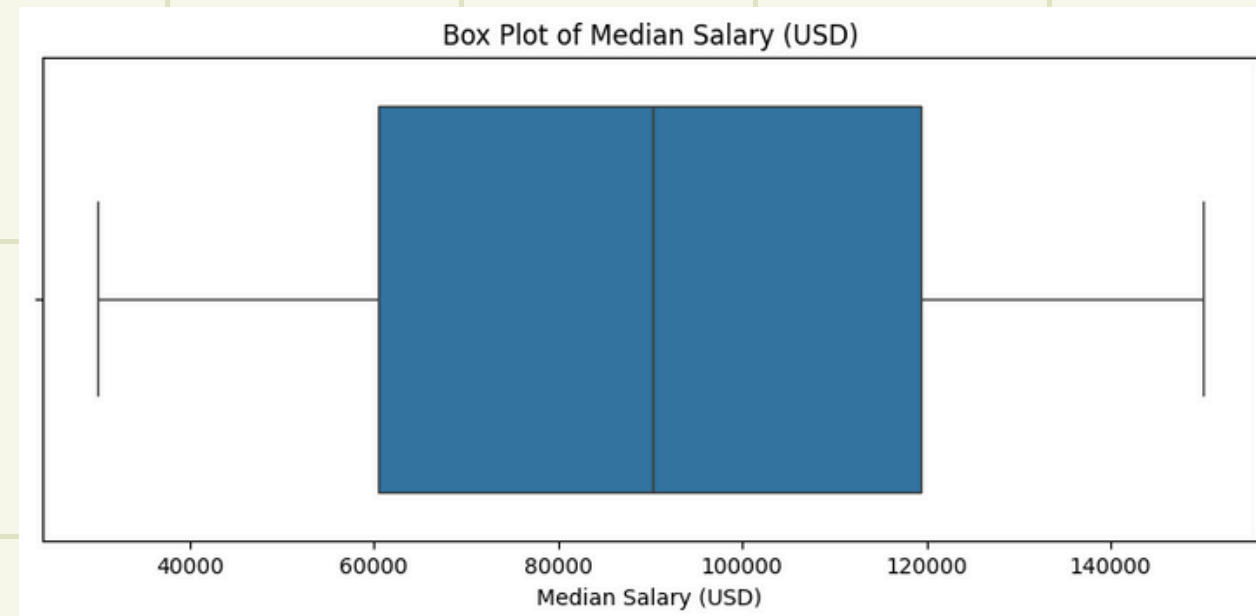
Jumlah nilai yang hilang per kolom:

Job Title	0
Industry	0
Job Status	0
AI Impact Level	0
Median Salary (USD)	0
Required Education	0
Experience Required (Years)	0
Job Openings (2024)	0
Projected Openings (2030)	0
Remote Work Ratio (%)	0
Automation Risk (%)	0
Location	0
Gender Diversity (%)	0
dtype: int64	

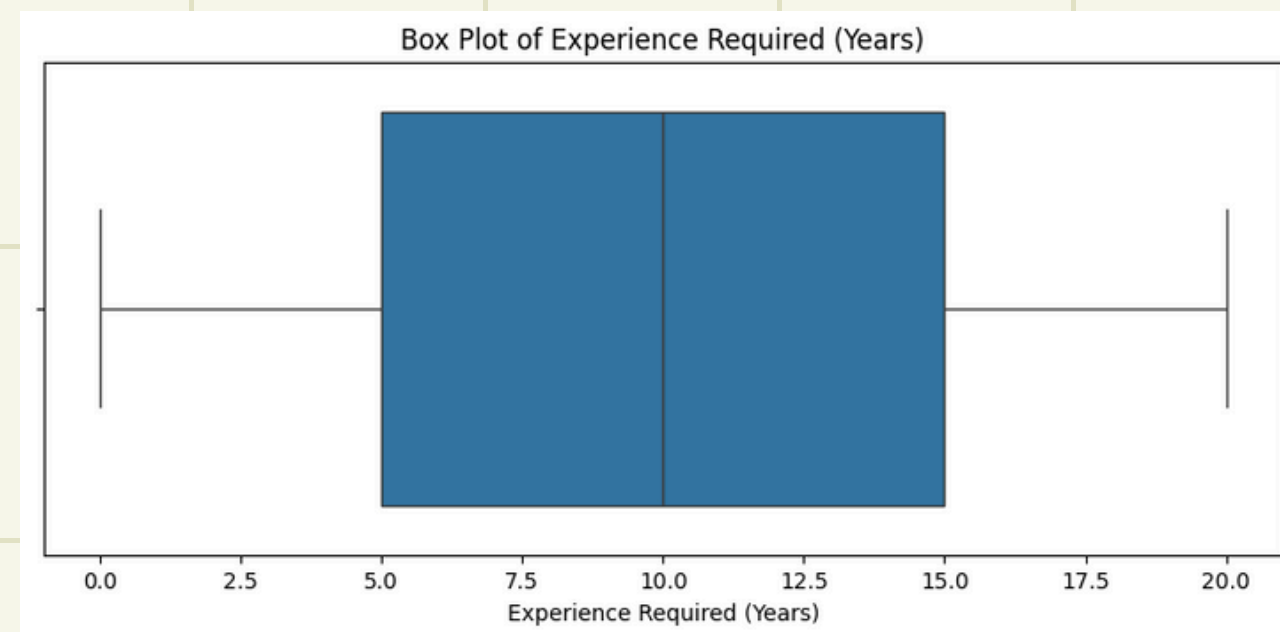
**berdasarkan hasil tersebut,  
menunjukkan bahwa tidak ada  
nilai yang hilang di kolom  
manapun dalam dataset.**



