

# TITANIC SURVIVAL



Natan Sisoiev (1706198)  
Ricard Urpí (1711326)  
Ferran Villarta (1704051)

	Title	Count
0	Mr	517
1	Miss	182
2	Mrs	125
3	Master	40
4	Dr	7
5	Rev	6
6	Col	2
7	Mlle	2
8	Major	2
9	Ms	1
10	Mme	1
11	Don	1
12	Lady	1
13	Sir	1
14	Capt	1
15	Countess	1
16	Jonkheer	1

	Title	Count
0	Mr	517
1	Miss	182
2	Mrs	125
3	Master	40
4	Dr	7
5	Rev	6
6	Col	2
7	Mlle	2
8	Major	2
9	Ms	1
10	Mme	1
11	Don	1
12	Lady	1
13	Sir	1
14	Capt	1
15	Countess	1
16	Jonkheer	1

	WithAge	MissingAge
Name		
Capt	1	0
Col	2	0
Countess	1	0
Don	1	0
Dr	6	1
Jonkheer	1	0
Lady	1	0
Major	2	0
Master	36	4
Miss	146	36
Mlle	2	0
Mme	1	0
Mr	398	119
Mrs	108	17
Ms	1	0
Rev	6	0
Sir	1	0



	Title	Count
0	Mr	517
1	Miss	182
2	Mrs	125
3	Master	40
4	Dr	7
5	Rev	6
6	Col	2
7	Mlle	2
8	Major	2
9	Ms	1
10	Mme	1
11	Don	1
12	Lady	1
13	Sir	1
14	Capt	1
15	Countess	1
16	Jonkheer	1

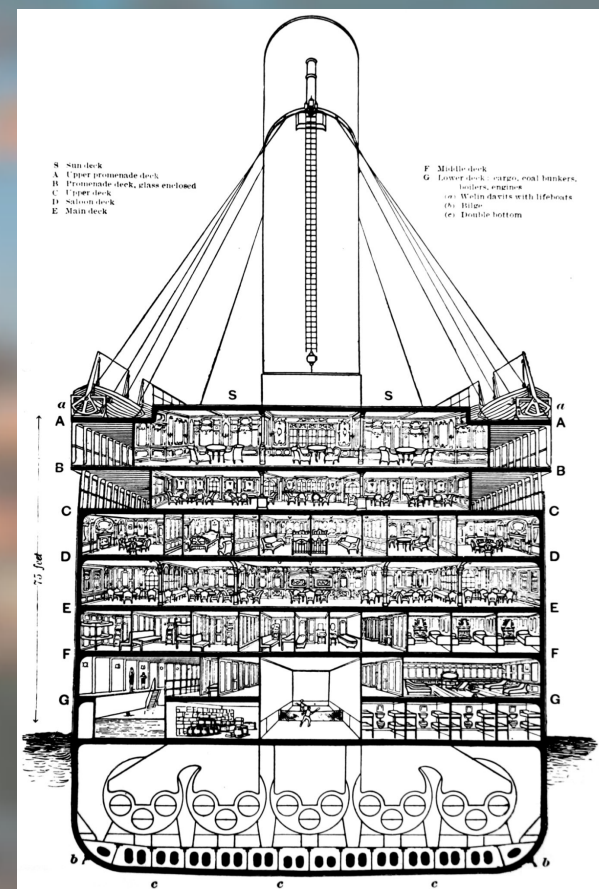
	WithAge	MissingAge
Name		
Capt	1	0
Col	2	0
Countess	1	0
Don	1	0
Dr	6	1
Jonkheer	1	0
Lady	1	0
Major	2	0
Master	36	4
Miss	146	36
Mlle	2	0
Mme	1	0
Mr	398	119
Mrs	108	17
Ms	1	0
Rev	6	0
Sir	1	0

