Model building:

Logistic regression score (logistic_model): 77.61194029850746

Decision tree score (tree_model): 75.0

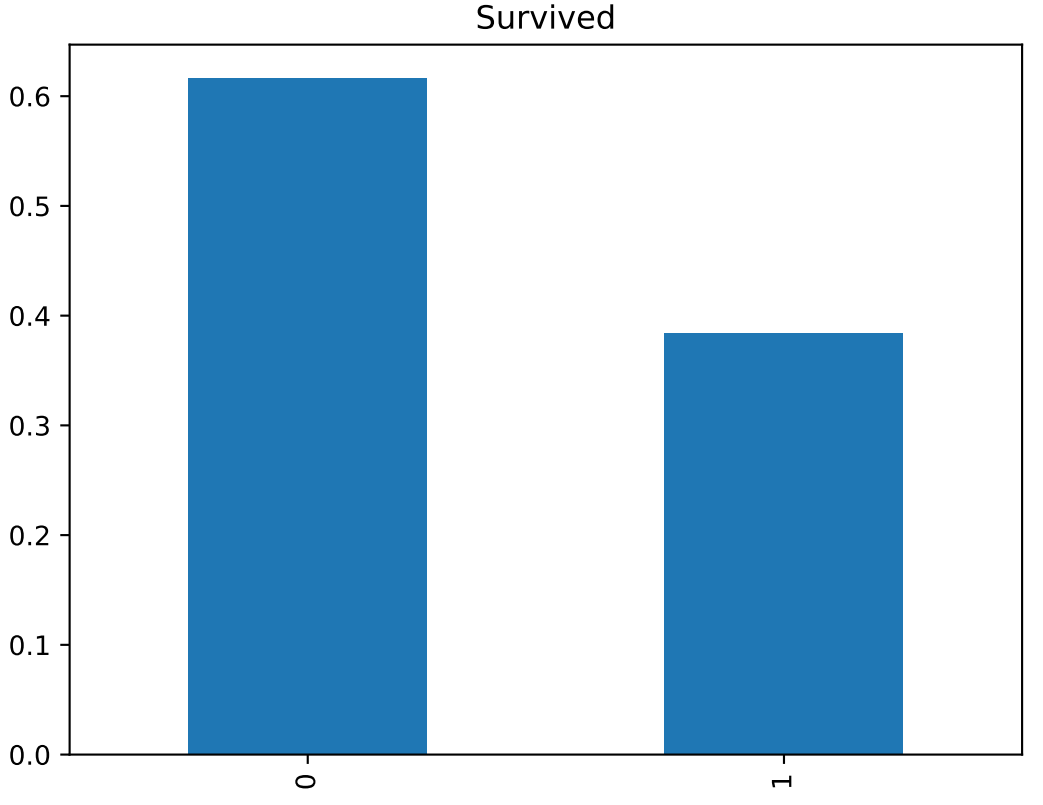
Random forest score (forest_model): 76.49253731343283

XGBoost score (xgb_model): 76.49253731343283

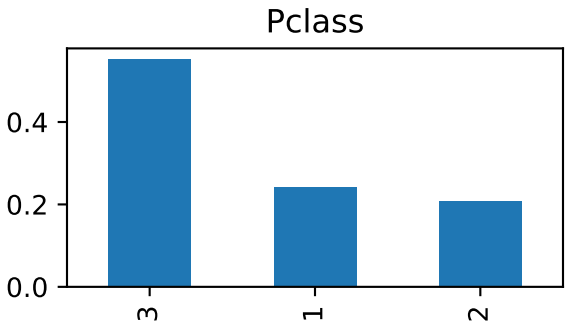
Model with highest accuracy is: logistic_model

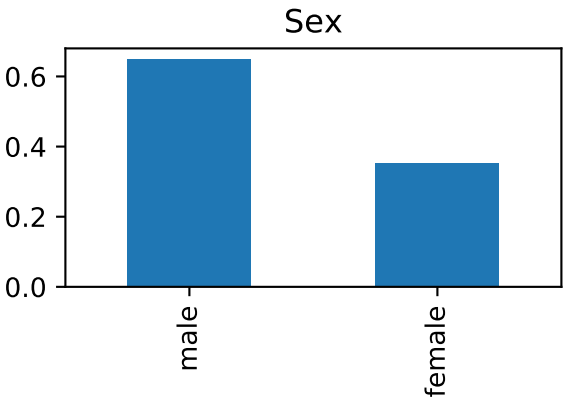
Using this model to predict the test data:

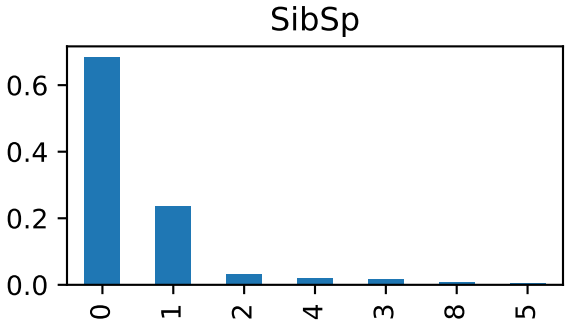
[0 0 0 0 1 0 1 0 1 0 0 0 1 0 1 1 0 0 1 0 0 0 1 1 1 0 1 0 0 0 0 0 1 1 1 0 1
 1 0 0 0 0 0 1 1 0 0 0 1 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 1 1 1 0 0 1 1 0 1 0
 1 0 0 1 0 1 0 0 0 0 0 0 1 1 1 1 1 0 1 0 1 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0 0
 1 1 1 1 0 0 1 0 1 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0
 0 0 1 0 0 0 0 0 1 1 0 1 1 0 1 0 0 1 0 0 1 1 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1
 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 0 1 0 1 1 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0
 1 0 1 1 0 1 0 0 0 1 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0 0 0 0 1
 0 0 0 1 1 0 0 0 0 1 0 0 0 1 1 0 1 0 0 0 0 1 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0
 1 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0
 1 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 1 1 0 0 0 1 0 1 0 0 1 0 1 1 0 1 1 0 1 1 0
 0 1 0 0 1 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 1 0 1 1 0 0 0
 0 1 1 1 1 1 0 1 0 0 0]

