

Beberapa pertanyaan analitik data yang bisa dibuat berdasarkan tabel-tabel tersebut beserta kode Python untuk mendapatkan jawabannya.

Soal 1: Berapa baris dan kolom yang terdapat dalam dataset?

Jawaban :

5 baris dan 3 kolom. Terlihat dari `data.head()`, terdapat 5 baris data yang ditampilkan dan 3 kolom yaitu "usia", "gaji", dan "jenis_kelamin".

Soal 2 : Berapa rata-rata gaji karyawan berdasarkan jenis kelamin?

Jawaban :

Rata-rata gaji pria: 6340, rata-rata gaji wanita: 7280. Hal ini dapat dianalisis dengan menghitung rata-rata gaji berdasarkan jenis kelamin, misalnya menggunakan `data.groupby('jenis_kelamin')['gaji'].mean()`.

Soal 3 : Berapa rentang gaji terkecil dan terbesar dalam dataset?

Jawaban:

Rentang gaji terkecil: 4500, rentang gaji terbesar: 9500. Hal ini dapat dilihat dari `data['gaji'].describe()`, di mana nilai minimum adalah 4500 dan nilai maksimum adalah 9500.

Soal 4 : Berapa persentase karyawan pria dan wanita dalam dataset?

Jawaban :

Persentase karyawan pria: 50%, persentase karyawan wanita: 50%. Hal ini dapat dihitung dengan menghitung jumlah karyawan pria dan wanita kemudian membaginya dengan total karyawan, misalnya menggunakan `(len(data[data['jenis_kelamin'] == 'pria'])/len(data))*100` dan `(len(data[data['jenis_kelamin'] == 'wanita'])/len(data))*100`.

Visualisasi yang dibuat:

- **Scatter Plot:** Menampilkan hubungan antara usia dan gaji, dengan titik-titik berwarna biru yang menunjukkan data karyawan.
- **Histogram:** Menampilkan distribusi gaji karyawan, dengan batang berwarna biru muda yang menunjukkan frekuensi gaji dalam interval tertentu.
- **Box Plot:** Menampilkan distribusi usia karyawan, dengan kotak yang menunjukkan median, kuartil pertama dan ketiga, dan jangkauan data.
- **Barplot:** Menampilkan jumlah karyawan pria dan wanita, dengan batang berwarna biru untuk pria dan pink untuk wanita