



ACL IN PRACTICE







ACL IN PRACTICE

Copyright © 2007 ACL Services Ltd.

All rights reserved.

No part of these materials may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means (photocopying, electronic, mechanical, recording, or otherwise), without permission in writing from the publisher, except by a reviewer who may quote brief passages in a review.

ACL 9

December 2007

**ACL Services Ltd.**

1550 Alberni Street  
Vancouver, BC  
Canada V6G 1A5

Telephone: +1-604-669-4225

E-mail: [info@acl.com](mailto:info@acl.com)

Web: [www.acl.com](http://www.acl.com)

Printed in Canada

ACL, the ACL logo, the ACL logo with the text, “ACL Data you can trust. Results you can see.” and Audit Command Language are trademarks or registered trademarks of ACL Services Ltd.

Microsoft, Windows and Windows Server are trademarks or registered trademarks of Microsoft Corporation. z/OS is a registered trademark of IBM Corporation. SAP, R/2 and R/3 are trademarks or registered trademarks of SAP AG. Crystal Reports is a trademark or registered trademark of Business Objects SA. All other trademarks are the property of their respective owners.

200711071500

# CONTENTS

<b>Chapter 1: Introducing Metaphor Corporation</b> .....	<b>1</b>
Tutorial objectives .....	2
About Metaphor Corporation .....	3
Metaphor credit card policies .....	4
Acceptable category codes .....	5
Metaphor data files .....	5
Plan your analysis .....	6
Preliminary investigation .....	6
General tips to consider .....	7
Chapter review .....	8
<b>Chapter 2: Examine employee data</b> .....	<b>9</b>
Examine the employee list .....	10
Open and view the employee data project .....	10
Get a statistical picture .....	11
Look for duplicate records .....	13
Examine employee salaries and bonuses .....	14
Find total bonuses paid in 2002 .....	15
View salaries in descending and ascending order .....	15
View employees hired in 2002 .....	16
Determine bonus as a percentage of salary .....	17
Exercises .....	18
Chapter review .....	19
<b>Chapter 3: Set up your project</b> .....	<b>21</b>
Start your project in ACL .....	22
Create a project in ACL .....	22
Import your data files as tables .....	23
Exercise .....	27

Check and verify the data .....	27
Check the data type and format .....	28
Verify the data in a table .....	30
Document your work .....	30
Add record notes to a table .....	31
Exercise .....	32
Chapter review .....	32
<b>Chapter 4: Begin Your Analysis .....</b>	<b>33</b>
Determine the total liability .....	34
Extract expiring card numbers .....	34
Find transactions by customer number .....	35
Check the credit card balances .....	37
Exercises .....	39
Chapter review .....	39
<b>Chapter 5: Examine expense patterns .....</b>	<b>41</b>
Find the total expenses for each category code .....	42
Summarize the total expenses for each category .....	42
Isolate transactions by category .....	43
Isolate particular recreational expenses .....	45
Exercises .....	46
List the amounts for each card number .....	46
Exercise .....	47
Chapter review .....	48
<b>Chapter 6: Analyze transactions .....</b>	<b>49</b>
Determine which transactions are unacceptable .....	50
About the Join and Relations commands .....	50
Check the code tables for obvious errors .....	51
Identify unacceptable transactions .....	53

Identify the employees .....	56
Identify the employees' departments .....	57
Create a list of employees .....	59
Report your findings .....	61
Exercises .....	62
Chapter review .....	62
<b>Chapter 7: Validate, correct, and extract data .....</b>	<b>63</b>
Import the May transaction data .....	64
Use folders to organize your project .....	64
Import transaction data for May 2003 .....	65
Compare and correct the table layouts .....	66
Combine and verify the tables .....	67
Combine the May transaction tables .....	67
Verify the new table .....	68
Check the combined table for duplicates .....	69
Exercises .....	69
Isolate and correct validity errors .....	70
Check the format of card numbers .....	70
Create a 16-digit card number field .....	71
Extract the corrected records to a new table .....	72
Export the combined transaction table to Excel .....	73
Create an error-free employee record table .....	73
Exercises .....	74
Ideas for further analysis .....	74
Chapter review .....	75





# INTRODUCING METAPHOR CORPORATION

*Experiment with real-world data analysis scenarios*

**In this chapter...**

Review the tutorial objectives

About Metaphor Corporation

Look at your files

Plan your project

*ACL in Practice* is a tutorial that examines human resources records and credit card transactions at a fictional company, Metaphor Corporation. As you work through the procedures and examples, you learn to apply *ACL*'s analysis and reporting capabilities to real-world data. When you finish, you will be able to apply the power of *ACL* to your own data.

## ■ Tutorial objectives

In this tutorial you will:

- Work with an existing data analysis project
- Create and work with a new *ACL* project from start to finish
- Report results of your *ACL* project
- Export the results to another file format

These steps are typical for many data analysis projects that you perform in *ACL*, both small and large.

### Chapter 2: “Examine employee data”

**Objective:** Examine employee profile data for Metaphor Corporation.

**Summary:**

- Examine the project and its contents
- Create a statistical overview
- Examine employee bonuses and salaries

### Chapter 3: “Set up your project”

**Objective:** Create the project and add the data.

**Summary:**

- Create an *ACL* project.
- Import the data into your project. Each data source is a separate table in the *ACL* project.
- Verify the data.
- Document your work.

### Chapter 4: “Begin Your Analysis”

**Objective:** Determine basic background information about the credit card numbers and transactions.

**Summary:**

- Determine the total credit limit exposure of the company.
- Determine which credit card numbers are due to expire within a given time period, and extract this information.
- Find which transactions are linked to a particular customer.
- Determine some basic information about the credit card balances.

### Chapter 5: “Examine expense patterns”

**Objective:** Find information about itemization of expenses and spending patterns.

**Summary:**

- Determine transactions per merchant category code.
- Separate the transactions for airfare, hotel stays, or rental cars.
- Determine expenses for drinking establishments and golf courses.
- Determine total transaction amounts per credit card number.
- Determine transaction amounts per customer code.

**Chapter 6: “Analyze transactions”**

**Objective:** Determine whether transactions are acceptable, unacceptable, or require approval.

**Summary:**

- Determine which transactions are unacceptable, and which employees made these transactions.
- Create a report that includes employees with unacceptable transactions, details of the transactions, and the employees’ departments.

**Chapter 7: “Validate, correct, and extract data”**

**Objective:** Expand your analysis to include more data files.

**Summary:** This portion of the tutorial extends the Metaphor Corporation scenario.

- Import the May 2003 credit card transaction files as tables.
- Clean up the data in the May 2003 credit card transaction tables.
- Extract the repaired data to new tables.
- Combine the two May 2003 credit card transaction tables into one.
- Using the techniques you have learned, analyze the May 2003 credit card transactions.

**■ About Metaphor Corporation**

Since 1995, Metaphor has sold software and IT services. Metaphor has several dozen consultants who are full-time employees. These consultants work out of Metaphor’s head office and they also travel the world in order to install software, provide training courses, and initiate and maintain sales relationships.

Metaphor also has administrative and professional staff members in the head office. In addition, Metaphor has sales agents in various locations around the world who are not employees.

## Metaphor credit card policies

Most employees have a company credit card. When an employee purchases products or services for the company, he or she must use a company credit card.

Employees entertain clients and prospective clients as appropriate, both in Metaphor’s home city and while travelling. They are also responsible for their own professional development, and are expected to charge expenses to the company credit card for courses, books, journal subscriptions, and professional memberships. Some employees are also responsible for purchasing supplies for the office.

Thus, Metaphor Corporation incurs significant expenses in several categories:

- Travel
- Entertainment
- Client service
- Professional development
- Internal office expenses

Metaphor’s policy is that all purchases must be assigned to a customer number. Although most of the customer numbers correspond to external clients, customer numbers that start with zero correspond to internal departments.

Employees are aware that certain categories of transactions are unacceptable. For example, employees must not use corporate credit cards for gambling. Although the company sometimes approves transactions in other categories, permission must be granted in advance. Some categories are prohibited entirely.

By using *ACL* to analyze the credit card transactions of Metaphor employees, you can find transactions that are contrary to policy and transactions that might be erroneous or fraudulent.

The following table shows allowable expenses.

Travel	Entertainment	Office and professional
Airfare	Restaurants	Dues for membership clubs
Hotel	Drinking places	Dues for professional organizations
Taxi	Boat charters/rentals	Dues for fraternal organizations
Restaurants	Caterers	Books and periodicals
Rental cars	Bakeries	Computers, computer parts, and software
Long-distance bus tickets	Sporting events	Office supplies
Travel agencies	Recreational services	
Toll and bridge fees	Amusement parks	
	Golf courses	
	Tourist attractions	
	Theater tickets	
	Tour buses	
	Musicians	
	Bowling alleys	

## Acceptable category codes

The credit card analysis in this tutorial relies on the Merchant Category Code list, which is used by major credit card providers. It contains thousands of four-digit numeric codes, including unique codes for many major airlines and hotel chains, and numerous generic codes. Every credit card transaction has an accompanying category code assigned to it.

In the real world, credit card companies occasionally assign incorrect codes, so if you obtain questionable results, contact the merchant for verification.

For the purposes of this analysis, you can compare transaction codes to the lists of acceptable and unacceptable codes. For example, if a Metaphor employee uses his credit card to gamble, the transaction code indicates an unacceptable transaction.

The **Policies.doc** document, which accompanies the tutorial data, discusses appropriate credit card usage policies. The data analyst uses this document to determine what types of transactions to flag.

## Metaphor data files

When you perform an analysis for your own company, you often obtain the data in whatever file format is available. In this tutorial, however, you continue an analysis project that has already been planned. Much of the preliminary work is already complete.

For this analysis project, you use these files:

- **Acceptable\_Codes.mdb**  
This Access database file has a list of acceptable codes from the Merchant Category Code list.
- **Credit\_Cards\_Metaphor.xls**  
This Excel worksheet lists credit card information.
- **Company\_Departments.txt**  
This tab-delimited file contains a list of departments and their internal numbers.
- **Employees.csv**  
This comma-delimited text file lists the employee names, employee numbers, and credit card numbers.
- **Trans\_April.xls**  
This Excel worksheet lists the credit card transactions for April 2003.
- **Unacceptable\_Codes.txt**  
This tab-delimited text file contains a list of unacceptable codes from the Merchant Category Code list.

## ■ Plan your analysis

Developing an analytical strategy is typically an iterative process. As you work with the data in *ACL*, some patterns will likely become apparent, which might lead to new strategies. Here is a typical cycle:

- Plan the analysis
- Analyze the data
- Reconsider the analysis
- Analyze the data again

As you develop and work through your analysis, you might decide that you need additional data. For example, you use hire date and salary information in your analysis, but once you assemble preliminary results, you decide to acquire employee absence records. You can always add more data to your project, but obtain as much raw data as you can before you begin your analysis.

An *ACL* analysis involves a project that contains one or more tables. These tables are imported from external data sources, such as Excel worksheets, Access databases, text files, Oracle databases, and so forth. In a typical corporation, various departments store and assemble data on a daily basis. You can use *ACL* to combine data from disparate sources to create a single coherent analysis.

Once you have your data in place, you can develop objectives for your analysis. Consider some possible approaches. Make a note of tests that you think might be useful.

## Preliminary investigation

Because the tutorial includes ready-made data, you do not have to perform any preliminary investigation of the data. In general, however, you should investigate the data you want to use before you proceed.

Consider these questions:

- What data currently exists, and what format is it in?
- Who entered the data?
- Who maintains the data?
- Which department owns this data?
- Is the data accurate and properly formatted? If not, can the data be modified to make it suitable?
- Can you combine the data from more than one source?
- Can you derive the data you need from the data you already have?

Before you import a file into your project, look at the file in its original format. For example, you can open a Microsoft Excel worksheet to see what it looks like, or you can open a text file in Notepad.

In a large file, such as a database that contains millions of records, it might be impractical to look at the data in its original format. In that case, you can ask your IS department for a description of the fields and the field sizes. If possible, review an excerpt of the file to get a better understanding of its content.

For example, if you want to analyze some aspects of customer invoices over the period 1999 to 2002, the data might exist in a comprehensive database that is in a form

that is not easily accessible. You could ask your IS department for information about specific fields, which might look like this:

Field	Comments
Company Name	30 characters, Text
Amount	12 characters, Text
Date of Invoice	10 characters, Date, DD/MM/YYYY
Date_Payment	10 characters, Date, MM/DD/YYYY

With this type of information, you know in advance how the data appears. For example, you know that the two date fields are formatted differently.

Although *ACL* is capable of handling difficult data, you can make the process easier for yourself by ensuring that your data is clean.

## General tips to consider

Consider the following steps when you begin to organize a new project.

- **Refine your strategy with test data.** There is little need to restrict yourself to using data samples because *ACL* can read data files of any size. However, small extracts of data might be useful when you are developing an analytical strategy.

For example, you can look at at one day's worth of data before you analyze five year's worth. Working with a thousand-record file is much quicker than working with a million-record file. When you have your strategy in place, you can apply it to a million-record file.
- **Eliminate errors from your data.** The quality of your analysis will improve with the quality of your data. Try to eliminate any errors from your data before bringing it into *ACL*. If that is not possible, you can use *ACL* to correct the data.
- **Experiment with your data.** If you aren't sure of what an *ACL* function or command will do, you can try it anyway. Your data remains unharmed.

## ■ Chapter review

In this chapter you learned:

- The tutorial objectives.
- The Metaphor credit card policy.
- How to plan a typical *ACL* project.



## EXAMINE EMPLOYEE DATA

*Analyze salaries, bonuses, and hiring dates*

**In this chapter...**

Examine the employee list

Examine salaries and bonuses

In this chapter, you look at some Metaphor Corporation employee profile data. You analyze salaries, bonuses, hiring dates, and other human resources information. In the course of this analysis, you learn the basics of *ACL*, practice navigating the program, and analyze data.

