

```
# Data analysis and wrangling
import pandas as pd
import numpy as np
import random as rng

# Visualization
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

# Machine Learning
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC, LinearSVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.linear_model import Perceptron
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier

In [54]:
# Load data
train_df = pd.read_csv('train.csv')
test_df = pd.read_csv('test.csv')
combine = [train_df, test_df]

In [55]:
# Analyze by describing data
print(train_df.columns.values)

PassengerId Survived Pclass Name Sex Age SibSp Parch Ticket Fare Cabin Embarked

In [56]:
# Preview the data
train_df.head()

Out[56]:
PassengerId Survived Pclass Name Sex Age SibSp Parch Ticket Fare Cabin Embarked
0 1 0 3 Brund, Mr. Owen Harris male 22.0 1 0 A/5 2171 7.2500 NaN S
1 2 1 1 Cunnage, Mrs. John Bradley/Powes-Edgar, R. female 38.0 1 0 AC/1550 71.2833 C86 C
2 3 1 3 Heikkinen, Mrs. Laina female 26.0 1 0 STON/O 30.00 842 S
3 4 1 1 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1 0 113803 53.1000 C123 S
4 5 0 3 Allen, Mr. William Henry male 35.0 0 0 773450 8.0500 NaN S

In [57]:
train_df.tail()

PassengerId Survived Pclass Name Sex Age SibSp Parch Ticket Fare Cabin Embarked
885 987 0 2 Monks, Rev. James male 27.0 0 0 21106 13.00 NaN S
887 988 1 2 Gopal, Mrs. Margaret Eden female 19.0 0 0 11060 30.00 842 S
889 989 0 3 Jomary, Mrs. Catherine Helen/Clarke female 26.0 1 2 W/C 6607 23.45 NaN S
889 990 1 1 Bethel, Mr. Karl Howell male 26.0 0 0 11393 10.00 C140 C
890 991 0 3 Douky, Mr. Patrick male 32.0 0 0 37076 7.75 NaN Q

In [58]:
train_df.info()
print('--*)')
test_df.info()

<class pandas.core.frame.DataFrame>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
# Column Non-Null Count Dtype
---
0 PassengerId 891 non-null int64
1 Survived 891 non-null int64
2 Pclass 891 non-null int64
3 Name 891 non-null object
4 Sex 891 non-null object
5 Age 714 non-null float64
6 SibSp 891 non-null int64
7 Parch 891 non-null int64
8 Ticket 891 non-null object
9 Fare 891 non-null float64
10 Cabin 91 non-null object
11 Embarked 891 non-null object
dtypes: float64(5), int64(4), object(3)
memory usage: 83.7+ KB

<class pandas.core.frame.DataFrame>
RangeIndex: 438 entries, 0 to 437
Data columns (total 11 columns):
# Column Non-Null Count Dtype
---
0 PassengerId 438 non-null int64
1 Pclass 438 non-null int64
2 Name 438 non-null object
3 Sex 438 non-null object
4 Age 332 non-null float64
5 SibSp 438 non-null int64
6 Parch 438 non-null int64
7 Ticket 438 non-null object
8 Fare 437 non-null float64
9 Cabin 91 non-null object
10 Embarked 438 non-null object
dtypes: float64(5), int64(4), object(2)
memory usage: 36.4+ KB

In [59]:
train_df.describe()
