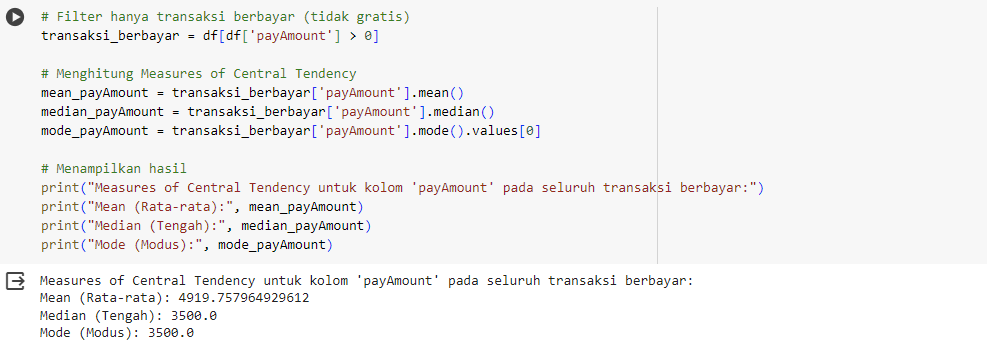
**Tugas Praktik Advanced Statistics**

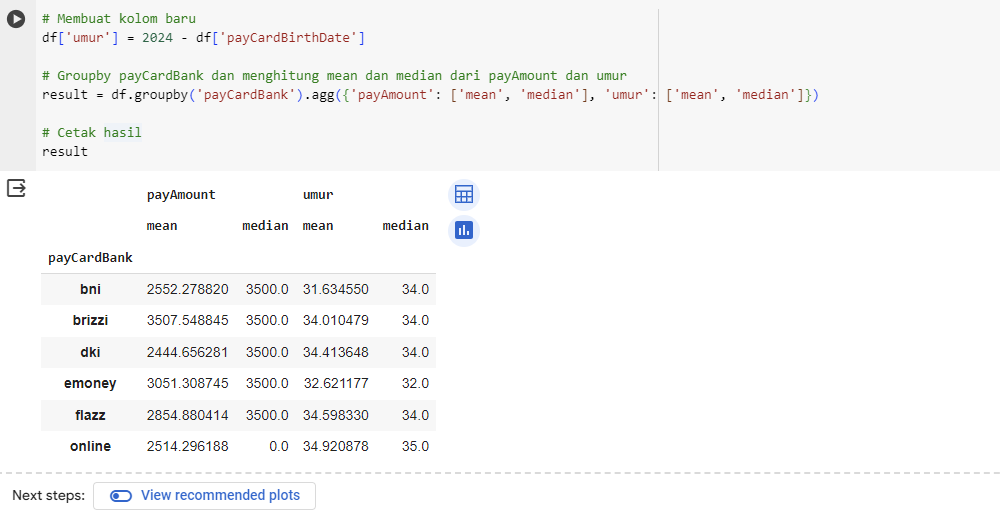
1. Hitunglah semua Measures of Central Tendency untuk kolom payAmountuntuk seluruh transaksi berbayar (tidak gratis)!

Berikut script dan hasilnya :



1. Buatlah satu kolom baru umur yang berisi umur penumpang di tahun 2024 ini. Dengan menggunakan groupby & aggregate function atau pivot table, hitunglah mean dan median payAmount dan umur berdasarkan kelompok payCardBank yang digunakan!

Berikut script dan hasilnya :



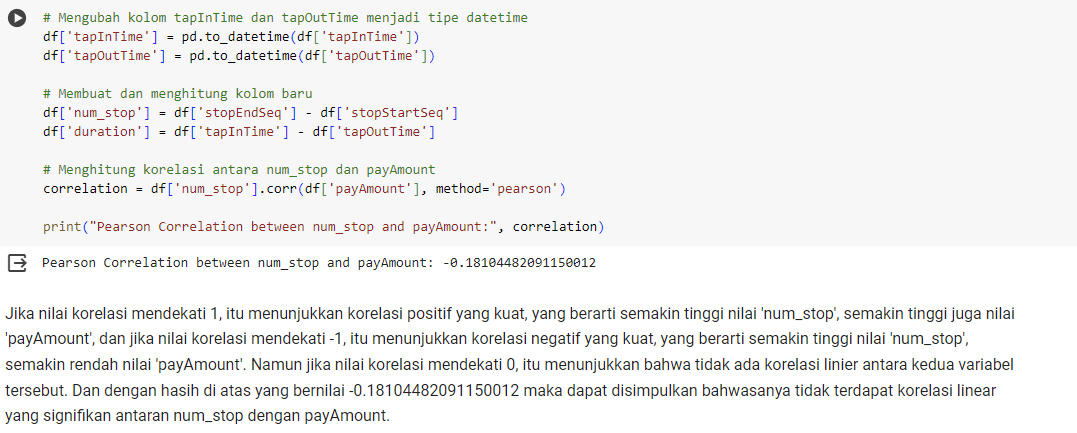
1. Buatlah beberapa kolom baru:

- num\_stop untuk menghitung jumlah pemberhentian (stopEndSeq - stopStartSeq)

- duration untuk menghitung lama perjalanan (tapInTime-tapOutTime)

kemudian gunakan method corr() pada dataframe (dokumentasi dan contoh penggunaan dapat dilihat pada link berikut) untuk menghitung pearson correlation antara num\_stop dan payAmount kemudian berikan penjelasan dari hasil nilai correlation tersebut.

Berikut script dan hasilnya :



1. Buatlah 2 dataframe baru yang berisi 600 sample data dengan kriteria berikut:

A. Sampling menggunakan random sample method sample()

B. Sampling menggunakan stratified sampling untuk kelompok payCardBank, hint : dapat menggunakan kombinasi groupby dan lambda

Kemudian hitung distribusi banyaknya jumlah transaksi (dalam percentage) di setiap kelompok payCardBank untuk dataframe original (sebelum di sampling), hasil random sampling, dan hasil stratified sampling. Dan tuliskan penjelasan apa yang dapat kamu simpulkan dari hasil tersebut.

Berikut script dan hasilnya :