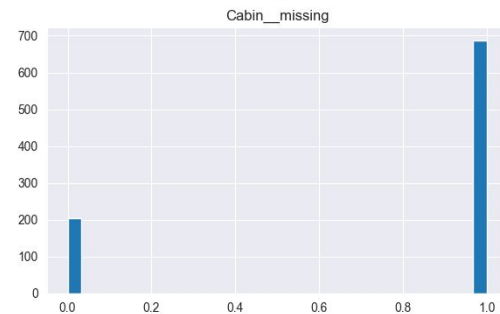
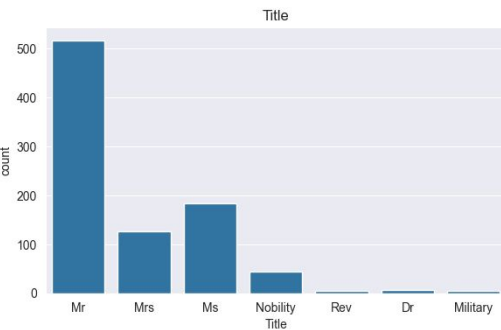
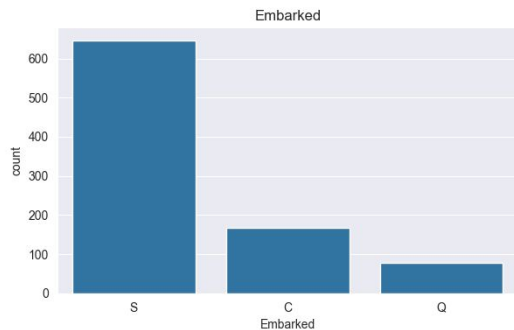
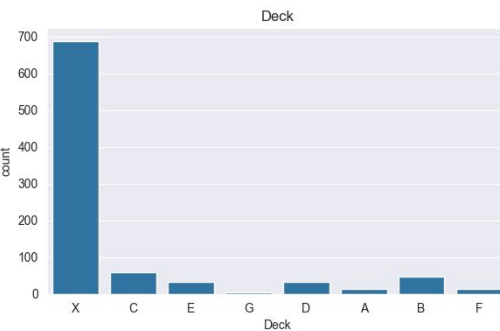
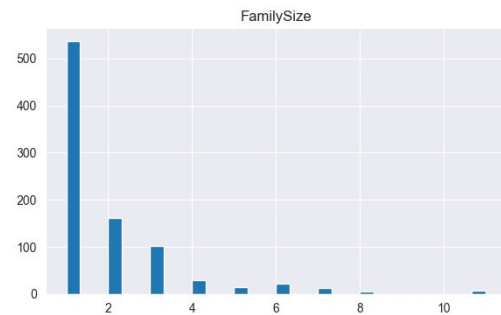
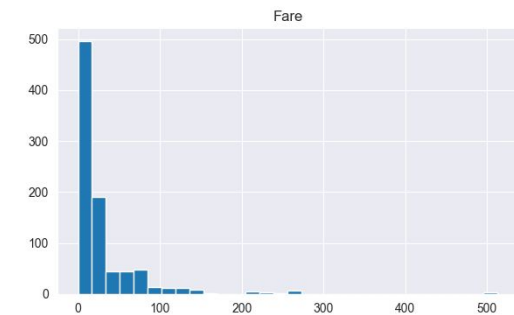
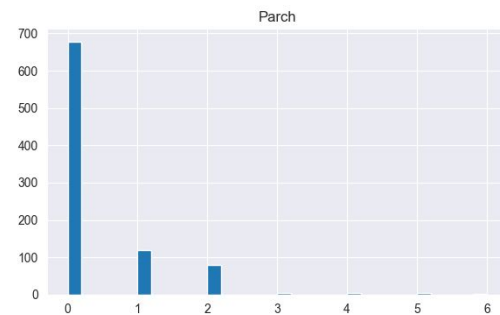
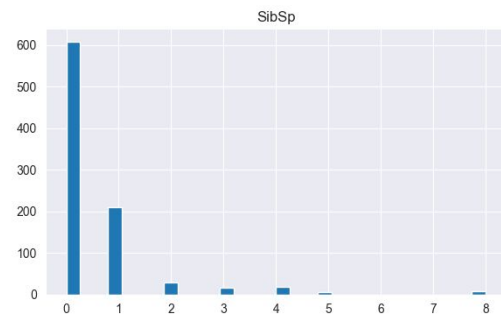
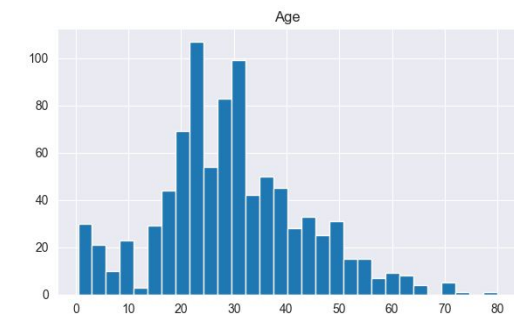
	Title	Count
0	Mr	517
1	Ms	185
2	Mrs	126
3	Nobility	45
4	Dr	7
5	Rev	6
6	Military	5