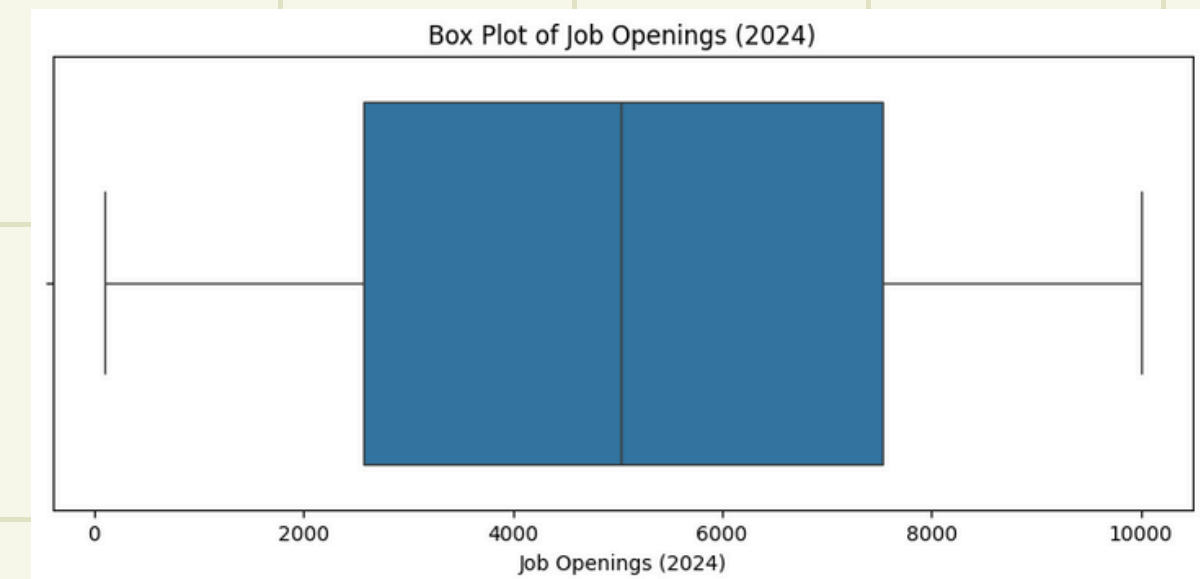
# PREPROCESSING DATA



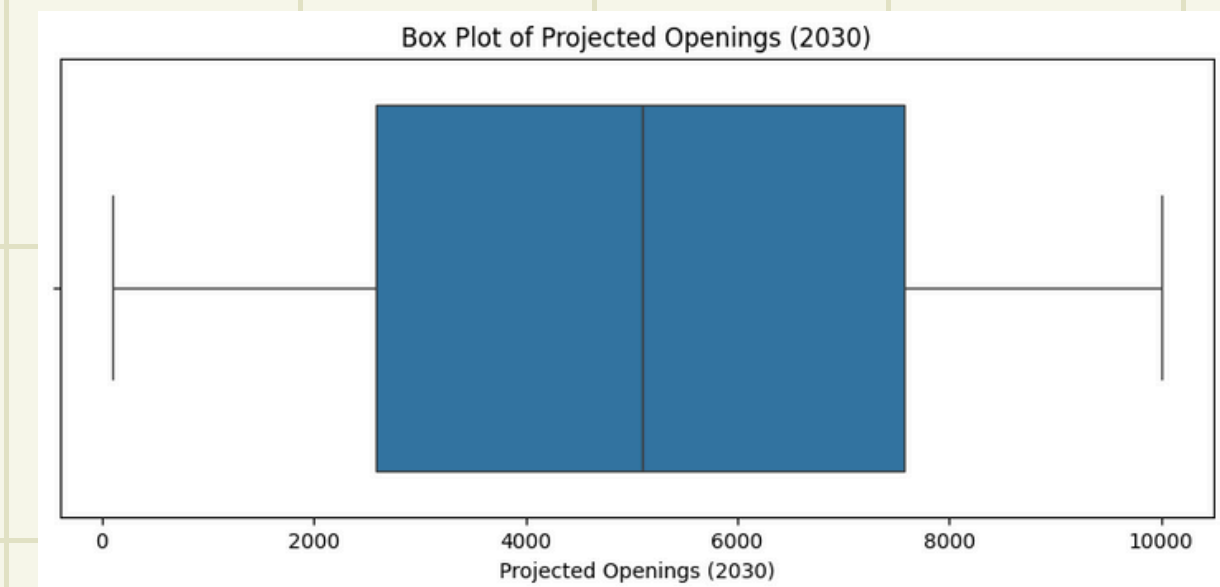
**Mayoritas pekerjaan memiliki gaji dalam rentang \$60.000 – \$120.000.**



