# 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: Shahina Athar**

**Batch Id: 10122020**

**Topic: Survival Analytics**

**Grading Guidelines:**

**1. An assignment submission is considered complete only when correct and executable code(s) are submitted along with the documentation explaining the method and results. Failing to submit either of those will be considered an invalid submission and will not be considered for evaluation.**

**2. Assignments submitted after the deadline will affect your grades.**

**Grading:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ans** | **Date** |  |  | **Ans** | **Date** |
| Correct | On time | A | 100 |  |  |
| 80% & above | On time | B | 85 | Correct | Late |
| 50% & above | On time | C | 75 | 80% & above | Late |
| 50% & below | On time | D | 65 | 50% & above | Late |
|  |  | E | 55 | 50% & below |  |
| Copied/No Submission |  | F | 45 |  |  |

* **Grade A: (>= 90):** When all assignments are submitted on or before the given deadline.
* **Grade B: (>= 80 and < 90):** 
  + When assignments are submitted on time but less than 80% of problems are completed.

(OR)

* + All assignments are submitted after the deadline.
* **Grade C: (>= 70 and < 80):** 
  + When assignments are submitted on time but less than 50% of the problems are completed.

(OR)

* + Less than 80% of problems in the assignments are submitted after the deadline.
* **Grade D: (>= 60 and < 70):**
  + Assignments submitted after the deadline and with 50% or less problems.
* **Grade E: (>= 50 and < 60):** 
  + Less than 30% of problems in the assignments are submitted after the deadline.

(OR)

* + Less than 30% of problems in the assignments are submitted before the deadline.
* **Grade F: (< 50):** No submission (or) malpractice.

**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).**

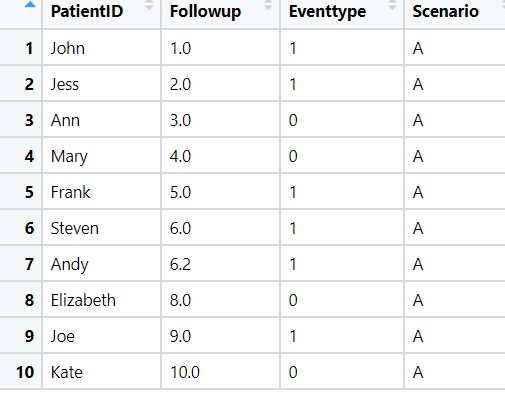
**4.2 Perform survival analytics on the given datasets.**

**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.



1. **Business Problem**
   1. **What is the business objective?**

**Ans:** The following dataset contains patient ID, follow up, event type, and scenarios.

Build a survival analysis model on the given data.

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

**Ans:** First of all we have to laod the dataset then, apply describe function, Followup is referring to time, Importing the KaplanMeierFitter model to fit the survival analysis, Initiating the KaplanMeierFitter model, Fitting KaplanMeierFitter model on Time and Eventtype, then finally we get Time-line estimations plot, Applying KaplanMeierFitter model on Time and Events for the group "A".

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

**Ans:** The following dataset contains patient ID, follow up, event type, and scenarios.

We can easily Build a survival analysis model on the given data.

**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.

A large room

Description automatically generated

1. **Business Problem**
   1. **What is the business objective?**

**Ans:** 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).

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

**Ans:** First of all we have to laod the dataset then, apply describe function, drop all rows that have any NaN values, survival\_time\_hr is referring to time, Importing the KaplanMeierFitter model to fit the survival analysis, Initiating the KaplanMeierFitter model, Fitting KaplanMeierFitter model on Time and Events, Time-line estimations plot, Applying KaplanMeierFitter model on Time and Events for the group "1",Applying KaplanMeierFitter model on Time and Events for the group "2",Applying KaplanMeierFitter model on Time and Events for the group "3".

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

**Ans:** 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). We can

easily build survival analysis on the dataset given below.