**Examples for Chapter 5 Math/Stat286**

1. Meddicorp Company sells medical supplies to hospitals, clinics and doctor’s offices. The company currently markets in three regions of the United States: the South, the West and the Midwest. These regions are each divided into smaller sales territories. Data for Meddicorp is contained in file: **MEDDICORP4.csv.** In the data, SALES are in thousands of dollars, ADV and BONUS are in hundreds of dollars. The management is concerned with the effectiveness of a new bonus program. This program is overseen by regional sales managers and provides bonuses to salespeople based on performance. Management wants to know if the bonuses paid in 2003 were related to sales. In determining whether this relationship exists, they also want to take into account the effects of advertising.
   1. Find the least square regression equation that describes the relationship between sales and advertising and bonus.
   2. Even when the effect of bonus is taken into account, is advertising related to sales? State the null and alternative hypotheses, get the p-value and make conclusion in both statistical language and plain English. (α = 5%)
   3. Even when the effect of advertising is taken into account, is bonus related to sales? State the null and alternative hypotheses, get the p-value and make conclusion in both statistical language and plain English. (α = 5%)
   4. If bonus is fixed, will increase advertising promote sales? State the null and alternative hypotheses, get the p-value and make conclusion in both statistical language and plain English. (α = 5%).
   5. Find the 95% confidence interval for the slope for advertising.
   6. What proportion of variation in sales has been explained by its relationship to advertising and bonus?
   7. Conduct the F test for overall fit of the regression.
   8. Do you think market share and the largest competitor’s sales are useful if you already taken into account of advertising and bonus? State the null and alternative hypotheses, get the p-value and make conclusion in both statistical language and plain English. (α = 5%).
   9. Use the reduced model to predict the **average** sales for companies who spend $50,000 in advertising and $25,000 in bonus. Find the 95% confidence interval for your prediction.
   10. Use the reduced model to predict the sales for a company who spend $50,000 in advertising and $25,000 in bonus. Find the 95% prediction interval for your prediction.
2. A company that provides transportation services uses a telemarketing division to help sell its services. The division manager is interested in the time spent on the phone by the telemarketers in the division. Data on the number of months of employment and the number of calls paced per day (an average for 20 working days) is recorded for 20 employees. The data are in the file: **TELEMARK5.csv**
   1. Look at the scatter plot of the data. Put the least-square line on the scatter plot. What’s the R2 of the regression that describes the relationship between CALLS and MONTHS? Get the residual plot for the regression that describes the relationship between CALLS and MONTHS.
   2. Get the relationship that describes the CALLS and MONTH. (polynomial)
   3. Test whether the second order is significant or not at 5% level of significance.
   4. Analyze the residual plot for your new model.
3. In this example, we are trying to find the model that describes the relationship between Weight and Height. Data set **HeightWeight.csv**
   1. Using the dummy variable “Female”, find the relationship between Weight and Height for different gender;
   2. Now use the dummy variable “Male”, find the relationship between Weight and Height for different gender. Compare your results with (a).
   3. Do you think the data provide strong evidence that the relationship between Height and Weight is the same for female and male? Test at 5% level of significance.
   4. What’s your final model?