The file “Assignment\_Linear\_Regression\_Data\_Train.csv” contains the data for 1000 randomly selected patients and is intended to be used as a training set. A test set of 300 randomly selected patients is saved in “Assignment\_Linear\_Regression\_Data\_Test.csv”. Each row in both files corresponds to the record for a patient. The data includes demographic information of the patients and the worst reading of some clinical variables during the patient’s first 24-hour stay in the intensive care unit (ICU). It is a common practice to use the worst reading in the first 24 hours of patient ICU stay to predict the total future length stay of the patient. The intuition is that the worst reading captures the initial condition of the patient, which may suggest the time needed to recover. The list of all the variables and some brief descriptions are provided below. You are encouraged to search online to find out more information related to each clinical variable.

**Age:** Patient’s age at the ICU admission

**Gender:** Patient’s gender

**Race**: Seven categories of race; disguised as “R1” to “R7”

**LOS**: Patient’s length of stay in the ICU (unit: hour)

**Temperature**: Body temperature (unit: Celsius); a vital sign

**RespirationRate**: Frequency of breath (unit: times/min); a vital sign

**HeartRate**: Heart rate (unit: beats/min); a vital sign

**SystolicBP**: Systolic blood pressure; maximum pressure in blood vessels during a heartbeat (unit: mmHg); a vital sign

**DiastolicBP**: Diastolic blood pressure; minimum pressure in blood vessels between two heartbeats (unit: mmHg); a vital sign

**SpO2**: Peripheral capillary oxygen saturation; the amount of oxygen carried by the red blood cells in the body’s arteries; a vital sign

**Haemoglobin**: Iron-containing oxygen-transport metalloprotein in the red blood cells (unit: g/dl)

**NumWBC**: Amount of white blood cell in the blood (unit: number )

**Creatinine**: Byproduct of muscle metabolism that is excreted unchanged by the kidneys, an important indicator of renal health (unit: )

**Urea**: Amount of urea nitrogen in blood (unit: mmol/L)

**ArterialpO2**: Partial pressure of oxygen in arterial blood (unit: mmHg)

**GCS**: Glasgow Coma Scale/Score; a neurological scale that records the conscious state of a person for initial as well as subsequent assessment (unit: point score from 3 (worst) to 15 (best))