

Data Preparation on Healthcare Workforce Mental Health Dataset

Disusun Oleh:

- Willy Jonathan Arsyad
- Agil Mughni
- Alfi Zamriza
- T.M. Fadlul Ihsan
- M. Arkan Haris



Deskripsi Dataset

1 Sumber dan Ukuran

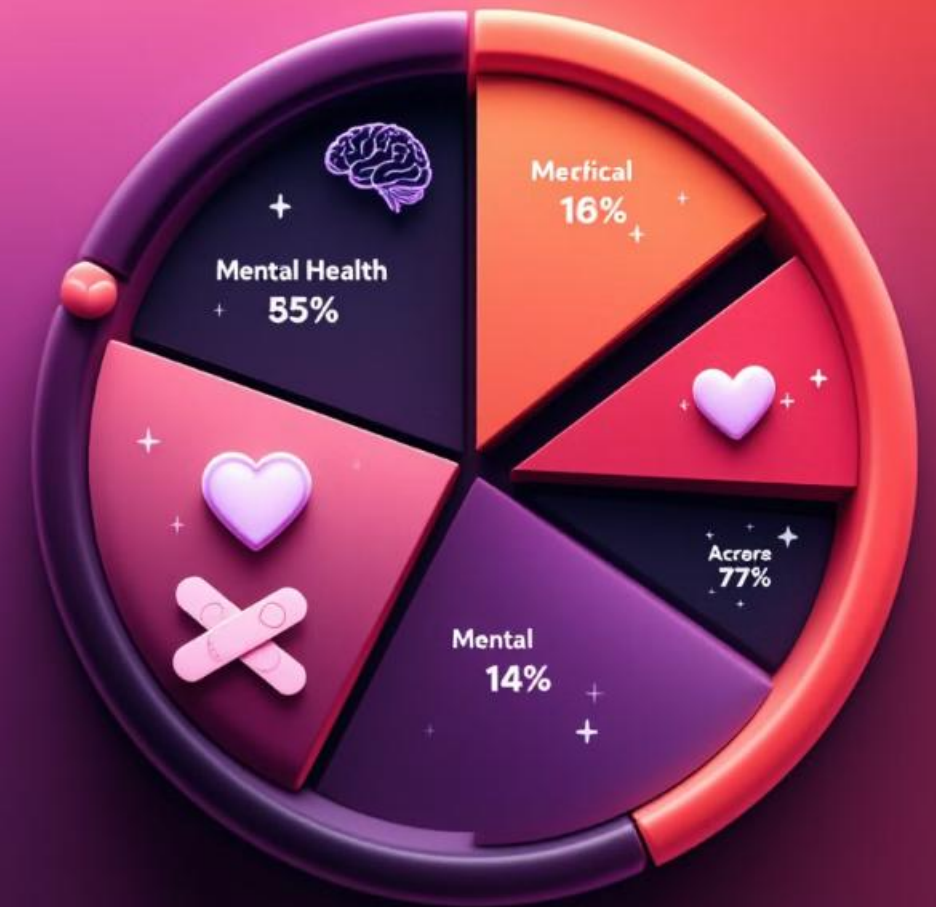
Dataset berasal dari Kaggle. Terdiri dari 5.000 baris dan 10 kolom dalam format CSV.

2 Atribut Utama

Kolom-kolom mencakup data identitas, peran, departemen, faktor stres, dan indikator kesehatan mental.

3 Variabel Target

Turnover Intention menjadi variabel target utama. Menunjukkan keinginan karyawan untuk berhenti.



Pemahaman Data Numerik

Tingkat Stres

Rata-rata tingkat stres karyawan mencapai 7,33 dari skala 10.

Menunjukkan tekanan kerja tinggi di sektor kesehatan.

