APPENDIX C

DATA MANAGEMENT WITH EXCEL

C-1 INTRODUCTION
Excel is not database management software; however, it provides basic data management features and commands that can organize a spreadsheet’s data in many forms for use in diverse decision-making activities. These basic features and commands include sort operations, search operations using the Filter command, statistical database functions, the Frequency function, and the pivot table.

C-2 SORT OPERATIONS
Sort operations are used to organize a table based on one or several of its columns (fields). Basically, a table can be organized around any or all of its columns. The first column chosen serves as the primary key, and the other columns serve as secondary keys.

To sort the table shown in Exhibit C.1 based on any of its columns, follow these steps:

1. Select the table (excluding the column headings).
2. Click Data.
3. Click Sort.
4. In the Sort dialog box (Exhibit C.2), choose the column around which you want to sort the table.
5. Choose the order (Smallest to Largest, Largest to Smallest, Custom List...).
6. Click OK.
Exhibit C.1
A sample table

Exhibit C.2
The Sort dialog box
In this example, we chose Income for the column and Largest to Smallest for the order. The result is shown in Exhibit C.3.

Exhibit C.3
The table in Exhibit C.1 is sorted based on Income in Largest to Smallest order.

Using the sample table from Exhibit C.1, we now want to sort first by Gender (A to Z) and then by Income (Z to A). To do this, follow these steps:
1. Select the table (excluding the column headings).
2. Click Data.
3. Click Sort.
4. Choose Gender (A to Z).
5. Click Add Level (in the upper-left corner of the Sort dialog box).
6. Choose Income (Z to A).
7. Click OK.
The result is shown in Exhibit C.4.
C-3 SEARCH OPERATIONS

When you perform a search operation, you are interested in specific records that meet certain criteria. For example, in a student grade table, you might want to search for all those students who have a GPA greater than 3.6; or in an employee table, you might want to search for all the employees who hold master’s degrees. To conduct a search operation, you click Data and then click Filter.

Using the sample table from Exhibit C.1, we now want to search for engineers. To do this, follow these steps:

1. Put the cursor in any of the table’s cells.
2. Click Data, then click Filter. The Filter command is displayed, as shown in Exhibit C.5. As you can see, a down-arrow icon is displayed above each field (column).
3. Click the down-arrow icon next to the word “Occupation.” You will see what is displayed in Exhibit C.6.
4. Uncheck all the occupations except Engineer, and then click OK. The result is shown in Exhibit C.7. As you see in Exhibit C.6, you can sort the filtered table in ascending or descending order.
Exhibit C.5
The Filter command

Exhibit C.6
Options under the Occupation field using the Filter command
Exhibit C.7
The table from Exhibit C.1 is filtered for Engineers

This table can be further filtered. For example, you may want to filter for only male engineers, female engineers, or engineers who meet certain age criteria.

C-4 CUSTOM FILTER

The basic Filter command allows you to filter a table when there are exact matches—for example, the job title “Engineer,” the gender “male,” or the income “$90,000.” To search for inexact matches—“less than,” “greater than,” “in between,” and so forth—you have to use the custom filter feature. Using the sample table from Exhibit C.1, we now want to search for employees with an income over $55,000. To do this, follow these steps:

1. Put the cursor in any of the cells in the Income column.

2. Click Data, then click Filter.

3. Click the down-arrow icon in the Income column. You will see what is shown in Exhibit C.8. Toward the middle of the pull-down menu, click Number Filters. You will see what is shown in Exhibit C.9.

4. Click Greater Than. You will see what is shown in Exhibit C.10.

5. In the Custom AutoFilter dialog box, to the right of “is greater than,” click the down-arrow icon and choose $55,000.

6. Click OK. The result (all employees with an income over $55,000) is shown in Exhibit C.11.
Exhibit C.8
Options under the Income field using the Filter command

Exhibit C.9
Options under Number Filters
Exhibit C.10
Custom AutoFilter dialog box

Exhibit C.11
Employees with income over $55,000
C-5 ADVANCED FILTER

The Advanced Filter feature allows you to combine several criteria. For example, in the sample table in Exhibit C.1, we are searching for all the employees who are female, are under the age of 40, and have an income over $50,000. Follow the steps below:

1. Copy the table headings to row #20 in order to establish the Criteria range. The Criteria range can be anywhere outside the table. We chose row #20.
2. Under Gender, type F; under Age, type <40; and under Income, type >50,000. Please notice that case and alignment could matter. To be safe, for nonnumeric data, use the copy command in order to create the exact data under the Criteria range. At this point, your screen should be similar to the one shown in Exhibit C.12.
3. Put the cursor on any of the table cells, click Data, then click Filter, and then click Advanced. The Advanced Filter Dialog box opens.
4. In the List range, enter A3:F16 (if it’s not already entered).
5. In the Criteria range, enter A20:F21. At this point your screen should be similar to the one shown in Exhibit C.13.
6. Click OK. You will see what is displayed in Exhibit C.14. Notice there are only two employees who meet all three conditions.

Exhibit C.12
Criteria range for female employees under 40 with incomes over $50,000
Exhibit C.13
Completed Advanced Filter for Exhibit C.12

<table>
<thead>
<tr>
<th>FIRST NAME</th>
<th>LAST NAME</th>
<th>ACHIEVEMENT</th>
<th>OCCUPATION</th>
<th>INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen</td>
<td>Brown</td>
<td>30</td>
<td>Engineer</td>
<td>$75,000</td>
</tr>
<tr>
<td>Mario</td>
<td>Lopez</td>
<td>45</td>
<td>Manager</td>
<td>$100,000</td>
</tr>
<tr>
<td>Grace</td>
<td>Smith</td>
<td>35</td>
<td>Analyst</td>
<td>$50,000</td>
</tr>
</tbody>
</table>

Exhibit C.14
Employees who meet all three conditions

<table>
<thead>
<tr>
<th>FIRST NAME</th>
<th>LAST NAME</th>
<th>AGE</th>
<th>GENDER</th>
<th>OCCUPATION</th>
<th>INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fay</td>
<td>Brown</td>
<td>30</td>
<td>F</td>
<td>Engineer</td>
<td>$75,000</td>
</tr>
<tr>
<td>Sue</td>
<td>Hayward</td>
<td>25</td>
<td>F</td>
<td>Engineer</td>
<td>$60,000</td>
</tr>
</tbody>
</table>

Created using Microsoft Excel®, used with permission from Microsoft.
C-6 CONDITIONAL FORMATTING

When you use the Filter option, you basically hide the information that does not meet the specified criteria. However, if you are interested to see the entire table and at the same time be able to highlight the information that meets specified criteria, then Conditional Formatting is the right alternative. Let’s say that in the sample table (Exhibit C.1) you want to highlight all the employees with an income over $50,000. This is how you do it:

1. Select the entire Income column by clicking over F (column letter).
2. Click the Home tab, and then click Conditional Formatting. Your screen should be similar to the one shown in Exhibit C.15.
3. From the pull-down menu, click Highlight Cells Rules, and then click Greater Than.
4. In the Greater Than dialog box, type 50,000, and then click OK. Your screen should be similar to the one shown in Exhibit C.16. As you see in this exhibit, all the employees with income over $50,000 are highlighted.

To turn off this feature, follow the steps below:

1. Click anywhere in the table.
2. Click Conditional Formatting.
3. Click Clear Rules.
4. Click Clear Rules from Entire Sheet.

Exhibit C.15
Conditional Formatting menu

Created using Microsoft Excel®, used with permission from Microsoft.
C-7 DATABASE STATISTICAL FUNCTIONS

In Appendix A, you studied statistical functions. The same functions can be used with database data to add more flexibility to your analysis. For example, in the sample table (Exhibit C.1), you can count the number of male engineers with income over $50,000 and perform other statistical analyses. In other words, database statistical functions allow you to be selective with the data in the worksheet and let you choose a subset of the data for various types of analyses. The general format of Excel statistical function is as follows:

\[ \text{DFunction(database, field, criteria)} \]

The elements of this formula are as follows:

- **D** stands for database and must precede all database statistical functions—for example, DAVERAGE, DCOUNT, and DMAX.

- **Database** is the range of cells that makes up the list or database. In the table in Exhibit C.1, the entire database is the range: A3:F16. You may choose to use a portion of this database for analysis. However, in any selection, the column headings must be included in your range.