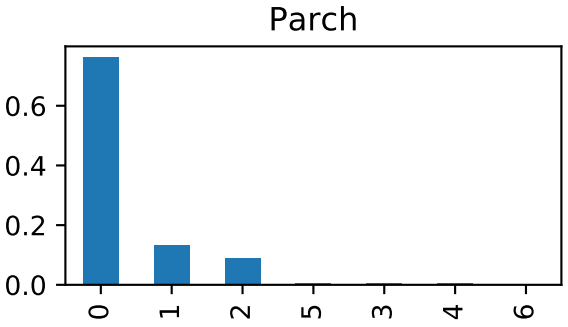

Survived
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1

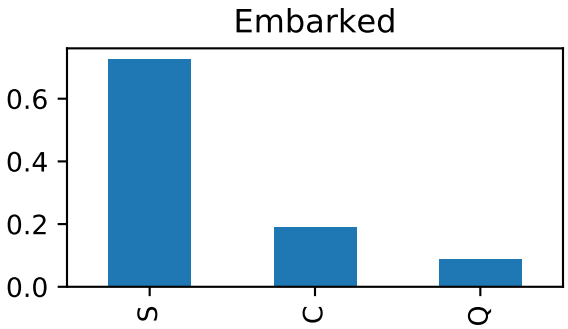
Analysing categorical variables



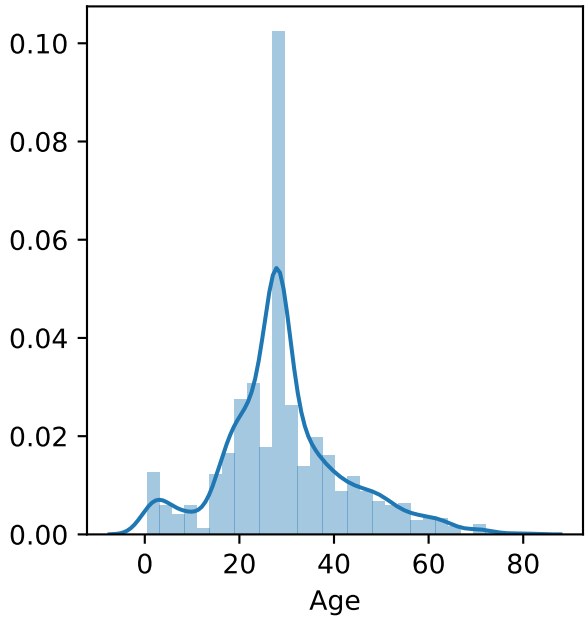


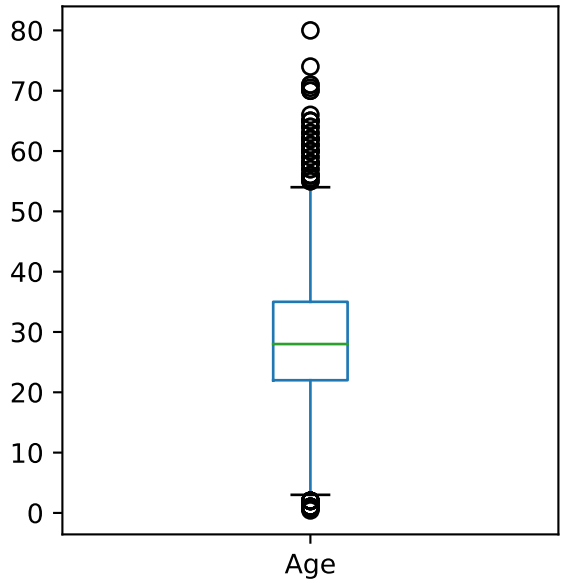


