LAB4. More Minitab Command for Regression Analysis

1. Some important command to manipulate data

1.1 Adding Labels to the Columns

You may wish to attach a more descriptive name to a column. A name in Minitab may be up to 8 characters long. The NAME command in Minitab assigns names to columns. The syntax is this:

```
NAME column= 'name' column='name'
```

The following example assigns the names TEST1,TEST2,TEST3, and TEST4 to columns C1 through C4.

```
MTB> NAME C1='ID' C2='TEST1' C3='TEST2' C4='TEST3'
MTB> NAME C5='TEST5'
```

The name of a column can be changed by another name command. To erase a name (but not the values in the column), assign a null name, like this:

```
MTB> NAME column=''
```

1.2 Deleting Rows

The DELETE command in MINITAB deletes rows in a worksheet. The syntax is:

```
DELETE row cn-cn
```

To delete row 5 of the worksheet, type:

```
MTB> DELETE 5 c1-c5
```

1.3 Deleting Columns

The ERASE command in MINITAB can delete columns and constants in a worksheet. The syntax is:

```
ERASE column
```

To delete column 1 of the worksheet, type:

```
MTB> ERASE C1
```

1.4 Creating a new column

You can calculate new columns based on old ones in Minitab. The LET command in Minitab is used to calculate a column. The syntax is this:

```
MTB> LET column= arithmetic expression
```

The following example calculates the improvement from TEST1 to TEST4.

```
MTB> LET C6=C5-C2
```

Since C5 and C2 are named, we could have also done this:

```
MTB> NAME C6= 'IMPROV'
```

```
MTB> LET 'IMPROV'='TEST4' -'TEST1'
```

Arithmetic operators include +, -, * (for multiply), / (for divide), and ** (for exponentiation). Also parenthesis can be used to clarify complex operations. The following example calculates column 7 as the square of column 6.

```
MTB> LET C7= C6**2
```

More complex operations are possible. The following example would compute a weighted mean if the first three tests were worth 50% of the grade and the final test were worth 50% of the grade.

```
MTB> LET C8=(.50*((C2+C3+C4)/3))+(.50*C5)
```

```
MTB> LET C8=(.50*((C2+C3+C4)/3))+(.50*C5)
```

1.5 Using Dummy Variables in Regression

An equivalent to using analysis of variance is to use what are called dummy variables in regression. Dummy variables are indicator variables. They are normally coded 0 or 1. In order to use our 'CLASS' variable as dummy variables in regression, we would need to create these dummy variables:

```
X1 = 1 if 'CLASS' is a 1 (freshman), 0 otherwise
X2 = 1 if 'CLASS' is a 2 (sophomore), 0 otherwise
X3 = 1 if 'CLASS' is a 3 (junior), 0 otherwise
```

It is necessary to have only 3 dummy variables even though there are 4 classes: if X1, X2, and X3 are all 0, then we have a senior.

The CODE statement in MINITAB allows us to create these variables from the existing 'CLASS' variable. Let's use columns 10,11, and 12 to store these new variables. The statements to create the new variables would be the following.

```
MTB> NAME C10= 'X1' C11= 'X2' C12= 'X3'
MTB> CODE (1)1 (2 3 4)0 'CLASS''X1'
```

```
MTB> CODE (2)1 (1 3 4)0 'CLASS' 'X2'
MTB> CODE (3)1 (1 2 4)0 'CLASS' 'X3'
```

The first code statement asks that we code the value 1 to a 1, and the values 2,3, and 4 to a 0, for the variable 'CLASS', and put the results in the variable 'X1'.

