

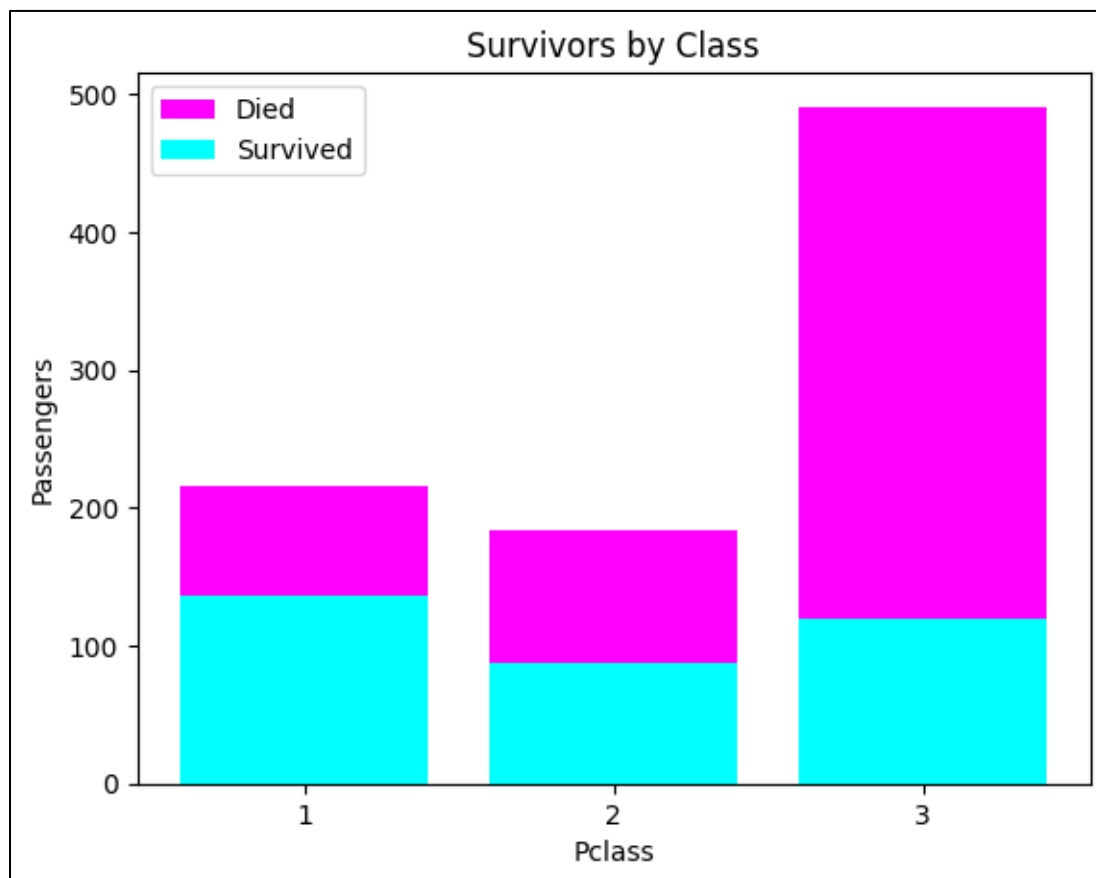
Assignment 3 Part 2

(3) Exploratory Data Analysis

(a) Survival rate associated to the class of passenger.

The table and graph below show the survival percentage of each class.

	Survived	Total	Percentage
Pclass			
1	136	216	62.96
2	87	184	47.28
3	119	491	24.24



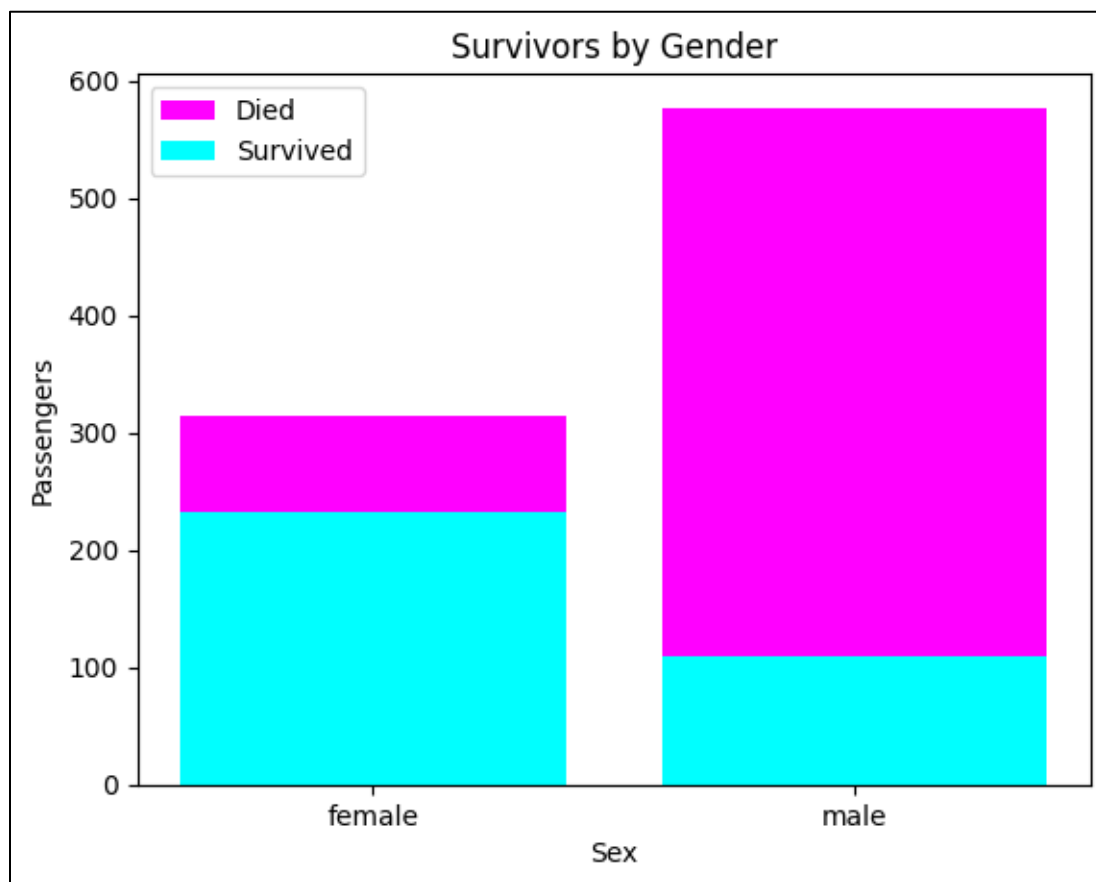
Explanation: As seen in the graph and from the data frame table – the 1st Class passengers had the highest rate of survival, followed by the 3rd class passengers, and the least survival rates was of 2nd class passengers. The passengers travelling in 3rd class were 491 which has the highest number yet only 24.24% survived. The graph also shows that the 3rd class passengers have the highest passengers and the

