Name :- Akash . A . Banage .
PRN :- 2024011200 75 Roll No :- CS86
Batch :- CS84

Page No. .
Date

Unit - 4

Significance of Pandas

a) working
b) Data manipulation .
c) Data exploration & analysis
d) Handling missing data .
e) Data integeration
f) time series analysis .
g) Data preparation for ML .
h) Data loading ⟶ csv / Json file .

↓

Ex. Code / sample code

i) Data storage ⟶ csv / Json / Excel file example

Ans a) working :- Pandas provide a fast and efficient dataframe structure for handeling and manipulating data, making it easier to work with structured datasets. It supports row and column-based operations for structured datasets .

b) Data Manipulation :- Pandas allows data cleaning, transformation and aggregation using powerful functions like filtering, sorting, merging and grouping. It supports label-based indexing for structured manipulation .

c) Data exploration & analysis :- It enables expolatory data analysis (EDA) by providing descriptive statistics, visualization support, and querying capabilities. It helps to identify patterns and trends in datasets.

d) Handeling missing data :- Pandas provides methods like .fillna(), .dropna() and .interpolate() to handle missing or null values in datasets efficiently. Missing

data handeling ensures better data integrity for analysis.

e) Data Integeration :- It supports merging, joining, and concatenation of multiple datasets making it easier to combine data from different sources. Pandas can integenate data from different sources like APIs and databases.

f) Time series analysis :- Pandas has built in support for time series data, including resampling, shifting, rolling windows, and handling datetime indexes. Time-based filtering is possible with timestamps and ranges. Pandas simplifies working with financial and temporal data.

g) Data preparation for Machine Learning :- Pandas helps in preprocessing and featuring converting raw data into structured formats suitable for ML models. Normalization and scaling of data can be done using built in functions. The .drop() function removes irrelevant or redundant features. Clean and structured data improves ML model accuracy.