**Project 1 conclusions:**

I checked the correlation between the survival and the following features:

1. Pclass – According to the Bar chart, the people from the first class had the highest chance to survive, and the people from the third class had the lowest chance to survive. The chance to survive of the people from the second class was in the middle. Logistic regression mean accuracy: 0.665
2. Fare – The pattern here is similar to the Pclass pattern: People who paid a higher fare were more likely to survive. Logistic regression mean accuracy: 0.673
3. Sex – The Bar chart shows that females had a higher chance to survive than males. The highest value of logistic regression mean accuracy: 0.794.
4. Relatives number– According to the bar chart, the people that had 1-3 relatives on the deck had the higher chance to survive and people that had no relatives or more than 3 relatives on the deck had lower chance to survive. Logistic regression mean accuracy: 0.617.

The Parch and SibSp features show almost similar pattern.

1. Embarked – The people that embarked from port C had a higher chance to survive than people that embarked from ports Q & S. Logistic regression mean accuracy: 0.607.
2. Age – The very young (age 0 – 10) had a higher chance to survive than the others. Logistic regression mean accuracy: 0.620

A separate check revealed that this pattern is stronger for males - Logistic regression mean accuracy: 0.811. No other interesting patterns were detected.

To summarize, The 'Sex' feature is the top driver for survival. The 'Pclass' and the 'Fare' are also driver for survival but in less accuracy. In males, the 'Age' feature has also a good correlation with survival.