In this scenario, you take over an *ACL* project from a coworker who has already started work on it. You analyze the data that your coworker assembled.

## ■ Examine the employee list

When you look at the employee data, you can get some ideas about how you might analyze this data. In this example, the project has already been assembled for you. When you work with your own projects, you decide what data to include. You can also add more data at any time to an existing project.

Your task is to determine various financial and human resource information, based on the data in your *ACL* project.

The project you work with in this chapter represents a small-scale, but fairly typical data analysis project. This project has two tables that contain information about Metaphor employees. You perform the analysis based on this information.

As you work through the tutorial, you might notice that the columns on your screen sometimes appear in a different order from the columns in the illustrations. To change the order of a column, click the field list heading in the dialog box with which you are working.

### ● ● ● ● ● ● ● ● ● ● Tasks

View the employee data project.

Get a statistical picture.

Check for duplicate employee records.

## Open and view the employee data project

You begin your examination of the employee data by viewing the contents of the project.

An *ACL* project is a holder for data that you import into *ACL* as tables. Once the data is imported into the project, it does not matter to *ACL* what the original data source was.

When you use *ACL*, you either create a new project to work with, or you open an existing project. You open, save, and close projects in *ACL* in much the same way that you manage documents in other Windows programs.

### ► Open the employee data project

- 1 Open *ACL*.
- 2 Select **File » Open Project**. Browse to locate the **C:\ACL Data\Sample Data Files** folder. Select **Metaphor\_Employee\_Data**. Click **Open**.

If your installation of *ACL* did not use the default folders, this file may be located elsewhere on your system.

Notice that the **Project Navigator** pane is at the left of the window, and the table view is at the right. On the **Overview** tab, you can access the tables, scripts, and log.

In this project there are two tables: **Agents\_Metaphor** and **Employee\_List**. The **Agents\_Metaphor** table lists all of Metaphor's independent sales agents around the world. The **Employee\_List** lists all of the Metaphor employees at the head office.

► View the contents of the employee data project

- 1 In the **Overview**, right-click **Metaphor\_Employee\_Data.ACL** and select **Properties**. Click the **Notes** tab.

Although this project contains no notes, you can add them at any time to record details of your work.

- 2 Close the **Project Properties** dialog box.
- 3 Double-click **Agents\_Metaphor** in the **Overview** to open the table.
- 4 Select **Analyze » Count Records**. The **Count** dialog box appears. Click **OK**.

The count of 41 records appears at the bottom of your screen. Although you cannot see a difference on the screen, *ACL* wrote the information to the log.

- 5 Open the **Employee\_List** table. Count the number of employees. How many employees are there? You should get 200.

## Get a statistical picture

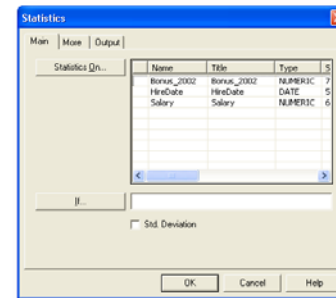
The **Statistics** and **Profile** commands give you a quick statistical snapshot. With larger tables, the information that you can get from **Statistics** and **Profile** is often useful.

The **Statistics** command works with numeric and date fields. **Profile** works with numeric fields only.

In this procedure, you generate statistics about the **Employee\_List** table, such as ranges of hiring dates and salaries.

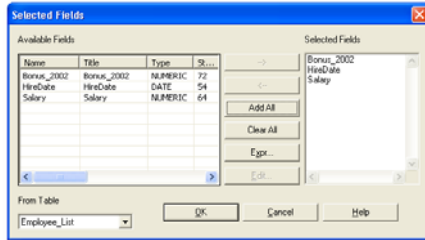
► Generate statistics on the employee list

- 1 With the the **Employee\_List** table active, select **Analyze » Statistical » Statistics**. The **Statistics** dialog box appears.

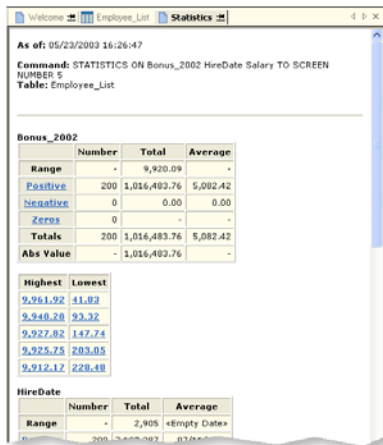


Before you run a command in *ACL*, ensure that no column in the view is selected. Otherwise, *ACL* will run the command on just the selected field.

- Click the **Statistics On** button. The **Selected Fields** dialog box appears. Click **Add All**.



- Click **OK**. On the **Output** tab, ensure that **Screen** is selected, Click **OK**. **ACL** displays the results.



If you scroll down, you can see that the earliest hiring date is 01/05/1995 and the latest hiring date is 12/19/2002. Some of the information provided is not useful, such as the total of the **HireDate** field values, so you can disregard this information.

If you prefer, you can run the Statistics command on each field separately, so that you see the results separately.

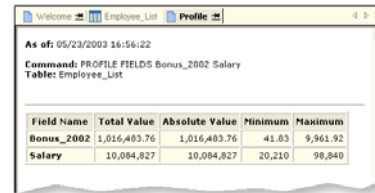
The Profile command also gives you an overview of the data that is more concise than that provided by Statistics.

- Use **Profile** to get information about a table

- With the **Employee\_List** table active, select **Analyze » Statistical » Profile**. The **Profile** dialog box appears.



- Select all of the fields and click **OK**. **ACL** displays the results.



Each of the values for each of the numeric table columns is shown on the tab.

## Look for duplicate records

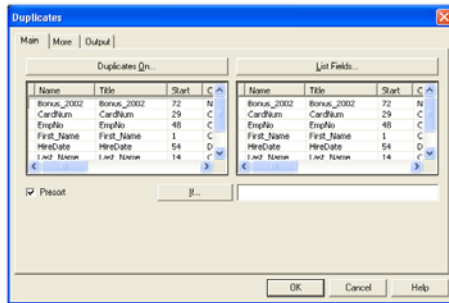
One way to ensure the integrity of the data is to check for duplicate records.

This might be a more complicated task than it seems at first glance. One employee might be listed under two names, or under the same name twice. Two employees might have been assigned the same employee number inadvertently. The type of duplicates you want to check depends on the data that you are working with.

In the next procedures, you check for duplicates on the employee last name and the employee number.

### ▶ Look for duplicate last names

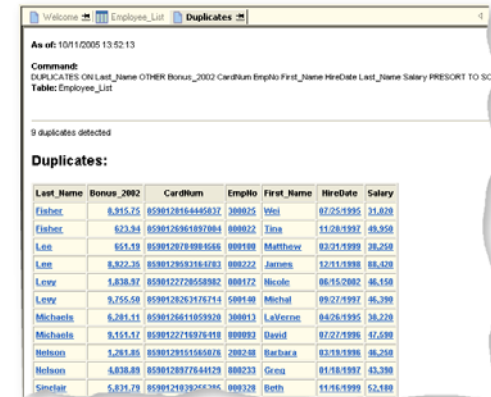
- 1 With the **Employee\_List** table active, select **Analyze » Look for Duplicates**. The **Duplicates** dialog box appears.



- 2 Click **Duplicates On**. The **Selected Fields** dialog box appears. Click **Last\_Name** and click the arrow button. Click **OK**.
- 3 Click the **List Fields** button. The **Selected Fields** dialog box appears. Click the **Add All** button. Click **OK**.

Although you are checking for duplicates in only one field, include all of the fields in the output screen to help you to interpret the results.

- 4 On the **Output** tab, select **Screen**. Click **OK**. **ACL** displays the results.



- 5 In the **Last\_Name** column, click a hyperlink. **ACL** displays a filtered view of the **Employee\_List**, showing that employee's details.

You next decide to check the **Employee\_List** to determine whether there are any duplicate employee numbers.

### ▶ Look for duplicate employee numbers

- 1 Click the **Remove Filter** button to remove the filter. The full **Employee\_List** table appears again.
- 2 Select **Analyze » Look for Duplicates**. The **Duplicates** dialog box appears.
- 3 Click **Duplicates On**. The **Selected Fields** dialog box appears. Select **EmpNo** and click the arrow button. Click **OK**.

- 4 On the **Output** tab, select **Screen**. Click **OK**.
- 5 View the **Duplicates** tab. No duplicate employee numbers are found.

## ■ Examine employee salaries and bonuses

You have been asked to find several things from the employee data:

- The total bonuses that were paid out in 2002.
- The salaries paid: from highest to lowest, and lowest to highest.
- The number of people who were hired in 2002.
- The ratio of bonus to salary for each employee in 2002.

### ● Tasks

- Find the total amount of paid bonuses.
- View salaries in ascending and descending order.

View employees hired in 2002.

Calculate bonus as a percentage of salary.

## Find total bonuses paid in 2002

Metaphor has a complex compensation structure. Because Metaphor is a project-oriented company, some employees have a compensation scheme that relies heavily on bonuses.

To find the total amount paid in bonuses in 2002, use the Total command.

### ► Use the Total command to find total bonuses paid in 2002

- 1 Open the **Employee\_List** table.
- 2 Select **Analyze » Total Fields**.
- 3 Click the **Total Fields** button. The **Selected Fields** dialog box appears. Select **Bonus\_2002** and click the arrow button. **Bonus\_2002** now appears in the **Selected Fields** list. Click **OK**. Click **OK**.
- 4 View the **Total** tab.

The figure shown on the **Total** tab is the total amount paid in bonuses in the year 2002. The total that appears should be **1,016,483.76**.

Suppose you wanted to see what command you just entered to achieve this result. You can look at the log.

- 5 Click the **Log** tab at the bottom left of the *ACL* window.

You can see the commands that you performed on this project.

A *command* is the name of a specific action taken in *ACL*, such as the *Statistics* command. To learn more about commands, select **Help » Index** and look up “commands”.

## View salaries in descending and ascending order

The Quick Sort feature lets you view the data in sequential order. Applicable only to the view, Quick Sort is a temporary means to view your records in ascending or descending order. When you are finished with Quick Sort, you can turn it off to return the view to the order in which the records exist in the table.

### ► View employee salaries in descending and ascending order

- 1 With the **Employee\_List** table active, right-click the **Salary** column header. Select **Quick Sort Descending**.

You can now see the employee salaries in descending order. You decide you would like to see them in ascending order.

- 2 Right-click the **Salary** column header. Select **Quick Sort Ascending**.


If you have a large file, Quick Sort operations can be lengthy. If you have small files, you might find Quick Sort useful. You can Quick Sort on any column, so you can get different perspectives on the data.

## View employees hired in 2002

To determine how many employees were hired in 2002, you apply a filter to the view. A filter is a logical *expression* that you can type in the view's Filter text box. *ACL* displays only those records that match the filter.

An *expression* is a set of operators and values that you use to perform calculations, specify conditions for a test, or create values that do not exist directly in the data. An *ACL* expression can be a combination of data fields or computed fields, operators, constants, functions, and variables.

### ► Create a filter to show employees that were hired in 2002

- 1 Click the **Edit View Filter**  button. The **Edit View Filter** dialog box appears. You can use this dialog box to create and edit filters that apply to the current table.
- 2 In the **Available Fields** list, double-click **HireDate**.
- 3 Click **>=** in the middle pane.
- 4 Click **Date**. Locate January 1, 2002 and click **OK**.

The expression should read:

```
HireDate >= `20020101`
```

Note that any filter that involves dates uses reverse single quotation marks. You cannot use typical quotation marks when you use dates, and you cannot use this special quote mark in other expressions.

Although you know that no employees have been hired in 2003, you want to ensure that the time period that you selected is correct. You want a list of all employees hired between January 1, 2002 and December 31, 2002.

- 5 Change the expression to read:

```
HireDate >= `20020101` AND HireDate <=  
`20021231`
```

- 6 Click **OK**.
- 7 View the filtered records.
- 8 Select **Analyze » Count Records** and click **OK**.

Look at the bottom of the *ACL* window. Note that there are 17 records showing up with this filter. This means 17 employees were hired in 2002, according to this list.




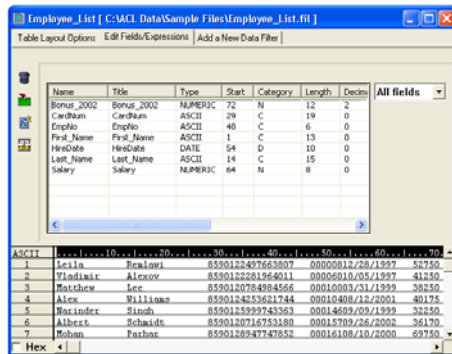
## Determine bonus as a percentage of salary


You want to determine the bonus as a percentage of salary for all employees. To do this, you add a computed field to the table layout.

A computed field is created with an expression. This expression can use one or more fields already in the table.

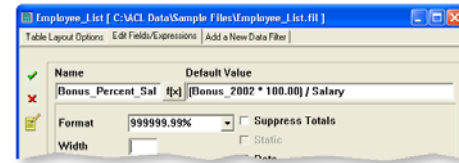
### ▶ Add a computed field to the table layout

- 1 With the **Employee\_List** table active, click the **Remove Filter**  button to remove the filter.
- 2 Select **Edit » Table Layout**. The **Table Layout** window opens with the **Edit Fields/Expressions** tab displayed.



- 3 Click the **Add a New Expression**  button on the left. The **Edit Fields/Expressions** tab changes so that you can enter the expression for the new computed field.
- 4 Enter **Bonus\_Percent\_Salary** in the **Name** field.
- 5 Click **f(x)**. The **Expression Builder** dialog box appears. The **Expression Builder** is the same as the **Edit View Filter** dialog box that you used already.