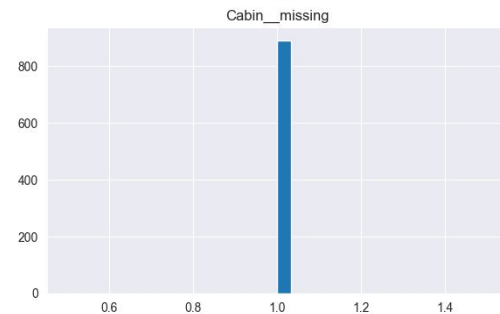
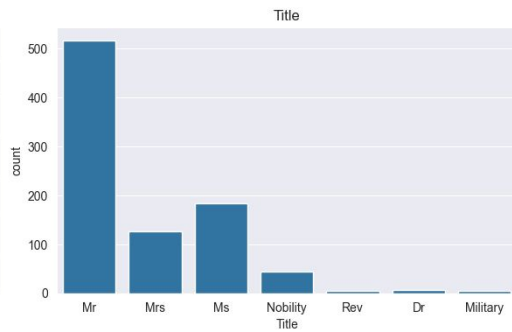
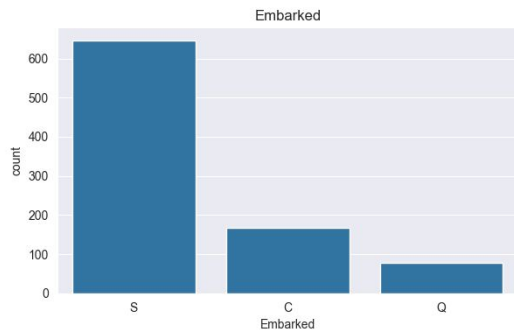
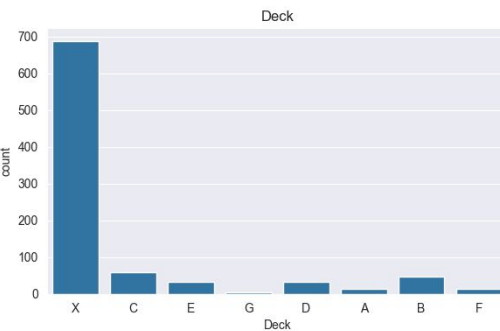
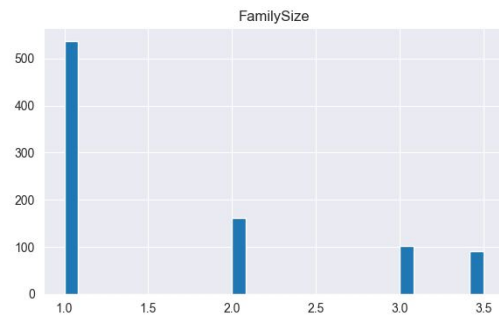
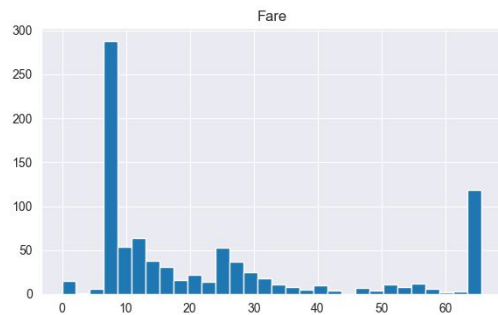
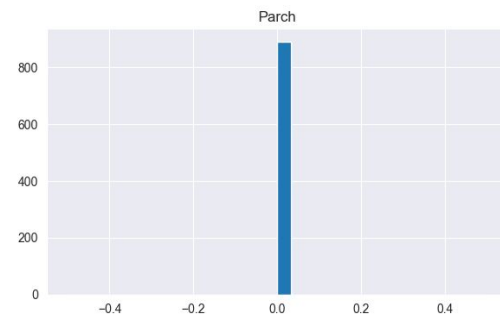
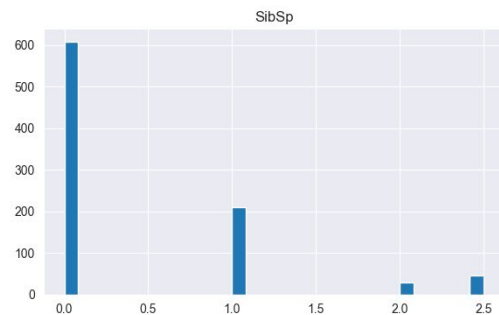
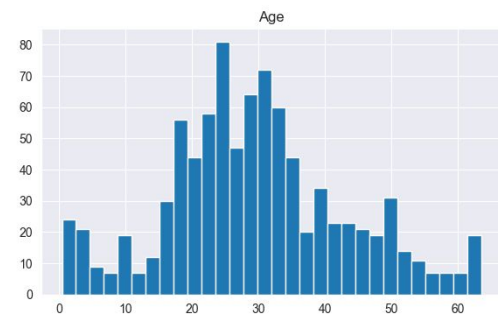
	Title	Count
0	Mr	517
1	Miss	182
2	Mrs	125
3	Master	40
4	Dr	7
5	Rev	6
6	Col	2
7	Mlle	2
8	Major	2
9	Ms	1
10	Mme	1
11	Don	1
12	Lady	1
13	Sir	1
14	Capt	1
15	Countess	1
16	Jonkheer	1

