**BASE DE DADOS STROKE (BASE 1)**

1. CONTEXTO

A base de dados de predição do AVC foi encontrada no site Kaggle.

De acordo com a World Health Organization (WHO) AVC é a 2.a causa de morte global, responsável por 11% do total de mortes. A base de dados disponibilizada é usada para predizer se o paciente tem predisposição para o AVC baseado em parâmetros de entrada como gênero, idade, diversas doenças e status de fumante. Cada linha dos dados prove informações relevantes por paciente.

1. DICIONÁRIO DE DADOS

* id: Identificador único do paciente
* gender: Gênero do paciente (Male/Female/Other)
* age: idade do paciente
* hypertension: Se o paciente é hipertenso ou não (0: não hipertenso, 1: hipertenso)
* heart\_disease: Se paciente possui doença do coração ou não (0: não possui, 1: possui)
* ever\_married: Se o paciente foi casado uma ou mais vezes ou se está casado (0: foi ou é casado, 1: não foi casado)
* work\_type: Tipo do trabalho exercido pelo paciente, por esse campo é possível identificar se é uma criança (children/Govt\_jov/Never\_worked/Private)
* Residence\_type: Tipo de residência, se rural ou urbana (Rural/Urban)
* avg\_glucose\_level: taxa média de glicose no sangue
* bmi: índice de massa muscular
* smoking\_status: Se o paciente fuma, já foi fumante ou não fuma (formely smoked/never\_smoked/smokes/Unknown)
* stroke: 1 se o paciente teve um AVC e 0 se nunca teve um AVC OBS: "Unknown" na variável smoking\_status significa que não se tem essa informação do paciente