# Review survival rates using 'percentiles' (25, 50) knowing our problem description mentions 38% survival rate.
# Review Parch distribution using 'percentiles' (75, 8)
# SibSp distribution (1, 68, 69)
# Age and Fare (1, 2, 3, 4, 5, 6, 7, 8, 9, 99)

Out[59]:
PassengerId Survived Pclass Age SibSp Parch Fare
count 891.000000 891.000000 891.000000 714.000000 891.000000 891.000000
mean 446.000000 0.383838 2.308642 29.696118 0.523008 0.381984 32.054208
std 257.353842 0.486692 0.836071 14.526497 1.102743 0.800557 49.093429
min 1.000000 0.000000 1.000000 0.420000 0.000000 0.000000 0.000000
75% 253.000000 0.000000 2.000000 0.250000 0.000000 0.000000 7.004000
90% 446.000000 0.000000 3.000000 26.000000 0.000000 0.000000 14.654200
75% 669.500000 1.000000 3.000000 30.000000 1.000000 0.000000 31.000000
max 891.000000 1.000000 3.000000 80.000000 0.000000 6.000000 512.329000

In [60]:
train_df.describe(include=[0])

Out[60]:
Name Sex Ticket Cabin Embarked
count 891 2 681 186 889
unique 891 2 681 147 8
top Brund, Mr. Owen Harris male 347082 B96 B88 S
freq 1 577 7 4 644

In [61]:
# Analyze by plotting features
train_df[['Pclass', 'Survived']].groupby(['Pclass'], as_index=False).mean().sort_values(by='Survived', ascending=False)

Out[61]:
Pclass Survived
0 1 0.29350
1 2 0.472826
2 3 0.242423

In [62]:
train_df[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean().sort_values(by='Survived', ascending=False)

Out[62]:
Sex Survived
0 male 0.743586
1 male 0.188908

In [63]:
train_df[['SibSp', 'Survived']].groupby(['SibSp'], as_index=False).mean().sort_values(by='Survived', ascending=False)

Out[63]:
SibSp Survived
1 1 0.535885
2 2 0.464266
0 0 0.543598
3 3 0.250000
4 4 0.166667
5 5 0.000000
6 6 0.000000

In [64]:
train_df[['Parch', 'Survived']].groupby(['Parch'], as_index=False).mean().sort_values(by='Survived', ascending=False)

Out[64]:
Parch Survived
3 3 0.560947
2 2 0.500000
0 0 0.343558
5 5 0.200000
4 4 0.000000
6 6 0.000000

In [65]:
# Analyze by visualizing data
# 0 = one factor (train_df, col='Survived')
# 1 = two factor (train_df, col='Survived')
# 2 = three factor (train_df, col='Survived')
# 3 = four factor (train_df, col='Survived')
# 4 = five factor (train_df, col='Survived')
# 5 = six factor (train_df, col='Survived')
# 6 = seven factor (train_df, col='Survived')
# 7 = eight factor (train_df, col='Survived')
# 8 = nine factor (train_df, col='Survived')
# 9 = ten factor (train_df, col='Survived')
# 10 = eleven factor (train_df, col='Survived')
# 11 = twelve factor (train_df, col='Survived')
# 12 = thirteen factor (train_df, col='Survived')
# 13 = fourteen factor (train_df, col='Survived')
# 14 = fifteen factor (train_df, col='Survived')
# 15 = sixteen factor (train_df, col='Survived')
# 16 = seventeen factor (train_df, col='Survived')
# 17 = eighteen factor (train_df, col='Survived')
# 18 = nineteen factor (train_df, col='Survived')
# 19 = twenty factor (train_df, col='Survived')
# 20 = twenty-one factor (train_df, col='Survived')
# 21 = twenty-two factor (train_df, col='Survived')
# 22 = twenty-three factor (train_df, col='Survived')
# 23 = twenty-four factor (train_df, col='Survived')
# 24 = twenty-five factor (train_df, col='Survived')
# 25 = twenty-six factor (train_df, col='Survived')
# 26 = twenty-seven factor (train_df, col='Survived')
# 27 = twenty-eight factor (train_df, col='Survived')
# 28 = twenty-nine factor (train_df, col='Survived')
# 29 = thirty factor (train_df, col='Survived')
# 30 = thirty-one factor (train_df, col='Survived')
# 31 = thirty-two factor (train_df, col='Survived')
# 32 = thirty-three factor (train_df, col='Survived')
# 33 = thirty-four factor (train_df, col='Survived')
# 34 = thirty-five factor (train_df, col='Survived')
# 35 = thirty-six factor (train_df, col='Survived')
# 36 = thirty-seven factor (train_df, col='Survived')
# 37 = thirty-eight factor (train_df, col='Survived')
# 38 = thirty-nine factor (train_df, col='Survived')
# 39 = forty factor (train_df, col='Survived')
# 40 = forty-one factor (train_df, col='Survived')
# 41 = forty-two factor (train_df, col='Survived')
# 42 = forty-three factor (train_df, col='Survived')
# 43 = forty-four factor (train_df, col='Survived')
# 44 = forty-five factor (train_df, col='Survived')
# 45 = forty-six factor (train_df, col='Survived')
# 46 = forty-seven factor (train_df, col='Survived')
# 47 = forty-eight factor (train_df, col='Survived')
# 48 = forty-nine factor (train_df, col='Survived')
# 49 = fifty factor (train_df, col='Survived')
# 50 = fifty-one factor (train_df, col='Survived')
# 51 = fifty-two factor (train_df, col='Survived')
# 52 = fifty-three factor (train_df, col='Survived')
# 53 = fifty-four factor (train_df, col='Survived')
# 54 = fifty-five factor (train_df, col='Survived')
