

Microsoft Excel Data Analysis ShortCourse Handout

July 2004

Copyright © 2004 Heide Mansouri, Technology Support, Texas Tech University. ALL RIGHTS RESERVED. Members of Texas Tech University or Texas Tech Health Sciences Center may print and use this material for their personal use only. No part of this material may be reproduced in any form without written permission from the [author](#).

Introduction

Microsoft Excel is an excellent data analysis tool. Simple functions such as Average, Median, and Mode are built into the program. However, you can add in pre-packed, statistical tools called the **Analysis ToolPak** to perform analytic analysis. You can use these tools to calculate Descriptive Statistics, create Histograms, create Rank-and-Percentile Tables, and more.

In this shortcourse, we'll explore statistical analysis functions built into Excel, as well as those included with the Analysis ToolPak. **You should already have a basic knowledge of Excel formulas, functions, and some experience with the science of statistics to gain the most from this shortcourse.**

Course Objectives

After completing this shortcourse, you should be able to:

- Calculate Descriptive Statistics for a set of data;
- Create a Histogram;
- Create Frequencies;
- Create a Pivot Table; and
- Make changes to a Pivot Table.

To Perform Data Analysis

- From the Tools menu, click Data Analysis. If Data Analysis is not available, load the Analysis ToolPak.
- On the Tools menu, click Add-Ins.
- In the Add-Ins available list, select the Analysis ToolPak box, and then click OK.

Data Entry in Excel List Format

The key point on data management in Excel is to enter or organize the data so that it is in Excel's "list format."

To Create Lists:

- Use the top row of the table to assign labels to each column.
- Keep column labels short (no more than 8 characters).
- Use numerical codes for any well defined variable, e.g. Gender:
0 = Female, 1 = Male.

Tools To Summarize Data in Excel

- PivotTables
- Descriptive Statistics

- Histograms

Descriptive Statistics

- The Descriptive Analysis tool generates a detailed list of statistics for a set of data, including mean, median, and mode.
- From the Tools menu, click Data Analysis > Choose Descriptive Statistics > Specify the Input and the Output range > OK.

PivotTable

A pivot table is an interactive table that provides summary information about data stored in an Excel Worksheet by using the row and column headings of your table. For example, you may use a Pivot table to count the number of occurrences of each category in a Data set.

Summary field Row field Column field

	A	B	C	D	E	F	G	H	I	J	K
1	Name	Status	Change of school			Count of Status	Name				
2	Tim	Enrolled	Transfer			Change of school	Alice	Mary	Tim	Tom	Grand Total
3	Mary	Probation	Computer			Business				1	1
4	Tom	Enrolled	Business			Computer		1			1
5	Alice	Enrolled	Transfer			Transfer	1		1		2
6						Grand Total	1	1	1	1	4

PivotTable Field List

Drag items to the PivotTable report

- Name
- Status
- Change of school

Add To: Row Area

PivotTable

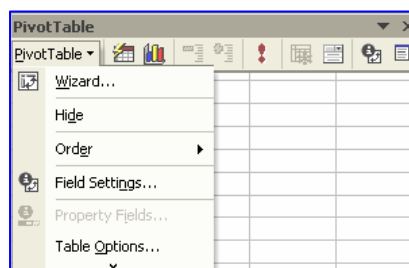
PivotTable

Creating a PivotTable

- Organize your data in a List.
- Remove all totals and subtotals created.
- Place your cursor inside the List (anywhere).
- Click **Data>PivotTable Wizard**.
- Follow the Wizard steps.

Making Changes to a PivotTable

- Select any data cell in the table; bring up the mouse shortcut menu (right-click) and choose the option you want.
- -Or- Select any data cell in the table.
- Click the **PivotTable>Field Setting**.
- Select the PivotTable (from the PivotTable Toolbar, Select)
- Click **Edit>Clear**.



To Delete a PivotTable

- Click inside the PivotTable.
- From the PivotTable Toolbar > Select > Entire Table.
- Click **Edit>Clear**.

Printing PivotTables

- Choose **File>Page Setup**.
- On the Sheet Tab, clear Rows To Repeat At Top and....
- Select any cell in your PivotTable.
- Choose **PivotTable>Table Options**.
- Select Set Print Titles.

