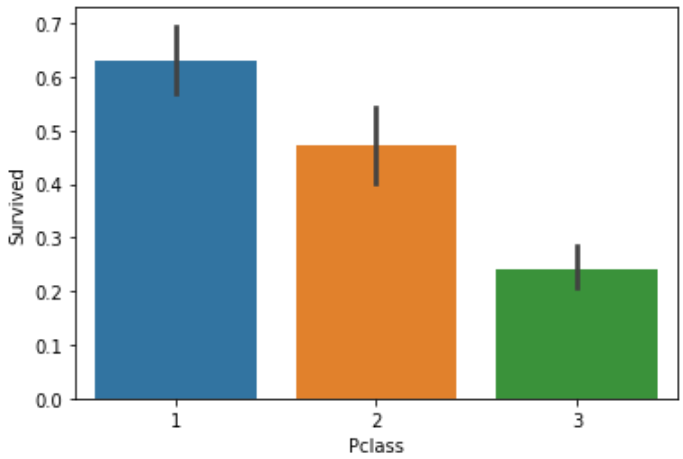
ASSIGNMENT 4: TITANIC

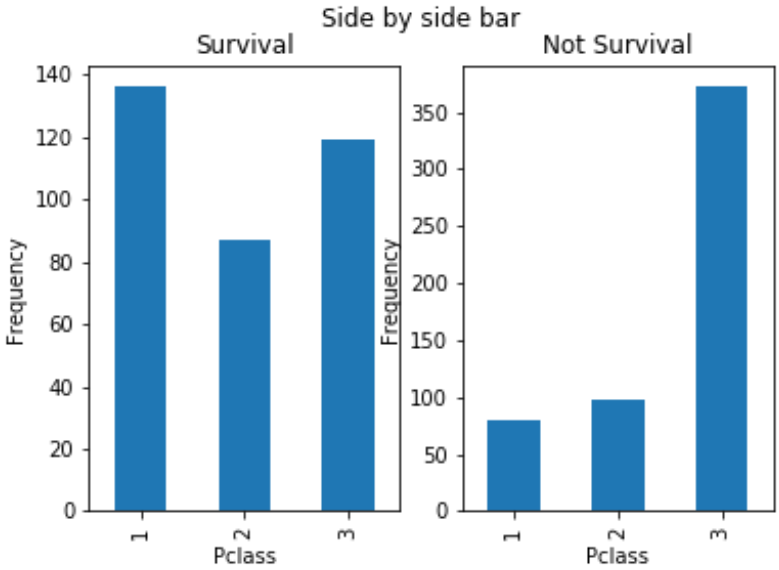
(Tinh Lo-100345588)

1/ Relation between Pclass and Survival.



Pclass	Survived
1	0.629630
2	0.472826
3	0.242363

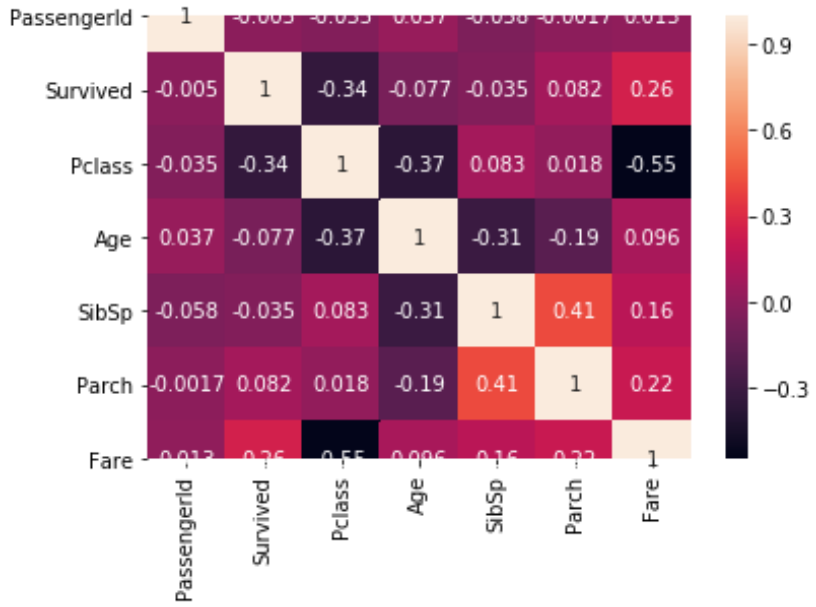
As can be seen from above barplot, we can see that first class passengers have highest rate of survival than other classes.