- 6 In the **Expression** box, enter the expression:  $(\text{BONUS\_2002} * 100.00) / \text{Salary}$   
This expression determines the 2002 bonus as a percentage of the employee's 2002 salary.
- 7 Click **OK**.
- 8 In the **Format** list, enter **999999.99%**.



To learn more about how ACL handles precision and accuracy in decimal calculations, select **Help » Index** and look up “controlling rounding in decimal expressions”.

- 9 To save the changes, click the green checkmark.
- 10 Close the **Table Layout** window.  
Although you have created the new field, it does not yet appear in the **Employee\_List** view.
- 11 Right-click the **CardNum** column header and select **Add Columns**. The **Add Columns** dialog box appears.
- 12 Select **Bonus\_Percent\_Salary** and click the arrow button. **Bonus\_Percent\_Salary** now appears in the **Selected Fields** list. Click **OK**.

Keep in mind that the ratio will be skewed for employees hired in 2002, because they did not work a full year at Metaphor. You can apply a filter to hide the records of employees hired in 2002.

When working in an *ACL* project, you might want to resize the widths of the columns. You can move the column separator to adjust the width of the column so that it occupies less space in the view. Alternatively, you can double-click the right side of the column heading to resize the column to the width of its visible display. To resize more than one column at a time, select multiple columns, then double-click the right side of a column heading.

- 13 Close the view and click **Yes** when *ACL* prompts you to save.

## Exercises

- 1 Can you run the Profile command on just the **HireDate** in the **Employee\_List**? Run the Statistics command on just the **HireDate**.  
What results do you get? When are the Statistics and Profile commands useful?
- 2 Quick Sort the **Employee\_List** on the **Bonus\_2002** column.
- 3 Quick Sort **Agents\_Metaphor** on the **Comm\_2002** column in ascending order, then in descending order. Run Statistics and Profile on this table.
- 4 Find the total commissions paid to Metaphor agents in 2002.
- 5 Make the view easier to read by adjusting column widths and dragging columns left or right. You can adjust the columns at any time while you follow the tutorial.

## ■ Chapter review

In this chapter you accomplished the following tasks:

- Opening a project
- Viewing the project contents
- Using Statistics and Profile commands to analyze data
- Using Quick Sort to sort the data in a specified order
- Creating and applying filters

You have completed a small demonstration of what you can do with *ACL*. You work through more complex examples in subsequent chapters.



## SET UP YOUR PROJECT

*Create the project and add the data*

**In this chapter...**

Manage projects in *ACL*

Import files into a project

Check and verify the data

Before you begin your analysis of Metaphor Corporation's credit card transactions, you must create a project and access the credit card data.

An *ACL* project functions as a container for data. You import the source data into your project from existing data sources, such as Excel worksheets, text files, Access databases, dBASE files, and so forth. You import each data source as a separate table. Once you import the data, you can then analyze it.

## ■ Start your project in ACL

Because you are starting a new data analysis project, you begin by creating a project in *ACL*.

### ● Tasks

- Create a project in *ACL*.
- Import your data files as tables.
- Other data import scenarios.

## ■ Create a project in ACL

To begin your work in *ACL*, create a project. You can save a project or close an *ACL* project just as you do with a word processing document or a spreadsheet.

### ▶ Create a project in ACL

- 1 Select **File » New » Project**.
- 2 In the **File name** box, enter `ACL_Tutorial_Metaphor` and click **Save**.

The **Data Definition Wizard** appears. You use this wizard to import data into your project.

If you need to change a previous selection when you use the **Data Definition Wizard**, click **Back** until you return to the screen that you want.

## Import your data files as tables

Now that you have begun a project in *ACL*, you add data to it. In each instance, you use the **Data Definition Wizard**, but the process varies slightly, depending on the type of file that you import.

You import data from each of the six files listed under “Metaphor data files” on page 5. *ACL* creates and saves a table for each file in a .fil file.

For the first file that you import, the Data Definition Wizard is already open because it appeared when you saved the new project. Start by importing the **Credit\_Cards\_Metaphor.xls** file, which *ACL* recognizes as an Excel file.

### ▶ Import the **Credit\_Cards\_Metaphor.xls** worksheet

- 1 On the **Data Definition Wizard** screen, click **Next** and proceed through the wizard until the **Select File to Define** dialog box appears. Browse to find the **C:\ACL Data\Sample Data Files** folder. If your installation did not use the default locations, this folder might be located elsewhere on your system.
- 2 Double-click **Credit\_Cards\_Metaphor.xls**.
- 3 Click **Next**. *ACL* displays the **Data Source** screen.
- 4 Select the worksheet. Enter a value in the **Maximum Character Field Length** box. Click **Next**. *ACL* displays the **Save Data File As** dialog box.
- 5 Enter **Credit\_Cards\_Metaphor**. Click **Save**.  
You get a listing of the fields with a description.
- 6 Click **Finish**. When a message appears that asks you to save the untitled table, click **OK**.

The **Credit\_Cards\_Metaphor** table opens in *ACL*.

In this example, the Excel workbook **Credit\_Cards\_Metaphor.xls** contains only one worksheet. If an Excel workbook has more than one worksheet, you select which worksheet to import on the **Data Source** screen of the **Data Definition Wizard**.

### ▶ Import the **Trans\_April.xls** worksheet

- 1 Select **File » New » Table**. The **Data Definition Wizard** appears. Click **Next** until the **Select File to Define** dialog box appears.
- 2 Locate and double-click **Trans\_April.xls**. Click **Next**.
- 3 Proceed through the wizard as you did with the previous table. Save this table as **Trans\_April**.
- 4 Click **Finish**. Click **OK**.

The procedures for importing the other data files as tables are similar to the procedure for importing Excel worksheets.

### ▶ Import the **Unacceptable\_Codes.txt** file

**Unacceptable\_Codes.txt** has two columns titled **Codes** and **Description**. The field entries are separated by tab characters.

It is best to give each field the correct data type when you import it into *ACL*. However, you can change the data type after you import the file.

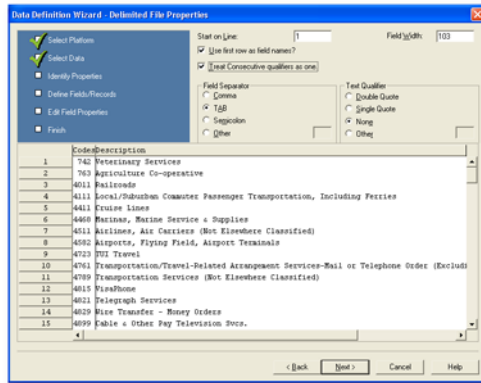
- 1 Select **File » New » Table**. Click **Next** until the **Select File to Define** dialog box appears.
- 2 Locate and double-click **Unacceptable\_Codes.txt**. Click **Next**.

- 3 On the **File Format** screen of the **Data Definition Wizard**, select **Delimited text file**, and click **Next**.

The **Delimited File Properties** screen of the **Data Definition Wizard** appears.

- 4 Select **TAB** in the **Field Separator** column. The data separates correctly into two fields.
- 5 Select **Use first row as field names**.

The first row of the file contains the field names, which are **Codes** and **Description** in this instance. You can see the format of the file as defined in the wizard.



- 6 Click **Next** and save the table as **Unacceptable\_Codes**.

The **Edit Field Properties** screen of the wizard appears. You can change the type of the data in each field. *ACL* has made assumptions about the data that are not necessarily correct.

- 7 On the **Edit Field Properties** screen of the wizard, the **Codes** field is selected by default. Select **ASCII Text** from the **Type** drop-down list.

Although the unacceptable codes are comprised of digits, you want *ACL* to treat these codes as character data, rather than numeric data.

- 8 Click the **Description** column header on the **Edit Field Properties** screen. It should have a **Type** of **ASCII Text**. You do not have to make any changes.

- 9 Click **Next**.

The wizard provides a summary of the file.

- 10 Click **Finish**. Click **OK** when you are prompted to save the table.

The **Unacceptable\_Codes** table appears in *ACL*. You can scroll through the contents.

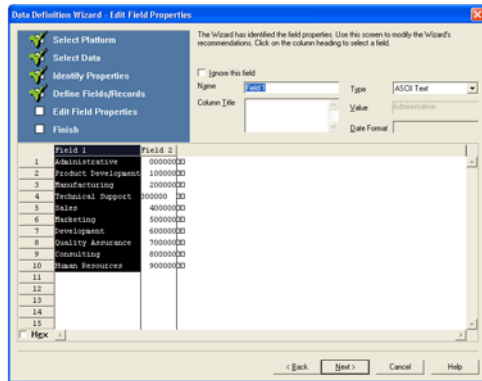
#### ► Import the **Company\_Departments.txt** file

This file is similar in format to the **Unacceptable\_Codes.txt** file, but it has no header row. You must add the field names in the **Data Definition Wizard**.

- 1 In the overview select the **ACL\_Tutorial\_Metaphor** project folder and right-click and select **New » Table** to open the **Data Definition Wizard**. Proceed through the wizard as you did for the previous file, but this time select **Company\_Departments.txt**.
- 2 Again, select **Delimited text file** on the **File Format** screen. Click **Next**.
- 3 On the **Delimited File Properties** screen, deselect **Use first row as field names** if necessary. Select **TAB** as the **Field Separator**.
- 4 Click **Next**. The **Save Data File As** dialog box appears.
- 5 Enter **Company\_Dept** as the **File name**. Click **Save**.



You can now edit the field properties.



- 6 On the **Edit Field Properties** screen, enter `Dept_Name` in the **Name** text box.
- 7 Click the **Field 2** column header. Enter `Dept_Code` in the **Name** field. From the **Type** drop-down list, select **ASCII Text**.
- 8 Click **Next**. Click **Finish**. Click **OK** to the message.

The **Company\_Dept** table opens in the view.

#### ► Import the Employees.csv file

- 1 Select **Data » External Data » Disk**. The **Select File to Define** window appears.
- 2 Double-click the **Employees.csv** file to open the **Data Definition Wizard**.
- 3 On the **File Format** screen, ensure that **ACL** imports the file as a **Delimited text file**. Click **Next**.
- 4 On the **Delimited File Properties** screen, ensure that:
  - **Comma** is selected as the **Field Separator**.
  - **Use first row as field names** is selected.

- 5 Click **Next**. Save the file as **Emp1oyees**.
- 6 On the **Edit Field Properties** screen:
  - Change the **CardNum** field and the **EmpNo** field to **ASCII Text** type.
  - Change the **HireDate** field to **Date** type. Specify the **Date Format**. If the file you define does not match the date format you will get blank dates in the view.
  - Ensure that **Salary** and **Bonus\_2002** are **Numeric (Formatted)** type.
  - You can change the column titles if you want more user-friendly titles. Changing the titles does not affect the field names.
- 7 Click **Next**. Save the table as **Emp1oyees**.
- 8 Click **Finish**. Click **OK** to the message.  
The **Employees** table opens in the view.

You next import the data from the **Acceptable\_Codes** database.

#### ► Import the Acceptable\_Codes Access database

- 1 Click **File » New » Table**. The **Data Definition Wizard** appears.  
If the welcome message also appears, select **Do not show this welcome screen again**, and click **Next**.
- 2 Select the **Acceptable\_Codes.mdb** file.
- 3 Continue to proceed through the wizard. Since this database includes only one table, **Acceptable\_Codes** is automatically selected in the **Select the Access Table/View** list.
- 4 Save the table as **Acceptable\_Codes**.

- 5 Click **Finish**. Click **OK** to the message. The **Acceptable\_Codes** table opens in the view.

As you can see in the **Overview** at the left of your project, the data from each file has been placed into separate tables. You should have six tables in your project:

- Acceptable\_Codes
- Company\_Dept
- Credit\_Cards\_Metaphor
- Employees
- Trans\_April
- Unacceptable\_Codes

### ► Save your project

- Click **File » Save Project**. You can also click the **Save the Open Project** button on the toolbar.

In order to work through this tutorial effectively, you must save your work as you move from one procedure to the next.

Just as with any Windows application, you should save your work in *ACL* every few minutes while you are working on a project. That way, you protect yourself from losing your work if there is a system crash. *ACL* automatically creates a backup copy of the working project as well. You should save your work when you add or remove data from your project.

Throughout the tutorial, you will be prompted to save changes to a table, particularly as you move from table to

table, or when you close the project. Click **Yes** to save the work you have done on each table.

You can stop following the tutorial at any point by closing and saving the project.

## Exercise

Save the existing project under a new name, **ACL\_Tutorial\_Metaphor\_Test**, using **File » Save As**. Continue working with the original project, **ACL\_Tutorial\_Metaphor**.

## ■ Check and verify the data

Before you proceed with analyzing your data, you should always verify it. For example, if a field is supposed to contain dates, but also contains names, your analysis may be meaningless. *ACL* has several commands that you can use to verify your data.

### ● ● ● ● ● **Tasks**

- Check the data type and format.
- Verify the data in a table.
- Document your work.

## Check the data type and format

Before you begin your analysis, check the table layout for each table to determine whether:

- The data types for each field are correct.
- The fields are properly defined.

The following table provides information about how the fields in the tables should look, and what data type *ACL* should assign to them.

The way the data should appear is a matter of preference. For example, you could decide that you wanted all card numbers to be formatted with hyphens after each four digits. In this project, the decision was to display a string sixteen characters long.

Files	Data type and formatting
Credit card numbers	Each entry should have 16 digits, without any extra characters. The field should have a data type of ASCII.
Dates	Each entry should be a date. Depending on the dates, you might want to determine if all the dates fall within a given time period. The field should have a data type of Date.
Transaction Amounts, Finance Charges, and Credit Limits	Each entry should be numeric, with two decimal places. The field should have a data type of Numeric.

Names	The entries should contain only letters, spaces, and hyphens. The field should have a data type of ASCII.
Merchant Category Codes	Each code should be 4 digits long, with no extraneous characters. The field should have a data type of ASCII.