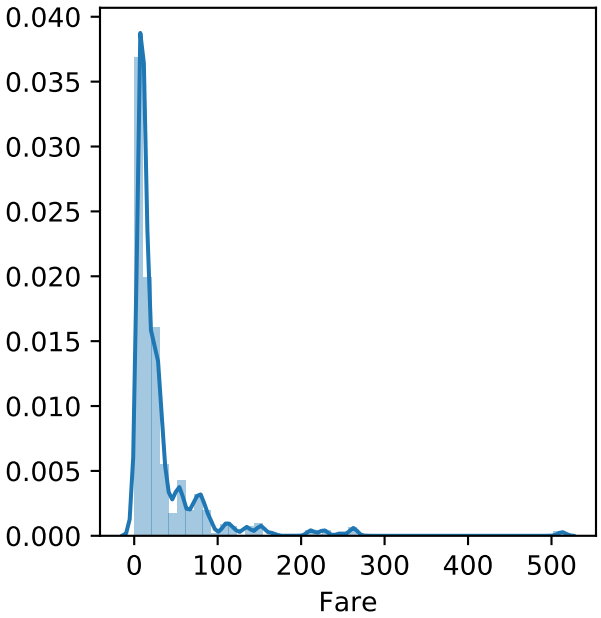


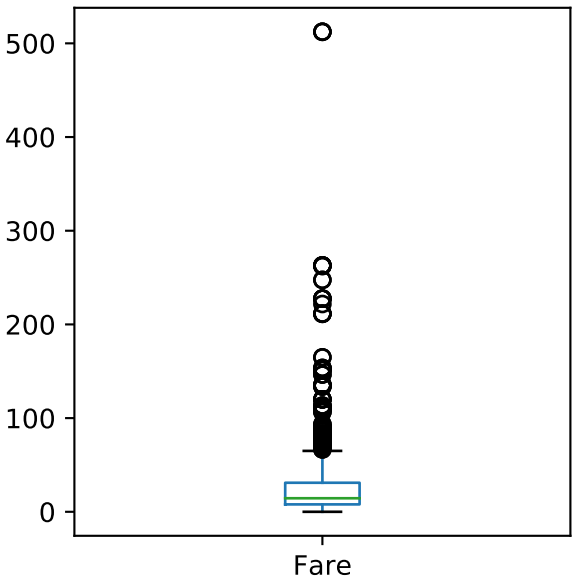


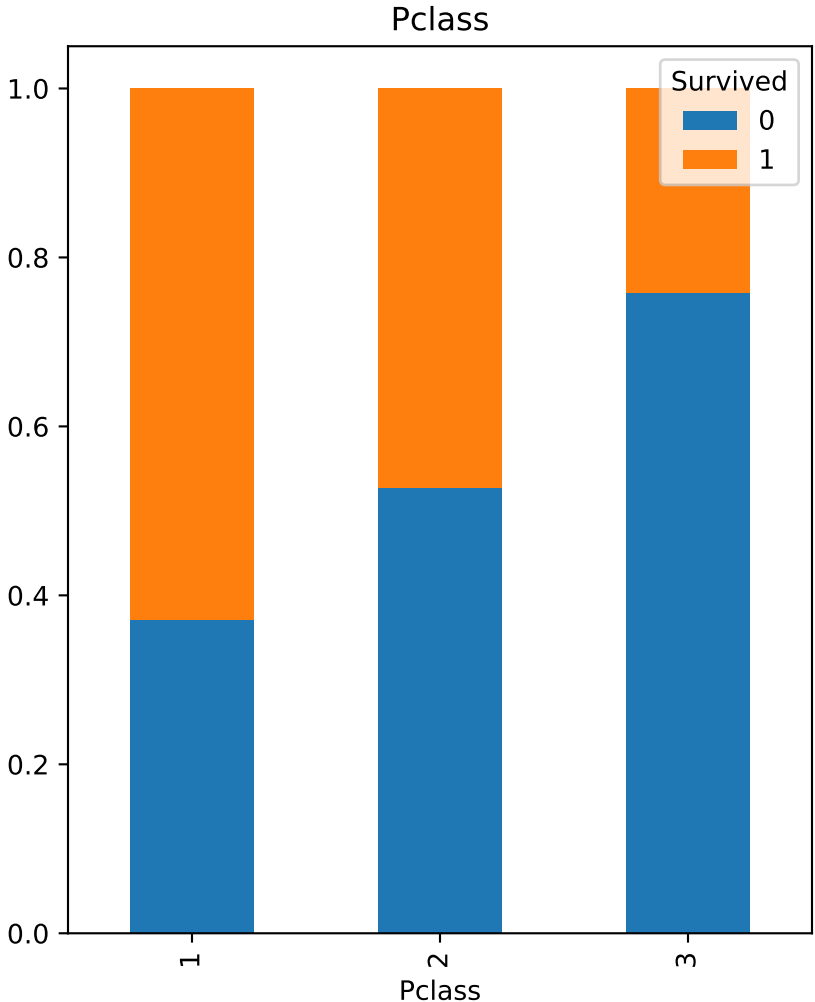
Analysing categorical variables

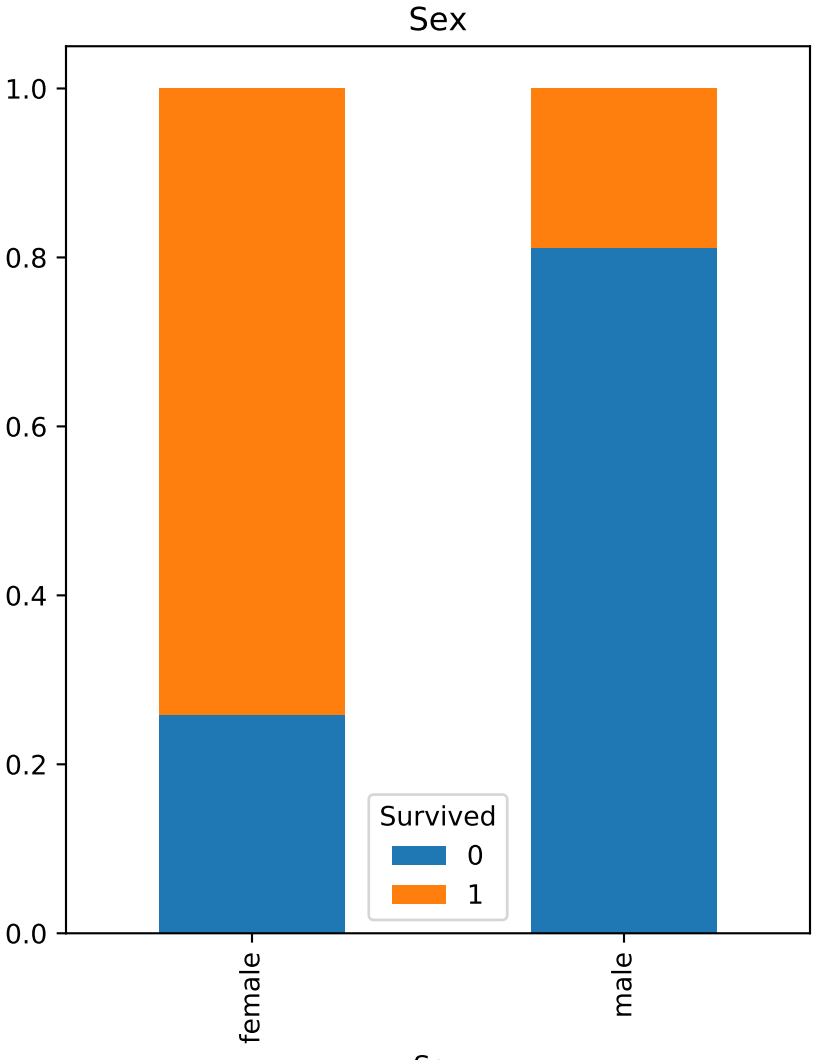


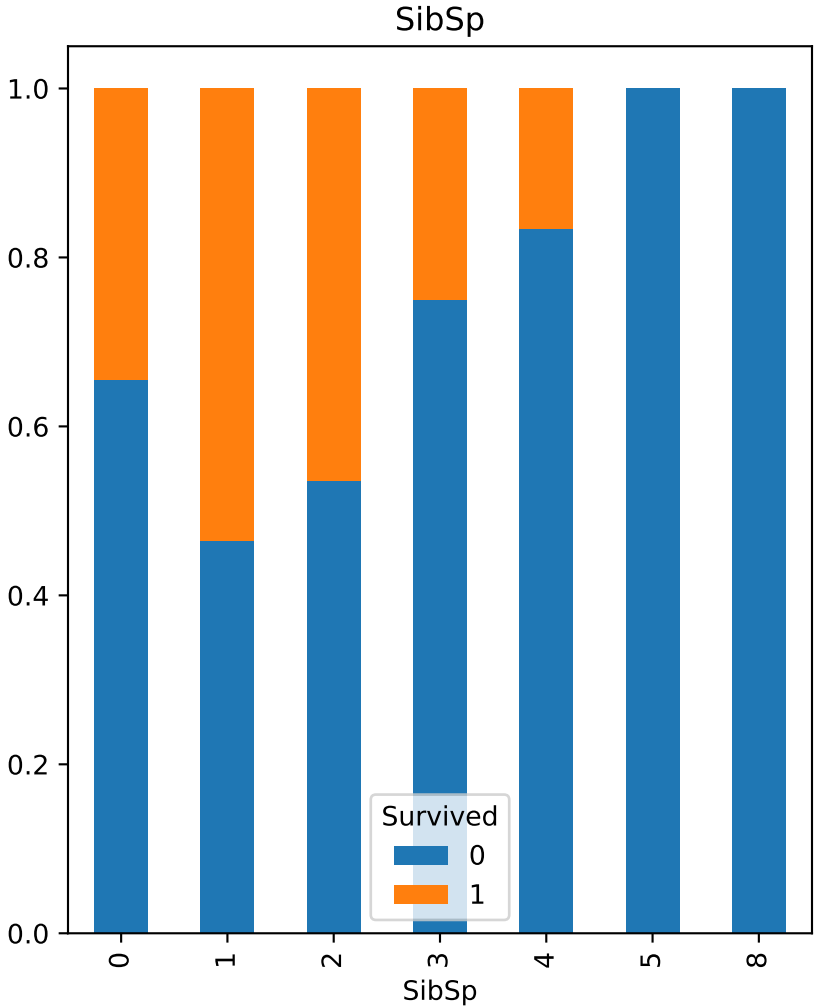