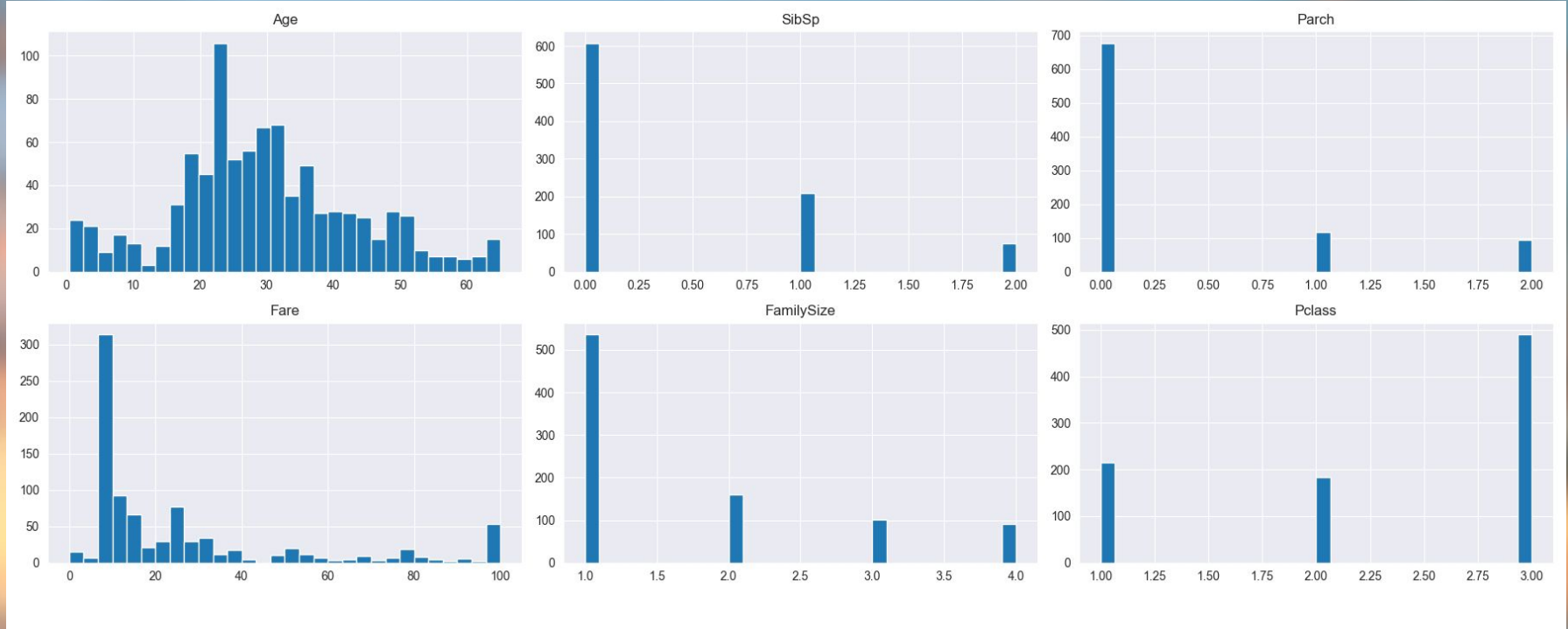
	WithAge	MissingAge
Name		
Capt	1	0
Col	2	0
Countess	1	0
Don	1	0
Dr	6	1
Jonkheer	1	0
Lady	1	0
Major	2	0
Master	36	4
Miss	146	36
Mlle	2	0
Mme	1	0
Mr	398	119
Mrs	108	17
Ms	1	0
Rev	6	0
Sir	1	0

