

McMaster University
Department of Mathematics and Statistics
STATISTICS 3A03: Applied Regression Analysis with SAS
Fall 2017
SAS Lab 4. October 3–6, 2017

This will be a revision lab to ensure you properly understand the features used in the previous three labs. The important things you need to know at this point are:

1. Importing datasets from text files.
2. Use of the **Data Step** to create new variables in a dataset which are functions of other variables.
3. Use of the **PROC GLOT** (or **PROC SGLOT**) to construct scatterplots of one variable against another.
4. Use of **PROC REG** to do simple linear regression of a single response variable against a single covariate.
5. Construction of scatter plot matrices using **PROC SGSCATTER**.
6. Use of **PROC REG** to do multiple regression.
7. Constructing plots and output datasets with **PROC REG**.
8. Estimation and prediction using **PROC REG**.

If you have any questions about any of these topics then please ask your TA as they circulate around the lab today.

In this lab I would like you to practice the above topics in SAS using the dataset **BGSgirls.txt** which gives the following variables on 70 girls from the Berekeley Guidance Study

WT2 Weight at age 2

HT2 Height at age 2

WT9 Weight at age 9

HT9 Height at age 9

LG9 Leg circumference at age 9

ST9 Strength measure at age 9

WT18 Weight at age 18

HT18 Height at age 18

LG18 Leg circumference at age 18

ST18 Strength measure at age 18

- (a) Construct scatterplots of weight against height at each of the three ages (2, 9, 18) for these girls.
- (b) Run simple regressions with weight as response and height as covariate separately at age 2, 9 and 18 and construct a scatterplot with the fitted line for each of the three regressions.
- (c) Create a dataset containing the new variable BMI18 defined as $WT18/(HT18/100)^2$.
- (d) Construct a scatterplot matrix of WT2, HT2, WT9, HT9 and BMI18.
- (e) Construct a multiple regression to predict BMI18 from the height and weight of the girls at ages 2 and 9.
- (f) Plot the fitted values of BMI18 from the above regression against the observed values.
- (g) Give 95% confidence intervals and prediction intervals for girls with the following values

WT2	HT2	WT9	HT9
12	82	25	125
14	84	30	135
15	86	35	135
15	90	40	140
16	93	45	140