DISCUSSION CASE

**Interpreting Regression Models**

This discussion case will help solidify your learning of the regression material we covered in class.

**Permissible variables in regression models**

The career services office in a business school on the west coast wants to identify the factors that are associated with the starting salary for its graduates. It has access to the following variables. Which would be permissible to include as independent variables in a regression model?

1. student id number
2. major (accountancy, finance, management, marketing)
3. whether student also had a minor (yes/no)
4. overall GPA
5. GPA in major
6. number of internships completed
7. participated in job fair (yes/no)
8. number of career center workshops participated in

**Interpreting regression models**

Customer service expense

A company that develops and maintains websites for businesses wants to develop estimates of how much developer time is required to maintain customer websites.

The dependent variable for the regression model is developer time per page on a customer’s site (in hundreds of hours, 00)

The independent variables are

• number of years the client has been a customer of the company

• whether the website includes a payment portal (1 = yes, 0 = no)

• company size (annual revenue in millions of dollars--$000,000)

The model is below.

Hours = 6.38 – .47 (customer years) + 1.14 (if payment portal) – .16 (annual revenue)

R = .41, R2 = .17

Employee longevity

A large insurance company has developed a model to identify the factors associated with employee turnover.

The dependent variable is number of years an employee stays with the company before leaving.

The independent variables are:

• whether the employee had received a promotion within the 12 months prior to leaving (1=yes, 0=no)

• the average of the employee’s last two performance reviews (10-point rating scale)

• number of people in the employees working group who had left the working group within the last 18 months, regardless of reason (promotion, termination, relocation, quitting, etc.)

The model:

Years = 2.73 + .91 (if promotion received) + .24 (performance review score) – .55 (# of work group departures)

R = .51, R2 = .26

Online plant purchases

Heavenly Gardens is a plant nursery that sells outdoor garden plants online to people in the U.S. The company used regression analysis of their sales records to identify the factors associated with the amount of purchases customers make.

The dependent variables is sales in dollars.

The independent variables are:

• latitude of customer’s residence (higher value means more north)

• population density at customer’s location (# of households per square mile)

• previous customer (1 if customer has purchased from Heavenly Gardens before, 0 if not)

The model is below.

Purchases = 206.5 – 24.7 (degrees latitude) – 9.3 (density) + 43.9 (if prior purchaser)

R = .62, R2 = .38