**Mayoritas pengalaman dalam rentang 5 tahun – 15 tahun**

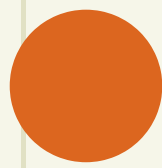


**Mayoritas lowongan kerja di tahun 2024 berada di kisaran 3.000 sampai 7.500 lowongan per bulan.**

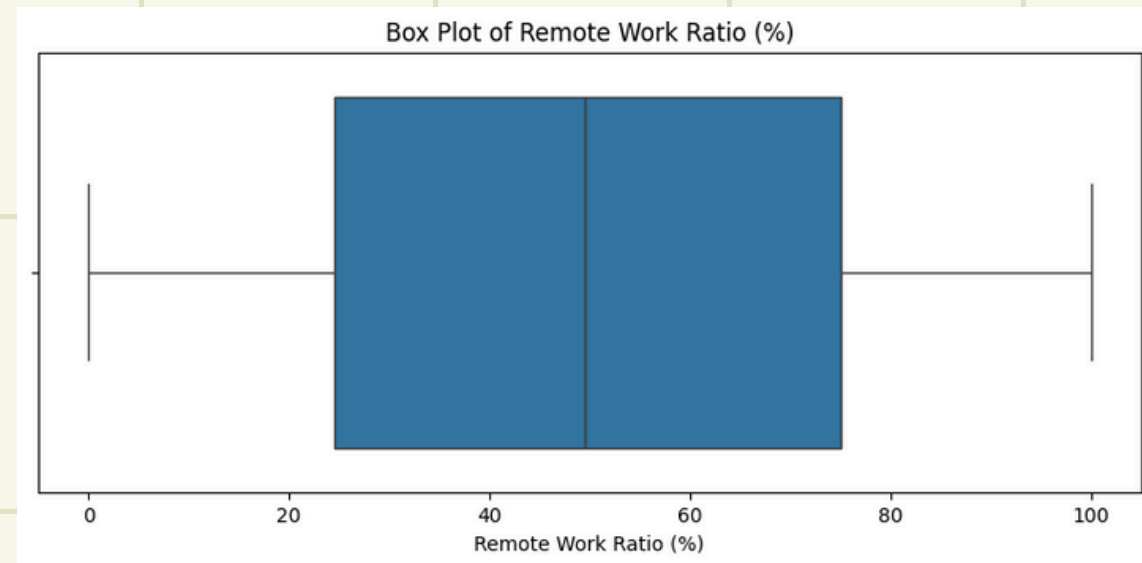


**Mayoritas lowongan kerja yang diproyeksikan pada tahun 2023 berada di kisaran 3000 – 7000 lowongan per bulan**

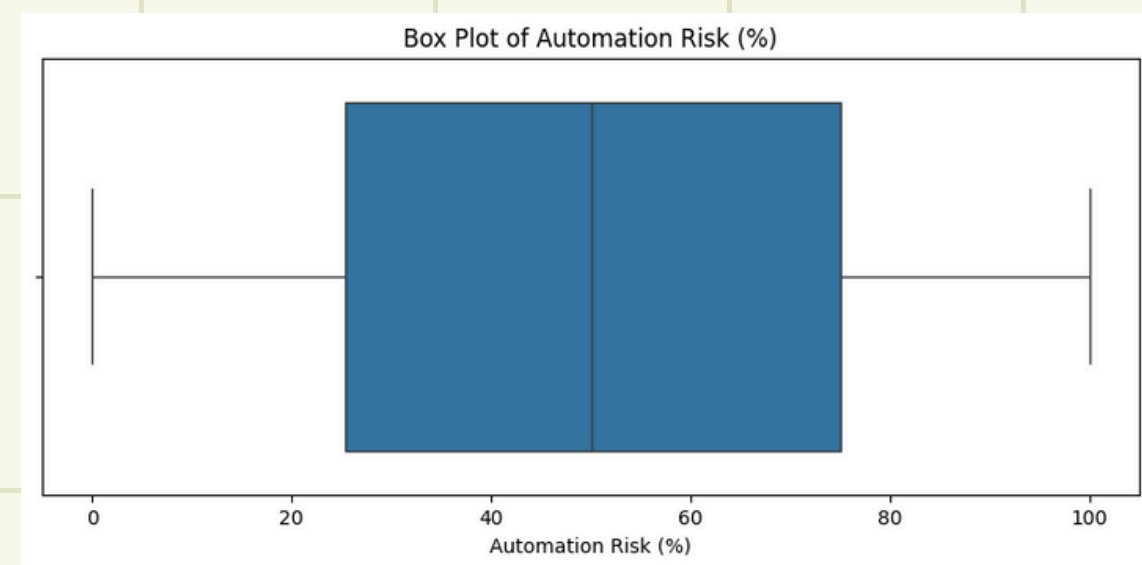




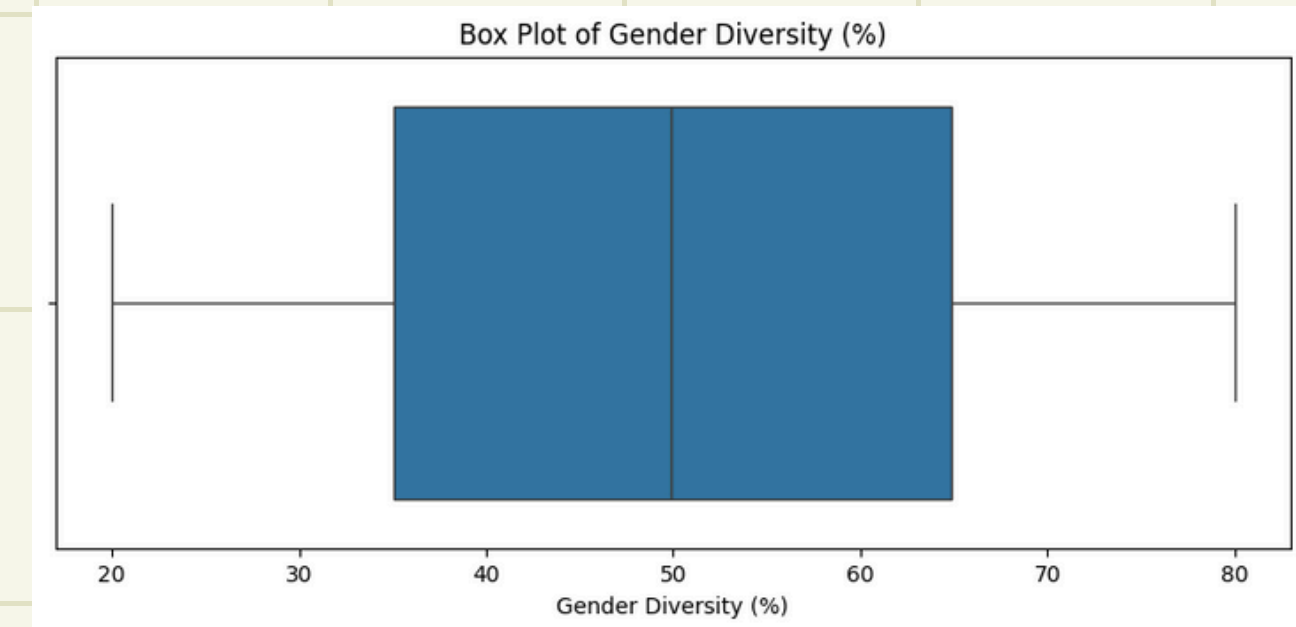
# PREPROCESSING DATA



**Mayoritas pekerja remote diantara 30% – 70%**



**Mayoritas