NURTURE
Education Solutions
TOMORROW'S HERE

MCA SEM-III
2023-2024

DATA SIENCE (Intermediate Level) –
DATA MANIPULATION WITH PANDAS

By,
Dr.Anupriya Kamble,
Senior Faculty-IT
ADYPU, Pune.

DATA MANIPULATION WITH PANDAS

INSTALLATION OF PANDAS

Create a new folder

Right click -> Open Powershell windows here

pip install pandas
pip install numpy

--upgrade pandas

pip install jupyter

DATA MANIPULATION WITH PANDAS

Jupyter

Python should be installed.
Jupyter will be launched and notebook server will be opened.

DATA MANIPULATION WITH PANDAS

Dictionary creation

```
import numpy as np
import pandas as pd
Dict={“roll no.”:[‘1’, ‘2’, ‘3’],
            “name”:[‘A’, ‘B’, ‘C’],
             “marks”:[’18’, ‘20’, ‘9’]}
marklist=pd.dataframe(dict)
//data frame works as an excel sheet. Numpy can be used. Analyse of big data is possible.

0, 1, 2, 3 is the index

//to export in csv
Marklist.to_csv(‘abc.csv’)

//In home of jupyter csv file will be created and the data entered in the dictionary is saved in this file.
```

DATA MANIPULATION WITH PANDAS

Dictionary creation

// If index is not to be visible again and again
Marklist.to_csv('abc.csv', index=False)

//Pandas is used for Data Analysis

//To view first one rows in the data frame
marklist.head(1)

//To view last one row in the data frame
marklist.tail(1)

//To view statistical analysis numerical columns
marklist.describe()

DATA MANIPULATION WITH PANDAS

Dictionary creation

// If index is not to be visible again and again
Marklist.to_csv('abc.csv', index=False)

//Pandas is used for Data Analysis

//To view first one rows in the data frame
marklist.head(1)

//To view last one row in the data frame
marklist.tail(1)

//To view statistical analysis numerical columns
marklist.describe()2

DATA MANIPULATION WITH PANDAS

Dropping Entries Indexing

```
// To drop an index from a dataframe
import pandas as pd

df=pd.read_csv(abc.csv)
df //Print

df.index //represents the index

df=df.style.hide_index()
df //Print
```

DATA MANIPULATION WITH PANDAS

Selecting rows in dataframes

```
// To select rows in a dataframe
import pandas as pd

df=pd.read_csv(abc.csv)
df.head//Print the columns in the dataframe

new_df=df[df[‘  ‘]>value] //select rows for greater than function
New_df //Print

new_df=df[df[‘column‘]<value] //select rows for smaller than function
New_df //Print

//Using Loc method
New_df=df.loc[df[‘column’]>value] //Greater
New_df=df.loc[df[‘column’]<value] //smaller
New_df=df.loc[df[‘column’]==value] //equal to
```

DATA MANIPULATION WITH PANDAS

Selecting rows in dataframes

```
// To select rows in a dataframe
import pandas as pd

df=pd.read_csv(abc.csv)
df.head//Print the columns in the dataframe

//isin method helps to select rows whose value is defined.
Options=[‘definevalue1’, ‘definevalue2’]
new_df=df[df[‘columnname’].isin(options)]
new_df


//isin method using loc
new_df=df.loc[df[‘columnname’].isin[(options)]
```

DATA MANIPULATION WITH PANDAS

References

1. https://www.geeksforgeeks.org/data-manipulattion-in-python-using-pandas/

THANK YOU