

# Data Analysis Report: Data Analysis

Report generated on: 2025-09-15 15:03:08

## 1. Dataset Summary

### General Information

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	891 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 Embarked      891 non-null    object
11 AgeGroup      891 non-null    category
dtypes: category(1), float64(2), int64(5), object(4)
memory usage: 77.8+ KB
```

## Descriptive Statistics

---

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	891.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.361582	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	13.019697	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	22.000000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	35.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

# Null Value Count

---

	null_count
--	------------

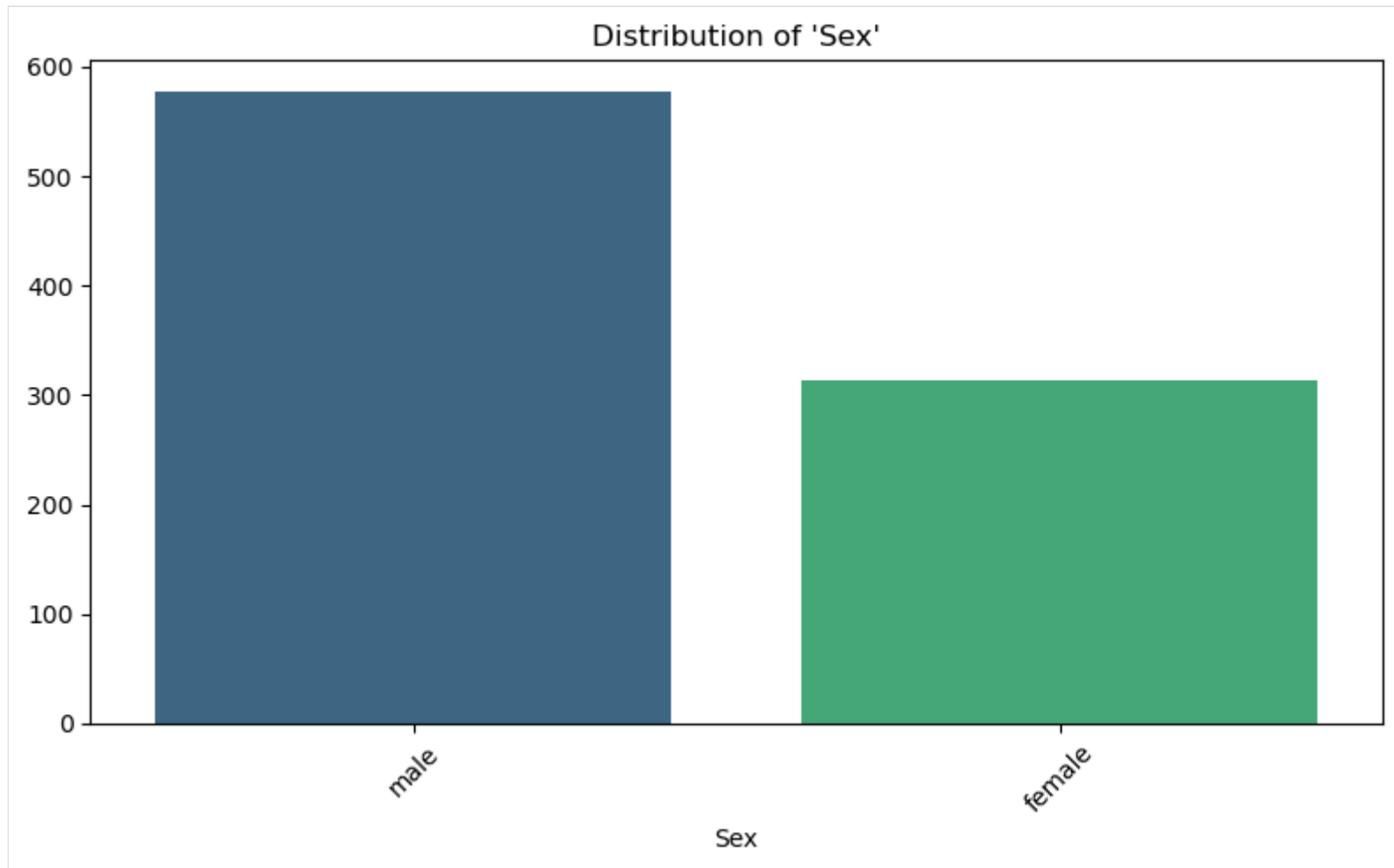