pekerjaan di data ini punya risiko otomatisasi sekitar 25 % sampai 75 %**



**Mayoritas pekerjaan hanya diisi oleh sekitar 35% sampai 65% dari satu gender, menunjukkan keragaman gender yang cukup tinggi**





# PREPROCESSING DATA

1. Box plot yang dihasilkan menunjukkan distribusi untuk setiap kolom numerik. Titik–titik di luar "whiskers" pada box plot mengindikasikan potensi outlier dalam data.
2. Terlihat ada outlier di beberapa kolom numerik, seperti 'Median Salary (USD)', 'Experience Required (Years)', 'Job Openings (2024)', 'Projected Openings (2030)', 'Remote Work Ratio (%)', 'Automation Risk (%)', dan 'Gender Diversity (%)'.
3. Keberadaan outlier ini berarti ada beberapa entri data yang memiliki nilai ekstrem yang jauh dari sebagian besar data lainnya di kolom–kolom tersebut



# MODELLING DATA

- RANDOM FOREST

Classification Report:

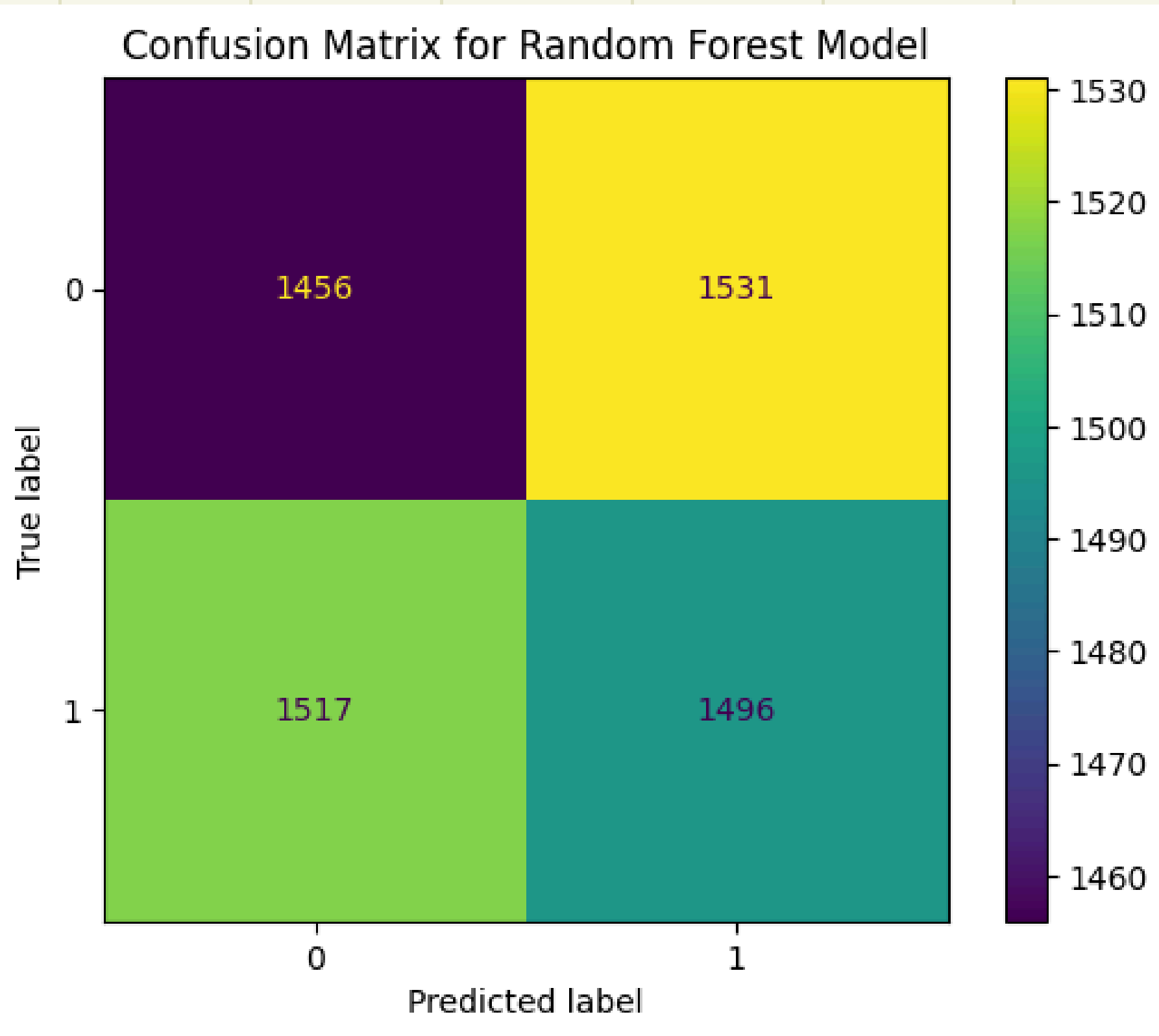
	precision	recall	f1-score	support
Decreasing	0.49	0.49	0.49	2987
Increasing	0.49	0.50	0.50	3013
accuracy			0.49	6000
macro avg	0.49	0.49	0.49	6000
weighted avg	0.49	0.49	0.49	6000