In this tutorial, you assign the ASCII data type to Merchant Category Codes and credit card numbers because you will not perform calculations with either field. You also require an ASCII data type for work that you do later in the tutorial. If you were performing mathematical tests on the card numbers, you would need a data type of Numeric.

Ideally, in a thorough data analysis, you verify the data in every field in every table of your project. If you like, you can set up a chart with the appropriate tests for each field, and keep a check list of the results. Each verification command is also shown in the log.

However, if you are confident that your data has been entered properly, you can skip the verification steps or verify only a few fields.

Often, you can use *ACL* to fix incorrectly formatted data. You can create computed fields that omit improper characters, for example.

In this chapter, you will modify the default data types that *ACL* selected, but the underlying data does not require repair.

► View and edit data definitions in a table.

- 1 Open the **Credit\_Cards\_Metaphor** table.
- 2 Click CTRL+I. The **Table Layout** window appears.

This window lists information about each of the data fields as it is defined in the project. Look at **Category** and **Type** to determine how the data has been defined. You may need to scroll the list horizontally to see all of the information. You can widen and rearrange columns to make the list easier to read.

Name	Title	St.	Category	Length	Decimals
CARDNUM	CARDNUM	1	N	19	0
CREDLIM	CREDLIM	20	N	12	0
CUSTNO	CUSTNO	32	C	6	0
EXPDT	EXPDT	38	D	10	0
FINCHG	FINCHG	48	N	11	2
INDPRV...	INDPRV...	59	N	10	2
NRWBAL	NRWBAL	69	N	10	2
PACTAL...	PACTAL...	79	N	15	2
PMTDLEDT	PMTDLEDT	94	D	10	0
PREVBAL	PREVBAL	104	N	12	2
RATE	RATE	116	N	9	2

ASCII	1	2	3	4	5	6	7
	855012249765807	900026235308/01/2003	0.00	10.00	37.;		
	8550122201264011	800051246503/01/2002	14.68	30.00	329.;		
	8550120784894566	690005159310/01/2002	113.20	125.00	6489.;		
	8550124253621744	720025840204/01/2002	101.41	0.00	5891.;		
	8550125939743163	960077868802/01/2002	0.00	0.00	384.;		
	8550120714753180	860077868801/01/2003	0.50	20.79	85.;		
	8550128947747852	1000025840206/01/2002	0.00	1.31	1.;		

Field lists like this one categorize data types as Character, Numeric, Date, or Logical. The ASCII data type, for example, belongs to the Character category. To learn more about the categories of data types, select Help » Index and look up “data type reference”.

- 3 Look through the list of fields to determine whether each field has the correct data type.

The **CardNum** field should be character data, but the **N** indicates that it is numeric data. Although it is comprised of digits, you will not be performing mathematical operations with the card numbers.

- 4 Double-click the **CardNum** entry.
- 5 In the **Type** drop-down list, select **ASCII**.
- 6 Click the green checkmark to save the changes.
- 7 Close the **Table Layout** window.

## Verify the data in a table

Once you establish that the data types and the data are correctly defined, you next determine whether the data matches its assigned data type. Use the Verify command to check whether the data in a table matches its data type.

For example, if you have letters in a field with a numeric data type, or decimal numbers in a field with a date data type, the Verify command indicates data validity errors.

### ► Verify the data in a table

- 1 Open the **Credit\_Cards\_Metaphor** table.
- 2 Click **Data » Verify**.
- 3 Click **Verify Fields**. The **Selected Fields** dialog box appears.
- 4 Click **Add All**. Click **OK**.

You should see a list of the fields you want to verify in the **Verify Fields** list.

- 5 Click **OK**.

You should have zero data validity errors.

You have now verified all fields in the table.

## Document your work

When you work with an *ACL* project, you want the methods that you use in your work to be clear. You can then justify the results of your work, or allow someone else to take over the analysis tasks if necessary. With proper documentation, you can see at a glance what work you have done in *ACL*. This will make your job easier, particularly if you are responsible for multiple analysis projects.

If you are taking over a project from someone else, documentation is particularly valuable. Similarly, if you later assign the project to someone else, that person will appreciate being able to read historical project notes.

### ► Add and edit project notes

- 1 Select the project name in the **Overview**.
- 2 Click **File » Properties**. The **Project Properties** dialog box appears.
- 3 In the **Notes** tab, enter the current date and time, and a short summary of what you plan to accomplish.
- 4 Click **OK** to close the **Project Properties** dialog box.

When you close the dialog box, *ACL* automatically saves the notes. However, if you accidentally erase the notes, you cannot get them back unless you have a file backup. For this reason, it is recommended that you use both a task sheet and the project notes for documentation.

If you find it helpful, you can adapt the following task sheet to your project to keep track of what analysis you have already done, and what analysis you plan to do. This task sheet is prepared and maintained outside of *ACL*.

Items	Description
Information you need	
Where to get this information	
Project folder location	
Related projects	
Verification (create a checklist for each verified field)	
Table names with extracted, validated data	
Description of each analytical task in <i>ACL</i>	
Analytical task #1 to <i>n</i> ... ( <i>ACL</i> steps)	
Report for each objective	
Reporting task #1 to <i>n</i> ... ( <i>ACL</i> steps)	
Location and description of exported data files	

## Add record notes to a table

You can add notes to specific records in a table. These notes can contain any information that might be helpful to someone viewing the project, and you can also link to other files containing information.

- 1 Open the **Employees** table. Note the numeric cells that indicate the record number at the left of the table.
- 2 Right-click record cell 3 and select **Edit Note**. The **Edit Note** dialog box appears.
- 3 Enter this note: This employee is leaving Metaphor Corporation by the end of the year.
- 4 Click **OK**.

Any notes you add to a table do not affect the table data.

## Exercise

Verify the data for at least one other table in the `ACL_Tutorial_Metaphor` project.

## ■ Chapter review

In this chapter you accomplished the following tasks:

- Creating a project.
- Importing files into your project.
- Saving your project.
- Verifying the data in your project.
- Documenting your work.

You have not actually done any analysis of the data yet, but you have set the project up so that it is ready.



## BEGIN YOUR ANALYSIS

*Start looking for patterns*

**In this chapter...**

- Determine total liability
- Extract expiring card numbers
- Find transactions for a customer
- Check credit card balances

So far, you looked at a project plan, worked on a simple project, started a new project, added data to the project, and verified the data. You now use *ACL* to perform simple analysis of your data.

In this chapter, you use simple procedures to obtain general information about your data. You look at the credit limits and the credit card balances to get information about Metaphor's financial liability and credit card usage patterns.

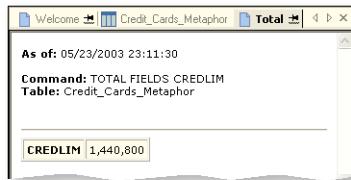
## ■ Determine the total liability

If every Metaphor employee reached the credit limit on his or her credit card, what would be the total debt for Metaphor? You can determine this figure by totalling the **CREDLIM** field.

You can use **Total** to add the total credit limits together to determine the company's total credit exposure.

### ► Total the CREDLIM field

- 1 Open the **Credit\_Cards\_Metaphor** table.
- 2 Select **Analyze » Total Fields**. The **Total** dialog box appears.
- 3 Select the **CREDLIM** field. You only want to total this field, so there is no need to select any other fields.
- 4 Click **OK**. View the **Total** tab.



What is your result? Your total should be 1,440,800.

You can export the results to another file format if you want to save the result. You can also view the results in the log if you come back to the project later.

## ■ Extract expiring card numbers

You plan to coordinate with the accounting department to ensure that there is no problem with the employees getting new credit cards. First, you filter all credit card numbers that have an expiry date of January 1, 2004 or earlier.

### ► Filter card numbers by expiry date

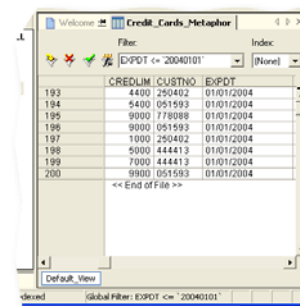
- 1 On the **Credit\_Card\_Metaphor** tab, click the **Edit View Filter** button. The **Edit View Filter** dialog box appears.

- 2 In the **Expression** box, enter this filter:

```
EXPDT <= `20040101`
```

If you prefer not to enter the expression, you can double-click the field in the **Available Fields** list, click the buttons for the expression symbols, and click the **Date** button to select the date.

- 3 Click **OK**.
- 4 Select **Analyze » Count Records**. Click **OK**. The status bar shows that 8/200 records fit the criteria of having an expiry date of January 1, 2004 or earlier.



You want to send the soon-to-expire credit card numbers that appear in the view to the clerk in charge of credit cards. In the next procedure, you export the credit card numbers to a text file.

► **Export soon-to-expire credit card numbers**

- 1 Click **Data » Export to Other Application**. The **Export** dialog box appears.
- 2 Select **View**. From the **Export As** drop-down list, select **Text**.
- 3 Enter a descriptive filename in the **To** text box, such as `CCN_To_Expire_by_January1.txt`. Click **OK**.

You can then attach the text file to an e-mail, print it out, or manage it in whatever way you want.

## ■ Find transactions by customer number

Because you have been warned about any expenditures related to customer number 444413, you want to find some information about the transactions associated with that customer number.

You can accomplish this task in at least two different ways:

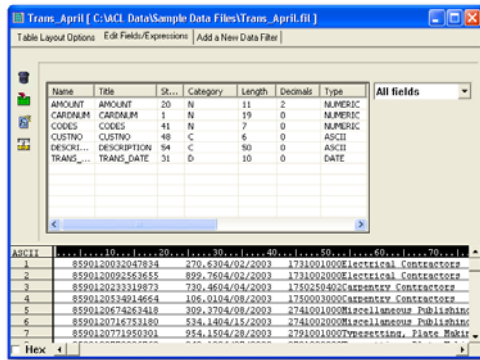
- Create a filter using the `FIND()` function that looks for all records with a **CUSTNO** value of 444413. You can use any of *ACL*'s functions to build a filter or computed field. Use `FIND()` to isolate specific records.
- Create a filter that sets **CUSTNO** equal to 444413.

You use functions to build either a filter or an expression for a computed field. You can see the functions in the **Expression Builder**.

Before you can do the search, however, check the data types assigned to the fields in the **Trans\_April** table. Because you imported the data from Excel, *ACL* made default assumptions about the data types for the fields. You might need to change some of the data types.

► **Edit the Trans\_April data types**

- 1 Open the **Trans\_April** table.
- 2 Click **Edit » Table Layout**. The **Table Layout** window appears.



- The **AMOUNT** field has the correct data type of **NUMERIC**.
- The **CARDNUM** and **CODES** fields should be edited to have a data type of **ASCII**.
- The **CUSTNO** and **DESCRIPTION** fields have the correct data type of **ASCII**.
- The **TRANS\_DATE** field has the correct data type of **DATE**.

3 Double-click **CARDNUM**. From the **Type** drop-down list, select **ASCII**. Click the green checkmark. Repeat for **CODES**.

4 Close the **Table Layout** window.

Now that you have checked and corrected the data types, you can proceed to use the **FIND( )** function.

► Use the **FIND( )** function to locate customer records

1 With the **Trans\_April** table active, click the **Edit View Filter** button. The **Edit View Filter** dialog box appears.

2 In the **Expression** box, enter this expression:

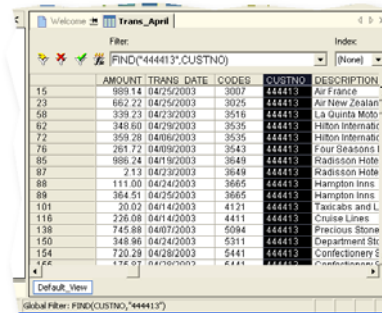
```
FIND("444413", CUSTNO)
```

Quotation marks are used around 444413 because the **CUSTNO** field has a field type of **ASCII**.

3 Click the **Verify** button. A message appears that indicates that the expression is valid. Click **OK**. If the expression is not valid, correct it and verify it again. You do not have to verify the expressions you create, but doing so is recommended.

4 Click **OK**. Scroll to the right to see that the **CUSTNO** field only contains values of 444413.




5 Select **Analyze » Count Records**. Click **OK**. 26/281 records have a **CUSTNO** value of 444413. In a typical analysis, you would likely try to find out who this customer is. However, the data files that come with the tutorial do not contain this information.



You can also use **FIND( )** to search for substrings within a field. For example, the filter **FIND("00", CUSTNO)** would

isolate all instances where 00 appears in the customer number field, whether at the beginning, middle, or end.

► **Setting a filter directly**

- 1 With the **Trans\_April** table active, click the **Remove Filter**  button to remove the filter.
- 2 Click the **Edit View Filter**  button. The **Edit View Filter** dialog box appears. Enter the expression:  
CUSTNO = "444413"
- 3 Click **OK**.  
You get exactly the same results with this filter as you did using the **FIND( )** function.
- 4 Click the **Remove Filter**  button to remove the filter.

When you apply a filter to data, and you want to save the results, you have several options. You can:

- Extract the filtered records to another table.
- Create a report containing the filtered records.
- Reapply the filter to the original data whenever you want to look at it again.

As long as you have the original data, you can always reapply the filter to see the filtered data again.


*ACL* typically provides multiple methods for performing any given task. If you want to be sure that your reasoning is correct, you can perform a task in more than one way and check to see if you get the same result both times.

## ■ Check the credit card balances

You might get a general overview of your data to alert you to possible areas of concern. For example, you might want to find the highest balances, or find balances over a certain limit. You can use this information as an indicator of potential problems with employee credit card usage.

One way to determine how many credit card users have high balances would be to filter the records to show those with a **NEWBAL** value greater than 2000 and a **NEWBAL** value greater than 1000. You can also sort the balances to see the balances in order of size.