## 2. Categorical Variable Analysis

---

Column: Sex

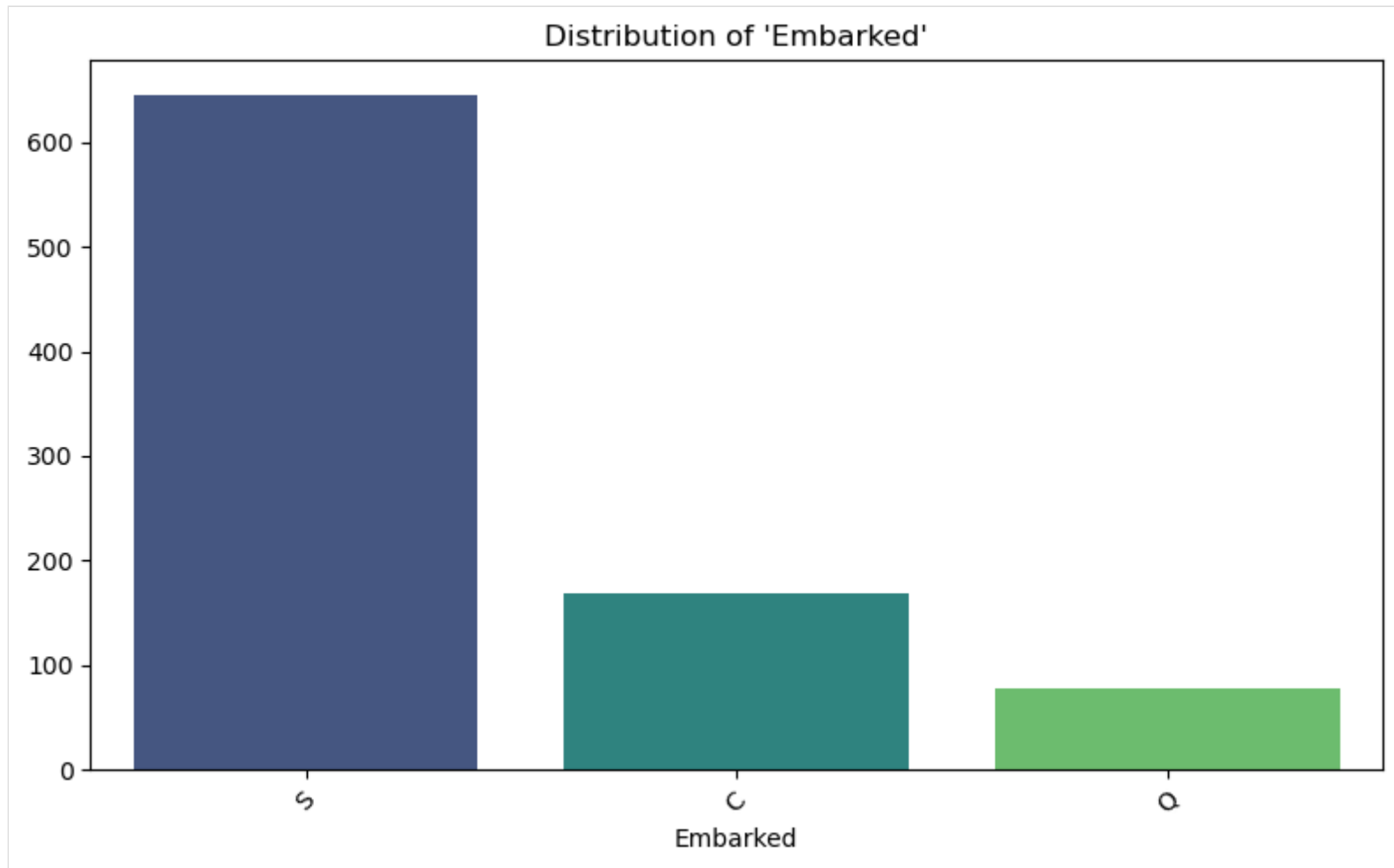
---

	count
Sex	
male	577
female	314



Column: Embarked

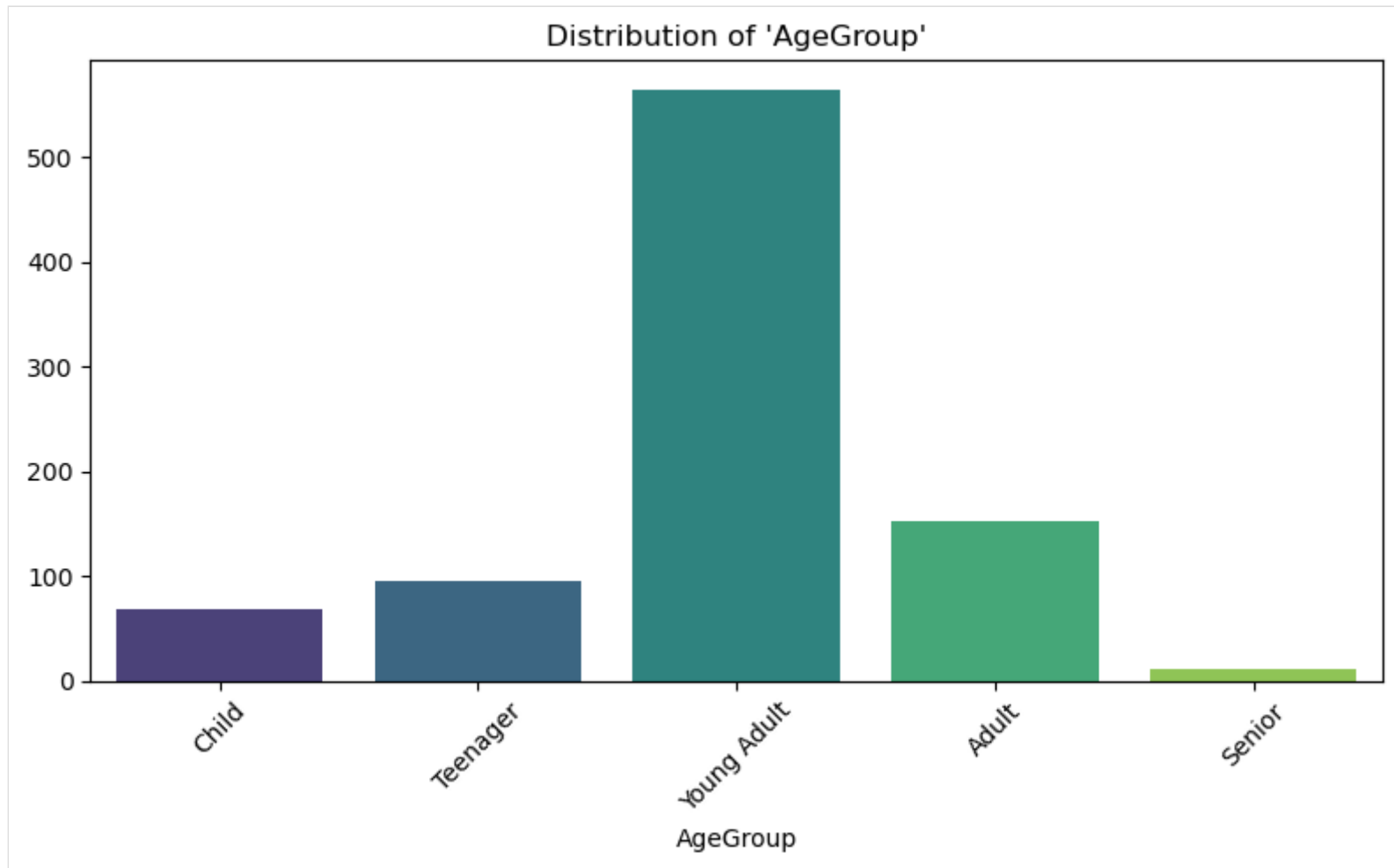
	count
Embarked	
S	646
C	168
Q	77



Column: AgeGroup

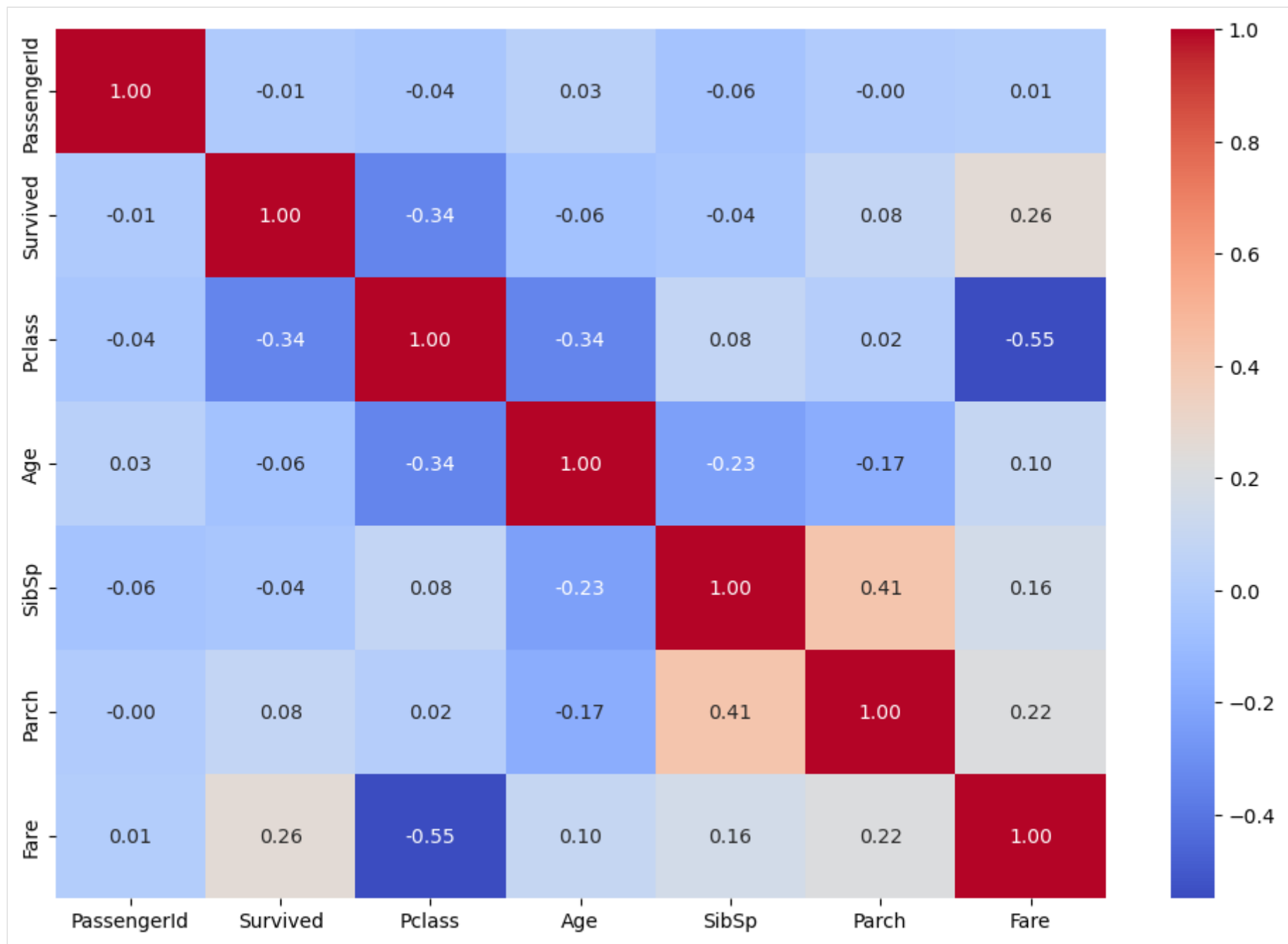
	count
AgeGroup	
Young Adult	564
Adult	152
Teenager	95
Child	69
Senior	11





### 3. Correlation Analysis

---



## Top 5 Correlations

---

Variable 1	Variable 2	Correlation
Fare	Pclass	-0.5495
Parch	SibSp	+0.4148
Age	Pclass	-0.3399
Survived	Pclass	-0.3385
Survived	Fare	+0.2573

## 4. Machine Learning Benchmark Results

### Classification Report

	precision	recall	f1-score	support
0	0.81	0.89	0.85	110
1	0.79	0.67	0.72	69
accuracy			0.80	179
macro avg	0.80	0.78	0.79	179
weighted avg	0.80	0.80	0.80	179

### Confusion Matrix

	Predicted Class 0	Predicted Class 1
Real Class 0	98	12
Real Class 1	23	46