

My Report

李承祐

2024-03-01

Table of contents

Dataset	1
Summary	2

Dataset

```
data <- read.csv("titanic.csv")
head(data)
```

PassengerId	Survived	Pclass						
1	1	0	3					
2	2	1	1					
3	3	1	3					
4	4	1	1					
5	5	0	3					
6	6	0	3					
			Name	Sex	Age	SibSp	Parch	
1	Braund, Mr. Owen Harris			male	22	1	0	
2	Cumings, Mrs. John Bradley (Florence Briggs Thayer)			female	38	1	0	
3	Heikkinen, Miss. Laina			female	26	0	0	
4	Futrelle, Mrs. Jacques Heath (Lily May Peel)			female	35	1	0	
5	Allen, Mr. William Henry			male	35	0	0	
6	Moran, Mr. James			male	NA	0	0	
Ticket		Fare	Cabin	Embarked				
1	A/5 21171	7.2500	S					

2	PC	17599	71.2833	C85	C
3	STON/02.	3101282	7.9250		S
4		113803	53.1000	C123	S
5		373450	8.0500		S
6		330877	8.4583		Q

Summary

```
data$Survived <- as.factor(data$Survived)
data$Pclass <- as.factor(data$Pclass)
data$Sex <- as.factor(data$Sex)
data$SibSp <- as.factor(data$SibSp)
data$Parch <- as.factor(data$Parch)
data$Embarked <- as.factor(data$Embarked)
summary(data)
```

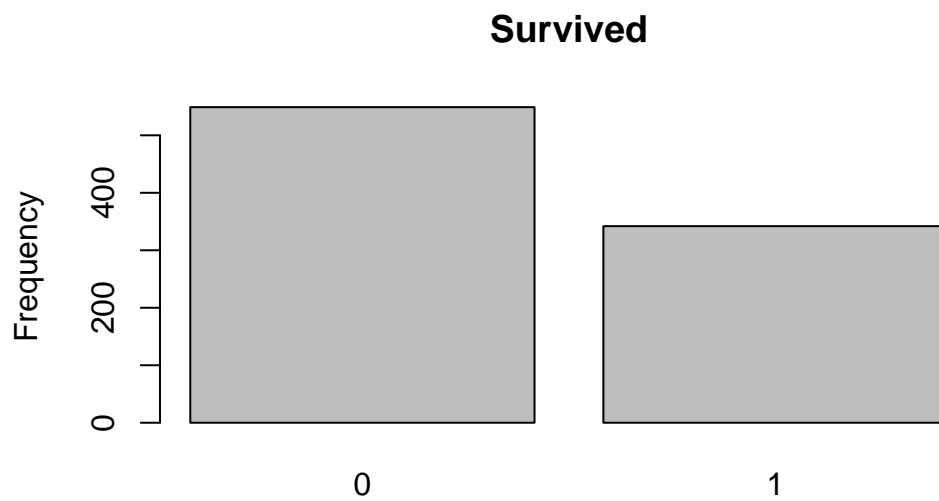
PassengerId	Survived	Pclass	Name	Sex
Min. : 1.0	0:549	1:216	Length:891	female:314
1st Qu.:223.5	1:342	2:184	Class :character	male :577
Median :446.0		3:491	Mode :character	
Mean :446.0				
3rd Qu.:668.5				
Max. :891.0				

Age	SibSp	Parch	Ticket	Fare
Min. : 0.42	0:608	0:678	Length:891	Min. : 0.00
1st Qu.:20.12	1:209	1:118	Class :character	1st Qu.: 7.91
Median :28.00	2: 28	2: 80	Mode :character	Median : 14.45
Mean :29.70	3: 16	3: 5		Mean : 32.20
3rd Qu.:38.00	4: 18	4: 4		3rd Qu.: 31.00
Max. :80.00	5: 5	5: 5		Max. :512.33
NA's :177	8: 7	6: 1		

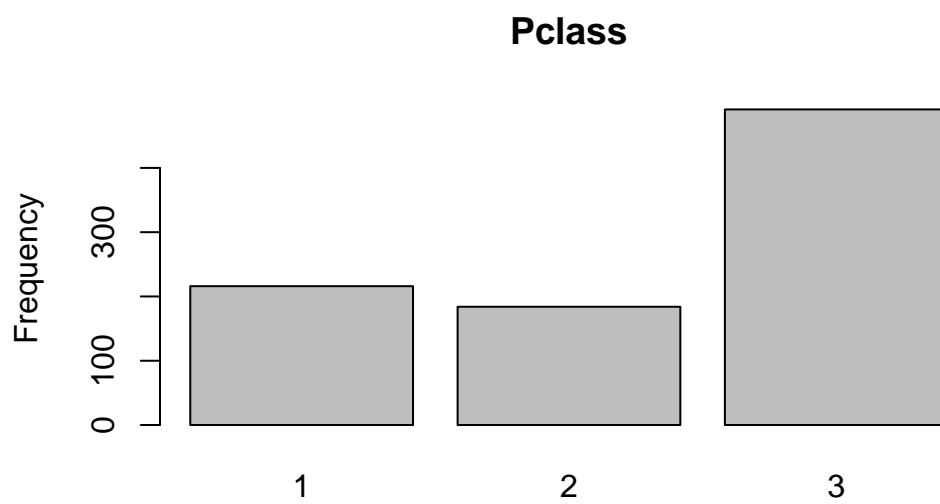
Cabin	Embarked
Length:891	: 2
Class :character	C:168
Mode :character	Q: 77
	S:644

```
#plot
```

```
barplot(table(data$Survived),ylab ="Frequency",main="Survived")
```



