GOVERNMENT COLLEGE
KASARAGOD

Open Data Lab
GCK

Micro-course in

DATA ANALYSIS

Session – 3
Introduction to NumPy & Pandas

HIRAN DEVA K

Open Data Lab
GCK

PANDAS

- Open source library
- Built on top of NumPy
- Offers various data structures and operations

Advantages of Pandas:

- Fast and efficient for manipulating and analyzing data
- Data from different file object can be loaded
- Easy handling of missing values
- Size mutability
- Data set merging and joining

- Importing library
- Types of data structures that Pandas package have:

1) **Series**: One dimensional array holding data of any type.
   (like a column in a table)
2) **Data frames**: Two dimensional data structure, like a two dimensional array, which can hold different data types.
   (like a table with rows and columns)

1. Creating a DataFrame using Dictionary
2. Extracting column
3. Locating row
4. Named indexes
5. Locating named indexes
6. Add a new column
7. Delete a column
8. Size: returns an int representing number of elements in the object
9. Exporting DataFrame
10. Loading files into DataFrame

print(‘Thank You’)

Thank You

Open Data Lab

GCK