- **Field** indicates which column is selected in the function that you are using. Enter the column label enclosed between double quotation marks, such as “GENDER” or “OCCUPATION,” or a number (without quotation marks) that represents the position of the column within the list: 1 for the first column, 2 for the second column, and so on. The sample table in Exhibit C.1 has six columns, so they are numbered from 1 to 6.

- **Criteria** is the range of cells that contains the conditions you specify. You can use any range for the criteria argument, as long as it includes at least one column label and at least one cell below the column label in which you specify a condition for the column.

Exhibit C.17 shows the results of performing database statistical analysis on female employees who are under 40 years of age and have incomes over $50,000.

You could enter the Dcount function in either of the following two ways:

\[ =\text{DCOUNT}(A3:F16,3,A20:F21) \]
\[ =\text{DCOUNT}(A3:F16,”AGE”,A20:F21) \]
Exhibit C.17
Database statistical functions

C-8 DISTRIBUTION ANALYSIS USING THE FREQUENCY FUNCTION

There are cases when you may be interested in classifying spreadsheet data into an orderly group. For example, you may want to classify the income of all the employees of Jack’s Manufacturing into nine groups or classify your customers into eight sales groups. The Frequency function performs this distribution analysis for you.

Exhibit C.18 shows an example of the Frequency function. To use this function, you first must define the range of values you want to classify. Then select two empty columns. The first one is used for your Bin Array. The Bin Array contains the numeric intervals you want to use for analysis. When you enter intervals, Excel shows how many values are less than or equal to a particular interval but greater than the preceding interval. The column to the right of the Bin Array must be blank—for the results. In our example, the empty column adjacent to the Bin Array column is used by Excel in calculating the frequency distribution. Follow the steps below:

1. Under Bin Array, enter 20,000, 40,000, 60,000, 80,000, 100,000 in range C3:C7.
2. Select range D3:D7. This is used for the results of the Frequency function.
3. While the range is still selected, in cell D3 enter =FREQUENCY(B3:B15,C3:C7).
4. After entering the above function, press and hold down Ctrl, Shift, and Enter simultaneously.
5. As shown in Exhibit C.18, numbers 1, 3, 5, 3, and 1 are displayed in cells D3 through D7. These numbers show the number of income groups in each particular interval. For example, there is only one employee with an income between $0–$20,000, three employees between $20,001–$40,000, and so forth.
C-9 PIVOT TABLE

A pivot table in Excel is a powerful tool that allows you to summarize data for decision-making purposes. Take a look at the table presented in Exhibit C.19.

The ABC Company sells different merchandise through different salespersons in different states. The sales manager might be interested in knowing who sold what, where, and how much. The pivot table does all of this and more in a very easy fashion.
Exhibit C.19
A sample table

1. Put the cursor in any of the cells in the table.
2. Click Insert, and then click PivotTable.
3. The Create Pivot Table dialog box opens (see Exhibit C.20). As you see in this dialog box, the table range is already selected.
4. Click OK to accept the default. This will generate the results in a different worksheet (see Exhibit C.21). This is the skeletal structure of the pivot table. As you see in this exhibit, the pivot table is displayed on the left side. In the upper-right corner, you see all fields in the table. In the lower-right corner, you see FILTERS, COLUMNS, ROWS, and ∑ VALUES (sum of values). These are the summaries that the Pivot Table feature can generate for you.
5. Click Amount Sold. You will see what is displayed in Exhibit C.22, showing 77100. This is the amount sold by all the salespersons in all regions.
6. Click Region. You will see what is displayed in Exhibit C.23. Notice that the total sales are now broken down into three regions.
7. Click Salesperson. You will see what is displayed in Exhibit C.24. Notice that this exhibit shows the performance of each salesperson in each region.
8. Click Merchandise. You will see what is displayed in Exhibit C.25. Notice that sales are now broken down by region, then by salesperson, and then by product.
9. Click Salesperson and drop it under FILTERS. Click Region and drop it under COLUMNS. You will see what is displayed in Exhibit C.26, which shows a two-dimensional table with sales broken down by region and types of merchandise. You can do many other types of analyses with just a few clicks of the mouse!
Exhibit C.20
Create Pivot Table dialog box