# 55 = fifty-six factor (train_df, col='Survived')
# 56 = fifty-seven factor (train_df, col='Survived')
# 57 = fifty-eight factor (train_df, col='Survived')
# 58 = fifty-nine factor (train_df, col='Survived')
# 59 = sixty factor (train_df, col='Survived')
# 60 = sixty-one factor (train_df, col='Survived')
# 61 = sixty-two factor (train_df, col='Survived')
# 62 = sixty-three factor (train_df, col='Survived')
# 63 = sixty-four factor (train_df, col='Survived')
# 64 = sixty-five factor (train_df, col='Survived')
# 65 = sixty-six factor (train_df, col='Survived')
# 66 = sixty-seven factor (train_df, col='Survived')
# 67 = sixty-eight factor (train_df, col='Survived')
# 68 = sixty-nine factor (train_df, col='Survived')
# 69 = seventy factor (train_df, col='Survived')
# 70 = seventy-one factor (train_df, col='Survived')
# 71 = seventy-two factor (train_df, col='Survived')
# 72 = seventy-three factor (train_df, col='Survived')
# 73 = seventy-four factor (train_df, col='Survived')
# 74 = seventy-five factor (train_df, col='Survived')
# 75 = seventy-six factor (train_df, col='Survived')
# 76 = seventy-seven factor (train_df, col='Survived')
# 77 = seventy-eight factor (train_df, col='Survived')
# 78 = seventy-nine factor (train_df, col='Survived')
# 79 = eighty factor (train_df, col='Survived')
# 80 = eighty-one factor (train_df, col='Survived')
# 81 = eighty-two factor (train_df, col='Survived')
# 82 = eighty-three factor (train_df, col='Survived')
# 83 = eighty-four factor (train_df, col='Survived')
# 84 = eighty-five factor (train_df, col='Survived')
# 85 = eighty-six factor (train_df, col='Survived')
# 86 = eighty-seven factor (train_df, col='Survived')
# 87 = eighty-eight factor (train_df, col='Survived')
# 88 = eighty-nine factor (train_df, col='Survived')
# 89 = ninety factor (train_df, col='Survived')
# 90 = ninety-one factor (train_df, col='Survived')
# 91 = ninety-two factor (train_df, col='Survived')
# 92 = ninety-three factor (train_df, col='Survived')
# 93 = ninety-four factor (train_df, col='Survived')
# 94 = ninety-five factor (train_df, col='Survived')
# 95 = ninety-six factor (train_df, col='Survived')
# 96 = ninety-seven factor (train_df, col='Survived')
# 97 = ninety-eight factor (train_df, col='Survived')
# 98 = ninety-nine factor (train_df, col='Survived')
# 99 = one hundred factor (train_df, col='Survived')
# 100 = one hundred and one factor (train_df, col='Survived')
# 101 = one hundred and two factor (train_df, col='Survived')
# 102 = one hundred and three factor (train_df, col='Survived')
# 103 = one hundred and four factor (train_df, col='Survived')
# 104 = one hundred and five factor (train_df, col='Survived')
# 105 = one hundred and six factor (train_df, col='Survived')
# 106 = one hundred and seven factor (train_df, col='Survived')
# 107 = one hundred and eight factor (train_df, col='Survived')
# 108 = one hundred and nine factor (train_df, col='Survived')
# 109 = one hundred and ten factor (train_df, col='Survived')
# 110 = one hundred and eleven factor (train_df, col='Survived')
# 111 = one hundred and twelve factor (train_df, col='Survived')
# 112 = one hundred and thirteen factor (train_df, col='Survived')
# 113 = one hundred and fourteen factor (train_df, col='Survived')
# 114 = one hundred and fifteen factor (train_df, col='Survived')
# 115 = one hundred and sixteen factor (train_df, col='Survived')
# 116 = one hundred and seventeen factor (train_df, col='Survived')
# 117 = one hundred and eighteen factor (train_df, col='Survived')
# 118 = one hundred and nineteen factor (train_df, col='Survived')
# 119 = one hundred and twenty factor (train_df, col='Survived')
# 120 = one hundred and twenty-one factor (train_df, col='Survived')
# 121 = one hundred and twenty-two factor (train_df, col='Survived')
# 122 = one hundred and twenty-three factor (train_df, col='Survived')
# 123 = one hundred and twenty-four factor (train_df, col='Survived')
# 124 = one hundred and twenty-five factor (train_df, col='Survived')
# 125 = one hundred and twenty-six factor (train_df, col='Survived')
# 126 = one hundred and twenty-seven factor (train_df, col='Survived')
# 127 = one hundred and twenty-eight factor (train_df, col='Survived')
# 128 = one hundred and twenty-nine factor (train_df, col='Survived')
# 129 = one hundred and thirty factor (train_df, col='Survived')
# 130 = one hundred and thirty-one factor (train_df, col='Survived')
# 131 = one hundred and thirty-two factor (train_df, col='Survived')
# 132 = one hundred and thirty-three factor (train_df, col='Survived')
# 133 = one hundred and thirty-four factor (train_df, col='Survived')
# 134 = one hundred and thirty-five factor (train_df, col='Survived')
# 135 = one hundred and thirty-six