# Cox Proportional-Hazard Model

The article is split into four components: Data, Model, Calibration. This article derives heavily from [*Applied Survival Analysis: Regression Modeling of Time-to-Event Data, Second Edition* by David W. Hosmer, Stanley Lemeshow and Susanne May](https://www.wiley.com/en-us/Applied+Survival+Analysis%3A+Regression+Modeling+of+Time+to+Event+Data%2C+2nd+Edition-p-9780471754992)*.*

The examples provided in this article corresponds to survival (i.e. non-default) of any given entity/individual.

## Data

The data used in survival analysis is referred to as survival data. Each observation of the survival data has a start time and an end time. The start time corresponds to the time at which the observation begins. The end time corresponds to the occurrence of an event of interest (eg. Death of an individual or Default by a company) or to the end of the observation window. The later scenario wherein the end time occurs due to the unavailability of the survival/death information is called censoring of the data, to be precise it is called right censoring. Censored data is unique to survival models and is unlike the data observed in standard regression analysis.

The start/end time of observation can be different for different individuals; the survival model is not dependent on the exact time but depends on the time difference between the start and the end point. This time difference is referred to as survival time.

## Model

### Kaplan-Meier estimator

Kaplan and Meier (1958) proposed the first estimator for estimating survival time.