► **Filter the NEWBAL field to find balances that exceed \$1000 or \$2000**

- 1 Open the **Credit\_Cards\_Metaphor** table. Click **Yes** if *ACL* prompts you to save the changes to the **Trans\_April** table.
- 2 Right-click the **NEWBAL** column header. You may need to scroll horizontally to locate the column. Select **Quick Sort Descending**.  
*ACL* lists the balances in descending order, with the top value first. By scrolling through the table, you can get an idea of what the various balances are.
- 3 Click the **Edit View Filter**  button. The **Edit View Filter** dialog box opens.
- 4 Enter this expression in the **Expression** box:  
NEWBAL > 2000
- 5 Click **OK**.

The balances greater than \$2000 are listed in descending order.

	FINCHG	MINPYMDUE	NEWBAL	PASTDUEAMT	PMTI
1	174.76	423.00	10221.00	209.00	05/11
2	73.00	191.00	6871.38	0.00	05/11
3	124.95	100.00	5833.98	0.00	04/20
4	140.34	445.00	7550.24	288.00	05/01
5	110.30	95.00	7045.71	0.00	05/11
6	109.60	41.00	7020.70	0.00	05/01
7	122.00	127.00	6378.02	0.00	05/11
8	95.47	33.00	6588.59	0.00	05/11
9	124.02	233.39	6463.56	98.38	05/01
10	113.20	126.00	6400.12	0.00	05/01
11	88.74	0.00	6264.42	0.00	05/01
12	101.41	0.00	5891.59	0.00	05/11
13	0.00	717.00	5792.57	598.00	05/01
14	0.00	116.00	5543.37	0.00	05/01
15	79.04	0.00	5510.02	0.00	05/01
16	78.31	0.00	5044.57	0.00	05/11

- 6 Select **Analyze » Count Records**. Click **OK**.

The status bar shows 53/200 records that have a **NEWBAL** amount greater than \$2000.

Next, you change the filter to look at records with a **NEWBAL** amount greater than \$1000.

- 7 In the **Filter** field at the top of the view, change the filter to **NEWBAL > 1000** and press **ENTER**.

- 8 Select **Analyze » Count Records**. Click **OK**.

You should have 65 records that fit the criteria.

You can have more than one condition in your filter. In addition to the transaction amounts, the amounts in the **PASTDUEAMT** and **FINCHG** fields might indicate a problem with credit card usage. According to the **Policies.doc** document, Metaphor credit card users must pay off their balances every month.

- 9 Change the filter to:

(NEWBAL > 1000) OR (PASTDUEAMT > 0) OR (FINCHG > 0)

Filters used in this session can be recalled from the drop-down list in the **Filter** field.

- 10 Count the filtered records. You should have 89.

If you want, you can extract the filtered records to a new table, or print a report.

You next decide to try another Quick Sort to see credit cards with past due amounts.

► **Quick Sort on PASTDUEAMT**

- 1 Click the **Remove Filter** button to remove the filter.
- 2 Right-click on **PASTDUEAMT** and select **Quick Sort Descending**.
- 3 Scroll through the records. Notice that many credit card holders have past due amount charges.

These procedures are minor examples of how *ACL* can analyze a large amount of data. As you gain more facility with *ACL*, you will undoubtedly think of new ways to analyze your data.

## ■ Exercises

- 1 Extract the records with a **NEWBAL** value greater than 1000 to a new table for later analysis.
- 2 What is the difference between the total number of transactions and the number of transactions with a **NEWBAL** value greater than 2000?
- 3 Of the credit card numbers that have a **NEWBAL** value greater than 2000, how many have a **PASTDUEAMT** greater than zero?
- 4 Check for card numbers with a zero **NEWBAL** value. Export these to a text file.
- 5 Check the **Credit\_Card\_Metaphor** table for all card numbers with a negative **NEWBAL** value. Export these to an Access file.
- 6 Under the terms of the credit card program, Metaphor can avoid finance charges if employees submit their receipts on time to the accounting department. Metaphor wants to determine how much money it loses unnecessarily from finance charges. Total the **FINCHG** field. Also, how many credit card holders have a **FINCHG** value that is greater than \$10.00?
- 7 Metaphor wants to know the total of all current balances. In the **Credit\_Card\_Metaphor** table, total the **NEWBAL** field.
- 8 Metaphor is considering reducing its upper credit limit. How many credit card numbers have credit limits exceeding \$10,000?

## ■ Chapter review

In this chapter, you:

- Calculated Metaphor's total credit card financial liability.
- Found the credit card numbers that are due to expire soon and extracted them to a text file.
- Found transactions linked to a particular customer.
- Checked the **NEWBAL** field to determine how many transactions exceeded particular limits.





## EXAMINE EXPENSE PATTERNS

*Determine information about expenses*

**In this chapter...**

Find the total expenses for each category code

Determine specific transaction information

In the previous chapters, you worked with *ACL* to gain familiarity with your data and you performed some preliminary analysis of the data. Now the accounting department at Metaphor Corporation wants you to provide information about spending patterns in the April credit card transactions. If you can develop a useful analysis plan, you will be asked to perform analyses for other months.

Once you provide information about the spending patterns, Metaphor might be able to economize by choosing a single vendor in some cases, or by obtaining bulk discounts, or by further limiting the acceptable categories of transactions. You will also look at expenses on a per-customer basis so that Metaphor can use the results to determine whether the money is being spent effectively.

## Find the total expenses for each category code

You begin your analysis by summarizing the total amount of all the transactions for each category code. Your plan is to look at these results to get ideas for further analysis.

### Tasks

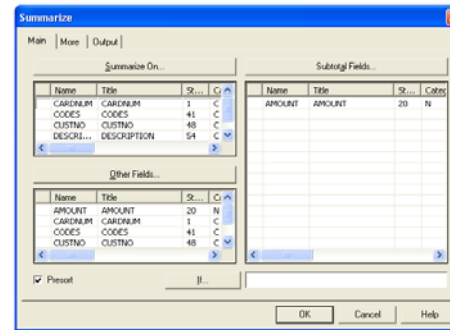
- Summarize the total expenses for each category code
- Isolate transactions for airfare, hotel stays, and rental cars
- Isolate particular recreational expenses

## Summarize the total expenses for each category

By summarizing the transaction amounts for each merchant category code, you can see how much money was spent on specific airlines and specific hotel chains. All restaurant expenses, on the other hand, will be categorized under a single code. Some Merchant Category Codes are specific, while others are broad.

### Summarize transaction amounts by category code

- 1 Open the **Trans\_April** table.
- 2 Select **Analyze » Summarize**. The **Summarize** dialog box appears.



- 3 In the **Summarize On** list, select **CODES**.
- 4 In the **Subtotal Fields** list, select **AMOUNT**. You want to total the amount spent on each category code.
- 5 In the **Other Fields** list, select **DESCRIPTION**. You want to see what the codes mean in the new table.
- 6 On the **Output** tab, select **File**. In the **Name** text box, enter **Trans\_by\_Code** and click **OK**.

The summarized results are extracted to the **Trans\_by\_Code** table.

View the new table. Note that the **COUNT** field indicates the number of transactions per code. You can do a **Quick Sort Descending** on the **AMOUNT** field to see where the most money was spent. You can also use this table for further analysis.

CODES	AMOUNT	COUNT	DESCRIPTION
1731	1170.39	2	Electrical Contractors
1750	836.47	2	Carpentry Contractors
2741	843.91	2	Miscellaneous Publishing and Printi
2791	1098.25	2	Typesetting, Photo Making, & Related
3000	1398.67	2	United Airlines
3001	965.49	2	American Airlines
3005	672.93	2	British Airways
3007	1695.06	2	Air France
3008	2197.87	3	Lufthansa
3009	2292.62	3	Air Canada
3025	2238.50	3	Air New Zealand
3032	2004.54	3	EI Al
3058	2084.27	3	Delta
3080	1483.01	2	Northwest
3083	740.04	2	Us Air
3086	2455.74	7	Southwest
3085	589.77	1	Midwest Express Airlines, Inc
3256	365.07	1	Alaska Airlines
3261	539.92	1	Air China
5145	794.34	1	American Intl Bank&Finar

This summary table is useful for showing the amounts spent on airlines, hotel chains, rental car agencies, and other categories. However, it does not show you the total amount spent on each of airfare, hotels, and rental cars, nor does it tell you the total amount spent on travel.

Often, when you first perform an analytical task, you do not get precisely the results you want. However, you can use those results to help you find a way to achieve more focused results.

## Isolate transactions by category

Metaphor wants you to look closely at the total amount spent on out-of-town stays in April, specifically on airfare, accommodation, and rental cars.

Hotels, motels, rental cars, and airfare fall exclusively in the code range 3000-3750. All transactions with these codes should relate to travel by Metaphor employees.

You can isolate these expenses by filtering and displaying all transactions with codes between and including 3000 to 3750.


One way to isolate this information is by using the **BETWEEN()** function on the **CODES** field.

**BETWEEN()** works on numeric, date, and character data, and it takes the following form:

```
BETWEEN( value, min, max)
```

You will use the **BETWEEN()** function to capture all transactions with codes in the range 3000 to 3750. In the **Trans\_April** table, the **CODES** field has extra spaces, so you have to use the **ALLTRIM()** function to trim the **CODES** field.

### ► Use the **BETWEEN()** function to find transactions in a given range

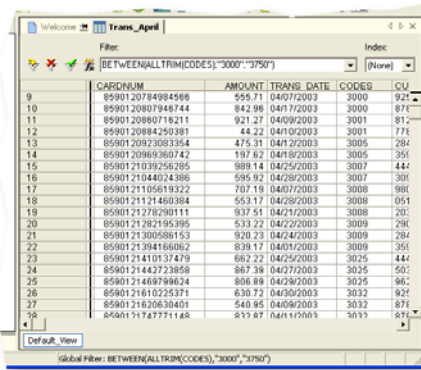
- 1 Open the **Trans\_April** table.
- 2 Click the **Edit View Filter**  button. The **Edit View Filter** dialog box appears.
- 3 Enter the following filter in the **Expression** box:  

```
BETWEEN(ALLTRIM(CODES), "3000", "3750")
```
- 4 Click **Verify**. A message box indicates that the expression is valid. Click **OK**.

## Chapter 5: Examine expense patterns

If the expression is not valid, look at it carefully again, particularly the quotation marks, and make the necessary changes.

- Click **OK**. The **Edit View Filter** dialog box closes. You now have a filtered view of the transactions that displays transactions with Merchant Category Codes between 3000 and 3750.



	CARDNUM	AMOUNT	TRANS_DATE	CODES	CU
9	0590120784984566	555.71	04/07/2003	3000	921
10	0590120807948744	842.96	04/17/2003	3000	871
11	059012080718211	921.27	04/09/2003	3001	812
12	0590120894250381	44.22	04/10/2003	3001	771
13	0590120923083354	475.31	04/12/2003	3005	284
14	0590120868360742	197.62	04/18/2003	3005	351
15	0590121039269295	989.14	04/25/2003	3007	444
16	0590121044074386	595.92	04/28/2003	3007	301
17	0590121105619322	707.19	04/07/2003	3008	981
18	059012117480384	553.17	04/29/2003	3008	051
19	0590121278290111	937.55	04/21/2003	3008	202
20	0590121282195395	533.22	04/22/2003	3009	291
21	0590121300568153	920.23	04/24/2003	3009	284
22	0590121384165882	939.17	04/01/2003	3009	351
23	0590121410137479	662.22	04/25/2003	3025	444
24	0590121442723858	867.38	04/27/2003	3025	501
25	0590121488788634	808.88	04/29/2003	3025	961
26	0590121610225371	830.23	04/09/2003	3032	852
27	0590121620630401	540.95	04/09/2003	3032	871
28	0590121747771148	873.87	04/11/2003	3032	871

- Select **Analyze » Count Records**. Click **OK**. How many records fit this criteria?
- Select **Analyze » Total Fields**. The **Total** dialog box appears. Select the **AMOUNT** field. Click **OK**.  
ACL displays the total of all transactions for airfare, accommodation, and rental cars. You should have a total of 47,248.27.  
To facilitate further analysis of these transactions, you can extract the filtered transactions to a new table.
- Select **Data » Extract Data**. The **Extract** dialog box appears.
- Ensure **Record** is selected. Enter **Trans\_April\_Travel** in the **To** field. Click **OK**.

You can now perform further analyses of these filtered transactions. For example, you can determine the total spent on airfare alone. Because the Merchant Category Codes for airfare are between 3000 and 3299, you can use the **BETWEEN()** function to find values between 3000 and 3299.


## Isolate particular recreational expenses

So far you have looked at travel-related expenses for April. Metaphor now wants you to look at the total expenses related to golf and drinking establishments. You need to isolate those transactions with a Merchant Category Code that is related to golf or drinking establishments.

The MATCH() function provides a way to apply multiple filters to a table. The filters must all relate to one field, which is CODES in this example. You could achieve the same results by using a filter with multiple OR statements, but the filter would be more difficult to set up and understand at a glance.

Again, you use ALLTRIM() to trim the codes before comparison.

### ► Use the MATCH() function to isolate recreational expenses

- 1 Open **Trans\_April**.
- 2 Click the **Edit View Filter**  button. The **Edit View Filter** dialog box appears.
- 3 In the **Expression** box, enter the following expression:
 

```
MATCH(ALLTRIM(CODES), "5813", "7992")
```

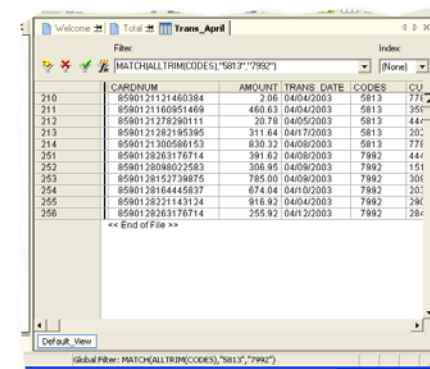
Note that the syntax for the MATCH() function is:

```
MATCH (comparison_value, test1, test2
<, test3 ...> )
```

5813 is the code for Drinking Establishments.  
7992 is the code for Public Golf Courses.

