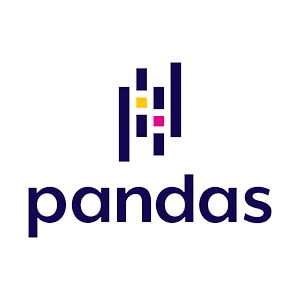
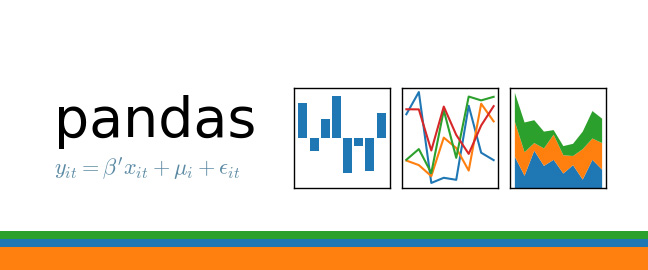


**Advanced Python Programming**

******Unit 5: Analyzing data with python**

**Aghera Kevin**

**---------------------------------------------------------------------------------------------------------------------------------------------------------**

In this chapter we are going to learn with basics of analyzing data using pandas and numpy module

**Getting data into and out of python** – Saving and loading data from csv, excel files.

**Preparing data** – cleaning and organizing data

**Visualizing data** – Line plot, Bar plot, Box plot, Histogram, Pie chart, Scatter plot

**---------------------------------------------------------------------------------------------------------------------------------------------------------**

For data analytics in python first we have to import pandas module inside our python file

Pandas

pip install pandas

For importing we will do **import pandas as pd**, generally pandas module was import as name pd

**How to read csv, excel, sql,…. etc files (importing data) in python?**

pd.read\_csv(filename)

pd.read\_excel(filename, sheet\_name=””)

df.head() first 5 records

df.tail() last 5 records

df.shape return tuple (rows,columns) count

to access column of dataframe

dataframe.column name e.g. df.rno

**pd.read\_csv(filename)** - From a CSV file

**pd.read\_table(filename)** - From a delimited text

file (like TSV)

**pd.read\_excel(filename)** - From an Excel file

**pd.read\_sql(query, connection\_object)** -

Reads from a SQL table/database

**pd.read\_json(json\_string)** - Reads from a JSON

formatted string, URL or file.

**pd.read\_html(url)** - Parses an html URL, string or

file and extracts tables to a list of dataframes

**pd.read\_clipboard()** - Takes the contents of your

clipboard and passes it to **read\_table()**

**pd.DataFrame(dict)** - From a dict, keys for

columns names, values for data as lists