	Title	Count
0	Mr	517
1	Ms	185
2	Mrs	126
3	Nobility	45
4	Dr	7
5	Rev	6
6	Military	5

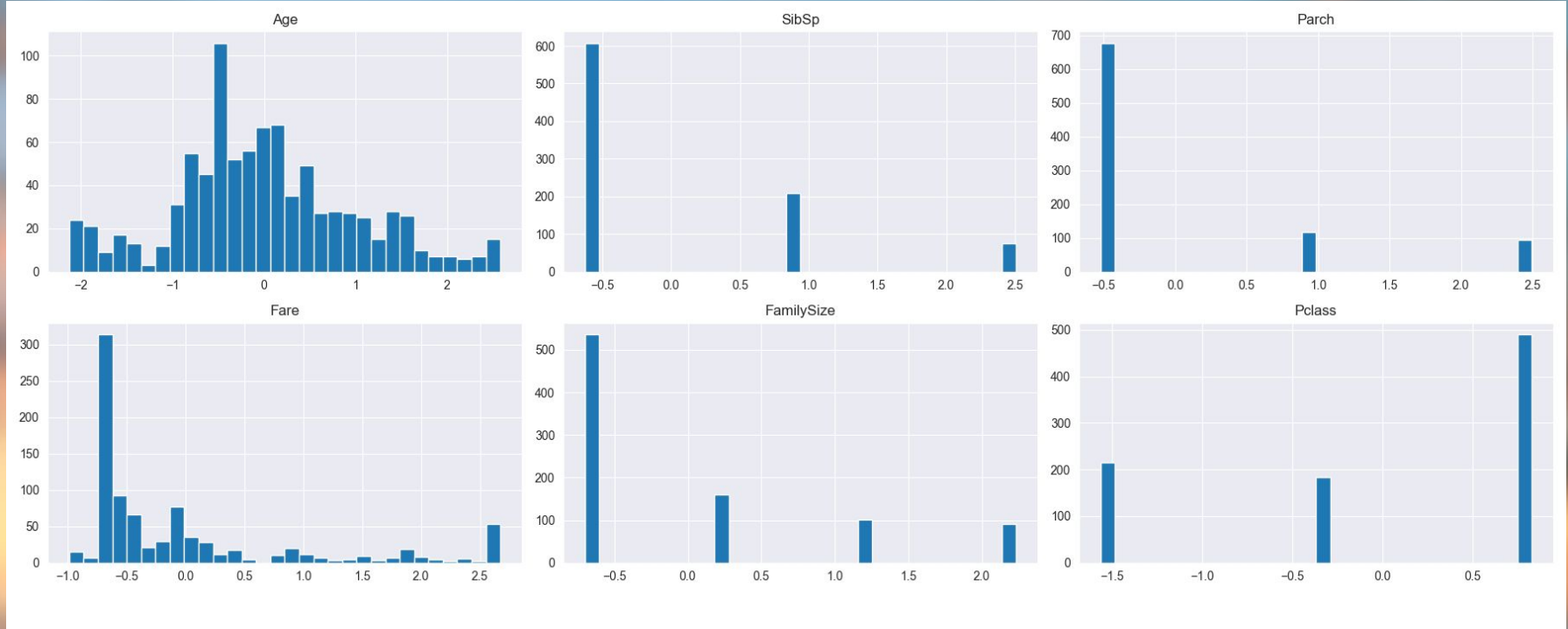


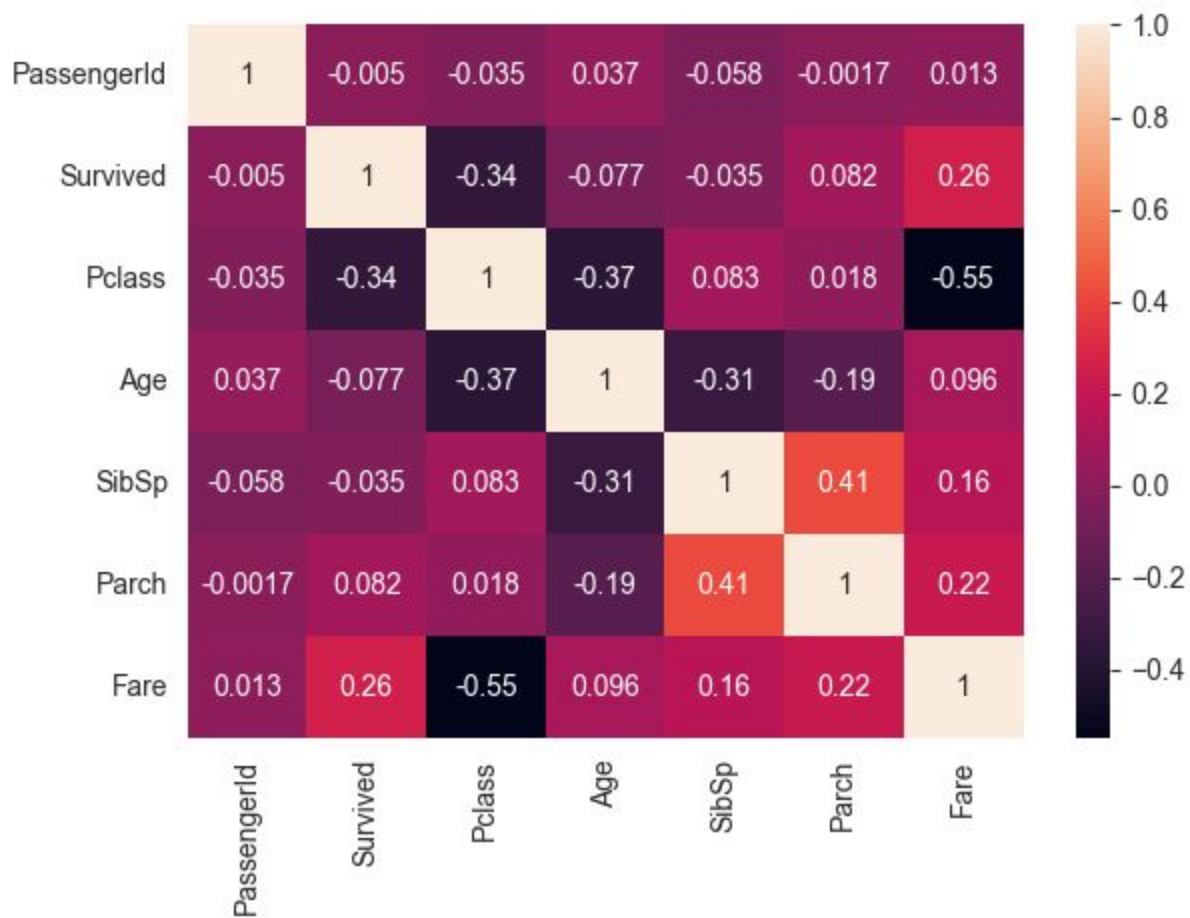












	Model	Features	F1	AUC-ROC	AUC_PR
0	Gradient Boosting	[Pclass, Fare, Sex_male]	0.747844	0.871349	0.844629
1	Extra Trees	[Pclass, Fare, Sex_male]	0.746277	0.836392	0.830462
2	Random Forest	[Pclass, Fare, Sex_male]	0.745747	0.872642	0.845793
3	Support Vector Machine	[Pclass, Fare, Sex_male]	0.724352	0.763182	0.659370



# MÈTRICA I COMPARACIÓ DE MODELS

ACCURACY → 38% - 62%

PRECISION

RECALL



# MÈTRICA I COMPARACIÓ DE MODELS

ACCURACY → 38% - 62%

PRECISION }  
RECALL } F1-SCORE

$$\text{F1-SCORE} = 2 \cdot \frac{\text{PRECISION} \cdot \text{RECALL}}{\text{PRECISION} + \text{RECALL}}$$

# MÈTRICA I COMPARACIÓ DE MODELS

ACCURACY → 38% - 62%

PRECISION-RECALL CURVE → roc-auc

PRECISION }  
RECALL } F1-SCORE

ROC CURVE → average precision

$$\text{F1-SCORE} = 2 \cdot \frac{\text{PRECISION} \cdot \text{RECALL}}{\text{PRECISION} + \text{RECALL}}$$

