***Wrangle data***

***In this data contain 3 tables (training data, testing data and gender submission). We must deal with it and clean it until it is entered into the machine learning algorithms in order to predict whether this person survived drowning or not. Therefore, in this report, how to deal with this data will be shown in order for us to clean it and extract correct data from it in order to extract data from it that will benefit us.***

***This data contains issues***

***1)***

***This is a good starting goal to execute.***

***By dropping features, we are dealing with fewer data points***

***Based on our assumptions and decisions we want to drop the Cabin and Ticket features***

***2)***

***We want to extract (Title) from column (name) and drop this column***

***We can replace many titles with a more common name or classify them as Rare***

***We can convert the categorical titles to ordinal***

***Now we can safely drop the Name feature from training and testing datasets. We also do not need the Passenger feature in the training and testing dataset***

***3)***

***Let us start by converting Sex feature to a new feature called Gender where female=1 and male=0.***

***4)***

***Let us start by preparing an empty array to contain guessed Age values based on Class x Gender combinations.***

***Let us create Age bands and determine correlations with Survived***

***Let us replace Age with ordinals based on these bands***

***Remove the Age Band feature.***

***5)***

***We can sum the two columns parch and sibs and this person we are predicting***

***We can create a new feature for Family Size which combines Parch and Sibs. This will enable us to drop Parch and SibSp from our datasets.***

***Let us drop Parch and Sibs features in favor of Family Size***

***6)***

***You must fill missing value in the Embarked feature takes S, Q, C values based on port of embarkation. Our training dataset has two missing values. We simply fill these with the most common occurrence***

***7)***

***We can now convert the Embarked feature to numeric***

***8)***

***You must fill missing value of the fare feature***

***We can now complete the missing value of the Fare feature***

***9)***

***Convert the Fare feature to ordinal values based on the Fare Band***