# Topic: Survival Analytics

**Instructions:**

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: RAKESH SETHU NP**

**Batch Id:11052022\_7.30PM**

**Topic: Survival Analytics**

**Hints:**

1. **Business Problem**
   1. **What is the business objective?**
   2. **Are there any constraints?**
2. **Work on each feature of the dataset to create a data dictionary as displayed in the below image:**



**2.1 Make a table as shown above and provide information about the features such as its Data type and its relevance to the model building, if not relevant provide reasons and provide description of the feature.**

1. **Exploratory Data Analysis (EDA):**
   1. **Summary.**
   2. **Univariate analysis.**
   3. **Bivariate analysis.**
2. **Model Building**

**4.1 Build the model on the scaled data (try multiple options).**

**4.2 Perform survival analytics on the given datasets.**

**4.3 Briefly explain the model output in the documentation.**

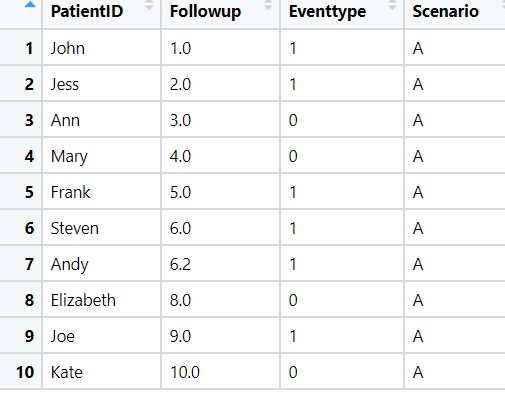
1. **Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?**

**Problem Statement:**

The following dataset contains patient ID, follow up, event type, and scenarios. Build a survival analysis model on the given data.

**Objective:** Maximize the accuracy in estimating time to event for a group of individuals. **Constraints (if any):** Kaplan meier estimate cannot be used for multivariate analysis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of feature** | **Description** | **Type** | **Relevance** |
| PatientID | Name of patient | Categorical | Relevant, Provides useful information. |
| Followup | No of follow up | Continuous, Ratio | Relevant, Provides useful information. |
| Eventtype | type of event | Discrete, count | Relevant, Provides useful information. |
| Scenario | Scenerio fo patient | Categorical | Relevant, Provides useful information. |



**Problem Statement: -**

ECG of different age groups of people has been recorded. The survival time in hours after the operation is given and the event type is denoted by 1 (if dead) and 0 (if alive). Perform survival analysis on the dataset given below and provide your insights in the documentation.

**Objective:** Maximize the accuracy in estimating time to event for a group of individuals. **Constraints (if any):** Kaplan meier estimate cannot be used for multivariate analysis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of feature** | **Description** | **Type** | **Relevance** |
| survival\_time\_hr | survival time per hour | Continuous, Interval | Relevant, Provides useful information. |
| alive | No of patient alive | Categorical,binary | Relevant, Provides useful information. |
| age | age of patient | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| pericardialeffusion | pericardialeffusion of patient | Categorical,binary | Irrelevant, Doesn’t Provides useful information. |
| fractionalshortening | fractionalshortening | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| epss | epss value | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| lvdd | lvdd value | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| wallmotion-score | wallmotion-score of patient | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| wallmotion-index | wallmotion-index of patient | Continuous, Ratio | Irrelevant, Doesn’t Provides useful information. |
| multi\_sensor | multi\_sensor value | Continuous, Interval | Irrelevant, Doesn’t Provides useful information. |
| name | Name of patient | Categorical | Irrelevant, Doesn’t Provides useful information. |
| group | group that which person belong to | Continuous, Ratio | Relevant, Provides useful information. |

A large room

Description automatically generated