factor (train_df, col='Survived')
# 136 = one hundred and thirty-seven factor (train_df, col='Survived')
# 137 = one hundred and thirty-eight factor (train_df, col='Survived')
# 138 = one hundred and thirty-nine factor (train_df, col='Survived')
# 139 = one hundred and forty factor (train_df, col='Survived')
# 140 = one hundred and forty-one factor (train_df, col='Survived')
# 141 = one hundred and forty-two factor (train_df, col='Survived')
# 142 = one hundred and forty-three factor (train_df, col='Survived')
# 143 = one hundred and forty-four factor (train_df, col='Survived')
# 144 = one hundred and forty-five factor (train_df, col='Survived')
# 145 = one hundred and forty-six factor (train_df, col='Survived')
# 146 = one hundred and forty-seven factor (train_df, col='Survived')
# 147 = one hundred and forty-eight factor (train_df, col='Survived')
# 148 = one hundred and forty-nine factor (train_df, col='Survived')
# 149 = one hundred and fifty factor (train_df, col='Survived')
# 150 = one hundred and fifty-one factor (train_df, col='Survived')
# 151 = one hundred and fifty-two factor (train_df, col='Survived')
# 152 = one hundred and fifty-three factor (train_df, col='Survived')
# 153 = one hundred and fifty-four factor (train_df, col='Survived')
# 154 = one hundred and fifty-five factor (train_df, col='Survived')
# 155 = one hundred and fifty-six factor (train_df, col='Survived')
# 156 = one hundred and fifty-seven factor (train_df, col='Survived')
# 157 = one hundred and fifty-eight factor (train_df, col='Survived')
# 158 = one hundred and fifty-nine factor (train_df, col='Survived')
# 159 = one hundred and sixty factor (train_df, col='Survived')
# 160 = one hundred and sixty-one factor (train_df, col='Survived')
# 161 = one hundred and sixty-two factor (train_df, col='Survived')
# 162 = one hundred and sixty-three factor (train_df, col='Survived')
# 163 = one hundred and sixty-four factor (train_df, col='Survived')
# 164 = one hundred and sixty-five factor (train_df, col='Survived')
# 165 = one hundred and sixty-six factor (train_df, col='Survived')
# 166 = one hundred and sixty-seven factor (train_df, col='Survived')
# 167 = one hundred and sixty-eight factor (train_df, col='Survived')
# 168 = one hundred and sixty-nine factor (train_df, col='Survived')
# 169 = one hundred and seventy factor (train_df, col='Survived')
# 170 = one hundred and seventy-one factor (train_df, col='Survived')
# 171 = one hundred and seventy-two factor (train_df, col='Survived')
# 172 = one hundred and seventy-three factor (train_df, col='Survived')
# 173 = one hundred and seventy-four factor (train_df, col='Survived')
# 174 = one hundred and seventy-five factor (train_df, col='Survived')
# 175 = one hundred and seventy-six factor (train_df, col='Survived')
# 176 = one hundred and seventy-seven factor (train_df, col='Survived')
# 177 = one hundred and seventy-eight factor (train_df, col='Survived')
# 178 = one hundred and seventy-nine factor (train_df, col='Survived')
# 179 = one hundred and eighty factor (train_df, col='Survived')
# 180 = one hundred and eighty-one factor (train_df, col='Survived')
# 181 = one hundred and eighty-two factor (train_df, col='Survived')
# 182 = one hundred and eighty-three factor (train_df, col='Survived')
# 183 = one hundred and eighty-four factor (train_df, col='Survived')
# 184 = one hundred and eighty-five factor (train_df, col='Survived')
# 185 = one hundred and eighty-six factor (train_df, col='Survived')
# 186 = one hundred and eighty-seven factor (train_df, col='Survived')
# 187 = one hundred and eighty-eight factor (train_df, col='Survived')
# 188 = one hundred and eighty-nine factor (train_df, col='Survived')
# 189 = one hundred and ninety factor (train_df, col='Survived')
# 190 = one hundred and ninety-one factor (train_df, col='Survived')
# 191 = one hundred and ninety-two factor (train_df, col='Survived')
# 192 = one hundred and ninety-three factor (train_df, col='Survived')
# 193 = one hundred and ninety-four factor (train_df, col='Survived')
# 194 = one hundred and ninety-five factor (train_df, col='Survived')
# 195 = one hundred and ninety-six factor (train_df, col='Survived')
# 196 = one hundred and ninety-seven factor (train_df, col='Survived')
# 197 = one hundred and ninety-eight factor (train_df, col='Survived')
# 198 = one hundred and ninety-nine factor (train_df, col='Survived')
# 199 = two hundred factor (train_df, col='Survived')
# 200 = two hundred and one factor (train_df, col='Survived')
# 201 = two hundred and two factor (train_df, col='Survived')
# 202 = two hundred and three factor (train_df, col='Survived')
# 203 = two hundred and four factor (train_df, col='Survived')
# 204 = two hundred and five factor (train_df, col='Survived')
# 205 = two hundred and six factor (train_df, col='Survived')
# 206 = two hundred and seven factor (train_df, col='Survived')
# 207 = two hundred and eight factor (train_df, col='Survived')
# 208 = two hundred and nine factor (train_df, col='Survived')
# 209 = two hundred and ten factor (train_df, col='Survived')
# 210 = two hundred and eleven factor (train_df, col='Survived')
# 211 = two hundred and twelve factor (train_df, col='Survived')
# 212 = two hundred and thirteen factor (train_df, col='Survived')
# 213 = two hundred and fourteen factor (train_df, col='Survived')
# 214 = two hundred and fifteen factor (train_df, col='Survived')
# 215 = two hundred and sixteen factor (train_df, col='Survived')
# 216 = two hundred and seventeen factor (train_df, col='Survived')
# 217 = two hundred and eighteen factor (train_df, col='Survived')
# 218 = two hundred and nineteen factor (train_df, col='Survived')
# 219 = two hundred and twenty factor (train_df, col='Survived')
# 220 = two hundred and twenty-one factor (train_df, col='Survived')
# 221 = two hundred and twenty-two factor (train_df, col='Survived')
# 222 = two hundred and twenty-three factor (train_df, col='Survived')
# 223 = two hundred and twenty-four factor (train_df, col='Survived')
# 224 = two hundred and twenty-five factor (train_df, col='Survived')
# 225 = two hundred and twenty-six factor (train_df, col='Survived')
# 226 = two hundred and twenty-seven factor (train_df, col='Survived')
# 227 = two hundred and twenty-eight factor (train_df, col='Survived')
# 228 = two hundred and twenty-nine factor (train_df, col='Survived')
# 229 = two hundred and thirty factor (train_df, col='Survived')
# 230 = two hundred and thirty-one factor (train_df, col='Survived')
# 231 = two hundred and thirty-two factor (train_df, col='Survived')
# 232 = two hundred and thirty-three factor (train_df, col='Survived')
# 233 = two hundred and thirty-four factor (train_df, col='Survived')
# 234 = two hundred and thirty-five factor (train_df, col='Survived')
# 235 = two hundred and thirty-six factor (train_df, col='Survived')
# 236 = two hundred and thirty-seven factor (train_df, col='Survived')
# 237 = two hundred and thirty-eight factor (train_df, col='Survived')
# 238 = two hundred and thirty-nine factor (train_df, col='Survived')
# 239 = two hundred and forty factor (train_df, col='Survived')
# 240 = two hundred and forty-one factor (train_df, col='Survived')
# 241 = two hundred and forty-two factor (train_df, col='Survived')
# 242 = two hundred and forty-three factor (train_df, col='Survived')
# 243 = two hundred and forty-four factor (train_df, col='Survived')
# 244 = two hundred and forty-five factor (train_df, col='Survived')
# 245 = two hundred and forty-six factor (train_df, col='Survived')
# 246 = two hundred and forty-seven factor (train_df, col='Survived')
# 247 = two hundred and forty-eight factor (train_df, col='Survived')
# 248 = two hundred and forty-nine factor (train_df, col='Survived')
# 249 = two hundred and fifty factor (train_df, col='Survived')
# 250 = two hundred and fifty-one factor (train_df, col='Survived')
# 251 = two hundred and fifty-two factor (train_df, col='Survived')
# 252 = two hundred and fifty-three factor (train_df, col='Survived')
# 253 = two hundred and fifty-four factor (train_df, col='Survived')
# 254 = two hundred and fifty-five factor (train_df, col='Survived')
# 255 = two hundred and fifty-six factor (train_df, col='Survived')
# 256 = two hundred and fifty-seven factor (train_df, col='Survived')
# 257 = two hundred and fifty-eight factor (train_df, col='Survived')
# 258 = two hundred and fifty-nine factor (train_df, col='Survived')
# 259 = two hundred and sixty factor (train_df, col='Survived')
# 260 = two hundred and sixty-one factor (train_df, col='Survived')
# 261 = two hundred and sixty-two factor (train_df, col='Survived')
# 262 = two hundred and sixty-three factor (train_df, col='Survived')
# 263 = two hundred and sixty-four factor (train_df, col='Survived')
# 264 = two hundred and sixty-five factor (train_df, col='Survived')
# 265 = two hundred and sixty-six factor (train_df, col='Survived')
# 266 = two hundred and sixty-seven factor (train_df, col='Survived')
# 267 = two hundred and sixty-eight factor (train_df, col='Survived')
# 268 = two hundred and sixty-nine factor (train_df, col='Survived')
# 269 = two hundred and seventy factor (train_df, col='Survived')
# 270 = two hundred and seventy-one factor (train_df, col='Survived')
# 271 = two hundred and seventy-two factor (train_df, col='Survived')
# 272 = two hundred and seventy-three factor (train_df, col='Survived')
# 273 = two hundred and seventy-four factor (train_df, col='Survived')
# 274 = two