

## EDA Analysis on Python

Promotes loose coupling between components by proving object

### Install Dependencies

```
Oct 20, 2025 (4s) 2 Python Generate (Ctrl + I)

#### To install and manage soft were packages
!pip install openpyxl -q

Note: you may need to restart the kernel using %restart_python or dbutils.library.restartPython() to use updated packages.
```

Assist offer specialized function

### Import libraries

We have the data of the viewership analysis of the movies and tv shows.

The total time watched by the viewers. The platform is the platform where the movies and tv shows are watched. The play event type is the type of event that is played. The video title is the title of the movie or tv show.

## meta data

```
Oct 20, 2025 (<1s) 5 Python

#### Importing panda help us process data
import pandas as dp

#### Importing numpy library help us to perform matematical operations
import numpy as np

Oct 20, 2025 (<1s) 6

#### To access the file
data_path = "/Workspace/Users/mavismbr@gmail.com/Viewership Analysis .xlsx"

Oct 20, 2025 (11s) 7

#### Attaching the excel file

survey_analysis = dp.read_excel(data_path)

survey_analysis: pandas.core.frame.DataFrame = [DateID: int64, CustomerID: object ... 4 more fields]

Oct 20, 2025 (2s) 8

#### display data given
```

▶

✓

Oct 20, 2025 (2s)

8

Python

🗑️

🔗

📄

⋮

#### display data given

display(survey\_analysis)

> [See performance \(1\)](#)

Generate (Ctrl + I)

Optimize

Table

+

🔍

🏠

📄

📄

|    | <div>DateID</div> | <div>CustomerID</div>          | <div>TotalTimeWatched</div> | <div>Platform</div> | <div>PlayEventType</div> | <div>VideoTitle</div>         |  |
|----|-------------------|--------------------------------|-----------------------------|---------------------|--------------------------|-------------------------------|--|
| 1  | 20201101          | EW1DENH0EC1J3M9WAOZF9LSV...    | 300                         | Leanback            | LiveTV                   | F1 '20: Emilia Romagna GP     |  |
| 2  | 20201101          | EW1DENH0EC1J3M9WAOZF9LSV...    | 300                         | Leanback            | LiveTV                   | F1 '20: Emilia Romagna GP     |  |
| 3  | 20201101          | 6TS2LLY0L3G66FVYB6Q0JEZE000K   | 360                         | Leanback            | Other                    | Chasing The Sun               |  |
| 4  | 20201101          | 6TS2LLY0L3G66FVYB6Q0JEZE000K   | 360                         | Leanback            | Other                    | Chasing The Sun               |  |
| 5  | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 120                         | Leanback            | LiveTV                   | Sonic The Hedgehog            |  |
| 6  | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 120                         | Leanback            | LiveTV                   | Frozen II                     |  |
| 7  | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 120                         | iOS                 | LiveTV                   | PL 20/21: Man Utd v Arsenal   |  |
| 8  | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 120                         | iOS                 | LiveTV                   | PL 20/21: Man Utd v Arsenal   |  |
| 9  | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 60                          | iOS                 | LiveTV                   | F1 '20: Emilia Romagna GP     |  |
| 10 | 20201101          | 6PMV67PLJ2S47S6BJ0Y30XFK003C   | 600                         | iOS                 | LiveTV                   | F1 '20: Emilia Romagna GP     |  |
| 11 | 20201101          | JMCVLXSA196YJ0KC6IDOIEVB000... | 1860                        | iOS                 | LiveTV                   | Betrayed                      |  |
| 12 | 20201101          | JMCVLXSA196YJ0KC6IDOIEVB000... | 960                         | iOS                 | Catch Up                 | The Bradshaw Bunch            |  |
| 13 | 20201101          | JMCVLXSA196YJ0KC6IDOIEVB000... | 1080                        | iOS                 | Catch Up                 | Mbau Reloaded: Always Rise    |  |
| 14 | 20201101          | JMCVLXSA196YJ0KC6IDOIEVB000... | 240                         | iOS                 | Catch Up                 | The Bradshaw Bunch            |  |
| 15 | 20201101          | JMCVLXSA196YJ0KC6IDOIEVB000... | 120                         | iOS                 | LiveTV                   | PL 20/21: Newcastle v Everton |  |

📄

📄

10,000+ rows | Truncated data | 1.70s runtime

Refreshed 13 days ago

▶

✓

Oct 20, 2025 (<1s)

9

#### display given number of rows and columns

survey\_analysis.shape

(118534, 6)

▶

✓

Oct 20, 2025 (<1s)

10

#### display Checking the data types for the columns

survey\_analysis.dtypes

DateID int64  
CustomerID object  
TotalTimeWatched int64  
Platform object  
PlayEventType object  
VideoTitle object  
dtype: object

▶

✓

Oct 20, 2025 (<1s)

