1. Linear Regression

Use NHANS2 data to develop a regression model that studies how various socio-economic and health factors affect individual 's systolic blood pressure (bpsystol). These socio-economic and health factors include gender (female) l , race (race), age (age), whether they live in a rural area or not (rural), body mass index (bmi), and whether or not they have diabetes (diabetes).

1. Write down the regression model in the form ofy — + ß1X1 + ß2X2 +. . .+ ßkXk + by using actual variables names instead ofy or Xi. Please see lecture slides for examples.
2. Given the model in (a), which estimation approach is most appropriate for this model? Explain.
3. Estimate the model in (a) using the regression approach chosen in step (b) above. Display regression table and discuss the statistical significance and meaning of each (statistically significant) coefficient from this regression. Please see lecture slides for examples.

2. Qualitative (binary) Data Analysis

Use NHANS2 data to develop a regression model that predicts whether or not someone has a diabetes (diabetes) based on various socio-economic and health factors. These socio-economic and health factors include gender (female)2, race (race), age (age), whether they live in a rural area or not (rural), body mass index (bmi), and blood pressure (bpsystol).

Repeat steps (a) — (c) from question l.

To receivefull credit, make sure that your work is professionally and aesthetically formatted andpresented. You may consult journal article publications or research reports to see how they are done.