highest rate of death. Proves to say that the 3rd class were not given much attention in the retrieval and rescue operation.

(b) Determine if the survival rate is associated with the gender.

The table and graph below show the survival percentage of each gender.

	Survived	Total	Percentage
Sex			
female	233	314	74.203822
male	109	577	18.890815



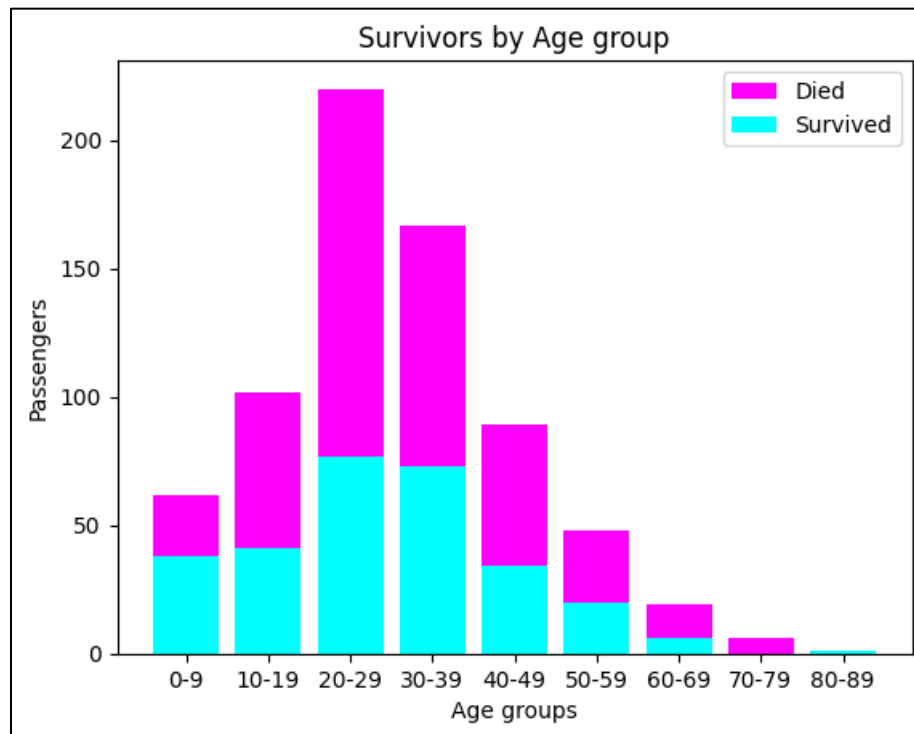
Explanation: We can see from the percentage table and graph above that female had the highest rate of survival. Female survival rate was 74.3% while male survival rate was 18.9% therefore, the female survival rate was about 4 times higher than male.

In conclusion, females were prioritized in rescue operations than male.

(c) Determine the survival rate is associated with age

The table and graph below show the survival percentage of each age group

	Survived	Total	SurvivedPercent	DiedPercent
AgeGroup				
0-9	38	62	61.29	38.71
10-19	41	102	40.20	59.80
20-29	77	220	35.00	65.00
30-39	73	167	43.71	56.29
40-49	34	89	38.20	61.80
50-59	20	48	41.67	58.33
60-69	6	19	31.58	68.42
70-79	0	6	0.00	100.00
80-89	1	1	100.00	0.00



Explanation: From the above graph and percentages, shockingly we can see that most who died were from 20-29 age group, when this is the age group that has the capability to save themselves.

Also, survival percentage of 0-9 age group is best - at 61.29% which means to say that children are more prioritized during rescue operations. But out of 891 passengers, there are 177 missing Age values, so this result might be less reliable.

From the analysis above, we would like to believe that passengers had higher chances of survival if they had a high-class ticket, they were women, and they were young. Also, there is missing data in the dataset especially the Age columns, thus we cannot rely on the survival rate result in terms of age.