Exhibit C.21
The skeletal structure of the pivot table
Exhibit C.22
Sum of Amount Sold

Exhibit C.23
Total sales broken down into three regions
Exhibit C.24
The performance of each salesperson in each region

Exhibit C.25
Sales broken down by region, then by salesperson, and then by product
Exhibit C.26
A two-dimensional table that shows sales performance by region and types of merchandise
1. What are two business applications of the sort operations?
2. How many keys (columns or fields) can be used when sorting a table?
3. What are two business applications of database statistical functions?
4. What are some business applications of the Pivot table?
5. Construct the table below and sort it by gender (Z to A), income (A to Z), and age (Z to A).

<table>
<thead>
<tr>
<th>FIRST NAME</th>
<th>LAST NAME</th>
<th>AGE</th>
<th>GENDER</th>
<th>OCCUPATION</th>
<th>INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randy</td>
<td>Smith</td>
<td>55</td>
<td>M</td>
<td>Professor</td>
<td>$90,000</td>
</tr>
<tr>
<td>Fay</td>
<td>Brown</td>
<td>22</td>
<td>F</td>
<td>Engineer</td>
<td>$75,000</td>
</tr>
<tr>
<td>Adam</td>
<td>Mills</td>
<td>37</td>
<td>M</td>
<td>Engineer</td>
<td>$55,000</td>
</tr>
<tr>
<td>Andrea</td>
<td>Harris</td>
<td>29</td>
<td>F</td>
<td>Teacher</td>
<td>$31,000</td>
</tr>
<tr>
<td>Moe</td>
<td>Santos</td>
<td>44</td>
<td>M</td>
<td>Officer</td>
<td>$49,000</td>
</tr>
<tr>
<td>Bob</td>
<td>Sakas</td>
<td>39</td>
<td>M</td>
<td>Engineer</td>
<td>$72,000</td>
</tr>
<tr>
<td>Vicki</td>
<td>Hashim</td>
<td>22</td>
<td>F</td>
<td>Teacher</td>
<td>$34,000</td>
</tr>
<tr>
<td>Paula</td>
<td>Schlosser</td>
<td>66</td>
<td>F</td>
<td>Nurse</td>
<td>$56,000</td>
</tr>
<tr>
<td>Jack</td>
<td>Negrete</td>
<td>51</td>
<td>M</td>
<td>Artist</td>
<td>$32,000</td>
</tr>
<tr>
<td>Mary</td>
<td>Fisher</td>
<td>27</td>
<td>F</td>
<td>Teacher</td>
<td>$33,000</td>
</tr>
<tr>
<td>Sue</td>
<td>Hayward</td>
<td>26</td>
<td>F</td>
<td>Engineer</td>
<td>$61,000</td>
</tr>
<tr>
<td>Jackie</td>
<td>Chipres</td>
<td>77</td>
<td>F</td>
<td>Engineer</td>
<td>$52,000</td>
</tr>
<tr>
<td>Lora</td>
<td>Heredia</td>
<td>26</td>
<td>F</td>
<td>Nurse</td>
<td>$48,000</td>
</tr>
</tbody>
</table>

6. In the above table (#5), filter all female engineers.
7. In the above table (#5), filter all employees who are making less than $40,000.
8. In the above table (#5) (using Advanced Filter), filter all employees who are male, under the age of 50, and have an income over $45,000.

9. In the above table (#5) using Conditional Formatting, highlight all the employees with incomes less than $50,000.
10. In the above table (#5) using database statistical functions, count all the female engineers, calculate their average salary, and list the oldest and the youngest in four separate rows of the worksheet.
11. Go through the Pivot Table presented in Exhibit C.19 step-by-step and confirm the screens and the numbers that have been generated. What are some other business applications of the Pivot Table?
12. To finalize the Frequency function you must press and hold down Ctrl, Shift, and Enter simultaneously. True or False?
13. In database statistical functions, Field indicates which row is selected in the function that you are using. True or False?
14. All of the following are examples of valid Excel database statistical functions except:
   a. =DTOTAL
   b. =DCOUNT
   c. =DMAX
   d. =DSUM

15. A table can be sorted by:
   a. One column
   b. Two columns
   c. As many columns as available in the table
   d. All of the above