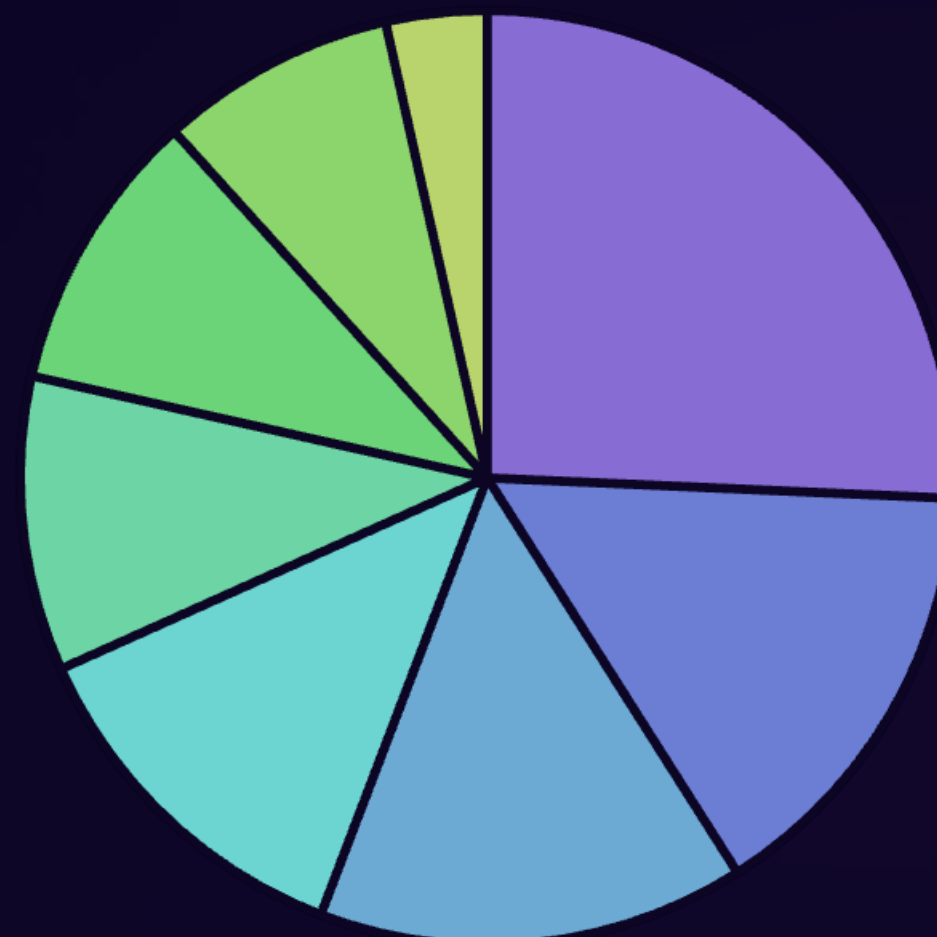
Kepuasan Kerja

Nilai rata-rata kepuasan kerja hanya 2,2 dari skala 5. Mengindikasikan rendahnya kepuasan kerja.

Ketidakhadiran

Rata-rata cuti karena masalah kesehatan mental mencapai 7,4 hari. Mencerminkan dampak signifikan terhadap produktivitas.

Analisis Variabel Kategorikal



■ Perawat ■ ICU ■ Pediatri ■ Medis Umum ■ Layanan Rawat Jalan ■ Laboratorium
■ Administrasi ■ Lainnya

Distribusi kategorikal menunjukkan beban kerja berat sebagai faktor utama (42,8%). Akses terhadap program bantuan karyawan tersedia bagi mayoritas (71,9%).

Korelasi dan Pola Utama

1

Tingkat Stres & Cuti Mental

Hubungan kuat teridentifikasi antara keduanya

2

Faktor Tempat Kerja & Stres

Beban kerja berat berkontribusi signifikan

3

Departemen & Tingkat Stres

ICU dan Medis Umum menunjukkan stres tertinggi

Langkah Persiapan Data

Pembersihan Data
Menghapus kolom tidak relevan
(Employee ID)

Validasi Hasil
Pemeriksaan keseimbangan dan
distribusi data akhir



Encoding Kategorikal
Transformasi variabel kategori
menggunakan LabelEncoder

Normalisasi
Penskalaan data numerik dengan
RobustScaler

Terima
Kasih

