

# Laporan Tugas Akhir

Cakap : Belajar Mengolah Data untuk Calon Data Engineer

Avif Maulana Azis

## Kode Python

```
# -*- coding: utf-8 -*-
"""
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).
size()
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
played'})
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

The screenshot displays a JupyterLab environment. The top section is the 'Variable Explorer' showing a list of variables and their details.

Name	Type	Size	Value
album	DataFrame	(35, 6)	Column names: Unnamed: 0, album_id, name, release_date, total_tracks, ...
analytics_album	DataFrame	(35, 6)	Column names: album_id, name, release_date, total_tracks, url_y, size
analytics_artist	DataFrame	(19, 4)	Column names: artist_id, name, url_y, size
artis	DataFrame	(20, 4)	Column names: Unnamed: 0, artist_id, name, url
top_5_albums	DataFrame	(5, 6)	Column names: album_id, name, release_date, total_tracks, url, time of ...
top_5_artists	DataFrame	(5, 4)	Column names: artist_id, name, url, time of played
track	DataFrame	(50, 10)	Column names: Unnamed: 0, unique_identifier, song_id, song_name, durat ...

Below the variable explorer is the 'Console 1/A' terminal window, which contains the following Python code:

```
In [294]: import numpy as np
In [295]: import pandas as pd
In [296]: track = pd.read_csv('songs.csv')
In [297]: album = pd.read_csv('albums.csv')
In [298]: artis = pd.read_csv('artists.csv')
In [299]: analytics_album = pd.merge(track, album, on='album_id')
In [300]: analytics_album = analytics_album.iloc[:, [8, 11, 12, 13, 14]]
In [301]: analytics_album = analytics_album.groupby(analytics_album.columns.tolist(), as_index=False).size()
In [302]: analytics_artist = pd.merge(track, artis, on='artist_id')
In [303]: analytics_artist = analytics_artist.iloc[:, [9, 11, 12]]
In [304]: analytics_artist = analytics_artist.groupby(analytics_artist.columns.tolist(), as_index=False).size()
In [305]: top_5_artists = analytics_artist.sort_values('size', ascending=False, kind='stable').head(5)
In [306]: top_5_albums = analytics_album.sort_values('size', ascending=False, kind='stable').head(5)
In [307]: top_5_albums = top_5_albums.rename(columns={'url_y': 'url', 'size': 'time of played'})
In [308]: top_5_artists = top_5_artists.rename(columns={'url_y': 'url', 'size': 'time of played'})
In [309]: top_5_albums.to_csv('top_5_albums.csv')
In [310]: top_5_artists.to_csv('top_5_artist.csv')
```

The bottom status bar indicates the environment is 'LSP Python: ready', using 'conda: base (Python 3.9.7)', at 'Line 30, Col 1', with 'UTF-8' encoding, 'CRLF' line endings, 'RW' permissions, and 'Mem 52%' usage.

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

top\_5\_albums.csv - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Cut Copy Paste Format Painter Clipboard

Calibri 11 A A B I U Font

Wrap Text Merge & Center Alignment

General Number \$ % , .0 .00

Conditional Formatting Format as Table Normal Calculation

GET GENUINE OFFICE Your license isn't genuine, and you may be a victim of software counterfeiting. Avoid interruption and keep your files safe with genuine Office today. Get genuine Office

J9

	A	B	C	D	E	F	G
1		album_id	name	release_date	total_tracks	url	time of played
2	8	0um9FI6BLBldL5POP4D4Cw	The Stone Roses	5/2/1989	12	https://open.spotify.com/album/0um9FI6BLBldL5POP4D4Cw	4
3	2	0H6TddUF2M63ZSHGvhk5yy	Disintegration (Deluxe Edition)	5/2/1989	44	https://open.spotify.com/album/0H6TddUF2M63ZSHGvhk5yy	3
4	4	0lj5MKZxQcYHhmPbJ9L5sk	Suicide is Murder	8/6/2021	1	https://open.spotify.com/album/0lj5MKZxQcYHhmPbJ9L5sk	2
5	12	1RTla2nSNLVISTN3EmsufZ	Live at St. Ann's Warehouse	2004	13	https://open.spotify.com/album/1RTla2nSNLVISTN3EmsufZ	2