You could add more MATCH() parameters to search for more codes if you wanted.
- 4 Click **OK**.

- 5 Count the records. The status bar indicates that 11/281 records fit the MATCH() parameters.



	CARDNUM	AMOUNT	TRANS_DATE	CODES	CO
210	85901211440384	2.06	04/04/2003	5813	71
211	8590121160951469	460.63	04/04/2003	5813	350
212	8590121278290111	20.78	04/05/2003	5813	44
213	8590121282195395	311.84	04/17/2003	5813	20
214	8590121300586153	830.32	04/06/2003	5813	77
251	8590128263176714	391.62	04/08/2003	7992	44
252	8590128098022583	306.85	04/09/2003	7992	151
253	8590128152739875	785.00	04/08/2003	7992	38
254	8590128164445837	674.04	04/10/2003	7992	20
255	859012822143124	916.92	04/04/2003	7992	29
256	8590128263176714	255.92	04/12/2003	7992	28

You can use the Total command to find the total spending on these two categories. You could apply further filtering. You could also extract the transactions to a new table.

## Exercises

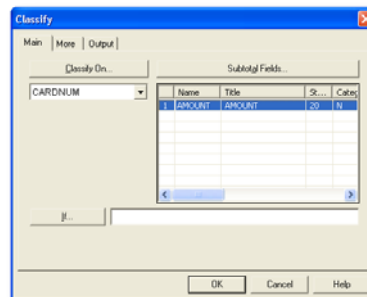
- 1 Total the expenses for airfare, hotel stays, and rental cars separately.
- 2 Total the expenses for all British Airway flights.
- 3 Total food-related expenses. How much money did the company spend on restaurants, fast-food restaurants, and catering, both separately and together?
- 4 Use MATCH() to find transactions billed to internal departments. Internal departments have customer numbers that begin with zero.
- 5 Determine what percentage of the total transaction amounts is billed to internal departments, and what percentage is billed to external clients.

## List the amounts for each card number

Now that you have a complete list of transactions, you want to determine the total transaction amounts for each credit card number for the month of April. You can get this information by using the Classify command.

### ► Use the Classify command

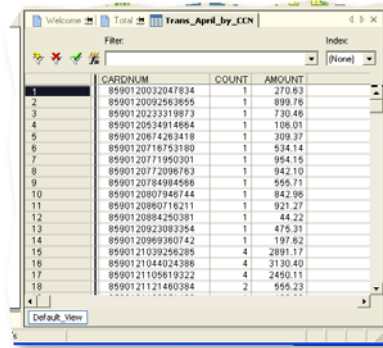
- 1 Open the **Trans\_April** table. Remove any filters, if applicable.
- 2 Select **Analyze » Classify**.
- 3 Select **CARDNUM** from the **Classify On** list. Click **AMOUNT** in the **Subtotal Fields** list.



- 4 On the **Output** tab, select **File**. Enter **Trans\_April\_by\_CCN** in the **Name** field. Click **OK**. View the new table.

The **Trans\_April\_by\_CCN** table lists the accumulated transaction amounts per card number. The **COUNT** field lists the number of transactions per card number.

- 5 Right-click the **Percent of Count** column and select **Remove Selected Columns**. Do the same for the **Percent of Field** column.



	CARDBNUM	COUNT	AMOUNT
1	8590120032047834	1	270.83
2	8590120092563855	1	899.76
3	8590120223310873	1	730.46
4	8590120954914064	1	108.01
5	8590120674263418	1	309.37
6	8590120716753180	1	534.14
7	8590120771950301	1	954.15
8	8590120772098763	1	942.10
9	8590120784984568	1	555.71
10	8590120807946744	1	842.96
11	8590120860716211	1	921.27
12	8590120884250381	1	64.22
13	8590120923083354	1	475.31
14	8590120969360742	1	197.62
15	8590121039256289	4	2891.17
16	8590121044024386	4	3130.40
17	8590121105819322	4	2450.11
18	8590121121460384	2	555.23

- 6 Close the view and click **Yes** when *ACL* prompts you to save.

## Exercise

Use the **Classify** command to create a table that lists the total transactions on a per-customer-number basis.

## ■ Chapter review

In this chapter you accomplished the following tasks:

- Determined the total transaction amounts for each merchant category code.
- Determined transaction amounts for specific expense categories.
- Determined the total transaction amounts per credit card number.



## ANALYZE TRANSACTIONS

*Identify unacceptable transactions*

**In this chapter...**

Determine which transactions are unacceptable

Identify the employees who made unacceptable transactions

Report your findings

So far you have performed a preliminary investigation and analysis of the Metaphor data in your *ACL* projects. You now determine whether Metaphor employees are using their company credit cards to make unapproved purchases, and whether they are using their company credit cards to buy specifically forbidden items and services.

## ■ Determine which transactions are unacceptable

Metaphor Corporation has created three categories of transaction codes, based on the Merchant Category Codes list. These include:

- Acceptable
- Unacceptable
- Conditional (requires prior approval)

One table lists the acceptable codes and one table lists the unacceptable codes. Codes that are not listed in either table are conditional.

To find all unacceptable transactions, proceed through the following tasks.

### ● Tasks

● Check the code tables for obvious errors.

● Identify transactions that have unacceptable codes.

Identify the employees who made the unacceptable transactions.

Create a report of your results.

## About the Join and Relations commands

When your project includes multiple data tables, you often want to find whether particular field values exist in more than one table. Examples:

- If you manage a travel agency, you might have a list of customers who purchased airfare and a list of customers who purchased hotel rooms. You can determine how many customers purchased both airfare and hotel rooms from your agency.
- If you work in marketing at a department store, you might have a list of customers who purchased over \$2000 last year at your store and a list of customers who obtained refunds of over \$1000. You can determine which customers purchased over \$2000 last year and also obtained a refund of over \$1000.

You can use the Join and Relations commands to compare tables and to combine data from tables. With these commands, you can:

- Find field values that exist in one table, but not in another.
- Find field values that exist in both tables.
- Combine specified fields from two source tables.

The Join command creates a new table that contains data from both of the original tables. The resulting table looks like any other data table in your project.

The Relations command does not physically combine the tables, but creates a virtual join among two or more tables. You can use the fields from the related tables as if all of the fields belonged to the same table. You can create expressions that use fields from any of the related tables. You can add fields from one related table to the view of

another related table so that you can see fields from both tables in a single view. If you want, you can then extract any or all of the fields from the related tables to a new table.

Because the Join command creates a new table based on the two original tables, you generally use the Relations command instead of Join if the original tables are very large.

To learn more about Join and Relations, select Help » Index and look up “combining tables”.

## Check the code tables for obvious errors

Before you can use the code tables to check the transactions, check the tables for obvious errors. You will use the Join command to determine whether the **Acceptable\_Codes** table and the **Unacceptable\_Codes** table have any codes in common.

Before you use the Join command, you must ensure that the key fields of both tables are character fields that are the same length.

### ► Compare the length and type of the key fields

- 1 Open the **Unacceptable\_Codes** table.
- 2 Select **Edit » Table Layout**. The **Table Layout** window appears.

The list of fields shows that the data type of the **Codes** field is **ASCII**. The field length is equal to either the length of the field name, or the length of the longest value in the field, whichever is longer.

- 3 Close the **Table Layout** window.
- 4 Open the **Acceptable\_Codes** table.
- 5 Select **Edit » Table Layout**. The **Table Layout** window appears.


The list of fields shows that the data type of the **Codes** field is **ASCII**, but the field length is 50.

The Merchant Category Codes are each four characters long. Each of these **Codes** fields is padded with extra spaces. You must harmonize the two fields by creating computed code fields in each table. These new fields will:

- Contain the four digits of the **Codes** field as character data.
- Have a field length of 4.

You can then perform the join successfully by using the computed fields as the key fields.

► **Create a computed field in each of the `Acceptable_Codes` and `Unacceptable_Codes` tables**

- 1 Open the `Acceptable_Codes` table, if not already open.
- 2 Open the **Table Layout** window, and click the **Add a New Expression**  button.

- 3 In the **Name** text box, enter `Codes_Sub_Acceptable`.

- 4 Click **f(x)** to open the **Expression Builder**. In the **Expression** text box, enter:

```
SUBSTR(ALLTRIM(Codes), 1, 4)
```

- The `ALLTRIM()` function trims any extraneous spaces that occur on the left or the right of the `Codes` field.
- The `SUBSTR()` function reads the first four characters of the `Codes` field.

- 5 Click the green checkmark. The new field, `Codes_Sub_Acceptable`, appears in the **Table Layout** window. Verify that the category is C for character and that the field length is 4.

- 6 Close the **Table Layout** window.

You can now add the `Codes_Sub_Acceptable` field to the view.

- 7 Right-click the `Codes` column header. Click **Add Columns**. The **Add Columns** dialog box appears. Double-click `Codes_Sub_Acceptable` and click **OK**.

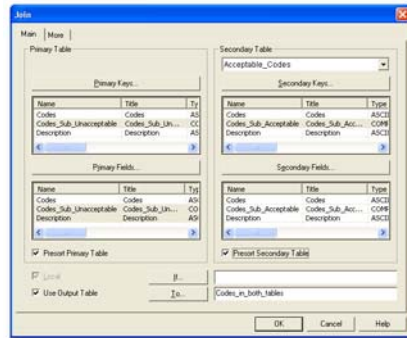
The `Codes_Sub_Acceptable` column appears in the view.

- 8 Close the view and click **Yes** when *ACL* prompts you to save.
- 9 Open the `Unacceptable_Codes` table. Repeat steps 2–7, but call the computed field `Codes_Sub_Unacceptable`.

Now that the key fields are harmonized, you can join the tables by using the `Codes_Sub_Unacceptable` field from the `Unacceptable_Codes` table and the `Codes_Sub_Acceptable` field from the `Acceptable_Codes` table.

► **Determine whether the tables have any codes in common**

- 1 Open the `Unacceptable_Codes` table.
- 2 Select **Data** » **Join Tables**. The **Join** dialog box appears.
- 3 In the **Secondary Table** list, select `Acceptable_Codes`.
- 4 Select `Codes_Sub_Unacceptable` as the **Primary Key**. Select `Codes_Sub_Acceptable` as the **Secondary Key**.
- 5 Select `Codes_Sub_Unacceptable` and **Description** in the **Primary Fields** list. If there are any codes that exist in both tables, you will be able to see their description.
- 6 Select **Presort Secondary Table**.
- 7 Enter `Codes_in_both_tables` in the **To** field. This is the name of the new table that will be created.



- 8 On the **More** tab, select **Matched Primary Records**, and ensure that no other **Join Categories** options are selected. Click **OK**.
- 9 ACL displays the new table. It should be empty. Although you can see that the table is empty, you should still count it in order to record the results in the command log and leave an audit trail.
- 10 Select **Analyze » Count Records**. Click **OK**.

Because there are no records in the **Codes\_in\_both\_tables** table, you know that no codes exist in both the **Unacceptable\_Codes** and **Acceptable\_Codes** tables. The original code tables have no obvious errors, so you can use them for your analysis.

## Identify unacceptable transactions

The **Trans\_April** table contains these fields:

- CARDNUM
- AMOUNT
- TRANS\_DATE
- CODES
- CUSTNO
- DESCRIPTION

The **Unacceptable\_Codes** table contains the following fields:

- Codes
- Description
- Codes\_Sub\_Unacceptable (a computed field)

To identify unacceptable transactions, you need to relate the **Trans\_April** table with the **Unacceptable\_Codes** table. However, you must first ensure that the key fields for each table both have a data type of ASCII and that they both have the same lengths.

In the previous section, you created a computed field called **Codes\_Sub\_Unacceptable** in the **Unacceptable\_Codes** table, and you know that the length of the **Codes\_Sub\_Unacceptable** field is 4 and that the category is C for character. You need to check the **CODES** field in the **Trans\_April** table to check its data type and length.


- ▶ **View the table layout for Trans\_April**
  - 1 Open the **Trans\_April** table.
  - 2 Select **Edit » Table Layout**.

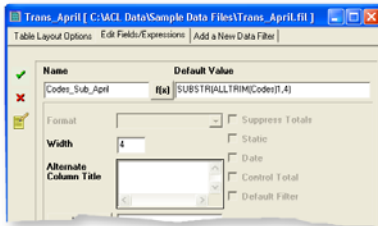
## Chapter 6: Analyze transactions

The list of fields shows that the **CODES** field has a category of **C**, but it has a field length of 7.

You must create a computed field in the **Trans\_April** table that trims the **CODES** field and has a **Length** of 4. You can then use this new computed field as the key field for relating **Unacceptable\_Codes** and **Trans\_April**.

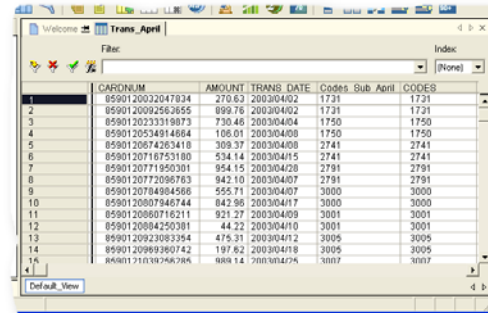
### ► Create a computed field in Trans\_April

- 1 In the **Table Layout** window, click the **Add a New Expression**  button.
- 2 Enter **Codes\_Sub\_April** as the name.
- 3 Click **f(x)**. The **Expression Builder** appears.
- 4 In the **Expression** box, enter:  
`SUBSTR(ALLTRIM(CODES),1,4)`
- 5 Click **OK**.