From side by side bar graph (or block diagram), we do not see it look alike that means Pclass and Survival are not independent.

Although it is generally not too conclusive, we could tell a rough conclusion that survival depends on passenger class among passengers on Titanic.

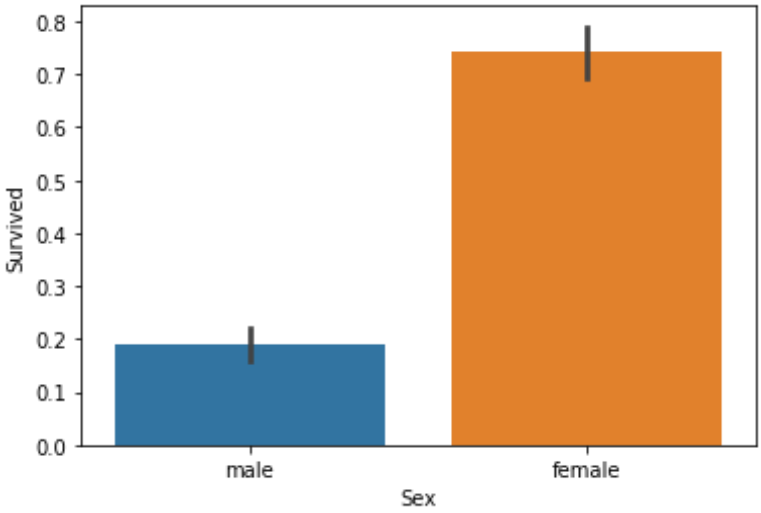
We need do Chi square test to identify relation between Pclass and Survival.



From heatmap above, we can see the correlation between Pclass and Survived is -0.34. This relation shows there is a quite strong relation between Pclass and Survived.

This number is negative, it means first class passengers could have higher rate of survival.