6	18	3P17levwUPzmFfLYdAK3A7	The King Of Limbs	2/18/2011	8	https://open.spotify.com/album/3P17levwUPzmFfLYdAK3A7	2

top\_5\_artist.csv - Excel

File

Home

Insert

Page Layout

Formulas

Data

Review

View

Help

Tell me what you want to do

Paste

Cut

Copy

Format Painter

Clipboard

Calibri

11

A

A

B

I

U

Font

Alignment

ab

Wrap Text

Merge & Center

General

\$

%

Number

Conditional Formatting

Format as Table

U25

	A	B	C	D	E	F	G	H
1		artist_id	name	url	time of played			
2	8	3UplbyXfGzmHG6TMH4dJEk	Aimee Mann	https://open.spotify.com/artist/3UplbyXfGzmHG6TMH4dJEk	8			
3	11	4Z8W4fKeB5YxbusRsdQVPb	Radiohead	https://open.spotify.com/artist/4Z8W4fKeB5YxbusRsdQVPb	7			
4	4	1lYT0A0LV5DUfxr6doRP3d	The Stone Roses	https://open.spotify.com/artist/1lYT0A0LV5DUfxr6doRP3d	6			
5	18	7bu3H8JO7d0UbMoVzbo70s	The Cure	https://open.spotify.com/artist/7bu3H8JO7d0UbMoVzbo70s	6			
6	3	12Chz98pHFMPJEknJQMWvI	Muse	https://open.spotify.com/artist/12Chz98pHFMPJEknJQMWvI	3			
7								

# Analisa perbedaan Sort\_values

**merge sort**

Index	album id	name	release date	total track	url v
8	0um9FI6BLBldLSP0P4D4Cw	The Stone Roses	1989-05-02	12	https://open.spotify.com/album/0um9FI6BLBldLSP0P4D4Cw
2	0H6TddUF2M63ZSHGvhk5yy	Disintegration (Deluxe Edition)	1989-05-02	44	https://open.spotify.com/album/0H6TddUF2M63ZSHGvhk5yy
4	0lj5MKZxQcYHhmPbJ9L5sk	Suicide is Murder	2021-08-06	1	https://open.spotify.com/album/0lj5MKZxQcYHhmPbJ9L5sk
12	1RTla2nSNLVLSTN3EmsufZ	Live at St. Ann's Warehouse	2004	13	https://open.spotify.com/album/1RTla2nSNLVLSTN3EmsufZ
18	3P17levwUPzmFfLYdAK3A7	The King Of Limbs	2011-02-18	8	https://open.spotify.com/album/3P17levwUPzmFfLYdAK3A7

**heap sort**

Index	album id	name	release date	total track	url v
8	0um9FI6BLBldLSP0P4D4Cw	The Stone Roses	1989-05-02	12	https://open.spotify.com/album/0um9FI6BLBldLSP0P4D4Cw
2	0H6TddUF2M63ZSHGvhk5yy	Disintegration (Deluxe Edition)	1989-05-02	44	https://open.spotify.com/album/0H6TddUF2M63ZSHGvhk5yy
31	6dVIqQ8qmQ5GBnJ9sh0YGE	OK Computer	1997-05-28	12	https://open.spotify.com/album/6dVIqQ8qmQ5GBnJ9sh0YGE
18	3P17levwUPzmFfLYdAK3A7	The King Of Limbs	2011-02-18	8	https://open.spotify.com/album/3P17levwUPzmFfLYdAK3A7
30	61RmMONNMRAvBaWiAJSu3z	Grace/Wastelands	2009-03-11	12	https://open.spotify.com/album/61RmMONNMRAvBaWiAJSu3z

**Stable**

Index	album id	name	release date	total track	url v
8	0um9FI6BLBldLSP0P4D4Cw	The Stone Roses	1989-05-02	12	https://open.spotify.com/album/0um9FI6BLBldLSP0P4D4Cw
2	0H6TddUF2M63ZSHGvhk5yy	Disintegration (Deluxe Edition)	1989-05-02	44	https://open.spotify.com/album/0H6TddUF2M63ZSHGvhk5yy
4	0lj5MKZxQcYHhmPbJ9L5sk	Suicide is Murder	2021-08-06	1	https://open.spotify.com/album/0lj5MKZxQcYHhmPbJ9L5sk
12	1RTla2nSNLVLSTN3EmsufZ	Live at St. Ann's Warehouse	2004	13	https://open.spotify.com/album/1RTla2nSNLVLSTN3EmsufZ
18	3P17levwUPzmFfLYdAK3A7	The King Of Limbs	2011-02-18	8	https://open.spotify.com/album/3P17levwUPzmFfLYdAK3A7

**Quick sort**

Index	album id	name	release date	total track	url v
8	0um9FI6BLBldLSP0P4D4Cw	The Stone Roses	1989-05-02	12	https://open.spotify.com/album/0um9FI6BLBldLSP0P4D4Cw
2	0H6TddUF2M63ZSHGvhk5yy	Disintegration (Deluxe Edition)	1989-05-02	44	https://open.spotify.com/album/0H6TddUF2M63ZSHGvhk5yy
34	7zJzNs8eVgbkVVSQSwKRtx	The Head On The Door	1985-08-26	28	https://open.spotify.com/album/7zJzNs8eVgbkVVSQSwKRtx
30	61RmMONNMRAvBaWiAJSu3z	Grace/Wastelands	2009-03-11	12	https://open.spotify.com/album/61RmMONNMRAvBaWiAJSu3z
18	3P17levwUPzmFfLYdAK3A7	The King Of Limbs	2011-02-18	8	https://open.spotify.com/album/3P17levwUPzmFfLYdAK3A7

Format: Resize ☒ Background color ☒ Column min/max Save and Close Close

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