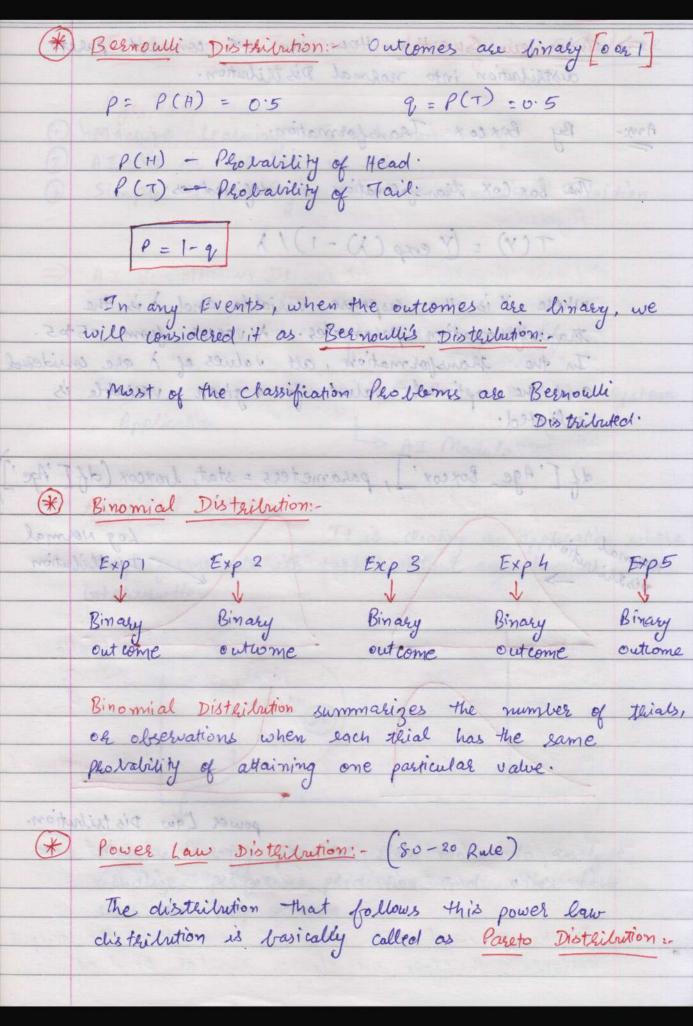
EDA PART-2. FEATURE ENGINEERING

	FEATURE ENGINEERING	
	Feature Engineering AND Preprocessing Ste	ps:-
-	Handle the missing value.	~
_	Handle the outlier.	(
-	Scalling of data. such sum Atim such as	4
-	Transformation [log, Box cox, square, Cube]	4
-	Encoding	
-	we can handle inbalance data.	(2)
-	Feature Selection.	
	We can do Dimension Reduction [PCA, tSME]	36
-	Duplicate value, Duplicate Column.	
-	split I merge I drop I Add column or Row.	
	Wester Root Handy of the story at the sail of the sail	
	Medical polyment	ex (
0	missing value Handle: - Method to handle the mis	sino
		lue.
0-		
0-	Fill Random value	3
0 -	Fill Random value. Forward Filling / backward filling	3
0-	Fill Random value. Forward Filling / backword filling Statical Apphoach.	-
0- 3-	Fill Random value. Forward Filling / bradeword filling Statical Apphoach. - Mean - Median - Mode	
(3)-	Fill Random value. Forward Filling / brackword Filling Statical Apphoach - Mean - Median - Mode	Tapas Single
(3)- (3)-	Fill Random value. Forward Filling / brackword Filling Statical Apphoach - Mean - Median - Mode End of the distribution.	They set for
(3)- (3)- (5)- (5)-	Fill Random value. Forward Filling / brackword Filling Statical Apphoach. - Mean - Median - Mode End of the distribution. Dhop that Low	The state of the s
(3)- (3)- (5)- (6)-	Fill Random value. Forward Filling / brackword Filling Statical Apphoach. - Mean - Median - Mode End of the distribution. Dhop that Low Knn - imputer.	Sand Street Stre
(3)- (3)- (5)- (5)-	Fill Random value. Forward Filling / tradeword filling Statical Apploach. - Mean - Median - Mode End of the distribution. Dhop that how Knn - imputer. Can we take that ML algorithm which handle	Constant of the second
(3)- (3)- (5)- (6)- (7)-	Fill Random value. Forward Filling I tradeword filling Statical Apphoach. - Mean - Median - Mode End of the distribution. Dhop that Low Knn - imputer. Can we take that ML algorithm which handle missing value	The state of the s
(3)- (3)- (5)- (6)-	Fill Random value. Forward Filling / tradeword filling Statical Apploach. - Mean - Median - Mode End of the distribution. Dhop that how Knn - imputer. Can we take that ML algorithm which handle	The state of the s
(3)- (3)- (5)- (6)- (7)-	Fill Random value. Forward Filling I tradeword filling Statical Apphoach. - Mean - Median - Mode End of the distribution. Dhop that Low Knn - imputer. Can we take that ML algorithm which handle missing value	
(3)- (3)- (5)- (6)- (7)- (8)-	Fill Random value Forward Filling / tradeword Filling Statical Apphoach - Me an _ Median - Mode End of the distribution Dhop that how Knn - imputer. Can we take that ML algorithm which handle missing value Own ML Model to handle missing value.	The state of the s