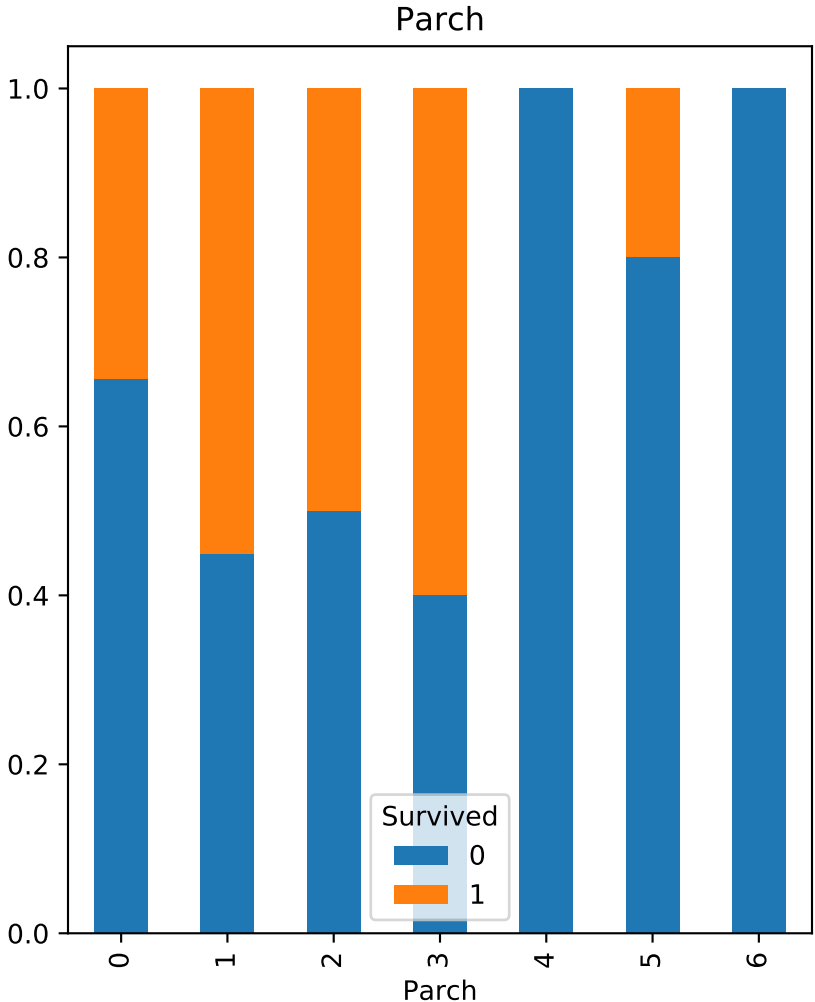


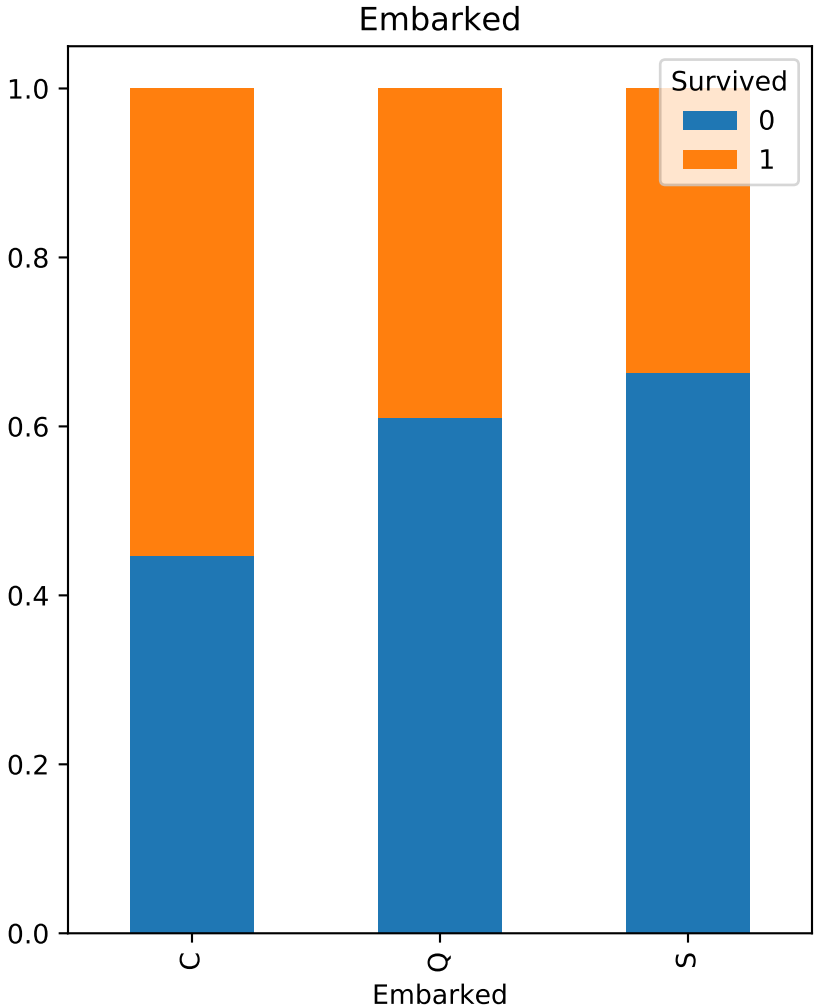



Pclass
Survived
0
1
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
Pclass

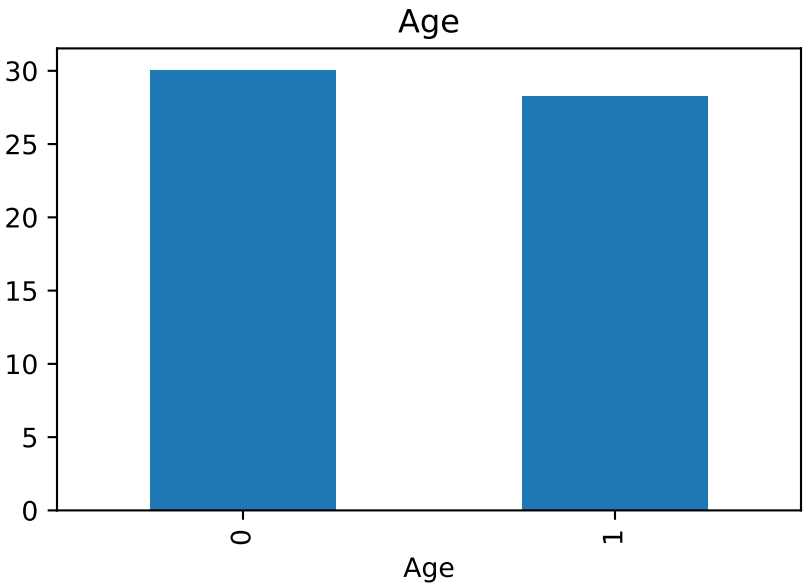

Sex
1.0
0.8
0.6
0.4
0.2
0.0
Survived
0
1
female
male

