

**BACKGROUND:**

The Data for a Study of Risk Factors Associated with Low Infant Birth Weight. Data were collected at Baystate Medical Center, Springfield, Massachusetts.

**DATA CATALOGUE:**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Description** | **Values/Coding** |
| LOW | Low Birth Weight | 0 = Not low, 1 = low |
| AGE | Age of the Mother | In years (continuous) |
| LWT | Weight at Last Menstrual Period | In pounds (continuous) |
| RACE | Race | 1 = White, 2 = Black, 3 = Other |
| SMOKE | Smoking Status During Pregnancy | 1 = Yes, 0 = No |
| PTL | History of Premature Labor | 0 = None, 1 = One, 2 = Two, etc. (count) |
| HT | History of Hypertension | 1 = Yes, 0 = No |
| UI | Presence of Uterine Irritability | 1 = Yes, 0 = No |
| FTV | Number of Physician Visits During First Trimester | 0 = None, 1 = One, 2 = Two, etc. (count) |

**ASSIGNMENT OBJECTIVE:**

Consider LOW as dependent variable and remaining variables listed above as independent variables.

**QUESTIONS:**

1. Import BIRTH WEIGHT data.

2. Cross tabulate dependent variable with each independent variable.

3. Develop a model to predict if birth weight is low or not using the given variables.

4. Generate three classification tables with cut-off values 0.4, 0.3 and 0.55.

5. Calculate sensitivity, specificity and misclassification rate for all three tables above. What is the recommended cut-off value?

6. Obtain ROC curve and report area under curve.