

R Homework for Chapter 17

Gross Domestic Product

The gross domestic product (GDP) per capita is a widely used measure of a country's (or state's) economy. It is defined as the total market value of all goods and services produced within a country (or state) in a specified period of time. The most common computation of GDP includes five items: consumption, gross investment, government spending, exports, and imports (which negatively impact the total). The Census Bureau reports the GDP for each state in the United States quarterly. The government also reports annual personal income totals (seasonally adjusted in \$millions) by state and each state's population. Let's examine how personal income is related to GDP at the state level.

You can find the data file on Blackboard. Download it and put it in the **same folder** as your R program file. Then, use the following command to read in the data

```
GDP <- read.table('GDP.txt', sep = '\\t', header = TRUE)
```

and answer the questions below.

1. Make a histogram and Q-Q plot of *Personal Income* to check the Normal Population Assumption. If the assumption is violated, propose an appropriate transformation from the Ladder of Powers for *Personal Income*, and justify it.
2. Make a scatterplot of the transformed *Personal Income* against *GDP*, and check the linearity assumption. If the assumption is not satisfied, propose an appropriate transformation from the Ladder of Powers on *GDP*, and justify it.
3. Fit a linear regression model using the transformed *GDP* as *x*-variable and the transformed *Personal Income* as *y*-variable, based on the transformations determined in the previous questions. Produce a residual plot against fitted values. Is the equal-variance assumption satisfied in the fitted model? Explain.
4. Is there any unusual observation according to the linear model you fitted above? If so, find the states they come from. Are they outliers? Do they have high leverages? Are they influential points? Explain.
5. Suppose there exists a state with $GDP = 300,000$. Can you predict the *Personal Income* for that state using the model from Question (3)? Please also show its 95% prediction interval.