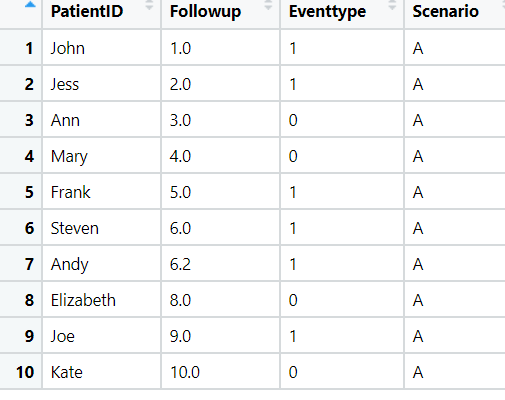
# Topic: Survival Analytics

Perform Kaplan Meier analysis for the given data and get the life table

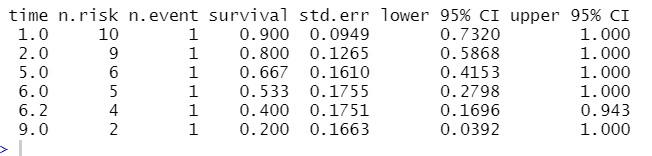


**Ans:**

* Loading the dataset.
* The dataset consists of 10 observations and 4 variables.
* Variable names are defined for Follow up and eventtype, where event takes the 2 values 0’s and 1’s which mean weather the patient have the disease or not

**Kaplan – Meier Non Parametric Analysis:**

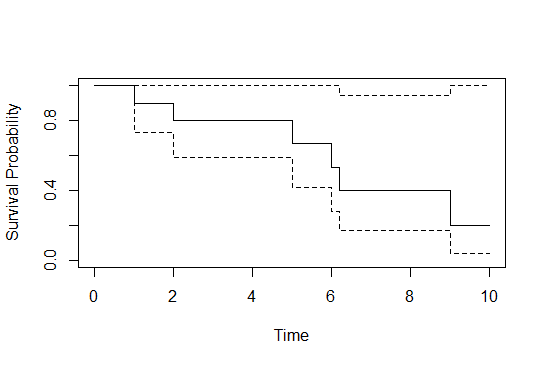
* This type of analysis, the time to a specific event, such as death or disease recurrence, is of interest and two (or more) groups of patients are compared with respect to this time.
* A non-parametric statistic is not based on the assumption of an underlying probability distribution, which makes sense since survival data has a skewed distribution.
* This statistic gives the probability that an individual patient will survive past a particular time t. At t = 0, the Kaplan-Meier estimator is 1 and with t going to infinity, the estimator goes to 0.

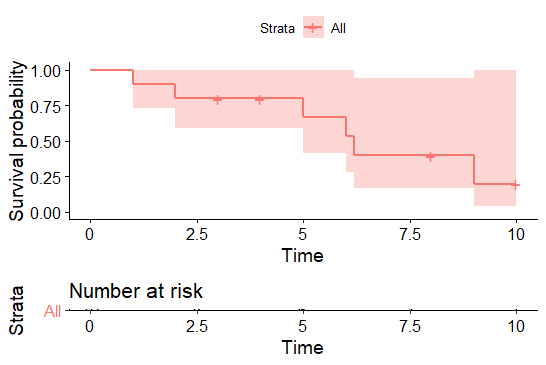


* From the above summary as the follow up of the patience increased the the risk of the patient has been decreased

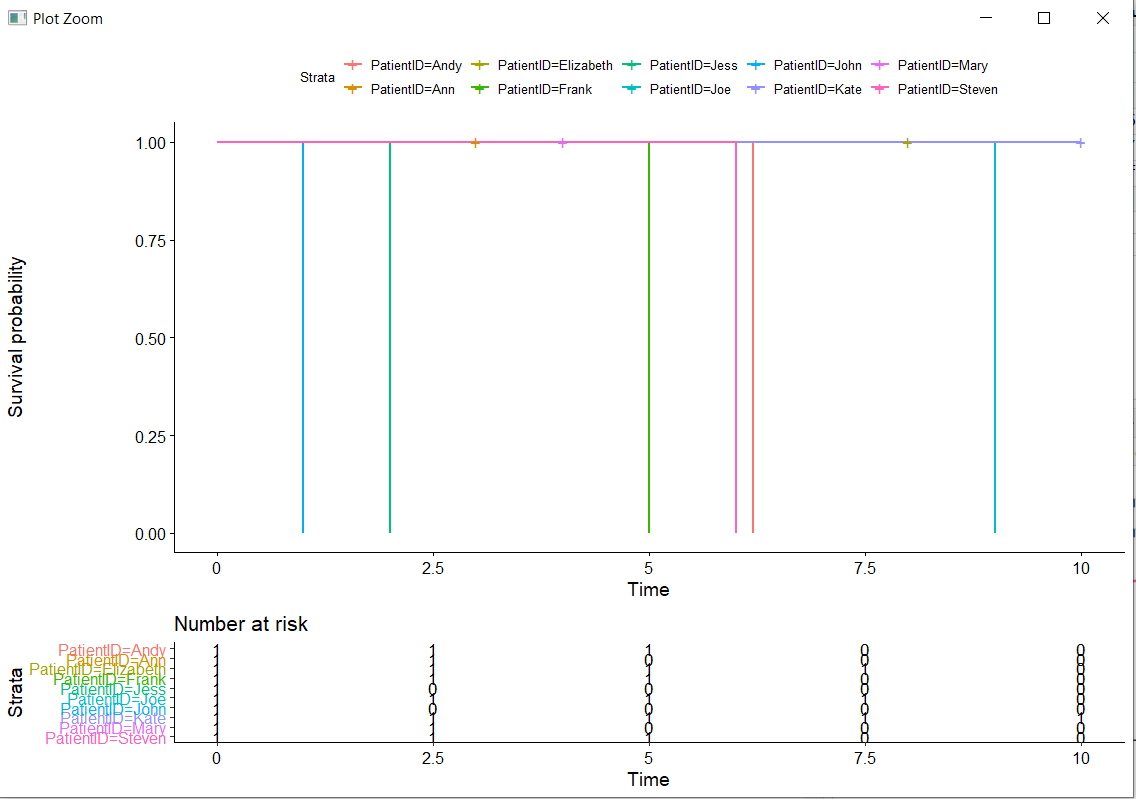
**Graphical Representation:**

* Below graph shows the relationship between time and event for the patients with the probability values and the risk range forming with group 1



****

* Below graphs represent the relation between event, time and patients



# Hints:

1. Business Problem
   1. Objective
   2. Constraints (if any)
2. Data Pre-processing

2.1 Data cleaning, Feature Engineering, EDA etc.

1. Model Building
   1. Partition the dataset
   2. Model(s) - Reasons to choose any algorithm
   3. Model(s) Improvement steps
   4. Model Evaluation
   5. Python and R codes
2. Result Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.

**Note:**

1. For each assignment the solution should be submitted in the format
2. Research and Perform all possible steps for improving the model(s) accuracy & reduce the RMSE (also evaluate errors like MAPE, MAE etc.)
3. All the codes (executable programs) are running without errors
4. Documentation of the module should be submitted along with R & Python codes, elaborating on every step mentioned here