hundred and seventy-five factor (train_df, col='Survived')
# 275 = two hundred and seventy-six factor (train_df, col='Survived')
# 276 = two hundred and seventy-seven factor (train_df, col='Survived')
# 277 = two hundred and seventy-eight factor (train_df, col='Survived')
# 278 = two hundred and seventy-nine factor (train_df, col='Survived')
# 279 = two hundred and eighty factor (train_df, col='Survived')
# 280 = two hundred and eighty-one factor (train_df, col='Survived')
# 281 = two hundred and eighty-two factor (train_df, col='Survived')
# 282 = two hundred and eighty-three factor (train_df, col='Survived')
# 283 = two hundred and eighty-four factor (train_df, col='Survived')
# 284 = two hundred and eighty-five factor (train_df, col='Survived')
# 285 = two hundred and eighty-six factor (train_df, col='Survived')
# 286 = two hundred and eighty-seven factor (train_df, col='Survived')
# 287 = two hundred and eighty-eight factor (train_df, col='Survived')
# 288 = two hundred and eighty-nine factor (train_df, col='Survived')
# 289 = two hundred and ninety factor (train_df, col='Survived')
# 290 = two hundred and ninety-one factor (train_df, col='Survived')
# 291 = two hundred and ninety-two factor (train_df, col='Survived')
# 292 = two hundred and ninety-three factor (train_df, col='Survived')
# 293 = two hundred and ninety-four factor (train_df, col='Survived')
# 294 = two hundred and ninety-five factor (train_df, col='Survived')
# 295 = two hundred and ninety-six factor (train_df, col='Survived')
# 296 = two hundred and ninety-seven factor (train_df, col='Survived')
# 297 = two hundred and ninety-eight factor (train_df, col='Survived')
# 298 = two hundred and ninety-nine factor (train_df, col='Survived')
# 299 = three hundred factor (train_df, col='Survived')
# 300 = three hundred and one factor (train_df, col='Survived')
# 301 = three hundred and two factor (train_df, col='Survived')
# 302 = three hundred and three factor (train_df, col='Survived')
# 303 = three hundred and four factor (train_df, col='Survived')
# 304 = three hundred and five factor (train_df, col='Survived')
# 305 = three hundred and six factor (train_df, col='Survived')
# 306 = three hundred and seven factor (train_df, col='Survived')
# 307 = three hundred and eight factor (train_df, col='Survived')
# 308 = three hundred and nine factor (train_df, col='Survived')
# 309 = three hundred and ten factor (train_df, col='Survived')
# 310 = three hundred and eleven factor (train_df, col='Survived')
# 311 = three hundred and twelve factor (train_df, col='Survived')
# 312 = three hundred and thirteen factor (train_df, col='Survived')
# 313 = three hundred and fourteen factor (train_df, col='Survived')
# 314 = three hundred and fifteen factor (train_df, col='Survived')
# 315 = three hundred and sixteen factor (train_df, col='Survived')
# 316 = three hundred and seventeen factor (train_df, col='Survived')
# 317 = three hundred and eighteen factor (train_df, col='Survived')
# 318 = three hundred and nineteen factor (train_df, col='Survived')
# 319 = three hundred and twenty factor (train_df, col='Survived')
# 320 = three hundred and twenty-one factor (train_df, col='Survived')
# 321 = three hundred and twenty-two factor (train_df, col='Survived')
# 322 = three hundred and twenty-three factor (train_df, col='Survived')
# 323 = three hundred and twenty-four factor (train_df, col='Survived')
# 324 = three hundred and twenty-five factor (train_df, col='Survived')
# 325 = three hundred and twenty-six factor (train_df, col='Survived')
# 326 = three hundred and twenty-seven factor (train_df, col='Survived')
# 327 = three hundred and twenty-eight factor (train_df, col='Survived')
# 328 = three hundred and twenty-nine factor (train_df, col='Survived')
# 329 = three hundred and thirty factor (train_df, col='Survived')
# 330 = three hundred and thirty-one factor (train_df, col='Survived')
# 331 = three hundred and thirty-two factor (train_df, col='Survived')
# 332 = three hundred and thirty-three factor (train_df, col='Survived')
# 333 = three hundred and thirty-four factor (train_df, col='Survived')
# 334 = three hundred and thirty-five factor (train_df, col='Survived')
# 335 = three hundred and thirty-six factor (train_df, col='Survived')
# 336 = three hundred and thirty-seven factor (train_df, col='Survived')
# 337 = three hundred and thirty-eight factor (train_df, col='Survived')
# 338 = three hundred and thirty-nine factor (train_df, col='Survived')
# 339 = three hundred and forty factor (train_df, col='Survived')
# 340 = three hundred and forty-one factor (train_df, col='Survived')
# 341 = three hundred and forty-two factor (train_df, col='Survived')
# 342 = three hundred and forty-three factor (train_df, col='Survived')
# 343 = three hundred and forty-four factor (train_df, col='Survived')
# 344 = three hundred and forty-five factor (train_df, col='Survived')
# 345 = three hundred and forty-six factor (train_df, col='Survived')
# 346 = three hundred and forty-seven factor (train_df, col='Survived')
# 347 = three hundred and forty-eight factor (train_df, col='Survived')
# 348 = three hundred and forty-nine factor (train_df, col='Survived')
# 349 = three hundred and fifty factor (train_df, col='Survived')
# 350 = three hundred and fifty-one factor (train_df, col='Survived')
# 351 = three hundred and fifty-two factor (train_df, col='Survived')
# 352 = three hundred and fifty-three factor (train_df, col='Survived')
# 353 = three hundred and fifty-four factor (train_df, col='Survived')
# 354 = three hundred and fifty-five factor (train_df, col='Survived')
# 355 = three hundred and fifty-six factor (train_df, col='Survived')
# 356 = three hundred and fifty-seven factor (train_df, col='Survived')
# 357 = three hundred and fifty-eight factor (train_df, col='Survived')
# 358 = three hundred and fifty-nine factor (train_df, col='Survived')
# 359 = three hundred and sixty factor (train_df, col='Survived')
# 360 = three hundred and sixty-one factor (train_df, col='Survived')
# 361 = three hundred and sixty-two factor (train_df, col='Survived')
# 362 = three hundred and sixty-three factor (train_df, col='Survived')
# 363 = three hundred and sixty-four factor (train_df, col='Survived')
# 364 = three hundred and sixty-five factor (train_df, col='Survived')
# 365 = three hundred and sixty-six factor (train_df, col='Survived')
# 366 = three hundred and sixty-seven factor (train_df, col='Survived')
# 367 = three hundred and sixty-eight factor (train_df, col='Survived')
# 368 = three hundred and sixty-nine factor (train_df, col='Survived')
# 369 = three hundred and seventy factor (train_df, col='Survived')
# 370 = three hundred and seventy-one factor (train_df, col='Survived')
# 371 = three hundred and seventy-two factor (train_df, col='Survived')
# 372 = three hundred and seventy-three factor (train_df, col='Survived')
# 373 = three hundred and seventy-four factor (train_df, col='Survived')
# 374 = three hundred and seventy-five factor (train_df, col='Survived')
# 375 = three hundred and seventy-six factor (train_df, col='Survived')
# 376 = three hundred and seventy-seven factor (train_df, col='Survived')
# 377 = three hundred and seventy-eight factor (train_df, col='Survived')
# 378 = three hundred and seventy-nine factor (train_df, col='Survived')
# 379 = three hundred and eighty factor (train_df, col='Survived')
# 380 = three hundred and eighty-one factor (train_df, col='Survived')
# 381 = three hundred and eighty-two factor (train_df, col='Survived')
# 382 = three hundred and eighty-three factor (train_df, col='Survived')
# 383 = three hundred and eighty-four factor (train_df, col='Survived')
# 384 = three hundred and eighty-five factor (train_df, col='Survived')
# 385 = three hundred and eighty-six factor (train_df, col='Survived')
# 386 = three hundred and eighty-seven factor (train_df, col='Survived')
# 387 = three hundred and eighty-eight factor (train_df, col='Survived')
# 388 = three hundred and eighty-nine factor (train_df, col='Survived')
# 389 = three hundred and ninety factor (train_df, col='Survived')
# 390 = three hundred and ninety-one factor (train_df, col='Survived')
# 391 = three hundred and ninety-two factor (train_df, col='Survived')
# 392 = three hundred and ninety-three factor (train_df, col='Survived')
# 393 = three hundred and ninety-four factor (train_df, col='Survived')
# 394 = three hundred and ninety-five factor (train_df, col='Survived')
# 395 = three hundred and ninety-six factor (train_df, col='Survived')
# 396 = three hundred and ninety-seven factor (train_df, col='Survived')
# 397 = three hundred and ninety-eight factor (train_df, col='Survived')
# 398 = three hundred and ninety-nine factor (train_df, col='Survived')
# 399 = four hundred factor (train_df, col='Survived')
# 400 = four hundred and one factor (train_df, col='Survived')
# 401 = four hundred and two factor (train_df, col='Survived')
# 402 = four hundred and three factor (train_df, col='Survived')
# 403 = four hundred and four factor (train_df, col='Survived')
# 404 = four hundred and five factor (train_df, col='Survived')
# 405 = four hundred and six factor (train_df, col='Survived')
# 406 = four hundred and seven factor (train_df, col='Survived')
# 407 = four hundred and eight factor (train_df, col='Survived')
# 408 = four hundred and nine factor (train_df, col='Survived')
# 409 = four hundred and ten factor (train_df, col='Survived')
# 410 = four hundred and eleven