(3)- (3)- (5)- (6)- (7)- (8)-	Fill Random value Forward Filling I tradeword filling Statical Apphoach - Mean - Median - Mode End of the distribution Dhop that how Knn - imputer. Can we take that ML algorithm which handle missing value Own ML model to handle missing value. Outliers: Some Steps to detett the outlier.	
(3)- (3)- (5)- (6)- (7)- (8)-	Fill Random value Forward Filling / brackword filling Statical Apphoach - Me an _ Median - Mode End of the distribution Dhop that how Knn - imputer. Can we take that mr algorithm which handle missing value Own Mr Model to handle missing value. Outliers:- Some Steps to detelt the outlier. Z-scoke.	

- 120	Some steps to handle the outlier.	la l
	He two check take is linear at not	
->	drop the outlier. and wind all whomas	-
\rightarrow		
\rightarrow	A A	
	Thip Thimming that past	
	Evication	
(3)	Thansformation: G Scaling:	
_	Box-Cox - Standardizat	
_	Power thanspormation min-max	
	log - Unit Scaling	
	Square most transformation.	
	Cule hoot thansformation.	
	Yeo Johnson	
	winter the small of bout one allowe the time of animal	10
7	ar Again Pensing V	
(3)	Encoding: B Handle Intralance	al Data
-	One hot encoding > Collect More data.	
_	label encoding > Undersampling	
_	Winnell Omending	
_	Binary encoding > Oversampling. Target Guided encoding > Claster based over Hash encoding	samplina
_	Hash encoding Huster based over	
	County gall	
	(2) Feature maneeling the madiately street men?	190
	an we take that me algorithm which handle	
	mulsing value	
	own the tracted to handle missing value	- (8)
	Think bulleauth commer at your and	
	و سادل ويديد	12
	some steps to detect the outlies.	4
	Z-Scote Scotter de	
	ISR Scatter plot	
	hox plat - molin plat	· ~



> Interview Suestion - How can we convert pareto distribution into normal Distribution. Ans:- By Box Cox Thansformation.

The Box Cox thansformation its defined as $T(Y) = (Y exp(\lambda) - 1)/\lambda$ where Y is the Response valiable and I is the thans formation parameter. I varies from -5 to 5. In the thansformation, all values of I are considered and the optimal value for a given variable is selected. of ['Age_Boxcox'], parameters = stat, boxcox (df ['Age']) pistlin ->

pistlin ->

moutro Log Normal
Disthibution Sine unial Distribution summarines the numbers of thinks of thinks Madelity of attaining one particular value. (oleg 02-03) - montrelist only sound (t) The distribution that dallows this power law dex technition is bosically collect as lacero materialitien