**Appendix B Variables in FRAMINGHAM dataset**

|  |  |
| --- | --- |
| **Variable name** | **Variable description** |
| Status | Person’s status (*Alive, Dead*) |
| Sex | Sex (*Female, Male*) |
| Height | Person’s height in inches |
| Weight | Person’s weight in kg |
| Diastolic | Person’s blood pressure in the arteries when the heart rests between beats. A normal diastolic blood pressure measurement is 80 or below. |
| Systolic | Person’s blood pressure in the arteries: when the heart beats, it contracts and pushes blood through the arteries to the rest of the body. A normal systolic blood pressure measurement is 120 or below. |
| Smoking | Number of cigarettes smoked per week |
| Death\_age | Age at death in years |
| Cholesterol | Person’s cholesterol measurement |
| Chol\_status | Cholesterol status (*borderline, desirable, high*) |
| Bpressure\_status | Blood pressure (*High, Normal, Optimal*) |
| Weight\_status | Weight status (*Underweight,* *Normal, Overweight)* |
| Smoke\_status | Smoking status (*Nonsmoker, Light, Moderate, Heavy, Very heavy)* |

**Note** Age of death is an interesting variable for exploratory data analyses (hence it’s inclusion) however it is recommended that this variable is rejected before applying directed data mining techniques.