Minitab Users Guide

to accompany

Statistics: Unlocking the Power of Data

by Lock, Lock, Lock, Lock, and Lock
Getting Started

NOTE: In Minitab, your worksheet is your dataset. Your project is all the analysis, including graphs. Be sure to save your worksheet (as a worksheet not a project) once you have entered your data!

To enter data:

Remember that each column is a different variable and the rows are the cases.

1. Download a dataset as a Minitab (.mtw) file, and double click to open it in Minitab.

2. If your data already exists as a spreadsheet (such as Excel or Google Doc), you can import it into Minitab by copying and pasting. Be sure, however, that titles of columns (names of variables) go in the title row only, which is the grey row directly below C1, C2, etc. Titles can never go below that, where only the data belongs.

3. If you are typing your data in yourself, you enter the values as you would into any spreadsheet, using arrow keys to move about the worksheet. Be sure that column titles (names of variables) go only in the row right below the row with C1, C2, etc. We recommend that you name your variables using something short and descriptive.

**Warning**: Minitab has quantitative variables, which can contain only numbers, and categorical variables, which can contain anything. *As soon as you enter anything other than a number in a column, Minitab assumes that the entire column is categorical*. This includes things like dollar signs, units, etc. If Minitab has concluded that a column is categorical, you will see a “-T” added to the column name at the top and Minitab will not ever allow you to do quantitative calculations with that column. It will stay a categorical variable forever, even after you delete the offending character.

If you enter something other than a number in a quantitative variable column by mistake, you will have to re-enter the values for that variable in a different column (without any non-numbers!)

Once you have entered or imported your data, be sure to SAVE the data as a Minitab worksheet so that you can use it again.

Selecting Variables for Analysis

To select a variable (for doing data analysis), click in the box where the variable is requested, then double click on the variable name.

If the variable you are looking for is not showing up, chances are that you are using a method that is only appropriate for quantitative variables, and the variable you are looking for has been deemed categorical (it will have a –T in the top row). If the variable should be quantitative, enter it in a new column with only numbers in the white cells.
Visualizing and Summarizing Data (Chapter 2)

One Quantitative Variable

Statistics for a single quantitative variable:
Stat ⇒ Basic Statistics ⇒ Display Descriptive Statistics
Select the variable: With the cursor in the Variables box, double-click on the variable desired.

Graphs for a single quantitative variable:
Graph ⇒ Histogram or Dotplot or Boxplot
Choose “Simple” for now.
Select the variable and click OK.
Note: If you want your boxplots to be horizontal, click on “scale” and then transpose value and category scales.

One Quantitative Variable by categories in One Categorical Variable

Statistics for a quantitative variable by categories in a categorical variable:
Stat ⇒ Basic Statistics ⇒ Display Descriptive Statistics
Select the quantitative variable for “Variables” and the categorical variable for “By”

Graphs for a quantitative variable by categories in a categorical variable:
Graph ⇒ Boxplot or Histogram or Dotplot
▪ For Histogram choose “With Groups”
▪ For Boxplot/Dotplot choose “One Y: With Groups”.
▪ Select the quantitative variable for the “Graph variable” and select the categorical variable for the “Categorical variable.”

Note: If you want your boxplots to be horizontal, click on the “scale” button and then click on the box to transpose scales.

Two Quantitative Variables

Statistics for two quantitative variables:

Correlation: Stat ⇒ Basic Statistics ⇒ Correlation
Select the two quantitative variables. (If you select more than two, you will get a correlation matrix.)

Regression: Stat ⇒ Regression ⇒ Fitted Line Plot
Select the response (Y) and explanatory (predictor, X) variables.

Graphs for two quantitative variables: Graph ⇒ Scatterplot
Choose either “Simple” (without the regression line) or “With Regression” if you want the regression line added. Enter the Y (response) variable and X (explanatory) variable.
Categorical Variables

Tables for categorical variables:

For a single categorical variable: **Stat ⇒ Tables ⇒ Tally Individual Variables**
With the cursor in the Variables box, double-click on the variable(s) you want. You can select Counts and/or Percents, then click OK.

For a two categorical variables: **Stat ⇒ Tables ⇒ Cross Tabulation and Chi Square**
Select a variable for rows and one for columns (you can ignore labels for now). Select Counts and/or Row Percents and/or Column Percents and/or Total Percents,

Graphs for categorical variables:

*If your data is in the form of a dataset (with each row as a different case):*

For a single categorical variable: **Graph ⇒ Bar Chart** or **Pie Chart**
For two categorical variables: **Graph ⇒ Bar Chart**, then choose **Cluster** or **Stack**

*If your data is in the form of a frequency table or two-way table:*

For a single categorical variable:

- Enter your table directly into open columns of a Minitab worksheet, with the category names in one column and counts in another column. For a variable with 25 in category A and 75 in category B, this would look like this:
  
  *Remember to only enter numbers in the white cells for the counts column.

- For a Bar Chart:
  - Use **Graph ⇒ Bar Chart**
  - Choose bars represent **Values from a table** from the drop-down menu
  - Choose **Simple**
  - Select the counts as the graph variable and the category names as the categorical variable

- For a Pie Chart:
  - Use **Graph ⇒ Pie Chart**
  - Select **Chart values from a table**
  - Select the category names as the Categorical variable and the counts as the Summary variable.

