

Feature (Column)	Delete?	Process NAN?	Transformation	
ABANDON_DATE	Y			
ABN_YEAR	Y			
APPDATE	N	Delete	to int	
APPMONTH	Y			
APPNUM	Y			
APPTYPE	Y			
APPYEAR	Y			
ASGCITY	Y			
ASGCOUNTRY	Y			
ASGNUM	N	Delete	already int	
ASGSEQ	Y			
ASGSTATE	Y			
ASSIGNEE	N	Delete	assign index to company names	
CATEGORY	N	change to "Others"	one-hot	
CLAIMS	N	Delete	already int	
CLASS	N	Delete	already int	
DISPOSAL_TYPE	Y			
EXAMINER_ART_UNIT	Y			
EXAMINER_ID	Y			
FILING_DATE	Y			
FILING_YEAR	Y			
FIRSTNAME	N	Delete	assign index to names	
GDATE	N	Delete	to int	
INVCITY	Y			
INVCOUNTRY	N	Delete	one-hot	
INVNUM	N	Delete	already int	
INVSEQ	N	Delete	already int	
INVSTATE	Y			
KIND	Y			
LASTNAME	N	Delete	assign index to names	
NBCITE	N	Delete	already int	

NFCITE	N	Delete	already int	
NUMAPP	N	Delete	already int	
NUMPAT	N	Delete	already int	
PATENT	Y			
RESIDENCE	Y			
SUBCLASS	N	Delete	assign index to subclass	
ABN	Y			
DES	Y			
UTL	Y			
US	Y			
CAT	Y			
PRIMINV	N	assign "0"	first number	
TEAM	N	assign "1"	first number	
NUMPRIM	N	Delete	already int	
LONE	Y			
NUMLONE	N	Delete	already int	
NUMCOINV	N	Delete	already int	
TOTAPP	N	Delete	already int	
USINV	Y			
INVCOUNTRY	Y			
TEAMSIZ	N	Delete	already int	