litanic import numpy as no import pandas as pol import Jealan as SIS SS. Set G'darkgrid') Style = from sklewn, ensemble import Random Forest Clasifier. from skleom. preprosessing import OneHot Encoder, Lobel Encoder, Standard Scaler from sklearn metrics import roc_curve, auc from Skleam Model_Selection import Stratified Ktob import String import wornings. wornings filter warnings (ignore') JEED = 42 def concat_df (train_data, test_data); Neturn pd. consot ([train_data, test_data], Jort = True). restet-index (dop = True) def divide-df (all-dota) reform all_dota 1, loc [:890], all-dota. loc [891:]. drop (['Survived' OXIS=1)

) f	df_train = pd. read_csu ('puth')
	df_test = pol. read_osu ('poth')
	df_all = Concat_df (df_tmin, df_teit)
	C See A Miles
	df_train. name = 'Traing det'
	df-test. nume = 'Test Set'
	df-all. name = 'All Det'
N	Lin stee mining int best fine sold first Date!
	Print (Number of training Sets : { 3', format (df-train, shape [0]))
	print (' " test " " "
	print (Training Sets X Stage : { 3', format(" x))
_	- HONE SHOW
	1. Exploratory Data Analysis.
	1. Exploratory Data Analysis. 1.1 Overview
	Print (df-train. Info())
	df_train, Sample (3)
	1.2 missing values.
	def display_missing (df):
1990	for cal in df. columns, tolistc):
	print ('{col, af [ol], isnoll, Jumo) 3 column misses value: { 3 "
	Print ('\n')
	for df in dfs:
V nat.	Print (Edf. name?)
	display - Missing (df)

1.2.1 Age. df-all-corr = df-all.corre s. absc s. unstacke s. Dort-values C+++ kind = 'quicksort', ascending = False), reset_index() df_all_corr. rename (calums = { '/evel_0' = 'Feotive /', --- 3, inplace = T) df_all-corr [df_all-corr [feature 1'] == Age] index & nedian 2 = age_by-pclass_sex = of_all, grouply (['Sex', 'pclass']), median()['Age'] 彭 经. df_all ['Age'] = df_all, grouply (['Sex', 'pdass']) ['Age'], apply (lambda x: X. Allna (X. Median ())) 1.22 Emburkal dfall [dfall ['Embarked']. isnull()] df-all ['Emborked'] = df-all ['Emborked']. fillna('S') 1.2.3 Fore. med_Fare = def_all_grouphy (['poliss', 'parch', 'sibsp']) ['Fare'], median() [3][0][0] [Fare'] 1.2.4 Cabin