factor (train_df, col='Survived')
# 411 = four hundred and twelve factor (train_df, col='Survived')
# 412 = four hundred and thirteen factor (train_df, col='Survived')
# 413 = four hundred and fourteen factor (train_df, col='Survived')
# 414 = four hundred and fifteen factor (train_df, col='Survived')
# 415 = four hundred and sixteen factor (train_df, col='Survived')
# 416 = four hundred and seventeen factor (train_df, col='Survived')
# 417 = four hundred and eighteen factor (train_df, col='Survived')
# 418 = four hundred and nineteen factor (train_df, col='Survived')
# 419 = four hundred and twenty factor (train_df, col='Survived')
# 420 = four hundred and twenty-one factor (train_df, col='Survived')
# 421 = four hundred and twenty-two factor (train_df, col='Survived')
# 422 = four hundred and twenty-three factor (train_df, col='Survived')
# 423 = four hundred and twenty-four factor (train_df, col='Survived')
# 424 = four hundred and twenty-five factor (train_df, col='Survived')
# 425 = four hundred and twenty-six factor (train_df, col='Survived')
# 426 = four hundred and twenty-seven factor (train_df, col='Survived')
# 427 = four hundred and twenty-eight factor (train_df, col='Survived')
# 428 = four hundred and twenty-nine factor (train_df, col='Survived')
# 429 = four hundred and thirty factor (train_df, col='Survived')
# 430 = four hundred and thirty-one factor (train_df, col='Survived')
# 431 = four hundred and thirty-two factor (train_df, col='Survived')
# 432 = four hundred and thirty-three factor (train_df, col='Survived')
# 433 = four hundred and thirty-four factor (train_df, col='Survived')
# 434 = four hundred and thirty-five factor (train_df, col='Survived')
# 435 = four hundred and thirty-six factor (train_df, col='Survived')
# 436 = four hundred and thirty-seven factor (train_df, col='Survived')
# 437 = four hundred and thirty-eight factor (train_df, col='Survived')
# 438 = four hundred and thirty-nine factor (train_df, col='Survived')
# 439 = four hundred and forty factor (train_df, col='Survived')
# 440 = four hundred and forty-one factor (train_df, col='Survived')
# 441 = four hundred and forty-two factor (train_df, col='Survived')
# 442 = four hundred and forty-three factor (train_df, col='Survived')
# 443 = four hundred and forty-four factor (train_df, col='Survived')
# 444 = four hundred and forty-five factor (train_df, col='Survived')
# 445 = four hundred and forty-six factor (train_df, col='Survived')
# 446 = four hundred and forty-seven factor (train_df, col='Survived')
# 447 = four hundred and forty-eight factor (train_df, col='Survived')
# 448 = four hundred and forty-nine factor (train_df, col='Survived')
# 449 = four hundred and fifty factor (train_df, col='Survived')
# 450 = four hundred and fifty-one factor (train_df, col='Survived')
# 451 = four hundred and fifty-two factor (train_df, col='Survived')
# 452 = four hundred and fifty-three factor (train_df, col='Survived')
# 453 = four hundred and fifty-four factor (train_df, col='Survived')
# 454 = four hundred and fifty-five factor (train_df, col='Survived')
# 455 = four hundred and fifty-six factor (train_df, col='Survived')
# 456 = four hundred and fifty-seven factor (train_df, col='Survived')
# 457 = four hundred and fifty-eight factor (train_df, col='Survived')
# 458 = four hundred and fifty-nine factor (train_df, col='Survived')
# 459 = four hundred and sixty factor (train_df, col='Survived')
# 460 = four hundred and sixty-one factor (train_df, col='Survived')
# 461 = four hundred and sixty-two factor (train_df, col='Survived')
# 462 = four hundred and sixty-three factor (train_df, col='Survived')
# 463 = four hundred and sixty-four factor (train_df, col='Survived')
# 464 = four hundred and sixty-five factor (train_df, col='Survived')
# 465 = four hundred and sixty-six factor (train_df, col='Survived')
# 466 = four hundred and sixty-seven factor (train_df, col='Survived')
# 467 = four hundred and sixty-eight factor (train_df, col='Survived')
# 468 = four hundred and sixty-nine factor (train_df, col='Survived')
# 469 = four hundred and seventy factor (train_df, col='Survived')
# 470 = four hundred and seventy-one factor (train_df, col='Survived')
# 471 = four hundred and seventy-two factor (train_df, col='Survived')
# 472 = four hundred and seventy-three factor (train_df, col='Survived')
# 473 = four hundred and seventy-four factor (train_df, col='Survived')
# 474 = four hundred and seventy-five factor (train_df, col='Survived')
# 475 = four hundred and seventy-six factor (train_df, col='Survived')
# 476 = four hundred and seventy-seven factor (train_df, col='Survived')
# 477 = four hundred and seventy-eight factor (train_df, col='Survived')
# 478 = four hundred and
```