3] BigData Fundamentals With Pyspark (M1 Introduction to Big Data analysis with Spark

Review of functional programming in Python (M1->SM3)

1) Download RAR on STAGING\_AREA

2) Extract the RAR

3) Put the .ipynb files in UBUNTU\_HOME/test-jupyter/P2/M1/SM3

Import in Jupyter Notebook and implement

* Code/Dataset

./day2/3\_ReviewOfFunctionalProgramming.zip

Time = 15 Mins.

4] Big Data Fundamentals with PySpark (M2 Programming in PySpark RDD’s

Abstracting Data with RDDs (M2->SM1)

1) Download RAR on STAGING\_AREA

2) Extract the RAR

3) Put the \*.ipynb files in UBUNTU\_HOME/test-jupyter/P2/M2/SM1

4) Import the notes in Jupyter Note book

5) Follow the instructions and give the solutions

* Code/Dataset

./day2/1\_AbstractingDatawithRDDs.zip

Time == 20 Mins

5] Big Data Fundamentals with PySpark (M2 Programming in PySpark RDD’s )

Basic RDD Transformations and Actions (M2 -> SM2)

1) Download RAR on STAGING\_AREA

2) Extract the RAR

3) Put the \*.ipynb files in UBUNTU\_HOME/test-jupyter/P2/M2/sm2

4) Import the notes in Jupyter Notebook

5) Follow the instructions and give the solutions

* Code/Dataset

./day2/2\_BasicRDDTransformationsandActions.zip

**TimeLine** = 20 Minutes