

Group Work - Week 11

1 The file “Galton-mother-daughter.csv” contains a subset 50 subjects from Galton’s mother/daughter height data.

- (a) Create a scatterplot of the data. Does there appear to be a linear relationship between mother's heights and daughter's heights?
- (b) Conduct a correlation hypothesis test at $\alpha = 0.05$ significance level. If there is significant correlation, how would you describe the strength of the correlation?
- (c) Find the estimated regression line for the relationship between mother's heights (predictor variable) and daughter's heights (response variable)? Is the slope significantly different than zero?
- (d) What is the best predicted daughter's height for a mother that is 56 inches tall? Is it appropriate to make such a prediction?

2 The file “MCA_scores_17.csv” on D2L contains average math MCA scores for 11th graders in 2017 by MN public school district, as well as percentage of 11th graders receiving free lunches in the district. Districts with missing data and charter schools are excluded.

(a) Create a scatterplot of the data. Does there appear to be a linear relationship between percentage of students receiving free lunch and average MCA scores?

(b) Conduct a correlation hypothesis test at $\alpha = 0.01$ significance level. If there is significant correlation, how would you describe the strength of the correlation?

(c) Find the estimated regression line for the relationship between percentage of students receiving free lunch (predictor variable) and average MCA scores (response variable)? Is the slope significantly different than zero?

(d) What is the best predicted average MCA score for a district that has 45% of 11th grade students receiving free lunch? Is it appropriate to make such a prediction?