

Consider the following linear model:

$$Y = 3.2 + 1.87X_1 + 2.1X_2$$

1. (3 points) What is the most important variable in terms of their estimated coefficients? Be specific.
2. (3 points) Let's assume the following:  $S_Y = 0.45$ ,  $S_{X_1} = 1.26$ , and  $S_{X_2} = 1.78$ . What is the most important variable in terms of the standardized regression coefficients? Be specific.
3. Consider the `NBA_2006_2007.csv` data file. This data file contains information on the 30 NBA teams from the 2006-2007 season. This file contains the following information:

- GP: games played
- W: wins
- L: losses
- Win\_pct: Win percentage
- min: minutes played
- EFG\_pct: Effective Field Goal Percentage
- FTA\_rate: Free Throw Attempt Rate
- TOV\_pct: Turnover Percentage
- OREB\_pct: Offensive Rebound Percentage
- OPP\_EFG\_pct: Opponent's Effective Field Goal Percentage
- OPP\_FTA\_pct: Opponent's Free Throw Attempted Rate
- OPP\_TOV\_pct: Opponent's Turnover Percentage
- OPP\_OREB\_pct: Opponent's Offensive Rebound Rate

**In Python**, answer the following:

- (a) (3 points) Using pandas, read the csv file and create a data-frame called `nba`.
- (b) (4 points) Make sure `EFG_pct`, `TOV_pct`, `OREB_pct`, and `FTA_rate` are on the same scale.
- (c) (15 points) Build a linear regression model with the four factors for team offense. That is, build the following model

$$W = \beta_0 + \beta_1 \times \text{EFG\_pct} + \beta_2 \times \text{TOV\_pct} + \beta_3 \times \text{OREB\_pct} + \beta_4 \times \text{FTA\_rate}$$

- What is the most and least important variables in terms of the estimated coefficients? Be specific.
- What is the most and least important variables in terms of the standardized regression coefficients? Be specific.
- What is the most and least important variables in terms of the contribution of the variable to  $R^2$ ? Be specific.