- 6 Click the green checkmark. Close the **Table Layout** window.

- 7 Right-click the **CODES** column header. Click **Add Columns**. The **Add Columns** dialog box appears. Double-click **Codes\_Sub\_April**. Click **OK**.



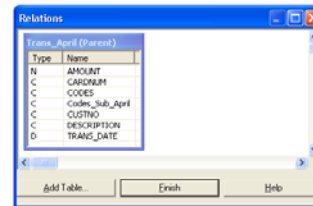
	CARONUM	AMOUNT	TRANS_DATE	Codes_Sub_April	CODES
1	8590120022047834	270.63	2003/04/02	1731	1731
2	8590120092562655	999.76	2003/04/02	1731	1731
3	8590120223319873	730.46	2003/04/04	1750	1750
4	8590120534914664	106.01	2003/04/08	1750	1750
5	8590120674263418	309.37	2003/04/08	2741	2741
6	8590120716753180	534.14	2003/04/15	2741	2741
7	8590120711959381	954.15	2003/04/20	2791	2791
8	8590120772068763	942.10	2003/04/07	2791	2791
9	8590120784984566	555.71	2003/04/07	3000	3000
10	8590120807946744	842.96	2003/04/17	3000	3000
11	85901208680716211	921.27	2003/04/09	3001	3001
12	8590120884250381	44.22	2003/04/10	3001	3001
13	8590120923083394	476.31	2003/04/12	3005	3005
14	8590120969260742	197.62	2003/04/18	3005	3005
15	8468121038746786	988.14	2003/04/24	3007	3007

Next, determine whether any of the April transactions have unacceptable codes. Your first step is to isolate unacceptable transactions by comparing the transaction codes in both tables.

### ► Use the Relations command to find unacceptable transactions

- 1 With the **Trans\_April** active, select **Data » Relate Tables**. The **Relations** dialog box appears.

If the **Getting Started** message also appears, select **Do not show this dialog again** and click **OK**.

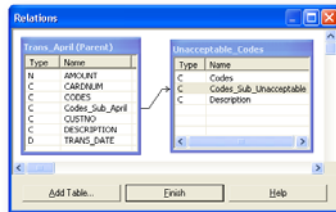


- 2 Click **Add Table**.

- Double-click **Unacceptable\_Codes**. The **Unacceptable\_Codes** table appears in the **Relations** dialog box.

To arrange the tables within the **Relations** dialog box, drag the tables with your mouse or right-click and select **Arrange Tables**.

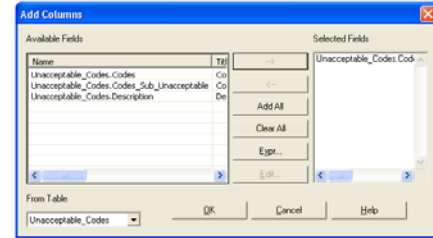
- Click and drag **Codes\_Sub\_April** from the **Trans\_April** table to **Codes\_Sub\_Unacceptable** in the **Unacceptable\_Codes** table.



If you link the wrong fields, right-click the line and select **Edit Relation**. Select the fields that you want to link, and click **OK**.

- Click **Finish**.  
Next, you will add the **Codes\_Sub\_Unacceptable** field from the **Unacceptable\_Codes** table to the view.
- Right-click the **Codes\_Sub\_April** column header. Click **Add Columns**. The **Add Columns** dialog box appears.
- In the **From Table** drop-down list, select **Unacceptable Codes**. The fields in the **Unacceptable\_Codes** table appear in the **Available Fields** list.

- Expand the **Name** column so that you can see the field names. Double-click **Unacceptable\_Codes.Codes\_Sub\_Unacceptable**.



- Click **OK**.

Now that you have the **Codes\_Sub\_Unacceptable** field from the **Unacceptable\_Codes** table in the view, the records without unacceptable codes display blanks in the **Codes\_Sub\_Unacceptable** field.

You can apply a filter to display unacceptable transactions only.

► **Apply a filter to display unacceptable transactions**

- Select and right-click any blank entry in the **Codes\_Sub\_Unacceptable** field. Select **Quick Filter » Not Equal**.  
You have filtered out all of the records that have blank entries in the **Codes\_Sub\_Unacceptable** field. Now **ACL** displays records that have unacceptable codes only.
- Remove the **Codes\_Sub\_April** and the **CODES** fields from the view.
- Select **Analyze » Count Records**. Click **OK**. There are 7/281 unacceptable transactions.  
Next, you extract the records to a new table.

► Extract the records to a new table

- 1 Select **Data » Extract Data**. The **Extract** dialog box appears.
- 2 In the **Extract** dialog box, select **Record**. In the **To** field, enter `Trans_April_Unacceptable`. Click **OK**. Click **Yes** when you are asked if you want to save the changes to `Trans_April`.

The new table appears in the view. Adjust the column widths as necessary and rearrange the columns as desired.

Use the **Count** command to count how many unacceptable transactions there are, in order to write this number to the log and create an audit trail. You should have seven unacceptable transactions.

By looking at the results, you can see that four unacceptable transactions took place on April 14.

You can create a date filter and count the number of records to write this information to the log.

- 3 Close the view and click **Yes** when **ACL** prompts you to save.

	CARDNUM	AMOUNT	TRANS_DATE	CODES	CUSTNO	DESCRIPTION
1	0590126009071523	314.33	04/14/2002	4411	204354	Cruise Lines
2	0590126089273535	226.00	04/14/2003	4411	444413	Cruise Lines
3	0590126061697004	745.00	04/07/2002	5094	444413	Precious Stone
4	0590127015016937	457.70	04/09/2002	5094	362353	Precious Stone
5	0590120245724706	642.03	04/14/2003	7994	070035	Video Game A
6	0590120346463420	439.52	04/11/2003	7995	444413	Betting (includ
7	0590120347406193	749.14	04/14/2003	7995	770000	Betting (includ

## ■ Identify the employees

Now that you have confirmed that some transactions do have unacceptable codes, you determine which employees made these purchases. You can also identify the department to which these employees belong.

You can obtain information from several tables and combine it in a single view. The goal is to report:

- Employee name
- Credit card number
- Department
- Transaction amount
- Transaction date
- Transaction code
- Transaction code description

### • Tasks

Determine the department of employees whose cards were used for unacceptable transactions.

Create a list of employees whose cards were used for unacceptable transactions.

Create a report of the employees whose cards were used for unacceptable transactions.



## Identify the employees' departments

Your first step is to associate each employee with his or her department.


- The **Company\_Dept** table contains department codes and names, such as 800000 for Consulting.
- The **Employees** table contains the **EmpNo** field, which begins with the first digit of the department code.

By comparing the first digit of the employee code to the first digit of the company department code, you can identify the department to which the employee belongs.

To make this comparison, you create a computed field in each table that takes the first digit of the **Dept\_Code** values. You then compare this digit to the first digit of the **EmpNo** in the **Employees** table.

You start by checking the data types of these fields on the **Edit Fields/Expressions** tab of the **Table Layout** window.

### ► Create a computed field that contains the first digit of the Dept\_Code field

- 1 Open the **Company\_Dept** table.
- 2 Select **Edit » Table Layout**. The **Table Layout** window appears. On the **Edit Fields/Expressions** tab, ensure that the **Dept\_Code** field has a data type of **ASCII**. Change the type if it is incorrect.
- 3 On the **Edit Fields/Expressions** tab, click the **Add a New Expression**  button.
- 4 Enter the name `Dept_Code_1`.

- 5 Click **f(x)**. The **Expression Builder** appears. Enter the following expression:

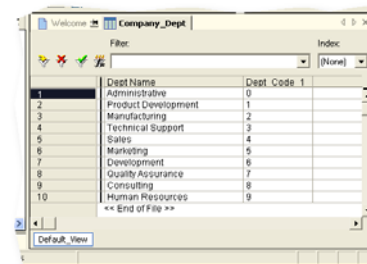
```
SUBSTR(ALLTRIM(Dept_Code),1,1)
```

- 6 Click **OK**. Click the green checkmark.
- 7 Close the **Table Layout** window.

You now have a new computed field called **Dept\_Code\_1**, which you can now add to the view.

### ► Add the new computed field to the view

- 1 Right-click the **Dept\_Code** column header. Select **Add Columns**.
- 2 Double-click **Dept\_Code\_1**. Click **OK**.  
The digit in the **Dept\_Code\_1** column matches the first digit of the **Dept\_Code** column.  
You can remove the **Dept\_Code** column because the new column has the relevant information.
- 3 Right-click the **Dept\_Code** column and select **Remove Selected Columns**. Click **Remove**.




	Dept Name	Dept Code	Dept Code_1
1	Administrative	0	0
2	Product Development	1	1
3	Manufacturing	2	2
4	Technical Support	3	3
5	Sales	4	4
6	Marketing	5	5
7	Development	6	6
8	Quality Assurance	7	7
9	Consulting	8	8
10	Human Resources	9	9

- 4 Close the view and click **Yes** when prompted to save.

Next, you create a computed field in the **Employees** table that uses the first digit of the employee number field, **EmpNo**.

► Create a computed field that displays the first digit of the EmpNo field

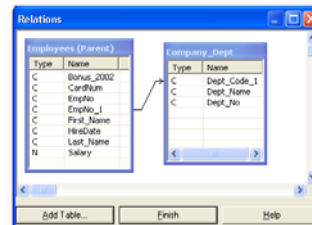
- 1 Open the **Employees** table.
- 2 Select **Edit » Table Layout**. The **Table Layout** window appears.
- 3 On the **Edit Fields/Expressions** tab, click the **Add a New Expression**  button.
- 4 Enter the name **EmpNo\_1**.
- 5 Click **f(x)**. The **Expression Builder** appears. In the **Expression** box, enter:  
`SUBSTR(ALLTRIM(EmpNo) , 1 , 1 )`
- 6 Click **OK**. Click the green checkmark.
- 7 Close the **Table Layout** window.  
 Now add this field to the view.
- 8 Right-click the **EmpNo** column header. Select **Add Columns**. The **Add Columns** dialog box appears. Double-click **EmpNo\_1**. Click **OK**.

The new **EmpNo\_1** column now appears in the view.

Now relate the **Employees** table to the **Company\_Dept** table, using the **EmpNo\_1** field and the **Dept\_Code\_1** field.

► Relate the **Employees** and **Company\_Dept** tables

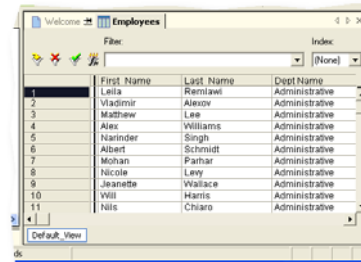
- 1 With the **Employees** table active, select **Data » Relate Tables**. The **Relations** dialog box appears.
- 2 Click **Add Table**. The **Add Table** dialog box appears.
- 3 Double-click **Company\_Dept**.
- 4 Drag **EmpNo\_1** in the **Employees** table to the **Dept\_Code\_1** field in the **Company\_Dept** table.



5 Click **Finish**.

Your next step is to add the **Dept\_Name** column to the view.

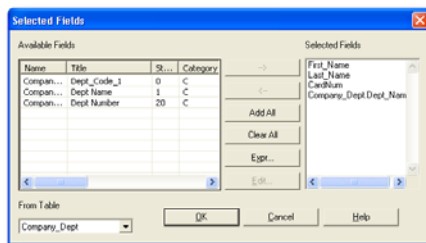
- 6 Right-click the **CardNum** column header and select **Add Columns**. The **Add Columns** dialog box appears.
- 7 In the **From Table** drop-down list, select **Company\_Dept**. Double-click the **Company\_Dept.Dept\_Name** field. Click **OK**.



The **Dept\_Name** column appears in the view. You are close to your goal of creating a report. Your next task will be to extract this view to a new table.

► Extract fields to a new table

- 1 With the **Employees** table active, select **Data » Extract Data**. The **Extract** dialog box appears.
- 2 Select **Fields**. Click the **Extract Fields** button. The **Selected Fields** dialog box appears.
- 3 Add the following fields to the **Selected Fields** list:
  - **First\_Name**
  - **Last\_Name**
  - **CardNum**
- 4 In the **From Table** drop-down list, select **Company\_Dept**. The fields in the **Company\_Dept** table appear.
- 5 Add **Company\_Dept.Dept\_Name** to the **Selected Fields** list.




- 6 Click **OK**.
- 7 In the **To** field, enter **Employees\_Dept**. Click **OK**. Click **Yes** when **ACL** prompts you to save. The new **Employees\_Dept** table opens in the view.

## Create a list of employees

Your next step is to relate the **Employees\_Dept** table to the **Trans\_April\_Unacceptable** table using the card number field in each table.

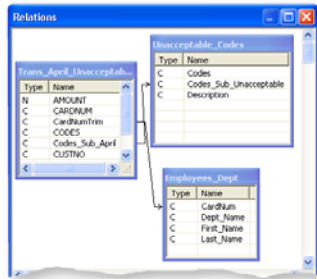
Before you can use the Relations command, you must harmonize the two card number fields so that you can use them as key fields. Check the card number fields in the **Table Layout** window for both tables.

► Compare and harmonize the card number fields

- 1 With the **Employees\_Dept** table active, select **Edit » Table Layout**. The list of fields indicates that the **CardNum** field has a field length of 16 and a category of C for character.
- 2 Close the **Table Layout** window.
- 3 Open the **Trans\_April\_Unacceptable** table.
- 4 Select **Edit » Table Layout**. The **CardNum** field has a field length of 19, which means that this field contains extra spaces. Therefore, you must trim this card number field.
- 5 Click the **Add a New Expression**  button.
- 6 Enter **CardNumTrim** as the **Name**.
- 7 In the **f(x)** box, enter this expression:  
`SUBSTR(ALLTRIM(CardNum), 1, 16)`
- 8 Click the green checkmark.
- 9 Close the **Table Layout** window.