Confusion Matrix:

```
[[1456 1531]
 [1517 1496]]
```

# MODELLING DATA

- RANDOM FOREST



1. True Negatives (Kiri Atas): 1456 Model dengan benar memprediksi 1456 pekerjaan berstatus 'Decreasing' yang sebenarnya memang 'Decreasing'.
2. False Positives (Kanan Atas): 1531 Model salah memprediksi 1531 pekerjaan berstatus 'Increasing', padahal sebenarnya pekerjaan tersebut 'Decreasing'. Ini adalah kesalahan Tipe I.
3. False Negatives (Kiri Bawah): 1517 Model salah memprediksi 1517 pekerjaan berstatus 'Decreasing', padahal sebenarnya pekerjaan tersebut 'Increasing'. Ini adalah kesalahan Tipe II.
4. True Positives (Kanan Bawah): 1496 Model dengan benar memprediksi 1496 pekerjaan berstatus 'Increasing' yang sebenarnya memang 'Increasing'.

# MODELLING DATA

- EVALUATE THE LIGHTGBM MODEL

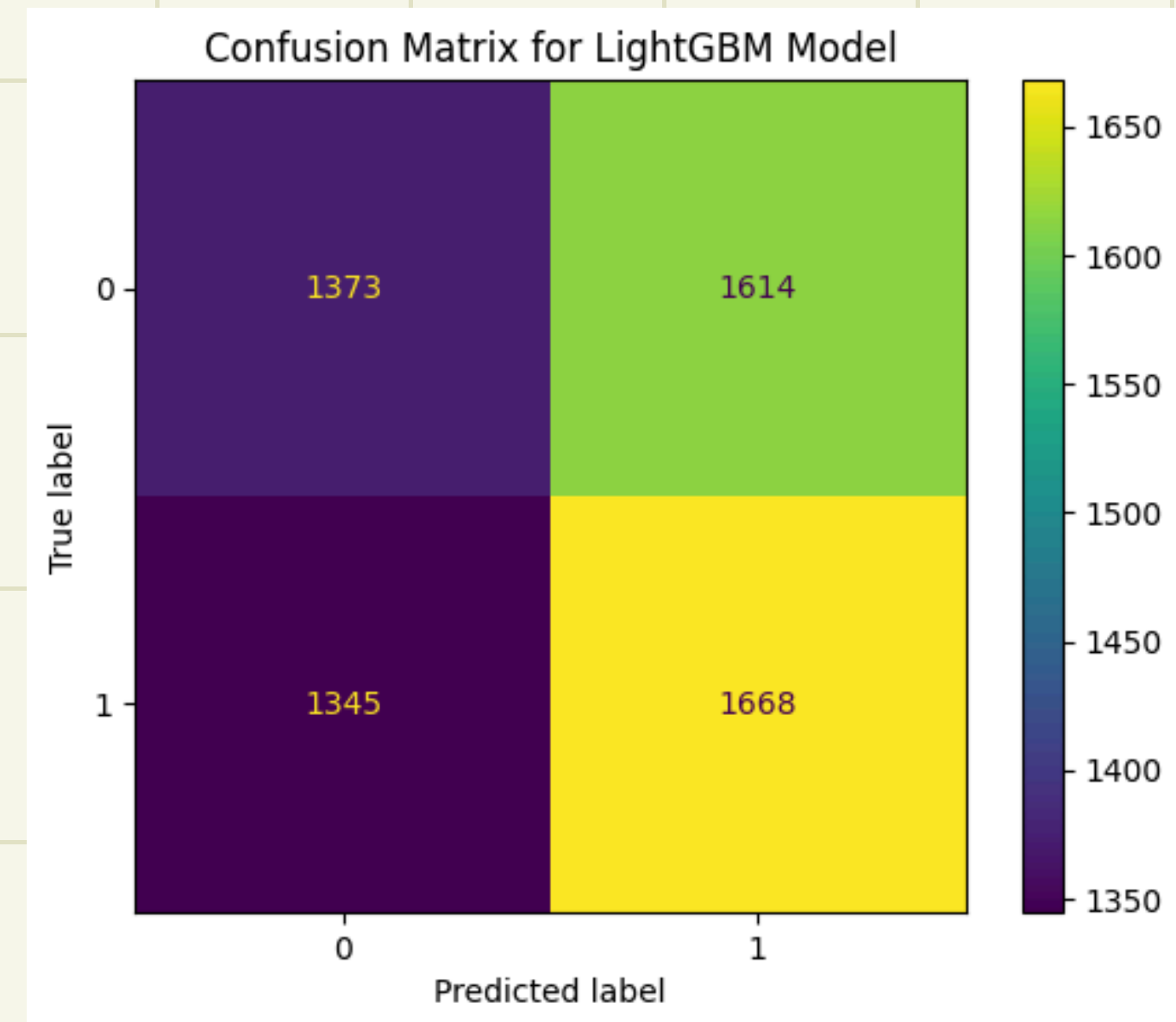
```
⇒ LightGBM Akurasi: 0.5068
LightGBM Presisi (weighted): 0.5067
LightGBM Recall (weighted): 0.5068
LightGBM F1-score (weighted): 0.5057
```

