University of California, Davis
Department of Statistics
Winter 2016

STA 135 **Final Project**

(Due on **Wednesday, March 16** with you in class final)
(This is a **group project**, and the maximum number of students in each group is **5**)

Diagnosis of Depression in Primary Care

In a study of diagnosis of depression in primary care, 400 patients were randomly selected and the following variables were recorded:

DAV: Diagnosis of depression in any visit during one year of care.

0 = Not diagnosed
1 = Diagnosed

PCS: Physical component of SF-36 measuring health status of the patient.

MCS: Mental component of SF-36 measuring health status of the patient

BECK: The Beck depression score.

GEND: Patient gender

0 = Female
1 = Male

AGE: Patient's age in years.

EDUCAT: Number of years of formal schooling.

The data are stored in the file **final.dat** and is available from the course web site. The goal of this project is to perform multivariate analyses of this data using SAS or any other statistical packages. This includes a comparison of the mean vectors of continuous measurements (PCS, MCS, BECK and AGE) between those diagnosed as depressed and those that have not been diagnosed, comparing the mean vectors of continuous measurements for the three education groups (those with ≤ 11 years of education, between 12 and 14 year including 12 and 14, and those with > 14 years of education), and constructing a multivariate regression of MCS and BECK on PCS, AGE and GEND. We also would like to do a principal component analysis on PCS, MCS and BECK. Finally, we would like to perform Fisher's discriminate analysis on PCS, MCS, BECK and AGE for classifying diagnosed and not diagnosed patients.

Please explain your findings in a 3 to 4- page report. Your report may include the following sections:

- *Introduction:* Statement of the problem.
- *Material and Methods:* Description of the data and methods that you used for the analysis.
- *Results:* Explain the results of your analysis in detail. You may cut and paste some of your computer outputs and refer to them in the explanation of your results.
- *Conclusion and Discussion:* Highlight the main findings and discuss.

Please cut and paste the computer outputs to your report and do not include any direct computer output as an attachment.