11

#### display the columns

survey\_analysis.columns

Index(['DateID', 'CustomerID', 'TotalTimeWatched', 'Platform', 'PlayEventType',  
'VideoTitle'],  
dtype='object')

▶

✓

Oct 19, 2025 (<1s)

12

#### display data information

survey\_analysis.info()

<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 118534 entries, 0 to 118533  
Data columns (total 6 columns):  
# Column Non-Null Count Dtype  
---  
0 DateID 118534 non-null int64  
1 CustomerID 118534 non-null object  
2 TotalTimeWatched 118534 non-null int64  
3 Platform 118534 non-null object  
4 PlayEventType 118534 non-null object  
5 VideoTitle 118534 non-null object  
dtypes: int64(2), object(4)  
memory usage: 5.4+ MB

▶

✓ Oct 19, 2025 (<1s)

13

#### display unique values for the column 'Platform'

survey\_analysis['Platform'].unique()

array(['Leanback', 'IOS', 'Web', 'Android'], dtype=object)

▶

✓ Oct 19, 2025 (<1s)

14

#### display unique values for the column 'PlayEventType'

survey\_analysis['PlayEventType'].unique()

array(['LiveTV', 'Other', 'Catch Up', 'Downloads'], dtype=object)

▶

✓ Oct 19, 2025 (<1s)

15

### display minimum time watched for the column 'TotalTimeWatched'

survey\_analysis['TotalTimeWatched'].min()

np.int64(1)

▶

✓ Oct 19, 2025 (<1s)

16

#### Check maximum time watched for the column 'TotalTimeWatched'

survey\_analysis['TotalTimeWatched'].max()

▶

✓ Oct 19, 2025 (<1s)

17

Python

#### display the summary of the data

survey\_analysis.describe

<bound method NDFrame.describe of DateID ... VideoTitle  
0 20201101 ... F1 '20: Emilia Romagna GP  
1 20201101 ... F1 '20: Emilia Romagna GP  
2 20201101 ... Chasing The Sun  
3 20201101 ... Chasing The Sun  
4 20201101 ... Sonic The Hedgehog  
...  
118529 20210416 ... New Day  
118530 20210416 ... Buying And Selling With Property...  
118531 20210416 ... Espresso  
118532 20210416 ... Unknown  
118533 20210416 ... Imbewu  
[118534 rows x 6 columns]>

▶

✓ Oct 19, 2025 (<1s)

18

#### display the first 5 rows

survey\_analysis.head (5)

|   | DateID   | CustomerID                   | TotalTimeWatched | Platform | PlayEventType | VideoTitle                |
|---|----------|------------------------------|------------------|----------|---------------|---------------------------|
| 0 | 20201101 | EW1DENH0EC1J3M9WAOZF9LSV004O | 300              | Leanback | LiveTV        | F1 '20: Emilia Romagna GP |
| 1 | 20201101 | EW1DENH0EC1J3M9WAOZF9LSV004O | 300              | Leanback | LiveTV        | F1 '20: Emilia Romagna GP |
| 2 | 20201101 | 6TS2LLY0L3G66FY86Q0JEZE000K  | 360              | Leanback | Other         | Chasing The Sun           |
| 3 | 20201101 | 6TS2LLY0L3G66FY86Q0JEZE000K  | 360              | Leanback | Other         | Chasing The Sun           |
| 4 | 20201101 | 6PMV67PLJ2S47S68J0Y30XFK003C | 120              | Leanback | LiveTV        | Sonic The Hedgehog        |

▶

✓ Oct 19, 2025 (<1s)

19

Python

#### display the last 5 rows

survey\_analysis.tail (5)

|        | DateID   | CustomerID                    | TotalTimeWatched | Platform | PlayEventType | VideoTitle                          |
|--------|----------|-------------------------------|------------------|----------|---------------|-------------------------------------|
| 118529 | 20210416 | DGXM5TJ18W1LANOLKJUN5MFT000G  | 600              | Leanback | LiveTV        | New Day                             |
| 118530 | 20210416 | DDMMVDBZFDT1F7WL4BGWQ5K0J000G | 5194             | Web      | LiveTV        | Buying And Selling With Property... |
| 118531 | 20210416 | 1JKBIAY7MEUH6BWX0SCU93GJ000I  | 47660            | Web      | LiveTV        | Espresso                            |
| 118532 | 20210416 | 1ADX18R3B33REZ270NK911E4002W  | 1335             | Leanback | LiveTV        | Unknown                             |
| 118533 | 20210416 | BO0REPHA6FP7C15B0K9N1GVO0050  | 360              | IOS      | LiveTV        | Imbewu                              |

▶

✓

Oct 19, 2025 (<1s)

