

TECH ALPHARETTA

2972 Webb Bridge Rd, Alpharetta, GA 30009

**Survival of a Patient**

A hospital in the province of Greenland has been trying to improve its care conditions by looking at historic survival of the patients. They tried looking at their data but could not identify the main factors leading to high survivals.

**Data**

##### **About the dataset**

The dataset contains the patient records collected from a hospital in Greenland. The "Survived\_1\_year" column is a target variable which has binary entries (0 or 1).

* **Survived\_1\_year == 0**, implies that the patient did not survive after 1 year of treatment
* **Survived\_1\_year == 1**, implies that the patient survived after 1 year of treatment
* **D\_Patient\_Care\_Situation**: Care situation of a patient during treatment
* **Diagnosed\_Condition**: The diagnosed condition of the patient
* **ID\_Patient**: Patient identifier number
* **Treatment\_with\_drugs**: Class of drugs used during treatment
* **Survived\_1\_year:** If the patient survived after one year (0 means did not survive; 1 means survived)
* **Patient\_Age**: Age of the patient
* **Patient\_Body\_Mass\_Index**: A calculated value based on the patient’s weight, height, etc.
* **Patient\_Smoker**: If the patient was a smoker or not
* **Patient\_Rural\_Urban**: If the patient stayed in Rural or Urban part of the country
* **Previous\_Condition**: Condition of the patient before the start of the treatment ( This variable is splitted into 8 columns - A, B, C, D, E, F, Z and Number\_of\_prev\_cond. A, B, C, D, E, F and Z are the previous conditions of the patient. Suppose for one patient, if the entry in column A is 1, it means that the previous condition of the patient was A. If the patient didn't have that condition, it is 0 and same for other conditions. If a patient has previous condition as A and C , columns A and C will have entries as 1 and 1 respectively while the other column B, D, E, F, Z will have entries 0, 0, 0, 0, 0 respectively. The column Number\_of\_prev\_cond will have entry as 2 i.e. 1 + 0 + 1 + 0 + 0 + 0 + 0 + 0 = 2 in this case. )

**Objective**

You are the best data scientist in Greenland and they've hired you to solve this problem. Now you are responsible for developing a model that will predict the chances of survival of a patient after 1 year of treatment (Survived\_1\_year).