

Studi Kasus Exercise: Analisis Data Pinjaman Konsumen (Consumer Loan Analysis)

Latar Belakang

Sebuah lembaga keuangan digital (fintech lending) ingin mengevaluasi dan melakukan segmentasi terhadap nasabah yang mengajukan pinjaman. Dengan meningkatnya jumlah aplikasi, perusahaan ingin memahami tren pinjaman, karakteristik nasabah, dan potensi risiko melalui data yang tersedia.

Deskripsi Dataset: **loan_customers.csv**

Kolom	Deskripsi
loan_id	ID unik pinjaman
customer_id	ID pelanggan
age	Umur nasabah
gender	Jenis kelamin
region	Wilayah domisili
loan_amount	Jumlah pinjaman yang diajukan
loan_purpose	Tujuan pinjaman (misalnya: konsumtif, pendidikan, kesehatan, usaha)
duration_months	Durasi pinjaman dalam bulan
interest_rate	Suku bunga (dalam persen)
approved	Status persetujuan pinjaman (True/False)
application_date	Tanggal pengajuan pinjaman

Tujuan Analisis

- Mengetahui tren pinjaman berdasarkan wilayah dan jenis kelamin.
 - Menganalisis hubungan antara usia, tujuan pinjaman, dan jumlah pinjaman.
 - Mengukur tingkat persetujuan dan pola penolakan pinjaman.
 - Mempersiapkan data untuk visualisasi dan modeling lanjutan.
-

Instruksi Exercise

Bagian 1: Data Preparation

1. Tampilkan informasi umum dataset menggunakan `.info()`, `.shape`, dan `.columns`.
 2. Ubah kolom `application_date` menjadi tipe data `datetime`.
 3. Tambahkan kolom baru `monthly_payment` dengan rumus berikut:
$$\text{monthly_payment} = (\text{loan_amount} * (1 + (\text{interest_rate} / 100))) / \text{duration_month}$$
 4. Periksa apakah ada nilai kosong (missing values) di kolom penting, dan tangani dengan strategi yang tepat.
 5. Hapus nasabah yang berusia di bawah 18 atau di atas 65 tahun.
-

Bagian 2: Analisis dan Agregasi

6. Hitung rata-rata dan median `loan_amount` untuk setiap `loan_purpose`.
 7. Wilayah mana yang memiliki tingkat persetujuan pinjaman (`approved == True`) tertinggi?
 8. Buat ringkasan total pinjaman (`sum`) dan rata-rata suku bunga berdasarkan kombinasi `region` dan `gender`.
 9. Tampilkan 5 nasabah dengan `monthly_payment` tertinggi.
 10. Bandingkan rata-rata bunga dan durasi pinjaman antara pinjaman konsumtif dan pinjaman usaha.
-

Bagian 3: Indexing dan Sorting

11. Jadikan kolom `loan_id` sebagai index.

12. Urutkan data berdasarkan `application_date` dan `loan_amount` secara menurun.
 13. Buat subset MultiIndex berdasarkan kolom `region` dan `loan_purpose`, lalu tampilkan data untuk salah satu kombinasi.
-

Bagian 4: Insight dan Rekomendasi

14. Apakah nasabah perempuan cenderung mengajukan pinjaman dengan jumlah lebih kecil?
 15. Apakah durasi pinjaman yang lebih panjang lebih sering disetujui?
 16. Berikan saran strategi produk pinjaman berdasarkan pola data yang ditemukan.
-

Output yang Diharapkan

- Jupyter Notebook atau Google Colab dengan analisis eksploratif dan dokumentasi markdown.
- Penjelasan setiap langkah dengan komentar kode.
- Penggunaan fungsi Pandas seperti:
 - `.groupby(), .agg(), .sort_values()`
 - `.set_index(), .reset_index()`
 - `.fillna(), .dropna(), .value_counts(), .mean(), .median()`