2/ Relation between Sex and Survival

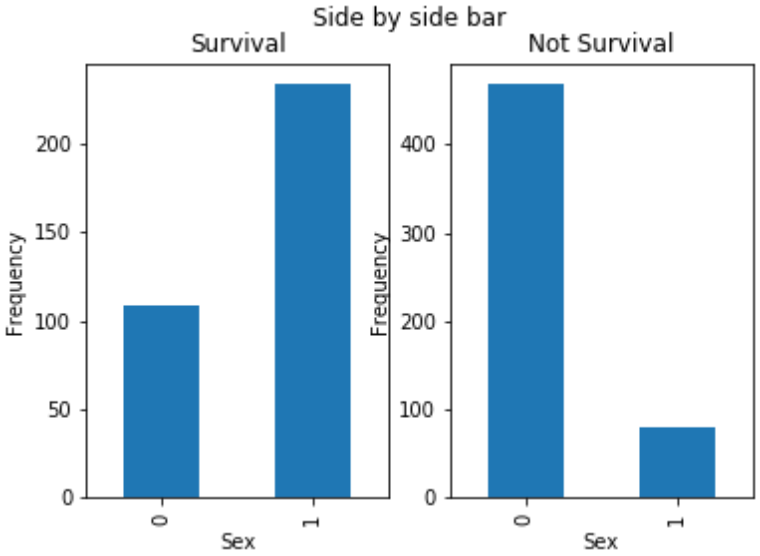


Sex	Survived
0	0.188908
1	0.742038

1: Female

2: Male

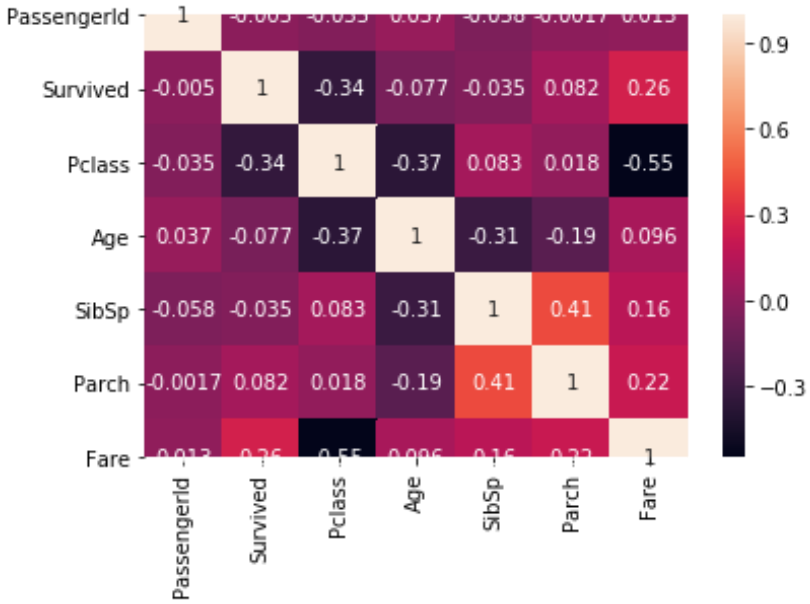
As we can see from the bar graph above, female passengers have far higher than man, it is easy to understand that female passengers were prioritized to be sent to lifeboat.



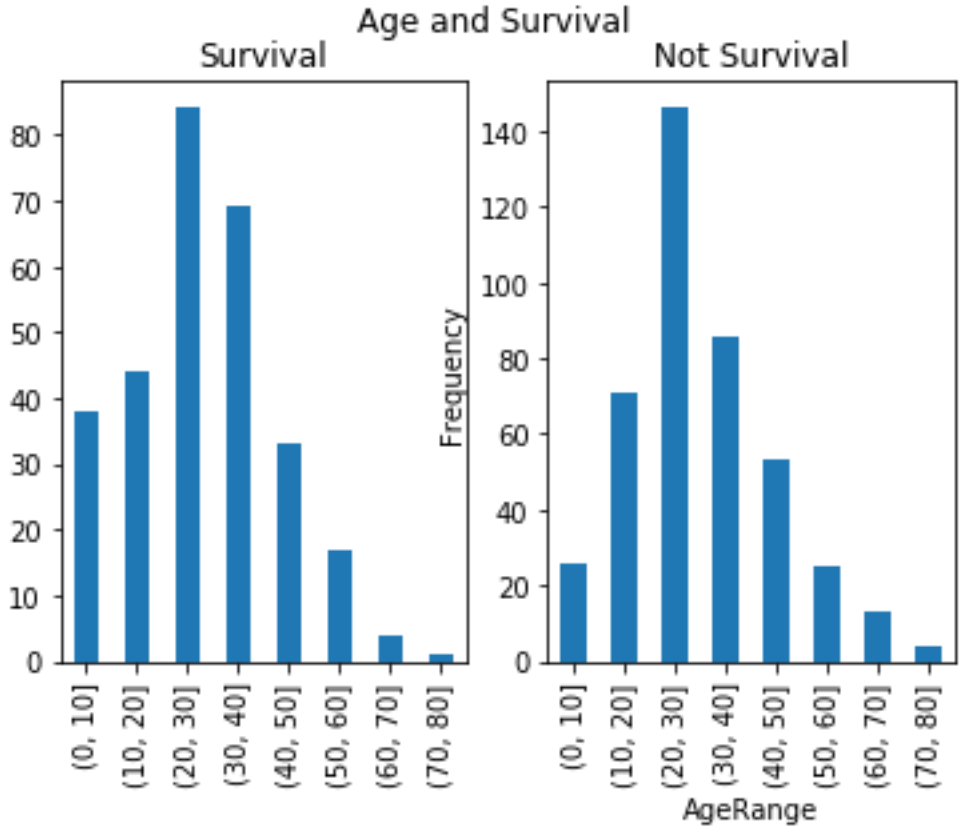
From side by side bar graph (or block diagram), we do not see it look alike that means Sex and Survival are not independent.

Although it is generally not too conclusive, we could tell a rough conclusion that survival depends on passenger Sex among passengers on Titanic.

3/ Age and Survival



From heatmap above, correlation between Age and Survival is -0.077. It means older passengers have less chance of survival, but this number is small (close to 0) that means it has a weak relation between Age and Survival.



From side by side bar above, we can see it looks alike. As a result, we could tell Survival is independent to Age among passengers on Titanic, this means survival rate is not different among age groups. And weak relation (r= -0.077) is an evidence to support the idea.

If we want to investigate more about this relation, we should use Chi square test.