For two categorical variables:

- Enter the table directly into open columns of a Minitab worksheet, like this:
  
  ![Table Example]

- Use **Graph ⇒ Bar Chart**
- Choose bars represent **Values from a table** from the drop-down menu
- Choose **Cluster** or **Stack** under **Two-way table**
- Enter all columns with counts as Graph variables and the column with row names as the Row labels.
Using Minitab in Chapter 6

Inference for Proportions: Intervals and Tests using the Normal Distribution

Interval and Test for a Proportion: Stat => Basic Statistics => 1 Proportion

If you have raw data: Choose “One or more samples, each in a column” and then select the variable. If you have summarized data: Choose “Summarized data” and enter the relevant numbers.

Interval: If you then click OK, you will get a 95% confidence interval. If you want a different confidence level, select “Options” and change the level.

Test: Check “Perform Hypothesis Test” and then enter the Hypothesized (null) proportion. If you want a one-tailed alternative, select “Options” and change the alternative.

Interval and Test for a Difference in Proportions: Stat => Basic Statistics => 2 Proportions

If you have raw data: Choose “Both samples are in one column” and then the relevant categorical variable for “Samples” and the relevant categorical variable for “Sample IDS” (the grouping variable). (Note that the variables can only have two possible categories for a difference in proportions test.)

If you have summarized data: Choose “Summarized data” and enter the relevant counts.

If you want a different confidence level or a one-tailed alternative hypothesis test, select “Options”.

Inference for Means: Intervals and Tests using the t-Distribution

Interval and Test for a Mean: Stat => Basic Statistics => 1-Sample t

If you have raw data: Choose “One or more samples, each in a column” and select the variable. If you have summarized data: Choose “Summarized data” and enter \( n, \bar{x}, \) and \( s \).

Interval: If you then click OK, you will get a 95% confidence interval. If you want a different confidence level, select “Options” and change the level.

Test: Check “Perform Hypothesis Test” and then enter the Hypothesized (null) proportion. If you want a one-tailed alternative, select “Options” and change the alternative.

Interval and Test for a Difference in Means: Stat => Basic Statistics => 2-Sample t

If you have raw data: Choose “Both samples are in one column” and then select the quantitative variable for Samples and the categorical variable for Sample IDS. (Note that the categorical variable can only have two possible categories for a difference in means test.)

If you have summarized data: Choose “Summarized data” and enter the relevant statistics.

If you want a different confidence level or a one-tailed alternative hypothesis test, select “Options”.
Using Minitab for Chapter 7

Chi-Square Tests

Chi-Square Goodness-of-Fit Test for a Single Categorical Variable

Stat => Tables => Chi-Square Goodness-of-Fit Test

If you have the observed counts, enter the counts in the “Observed counts” box, separated by spaces. Enter the corresponding category names in the “Category names” box. If your categorical data is in a column of a dataset, select “Categorical Data” and click in the box and select your categorical variable.

If you are testing that the categories are equally likely, leave the “Equal Proportions” option checked. If, instead, you have different assumed proportions for the different categories, select “Specific Proportions” and change the “Input column” box to “Input constants”. A window will open allowing you to enter the null hypothesis assumed proportions. Then click OK.

The test gives the observed counts, the expected counts, the contribution of each cell to the chi-square statistic, the chi-square statistic, and the p-value. It also opens two graphs related to the test.

Chi-Square Test for Association for Two Categorical Variables

Stat => Tables => Chi-Square Test for Association

If your categorical variables are in two columns of a dataset, choose “Raw data” and enter them in the boxes “For rows” and “For columns”. (It doesn’t matter which is which.)

If you just have the two-way table, enter the counts from the two-way table (not including totals) directly into blank columns of your worksheet like this: Choose “Summarized data in a two-way table” and select the columns containing the counts and the column containing the row labels.

By default, you get the observed counts, expected counts, chi-square statistic, and p-value. If you also want to see each cell’s contribution to the chi-square statistic, before clicking OK click “Statistics” and choose “Each cell’s contribution to chi-square”.

Using Minitab for Chapter 8

Analysis of Variance for Difference in Means

Stat => ANOVA => One-Way

Choose “Response data are in the one column for all factor levels” then select the quantitative variable as the response and the categorical variable as the factor. You will see the analysis of variance table, including the p-value, as well as some summary statistics for each of the groups. You will also see 95% confidence intervals for each of the group means based on the pooled standard deviation, as described in Section 8.2, displayed in the output and a plot.
Using Minitab for Chapters 9 and 10

Correlation

Stat => Basic Statistics => Correlation

If you enter two quantitative variables in the box and click OK, you will see the correlation between the two variables and the p-value for a two-tailed test of that correlation. (Make sure “Display p-values” is checked in order to see the p-value.)

If you enter more than two quantitative variables in the box and then click OK, you will see a correlation matrix. In this matrix, the rows and columns are labeled with the variables, and the correlation (with p-value below it) is shown in the matrix entry for that row and that column. Again, make sure “Display p-values” is checked in order to see the p-values.

Scatterplot

Graph => Scatterplot

Choose either “Simple” (without the regression line) or “With Regression” if you want the regression line added. Enter the Y (response) variable and X (explanatory) variable.

Linear Regression

Stat => Regression => Regression

Enter the Response variable in the “Response” box and the explanatory variable(s) in the “Predictors” boxes, and click OK. You will see:

- Analysis of Variance
  - The p-value for the overall significance of the model is the top p-value
- Model Summary:
  - the standard deviation of the residuals (S)
  - R² (R-sq)
- Coefficients:
  - slope coefficients (Coef) for each explanatory variable and the intercept (Constant)
  - standard errors for each coefficient (SE Coef)
  - t-statistic for each coefficient (T-Value)
  - p-value for each coefficient (P-Value)
- Regression equation