20

Python

🗑️

🔗

📄

⋮

Generate (Ctrl + I)

```
#### display the duplicate row in the data

survey_analysis.duplicated
```

|        | bound    | method | DataFrame.duplicated | of                                  | DateID | ... | VideoTitle |
|--------|----------|--------|----------------------|-------------------------------------|--------|-----|------------|
| 0      | 20201101 | ...    |                      | F1 '20: Emilia Romagna GP           |        |     |            |
| 1      | 20201101 | ...    |                      | F1 '20: Emilia Romagna GP           |        |     |            |
| 2      | 20201101 | ...    |                      | Chasing The Sun                     |        |     |            |
| 3      | 20201101 | ...    |                      | Chasing The Sun                     |        |     |            |
| 4      | 20201101 | ...    |                      | Sonic The Hedgehog                  |        |     |            |
| ...    | ...      | ...    |                      | ...                                 |        |     |            |
| 118529 | 20210416 | ...    |                      | New Day                             |        |     |            |
| 118530 | 20210416 | ...    |                      | Buying And Selling With Property... |        |     |            |
| 118531 | 20210416 | ...    |                      | Espresso                            |        |     |            |
| 118532 | 20210416 | ...    |                      | Unknown                             |        |     |            |
| 118533 | 20210416 | ...    |                      | Imbewu                              |        |     |            |

[118534 rows x 6 columns]>

▶

✓

Oct 19, 2025 (<1s)

21

Python

🗑️

🔗

📄

⋮

Generate (Ctrl + I)

```
#### display the duplicate row in the data if its true/false

survey_analysis.duplicated()
```

|        |       |
|--------|-------|
| 0      | False |
| 1      | True  |
| 2      | False |
| 3      | True  |
| 4      | False |
| ...    | ...   |
| 118529 | False |
| 118530 | False |
| 118531 | False |
| 118532 | False |
| 118533 | False |

Length: 118534, dtype: bool

▶

✓

Oct 19, 2025 (<1s)

22

Python

🗑️

🔗

📄

⋮

Generate (Ctrl + I)

```
#### drop all the duplicate row

survey_analysis.drop_duplicates()
```

|        | DateID   | CustomerID                   | TotalTimeWatched | Platform | PlayEventType | VideoTitle                          |
|--------|----------|------------------------------|------------------|----------|---------------|-------------------------------------|
| 0      | 20201101 | EW1DENH0EC1J3M9WAOZF9LSV004O | 300              | Leanback | LiveTV        | F1 '20: Emilia Romagna GP           |
| 2      | 20201101 | 6TS2LLYDL3G66FVY86QQJEZE000K | 360              | Leanback | Other         | Chasing The Sun                     |
| 4      | 20201101 | 6PMV67PLJ2S47S68J0Y30XFK003C | 120              | Leanback | LiveTV        | Sonic The Hedgehog                  |
| 5      | 20201101 | 6PMV67PLJ2S47S68J0Y30XFK003C | 120              | Leanback | LiveTV        | Frozen II                           |
| 6      | 20201101 | 6PMV67PLJ2S47S68J0Y30XFK003C | 120              | iOS      | LiveTV        | PL 20/21: Man Utd v Arsenal         |
| ...    | ...      | ...                          | ...              | ...      | ...           | ...                                 |
| 118529 | 20210416 | DGXM5TJ18W1LANOLKJUN5MFT000G | 600              | Leanback | LiveTV        | New Day                             |
| 118530 | 20210416 | DMMVDBZFD1F7WL4BGWQ5KOJ000G  | 5194             | Web      | LiveTV        | Buying And Selling With Property... |
| 118531 | 20210416 | 1JKBIAY7MEUH6BWX0SCU93GJ000I | 47660            | Web      | LiveTV        | Espresso                            |
| 118532 | 20210416 | 1ADX18R3B33REZ2T0NK911E4002W | 1335             | Leanback | LiveTV        | Unknown                             |
| 118533 | 20210416 | B00REPHa6FP7C15B0K9N1GVO0050 | 360              | iOS      | LiveTV        | Imbewu                              |

108661 rows x 6 columns

▶

✓

Oct 20, 2025 (<1s)

23

Python

🗑️

🔗

📄

⋮

Generate (Ctrl + I)

```
#### Display the null value

survey_analysis.isnull()
```

|        | DateID | CustomerID | TotalTimeWatched | Platform | PlayEventType | VideoTitle |
|--------|--------|------------|------------------|----------|---------------|------------|
| 0      | False  | False      | False            | False    | False         | False      |
| 1      | False  | False      | False            | False    | False         | False      |
| 2      | False  | False      | False            | False    | False         | False      |
| 3      | False  | False      | False            | False    | False         | False      |
| 4      | False  | False      | False            | False    | False         | False      |
| ...    | ...    | ...        | ...              | ...      | ...           | ...        |
| 118529 | False  | False      | False            | False    | False         | False      |
| 118530 | False  | False      | False            | False    | False         | False      |
| 118531 | False  | False      | False            | False    | False         | False      |
| 118532 | False  | False      | False            | False    | False         | False      |
| 118533 | False  | False      | False            | False    | False         | False      |

118534 rows x 6 columns