Frequency

- Frequency function calculates how often values occur within a range of values and then returns a vertical array of numbers. An array is a collection of items (an orderly arrangement). For example, use FREQUENCY to count the number of test scores that fall within ranges of scores.
- Because FREQUENCY returns an array, it must be entered as an array formula.
- Array formula acts on two or more sets of values called array arguments, returning multiple results (an array).
- Syntax: FREQUENCY(data_array,bins_array) Where:
 - o **Data_array** is an array of or reference to a set of values for which you want to count frequencies. If data_array contains no values, FREQUENCY returns an array of zeros.
 - o **Bins_array** is an array of or reference to intervals into which you want to group the values in data_array.

Array Formula Rules

- Select the cell or range that will contain the results.
- Type your formula into the formula bar.
- Press Ctrl + Shift + Enter (to lock the formula).

The Histogram Analysis Tool

- The Histogram Analysis tool calculates individual and cumulative frequencies for a cell range of data and data bins.
- You should determine the intervals of the distribution before using this tool.

Creating a Histogram

- Select **Tools>Data Analysis>Histogram** from the menu bar, and a dialog box will appear.
- Insert the input range by either entering the reference by keyboard or highlighting the input range on the worksheet.
- Set up the ranges for the histogram divisions.
- Check the chart output box to obtain the histogram and click on the OK button.

Where to Get Help

If you need help with your project, you may contact Help Central Office at 742-HELP.

If you need help from your instructor, you may e-mail heide.mansouri@ttu.edu.

Exercise 1

Find the Descriptive Statistics for the following data:

Vehicle mileage: 27, 29, 33, 21, 21, 12, 16, 25, 8, 17, 24, 34, 38, 15, 19, 19, 41

Steps:

1. Enter the data in a single column in a worksheet.
2. Click in a cell where you want the results to go (C4 for example).
3. From the Tools menu, choose the Data Analysis.
4. Double-click the Descriptive Statistics.
5. Enter the range of the cells containing the data, including the label (drag over the cells or type).
6. Click columns for "Grouped By" options.
7. Check box Labels in the first row.
8. Check-mark the summary statistics and other options.

Exercise 2

10 people were asked to rate 3 restaurant owners' business outlooks (1=very unfavorable, to 6=very favorable). The following data was obtained:

ID	Outlook	Owner
261	6	1
262	4	2
263	1	3
264	3	2
265	3	2
266	2	1
267	1	3
268	5	2
269	4	3
270	4	2

Tally the results using the PivotTable Wizard.

Exercise 3

Using the PivotTable Wizard, find the percentages of each academic major for the following data:

ID	Major
1	Finance
2	Accounting
3	Finance
4	Accounting
5	Computer
6	Accounting
7	Computer
8	Finance
9	Computer
10	Computer

Exercise 4

Using the PivotTable Wizard, tally the following data (construct a contingency table):

Name	Status	Change of school
Tim	Enrolled	Transfer
Mary	Probation	Computer
Tom	Enrolled	Business
Alice	Enrolled	Transfer

Exercise 5

Academic rank of 20 professors is tabulated in the following table. Find the frequency distribution of the data, making sure to create a bin array first.

Professor-ID	Academic rank (1-5)
1	2
2	2
3	1
4	4
5	5
6	3
7	2
8	3
9	4
10	2
11	4
12	5
13	5
14	2
15	5
16	2
17	3
18	4
19	2
20	5

Exercise 6

SAT test scores of 20 students are given in the following table. Using the bin values of 850, 900, 950, 1000, 1050, 1100, and 1150:

1. Construct a frequency distribution table of the total SAT scores.
2. Construct a histogram, checking the Chart Output box in the Output Options of the Histogram dialog box.
3. Format your histogram:
 - a. Change the background color of the chart.
 - b. Type SAT Scores for “bin” on the x-axis of the chart.
 - c. Perform more charts editing, if you can.

Student -ID	Verbal	Math	Total
1	418	518	936
2	465	555	1020
3	466	556	1022
4	533	570	1103
5	476	566	1042
6	520	554	1074
7	470	499	969
8	468	545	1013
9	503	549	1052
10	551	519	1070
11	414	488	902
12	497	477	974
13	444	502	946
14	420	524	944
15	520	533	1053
16	511	489	1000
17	411	477	888
18	509	540	1049
19	521	530	1051
20	412	560	972

Please e-mail your comments or suggestions to: heide.mansouri@ttu.edu