



Vellore Institute of Technology

## SCOPE

### FDA (CSE1006) - Slot - L22+L23

#### DA -II

1. A hospital maintains a patients dataset with following data.

PatientID	Name	Age	Gender	Disease	AdmissionDate	TreatmentCost
101	John Doe	25	M	Flu	2023-01-15	1000
102	Jane Smith	45	F	Diabetes	2022-11-10	2500
103	NA	60	M	Flu	2023-05-20	
104	Chris Evans	30	M	Cancer	2022-12-05	5000
105	Emily Davis	50	F		2023-07-30	3000
102	Jane Smith	45	F	Diabetes	2022-11-10	2500
106	Mike Ross	70	M	Hypertension	2021-06-18	7000
107	Sarah Lee	22	F	Asthma	2023-02-14	4000
108	David Kim	NA	M	NA	2023-08-22	NA
109	NA	40	M	Covid-19	2023-04-10	1500

- Create a dataframe for above patients data.
- Sort the dataset by AdmissionDate (oldest to newest).
- Find and remove duplicate patient records based on PatientID.
- Replace empty strings and "NA" as text with actual NA values.
- Convert AdmissionDate to Date format and TreatmentCost to numeric.
- Convert Age column to categorize patients into Young (0-30), Middle-aged (31-60), and Senior (61+).

Note:

- ✓ Mention your name, roll number, and assessment subject details.
- ✓ Document should be in PDF/word with your registration number.
- ✓ In program execution shot 1<sup>st</sup> display your name and reg no followed by output.

\*\*\*\*\*