

Question 1

1. What are the steps for handling missing data?
2. How to deal with missing values in the following dataset apply all above steps. Write the code to clean the entire dataset using the appropriate methods. Also explain why these methods are used for a particular missing value or feature.

S. No	Name	State	Gender	Score
0	Kamran	Arizona	M	63
1	Imran	NaN	M	48
2	Micheal	Georgia	NaN	56
3	Irfan	Newyork	M	75
4	Zeeshan	Indiana	M	63
5	Adnan	Florida	M	NaN
6	Farman	California	M	63
7	NaN	NaN	NaN	NaN
8	Shaighan	Lahore	M	23

Question 2

What is binning? Explain the concept and write python code for binning, by making two bins of the following data set. (hint: if the passing marks are 60)

Names	Marks
George	50
Andrea	65
Micheal	78
Maggie	25
Ravi	36
Xien	38
Jalpa	42
Joy	22

Question 3

What is data normalization? why it is important? Write a python code to normalize the data of the following dataset.

House	Sale Price (\$100)	Size (sqft)
Avalon	2050	2650
Cross Winds	2080	2600
The White House	2150	2554
The Rectory	2150	2921
Larchwood	1999	2580
Orchard House	1900	2580
Shangri-La	1800	2774
The Stables	1560	1920
Cobweb Cottage	1450	2150
Narin House	1449	1710