Example) Gender Dummy Variable

C1-T	C2
MALE	1
FEMAIL	0
MALE	1
FEMAIL	0

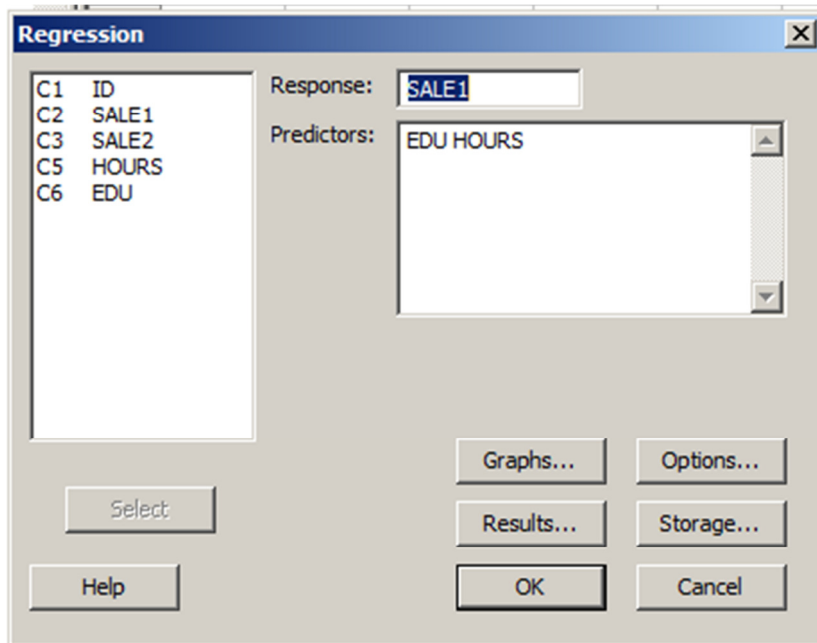C1 is a string variable including the gender information. To create a dummy variable to C2, use

```
MTB >  CODE ('MAIL')1 ('FEMAIL')0 C7 C8
```
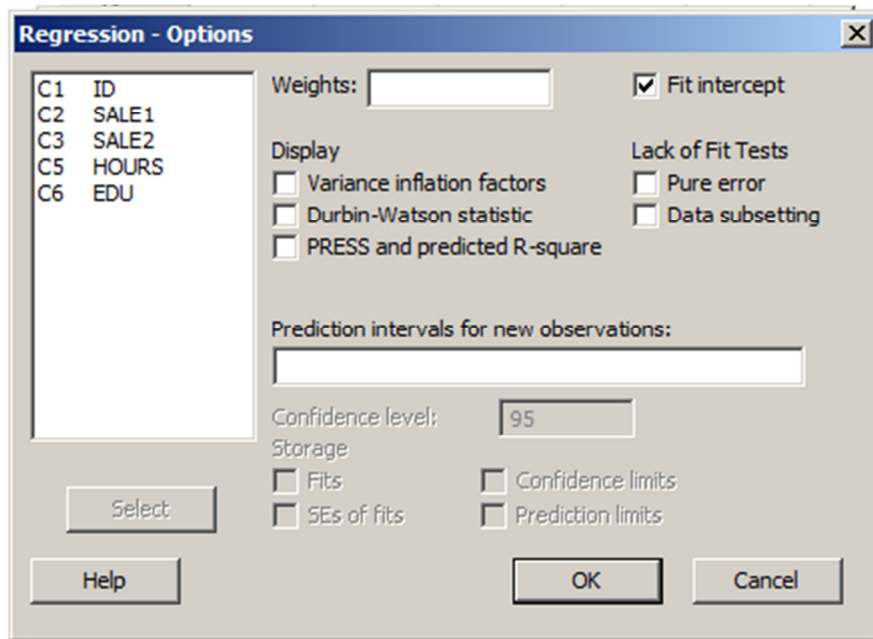
2. Regression Analysis

The command for regression is REGR with dependent variable and follows a number indicates the number of independent variables.

```
MTB >  REGR C2 1 C5
MTB >  REGR C2 2 C5 C5
```

To use menu STAT - > Regression -> Regression ->

To add some more statistics, you can click **Options** button on the bottom right side.

The Variance Inflation Factors (VIF) and the Durbin-Watson test are available under this options.