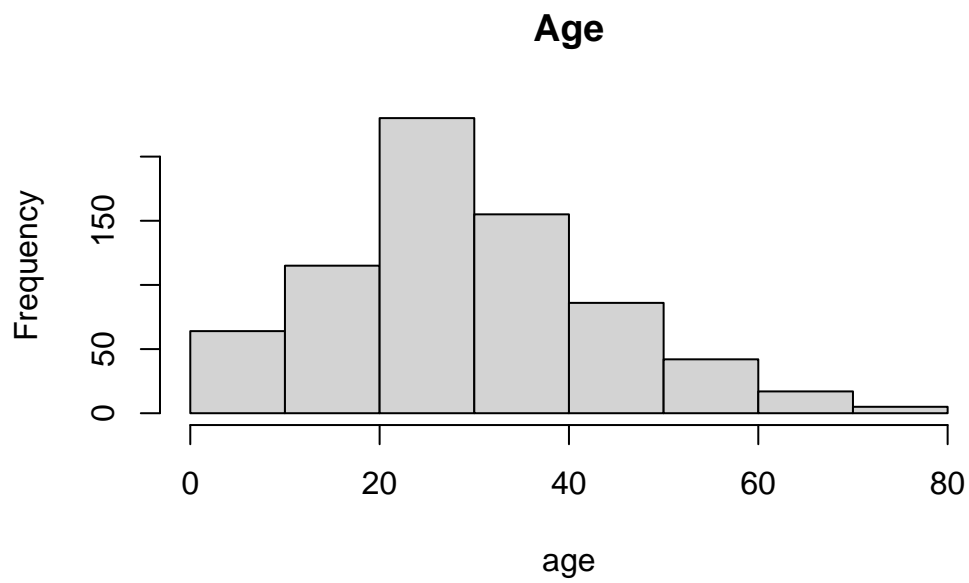
```
barplot(table(data$Pclass),ylab ="Frequency",main="Pclass")
```



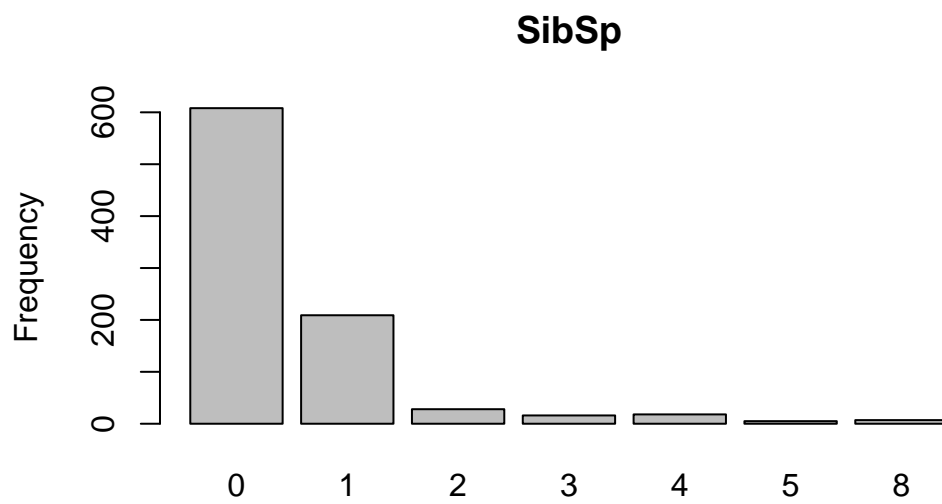
```
barplot(table(data$Sex),ylab = "Frequency",main="Sex")
```



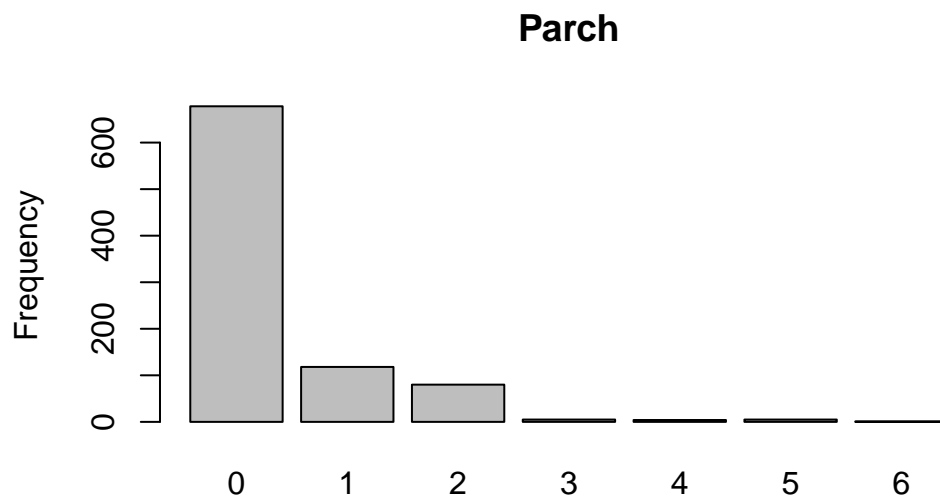
```
hist(data$Age,xlab = "age",main="Age")
```



```
barplot(table(data$SibSp),ylab = "Frequency",main="SibSp")
```



```
barplot(table(data$Parch),ylab ="Frequency",main="Parch")
```



```
barplot(table(data$Embarked),ylab ="Frequency",main="Embarked")
```