*Random Forest*

*Extra Trees*

*Gradient Boosting*

*Logistic Regression*

*LinearSVC*

*KNN*

*Naive Bayes*

*Gradient Descent*

*SVM*

*Perceptron*

*Passive Aggressive*

*Neural Network*

*Random Forest*

*Extra Trees*

*Gradient Boosting*

*Logistic Regression*

*LinearSVC*

*KNN*

*Naive Bayes*

*Gradient Descent*

*SVM*

*Perceptron*

*Passive Aggressive*

*Neural Network*

# Repeated Stratified K-Fold

*Random Forest*

*Extra Trees*

*Gradient Boosting*

*Logistic Regression*

*LinearSVC*

*KNN*

*Naive Bayes*

*Gradient Descent*

*SVM*

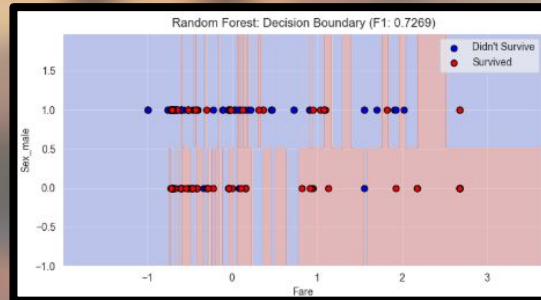
*Perceptron*

*Passive Aggressive*

*Neural Network*

# Repeated Stratified K-Fold

## GridSearchCV





Random Forest

Extra Trees

Gradient Boosting

Logistic Regression

LinearSVC

KNN

Naive Bayes

Gradient Descent

SVM

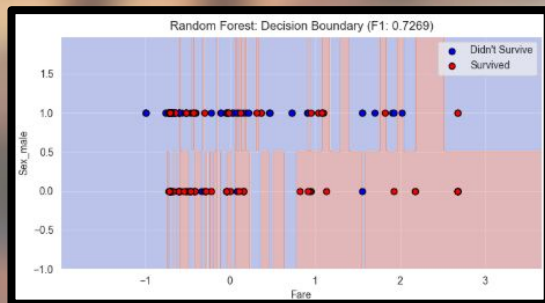
Perceptron

Passive Aggressive

Neural Network

# Repeated Stratified K-Fold

## GridSearchCV



Model	Mean Test Score	Best Parameters
Logistic Reg.	0.7106	C=1, max_iter=1000, penalty=l2
KNN	0.7098	metric=euclidean, n_neighbors=9, weights=uniform
Grad. Descent	0.7109	alpha=0.001, eta0=0.01, lr=optimal
Random Forest	0.7042	max_depth=5, n_estimators=101
Extra Trees	0.7101	max_depth=2, n_estimators=51
SVM	0.7101	C=0.1, gamma=scale, kernel=linear
LinearSVC	0.7101	C=0.1, loss=hinge, max_iter=1000
<b>Grad. Boost.</b>	<b>0.7265</b>	lr=0.1, max_depth=3, n_estimators=201
Naive Bayes	0.7164	var_smoothing=1e-12
Perceptron	0.6108	alpha=1e-05, eta0=0.1, max_iter=1000
Passive Agg.	0.7101	C=0.001, max_iter=1000
Neural Net	0.7093	act=tanh, alpha=1e-05, lr=adaptive, max_iter=1000

