Laporan Tugas Akhir

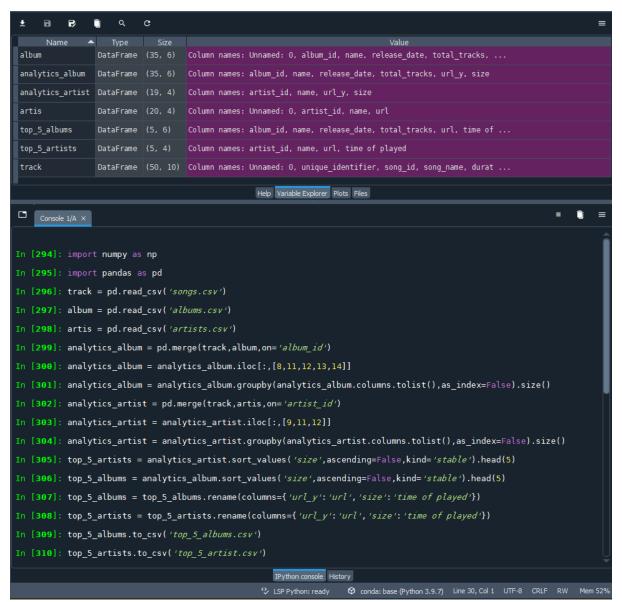
Cakap : Belajar Mengolah Data untuk Calon Data Engineer

Avif Maulana Azis

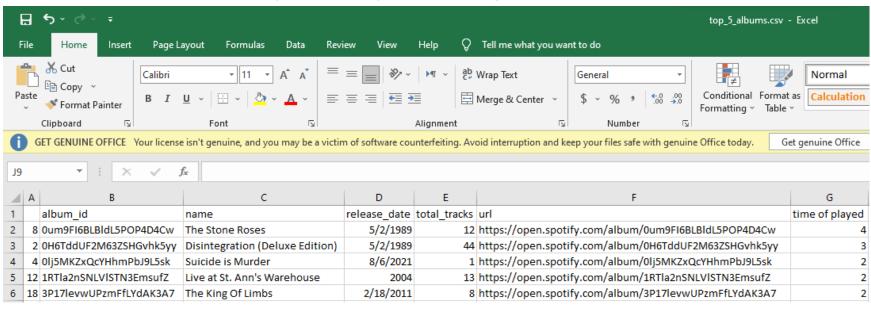
Kode Python

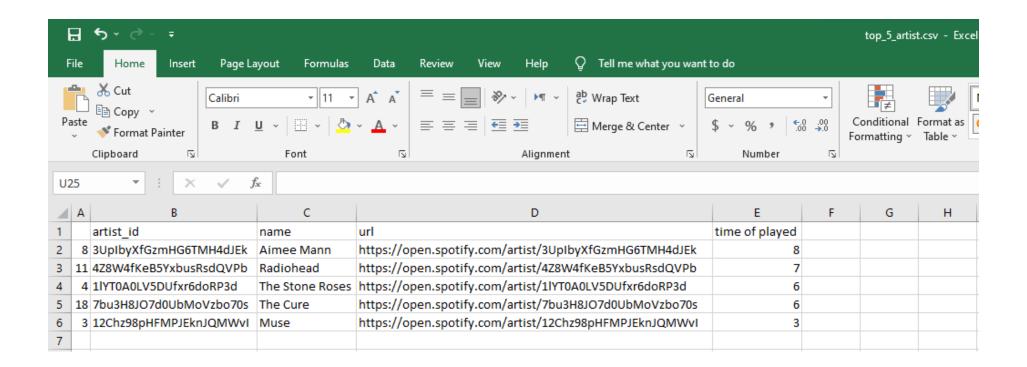
```
Created on Fri Nov 11 06:13:45 2022
@author: avifa
import numpy as np
import pandas as pd
track = pd.read csv('songs.csv')
album = pd.read_csv('albums.csv')
artis = pd.read csv('artists.csv')
analytics_album = pd.merge(track,album,on='album id')
analytics album = analytics album.iloc[:,[8,11,12,13,14]]
analytics album =
analytics album.groupby(analytics album.columns.tolist(),as index=False).
analytics artist = pd.merge(track,artis,on='artist id')
analytics artist = analytics artist.iloc[:,[9,11,12]]
analytics artist =
analytics_artist.groupby(analytics_artist.columns.tolist(),as index=False
).size()
top 5 artists =
analytics_artist.sort_values('size',ascending=False,kind='stable').head(5
top 5 albums =
analytics album.sort values('size',ascending=False,kind='stable').head(5)
top 5 albums = top 5 albums.rename(columns={'url y':'url','size':'time of
top 5 artists = top 5 artists.rename(columns={'url y':'url','size':'time
of played'})
top 5 albums.to csv('top 5 albums.csv')
top 5 artists.to csv('top 5 artist.csv')
```

Screenshot Terminal



File CSV (tidak bisa input banyak data, jadi diScreenshot disini)





Analisa perbedaan Sort_values



Karena sejauh ini yang paling stabil adalah jenis stable dan margesort, maka saya menggunakan jenis stable dalam mengurutkan data mengingat ada nilai yang sama (jumlah kemunculan yang sama).