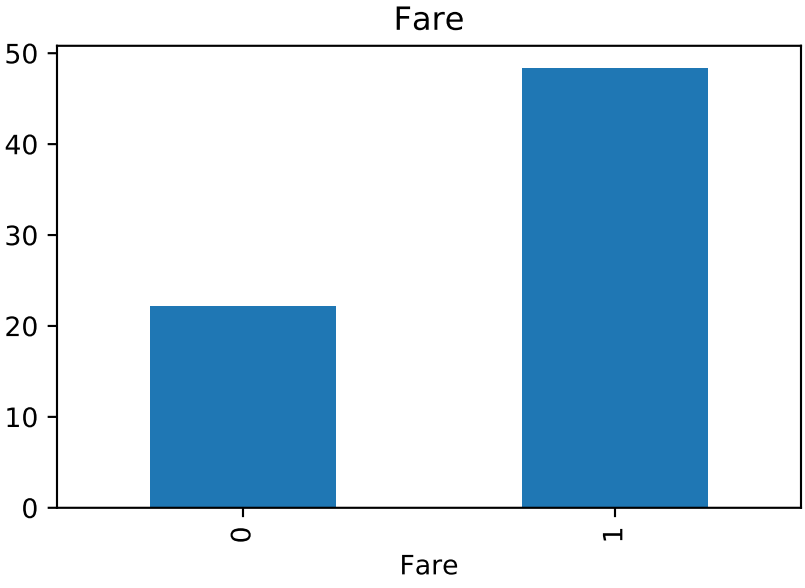

SibSp
1.0
0.8
0.6
0.4
0.2
0.0
Survived
0
1
0
1
2
3
4
5
8
SibSp

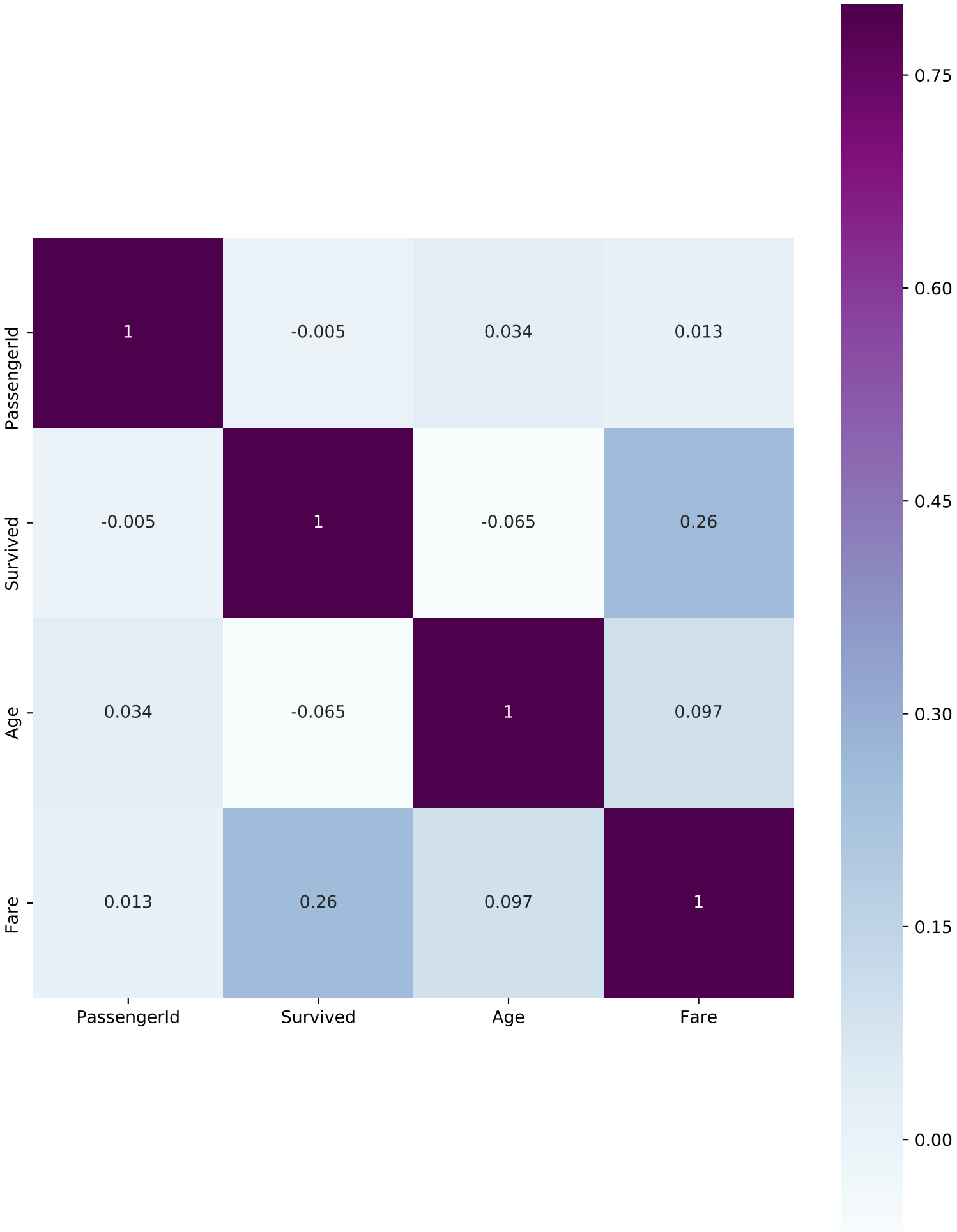

Parch
1.0
0.8
0.6
0.4
0.2
0.0
Survived
0
1
0
1
2
3
4
5
6
Parch

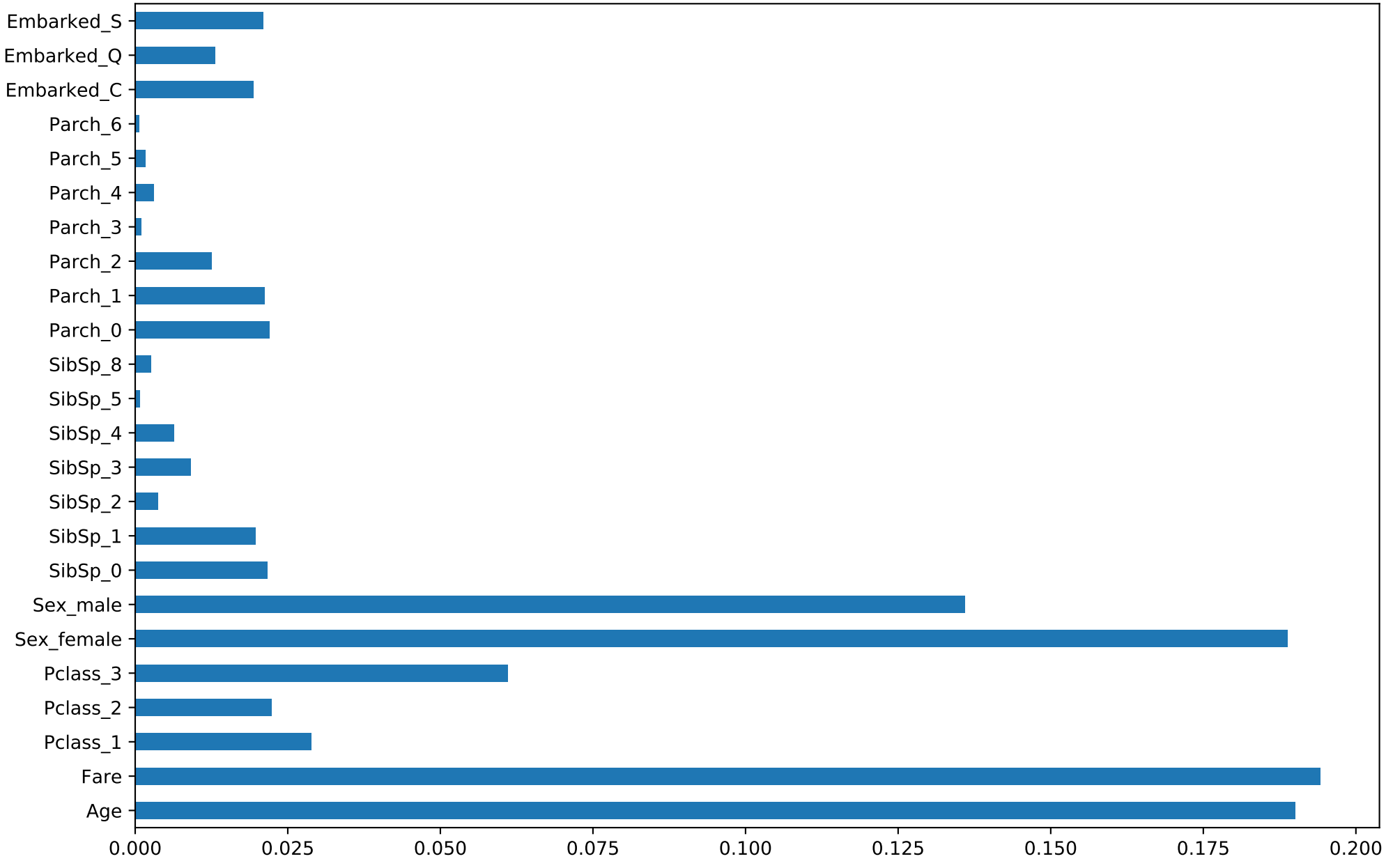

Embarked
Survived
0
1
1.0
0.8
0.6
0.4
0.2
0.0
C
Q
S
Embarked

Analysing numerical variables vs decision variable






PassengerId
Survived
Age
Fare
1
-0.005
0.034
0.013
-0.005
1
-0.065
0.26
0.034
-0.065
1
0.097
0.013
0.26
0.097
1
PassengerId
Survived
Age
Fare
0.75
0.60
0.45
0.30
0.15
0.00


Embarked_S
Embarked_Q
Embarked_C
Parch_6
Parch_5
Parch_4
Parch_3
Parch_2
Parch_1
Parch_0
SibSp_8
SibSp_5
SibSp_4
SibSp_3
SibSp_2
SibSp_1
SibSp_0
Sex_male
Sex_female
Pclass_3
Pclass_2
Pclass_1
Fare
Age
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200