**GRADIENT BOOSTING**

**0.7265**

**1r=0.1, max\_depth=3, n\_estimators = 201**

Random Forest

Extra Trees

Gradient Boosting

Logistic Regression

LinearSVC

KNN

Naive Bayes

Gradient Descent

SVM

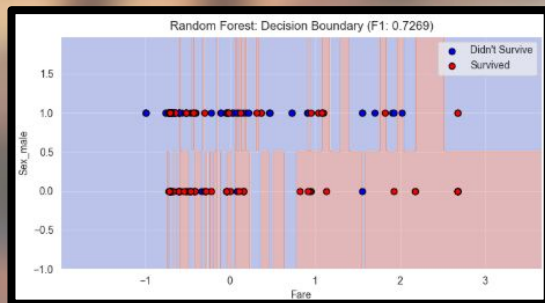
Perceptron

Passive Aggressive

Neural Network

Repeated  
Stratified  
K-Fold

GridSearchCV

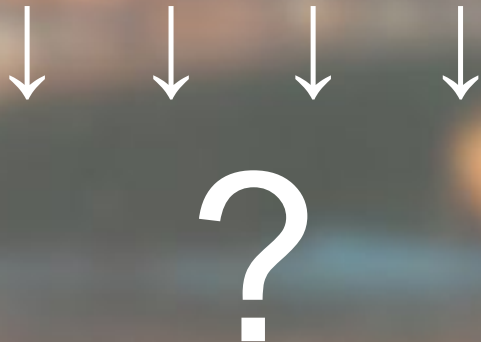


Model	Mean Test Score	Best Parameters
Logistic Reg.	0.7106	C=1, max_iter=1000, penalty=l2
KNN	0.7098	metric=euclidean, n_neighbors=9, weights=uniform
Grad. Descent	0.7109	alpha=0.001, eta0=0.01, lr=optimal
Random Forest	0.7042	max_depth=5, n_estimators=101
Extra Trees	0.7101	max_depth=2, n_estimators=51
SVM	0.7101	C=0.1, gamma=scale, kernel=linear
LinearSVC	0.7101	C=0.1, loss=hinge, max_iter=1000
<b>Grad. Boost.</b>	<b>0.7265</b>	lr=0.1, max_depth=3, n_estimators=201
Naive Bayes	0.7164	var_smoothing=1e-12
Perceptron	0.6108	alpha=1e-05, eta0=0.1, max_iter=1000
Passive Agg.	0.7101	C=0.001, max_iter=1000
Neural Net	0.7093	act=tanh, alpha=1e-05, lr=adaptive, max_iter=1000

GRADIENT BOOSTING

0.7265

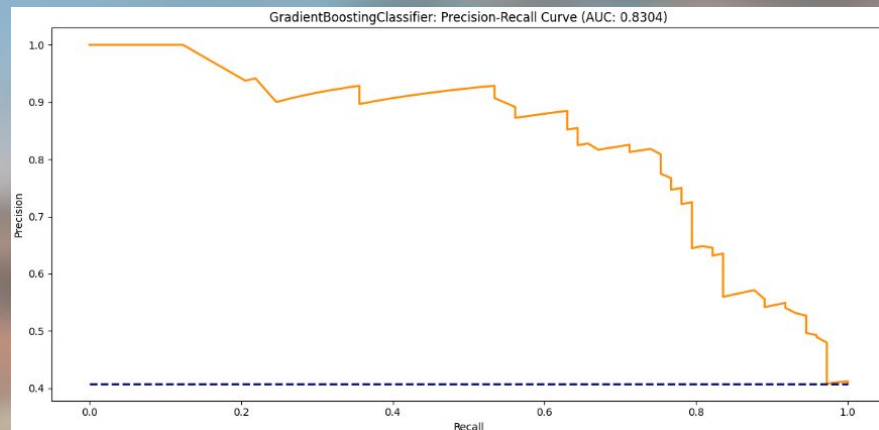
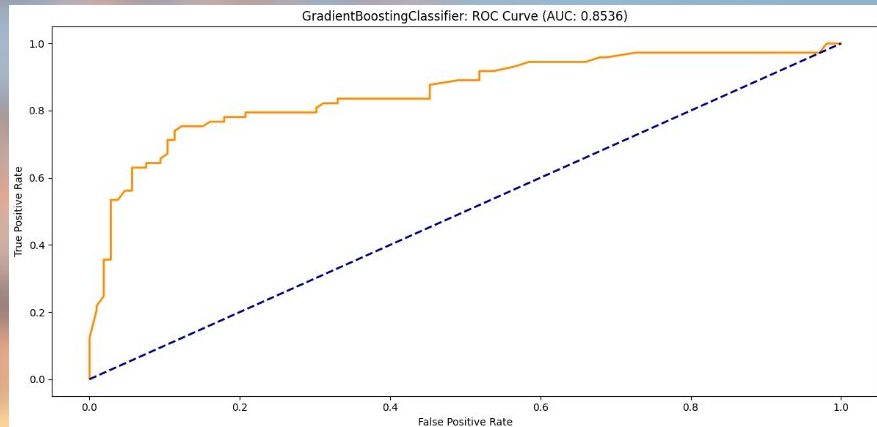
1r=0.1, max\_depth=3, n\_estimators = 201



**[PCLASS, FARE, SEX\_MALE]**



## [PCLASS, FARE, SEX\_MALE]



Accuracy	Precision	Recall	F1-score	AUC-ROC	AUC-PR
0.8044	0.8392	0.6438	0.7496	0.8535	0.8299

Mètriques finals del model Gradient Boosting