TREATMENT OF OUTLIERS whole of the state of the state of 8) should we drop or not to drop Outliers? Ans - Dropping data is always a horsh step and should be taken when we are sure that outlier is a measurement enough, which we generally - when we drop the data, we loose information in terms of variability in data. Suppose, we are doing segmentation, variability of data becomes major role because if there is no variation then we cannot divide segment the data into groups. - so when we have too many observations and outliers are very very few, we can think of dropping the observations, Example - Regression Line without outlines. with outliers What are the steps to remove / deal with outliers? Ans -) Capping 11) Log-Scale Transformation. 11) Binning) Capping -> This methods involve setting the extreme values of an attribute to some specified value. For example, bottom 5% of value are set equal to minimum value in 5th percentile, while the upper 5% of values are set equal to the maximum value in 95th percentile. 11) Log scale transformation -> This method is often used to reduce the variability of data including outlying observations. Here value is change with log i.e, y > log (y). It's often preferred when the response variable follows exponential distribution/right shewwed.

Variable * (Histogram) Histogram log(x) > This is controversial steps as It does not reduce variance, only thing is it change the scale 60.80,100 -> 0.6,0.8,1.0

Binning > This refers to dividing a list of continuous variable into groups. We do this to discover set of patterns in continuous variable, which is difficult to analyze orthonorise.

But it also lead to loss of information and loss of power because suppose if we group them, then even if they are have variability they are one group.

4) Which are the algorithms affected by outliers?

Ans - ML models like linear 2 logistic regression are easily affected by the outliers.

Random forest Idecision tree are not affected by outliers.

& selften with loady seman of agots with a restainty

allowed to the species of a toologic

More of Copies and Described Transportation of Melicing

to the wider from the soft with the sollowing about the Soft to single (

because the front of the following the skinks

occurrent of longs to sto solow to all morted askings sol

moreover possess region of state states and the states

de work & Representation for