LightGBM Classification Report:

	precision	recall	f1-score	support
Decreasing	0.51	0.46	0.48	2987
Increasing	0.51	0.55	0.53	3013
accuracy			0.51	6000
macro avg	0.51	0.51	0.51	6000
weighted avg	0.51	0.51	0.51	6000

LightGBM Confusion Matrix:

```
[[1373 1614]
 [1345 1668]]
```





# MODELLING DATA

- DISTRIBUSI NILAI BINER

```
➡ Job Status Job Status_Binary
0 Increasing 1
1 Increasing 1
2 Increasing 1
3 Increasing 1
4 Increasing 1
```

Distribusi nilai biner:

Job Status\_Binary

1 15136

0 14864

Name: count, dtype: int64

```
➡ LightGBM Akurasi (target biner): 0.5068
LightGBM Presisi (target biner): 0.5082
LightGBM Recall (target biner): 0.5536
LightGBM F1-score (target biner): 0.5299
```

LightGBM Classification Report (target biner):

	precision	recall	f1-score	support
0	0.51	0.46	0.48	2987
1	0.51	0.55	0.53	3013
accuracy			0.51	6000
macro avg	0.51	0.51	0.51	6000
weighted avg	0.51	0.51	0.51	6000

LightGBM Confusion Matrix (target biner):

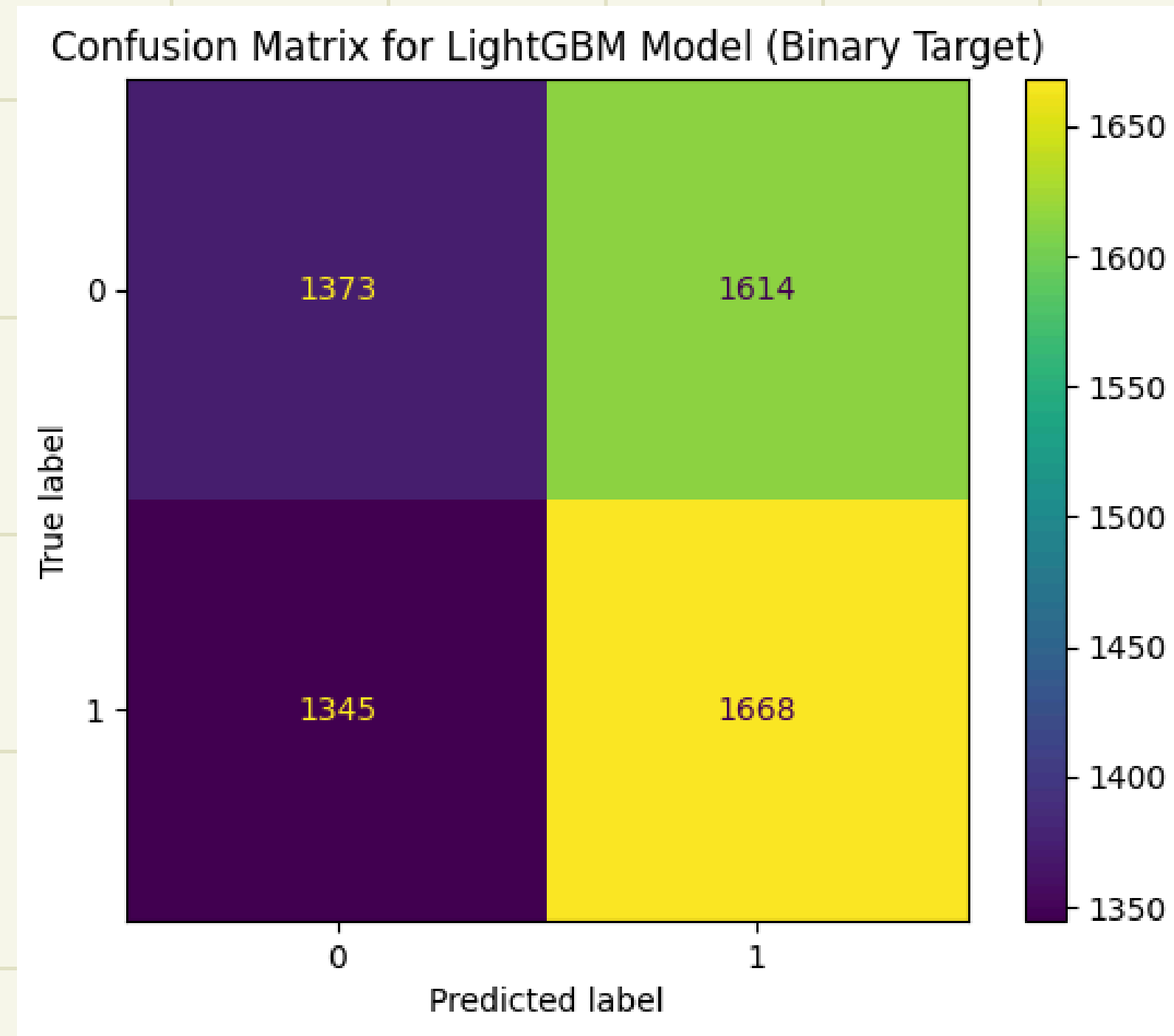
```
[[1373 1614]
 [1345 1668]]
```

UKURAN SET PELATIHAN (DENGAN TARGET BINER): (24000, 12) (24000,) UKURAN SET  
PENGUJIAN (DENGAN TARGET BINER): (6000, 12) (6000,)



# MODELLING DATA

- DISTRIBUSI NILAI BINER



# KESIMPULAN

**Mayoritas pekerjaan memiliki gaji \$60.000–\$120.000 dan membutuhkan pengalaman 5–15 tahun. Lowongan kerja 2024 dan proyeksi 2030 berkisar 3.000–7.500 per bulan. Pekerjaan umumnya mendukung sistem kerja remote 30%–70%, dengan risiko otomatisasi 25%–75%. Gender diversity cukup baik (35%–65%). Terdapat outlier pada variabel seperti gaji dan automation risk, menunjukkan keberagaman karakteristik antar pekerjaan.**

# REKOMENDASI

- **Pekerja dan pelajar disarankan meningkatkan keterampilan digital dan adaptasi teknologi untuk menghadapi risiko otomatisasi.**
- **Pencari kerja sebaiknya fokus pada bidang dengan tren peningkatan lowongan dan gaji menengah–tinggi.**
- **Perusahaan dapat memperluas kebijakan kerja fleksibel dan mendorong keragaman gender untuk menciptakan lingkungan kerja inklusif.**
- **Pembuat kebijakan perlu memperhatikan jenis pekerjaan dengan risiko tinggi otomatisasi dan menyiapkan program reskilling.**



# THANK YOU

KELOMPOK 1