► Create a detailed view of the unacceptable transactions

- 1 With the **Trans\_April\_Unacceptable** table active, select **Data » Relate Tables**.  
**Trans\_April\_Unacceptable** already has a relation with **Unacceptable\_Codes**.
- 2 Select **Add Table**. The **Add Table** dialog box appears. Double-click **Employees\_Dept**.
- 3 Drag **CardNum** from the **Employees\_Dept** table to **CardNumTrim** in the **Trans\_April\_Unacceptable** table.



4 Click **Finish**.

Three tables, **Trans\_April\_Unacceptable**, **Unacceptable\_Codes**, and **Employees\_Dept**, are now related.

In the view, you now add all of the columns that you want to include in the final report.

- 5 Right-click the **CardNum** column header. Select **Add Columns**. The **Add Columns** dialog box appears.
- 6 In the **From Table** drop-down list, select **Employees\_Dept**. Add these fields to the **Selected Fields** list:
  - **Employees\_Dept.First\_Name**
  - **Employees\_Dept.Last\_Name**

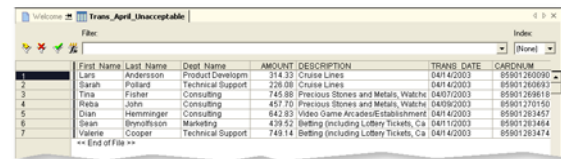
■ **Employees\_Dept.Dept\_Name**

Click **OK**.

7 Arrange the fields in the view so that they appear in the following order:

- **First\_Name**
- **Last\_Name**
- **Dept\_Name**
- **AMOUNT**
- **DESCRIPTION**
- **TRANS\_DATE**
- **CARDNUM**
- **Codes\_Sub\_Unacceptable**
- **CUSTNO**

You can now see the list of employees whose credit cards were used to charge unacceptable transactions. You can rearrange and adjust the fields as necessary so that the view is attractive.



Of course, you cannot assume immediately that the employees listed are responsible for the unacceptable transactions because it is possible someone else used their credit cards. Nevertheless, you can create a report of the results, and then the staff at Metaphor Corporation can investigate the matter further.

## Report your findings

ACL provides two ways to create a report. You can generate a standard ACL report, or you can create a report that you can view in Crystal Reports®. In this procedure, you use ACL's standard report capability.

To learn more about using Crystal Reports to generate a report, select **Help » Index** and look up "Crystal Reports".

### ► Create a standard ACL report

#### 1 Select **Data » Report**.

The **Report** dialog box appears.

#### 2 Enter a descriptive **Header**, such as:

Unacceptable Transactions at Metaphor Corporation in April 2003

#### 3 On the **Output** tab, select **Screen**. Click **OK**.

The report appears on the **Do** tab.

First_N	Last_Name	Dept_Name	AMOUNT	DESCRIPTION	TRANS_DATE	CASHNUM	CODES	CUSTNO
Lars	Andersson	Product Dev	314.33	Cruise Lines	04/14/2003	8590126009	44	444413
Sarah	Pollard	Technical S	226.00	Cruise Lines	04/14/2003	8590126009	44	444413
Tina	Fisher	Consulting	745.88	Precious Stones and He	04/07/2003	8590126961	50	444413
John	John	Consulting	457.70	Precious Stones and He	04/09/2003	8590127015	50	962252
Dian	Heminger	Consulting	641.83	Video Game Arcades/Exc	04/14/2003	8590128345	79	878035
Sean	Brynsolfs	Marketing	439.52	Betting (including Lot	04/11/2003	8590128346	79	444413
Valerie	Cooper	Technical S	749.14	Betting (including Lot	04/14/2003	8590128347	79	778088
			2575.49					

Next, you save the report as an HTML file.

#### 4 Select **Data » Report**.

#### 5 On the **Output** tab, select **File**. Select **HTML Text File** from the **File Type** drop-down list.

6 In the **Name** text box, enter `Unacceptable_Trans_Apr_2003`. Click **OK**.

7 In Windows Explorer, locate the folder `C:\ACL Data\Sample Data Files` and double-click `Unacceptable_Trans_Apr_2003.html`. The report opens in your Web browser.

## Exercises

- 1 Create a report that lists all acceptable transactions. The procedures will be the same as for the report that lists unacceptable transactions, except you will use the **Acceptable\_Codes** list as the basis for comparison.

Using Relations, create a view that displays the following information:

- Employee name
- Credit card number
- Department
- Acceptable transaction amount
- Transaction date
- Transaction code
- Transaction code description

Extract this view to a new table **Acceptable\_Trans\_by\_Emp\_April**, and export the new table to an Excel file.

- 2 Remember that some transactions have neither acceptable nor unacceptable codes. These are the transactions that require approval, but are not forbidden. Create a table of transactions that are neither acceptable nor unacceptable. How many transactions are in this category?
- 3 Extract all of the transactions that exist in the approval-required category to another table.
- 4 Refer to the **Trans\_April**, **Trans\_April\_Unacceptable**, and **Employees** tables. What percentage of employees with credit cards made unacceptable transactions?
- 5 How would you use *ACL* to e-mail the results of your report?

## ■ Chapter review

In this chapter you accomplished the following tasks:

- Examined the tables that contain acceptable and unacceptable Merchant Category Codes to ensure that no codes are listed in both tables.
- Created a table with unacceptable transactions only.
- Linked the employees who made unacceptable transactions to the employees' departments.
- Created a report that provides a detailed description of all unacceptable transactions in April.

## VALIDATE, CORRECT, AND EXTRACT DATA

*Repair your data so it can be analyzed*

### **In this chapter...**

Organize your project with folders

Expand your project to include May transactions

Combine May transactions into a single table

Correct card number and employee lists

So far, you have planned a project, created a project, added data to a project, and analyzed it. You based your analysis on Metaphor's April 2003 credit card transactions.

Now you have been asked to analyze the May 2003 transactions.

Unfortunately, these transactions are in two separate files, and the data is not as neat as the April 2003 transactions. You must perform additional steps to get the data into a state that you can analyze readily. Once you repair the transaction data, you extract the data to new tables.

## ■ Import the May transaction data

You already have a project for Metaphor analysis in *ACL*. You created a project that contains all of the April transactions. Now you plan to add May transactions too.

### ● Tasks

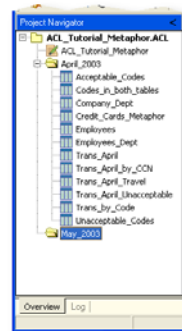
- Create folders for April and May transactions.
- Import worksheets from the *Trans\_May.xls* Excel workbook.
- Compare layouts of the two transaction tables.

## Use folders to organize your project

To keep your project organized, you will move your April work into another folder. You can then set up another folder for your May work. All of the work can then be viewed as needed, but the **Overview** will not be cluttered with items that you are not using.

### ► Create a folder called April 2003

- 1 Select **File » New » Folder**. A folder appears. Enter the name *April 2003* for the folder. Press **ENTER**.
- 2 Select and drag all of the tables to the **April 2003** folder, as you would in Windows Explorer.
- 3 Click **ACL\_Tutorial\_Metaphor.ACL**. Select **File » New » Folder**. Name the folder *May 2003* and press **ENTER**.





## Import transaction data for May 2003

The **Trans\_May.xls** workbook contains two worksheets with your transaction data for May 2003. Refer to Chapter 3: “Set up your project” if you are uncertain about how to import data.

### ► Import the **Trans\_May.xls** Excel workbook

- 1 Right-click the May 2003 folder. Select **New » Table**.  
The **Data Definition Wizard** appears. Continue through the wizard, and select **Trans\_May.xls**.
- 2 Because **ACL** saves the data from each worksheet as a separate table, you need to import **Trans\_May.xls** once for each worksheet. On the **Data Source** screen select the first worksheet and enter a value in the **Maximum Character Field Length** box.
- 3 Save the table as **Trans1\_May**. Return to step 1 and save the second worksheet as **Trans2\_May**.

### ► Edit the table layout for both new tables

- 1 With the **Trans1\_May** table active, select **Edit » Table Layout**. The **Table Layout** window appears.
- 2 Edit the **AMOUNT**, **CARDNUM**, and **CODES** fields:
  - Double-click the **AMOUNT** field. Select **NUMERIC** from the **Type** drop-down list. Enter 2 in the **Dec** field so that **AMOUNT** has two decimal places. Select **-999999.99** from the **Format** drop-down list. Click the green checkmark.
  - Double-click the **CODES** field, change the data type to **ASCII**, and click the green checkmark.
  - Double-click the **CARDNUM** field, change the data type to **ASCII**, and click the green checkmark.

- 3 Close the **Table Layout** window.
- 4 Select **Data » Verify**, then verify all of the fields. You should have zero validity errors.
- 5 Repeat the same procedures for the **Trans2\_May** table.

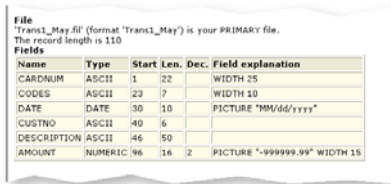
## Compare and correct the table layouts

Before you can combine your tables, the **Start**, **Type**, and **Length** values for each field in each table must be exactly the same. To compare the tables, you can use `DISPLAY` on the command line.


Most activities in *ACL* can also be accomplished by typing ACLScript commands on the command line. ACLScript is the language for writing scripts and executing commands in *ACL*.

### ► Compare the layout of the transaction tables

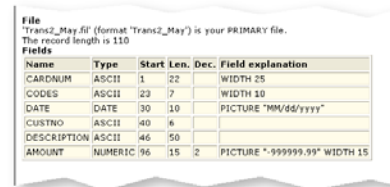
- 1 Open the **Trans1\_May** table. Select **Window » Show Command Line**. The command line appears.
- 2 In the **Command Line** text box, enter `DISPLAY`. Press `ENTER`. *ACL* displays the table layout.



Name	Type	Start	Len.	Dec.	Field explanation
CARDNUM	ASCII	1	22		WIDTH 25
CODES	ASCII	23	7		WIDTH 10
DATE	DATE	30	10		PICTURE 'MM/DD/YYYY'
CUSTNO	ASCII	40	6		
DESCRIPTION	ASCII	46	50		
AMOUNT	NUMERIC	96	15	2	PICTURE '-999999.99' WIDTH 15

- 3 On the **Display** tab, click the push-pin  button to ensure that *ACL* displays subsequent command results in addition to the “pinned” results.
- 4 Open the **Trans2\_May** table.

- 5 In the **Command Line** text box, enter `DISPLAY`. Press `ENTER`. *ACL* displays the table layout.



Name	Type	Start	Len.	Dec.	Field explanation
CARDNUM	ASCII	1	22		WIDTH 25
CODES	ASCII	23	7		WIDTH 10
DATE	DATE	30	10		PICTURE 'MM/DD/YYYY'
CUSTNO	ASCII	40	6		
DESCRIPTION	ASCII	46	50		
AMOUNT	NUMERIC	96	15	2	PICTURE '-999999.99' WIDTH 15

- 6 Compare the contents of both **Display** tabs. The table layouts are identical except for one inconsistency: in **Trans1\_May**, **AMOUNT** has a field length of 16, but in **Trans2\_May**, **AMOUNT** has a field length of 15. You can resolve this problem by correcting the layout of **Trans1\_May**.

### ► Correct the layout of Trans1\_May

- 1 Open **Trans1\_May**.
- 2 In the view, right-click the **AMOUNT** column and select **Quick Sort Descending**.

The view shows that the highest **AMOUNT** value you have in **Trans1\_May** is 1983.04. This value does not require 16 characters. You can safely change the **AMOUNT** field length to 15 so that it matches the **AMOUNT** field in the **Trans2\_May** table. **AMOUNT** starts at position 96, and no field comes after **AMOUNT**, so you will not affect the positions of the other fields by changing **AMOUNT**.

- 3 Select **Edit » Table Layout**. The **Table Layout** window appears. Double-click **AMOUNT**. Change **Len** to 15. Click the green checkmark.
- 4 Close the **Table Layout** window.

## ■ Combine and verify the tables

In this section, you combine the transactions from both May tables into one table.

You can combine information from two or more tables by using the following methods:

- **Join** – The Join command creates a new table that includes selected fields from both files.
- **Relations** – The Relations command does not create a new table, but the fields in each of the related tables can be analyzed as if they were in the same table.
- **Extract-and-append** – You can extract records and append them to an existing table.
- **Merge** – You can merge fields from two tables into one table.

You have two credit card transaction tables, **Trans1\_May** and **Trans2\_May**, with exactly the same layout. To analyze them as one unit, you combine them.

### ● Tasks

- Use extract-and-append to combine the May transaction tables.
- Verify the combined table.
- Check the table for duplicate transactions.

## Combine the May transaction tables

Using the extract-and-append method, you can extract **Trans1\_May** into a new table.

You can then extract the information in **Trans2\_May** and add it to the end of the new table. The new table will contain the combined information from both tables.

You could use other methods to combine the data, but the extract-and-append method is ideal for tables with identical layouts.

### ► Use the extract-and-append method to combine two tables

- 1 With the **Trans1\_May** table active, select **Data » Extract Data**.
- 2 Select **Fields**. Click **Extract Fields**. The **Selected Fields** dialog box appears.
- 3 In the **Available Fields** list, click the **Start** column header so that all of the fields are listed in sequential order starting with **CARDNUM**.
- 4 Click **Add All**. Click **OK**. You are extracting all of the fields in the order listed.
- 5 Enter `Trans_May_All` in the **To** field. Click **OK**. *ACL* creates a new table, **Trans\_May\_All**, which contains the data from **Trans1\_May**.  
Now append the **Trans2\_May** data to the **Trans\_May\_All** table.
- 6 Open **Trans2\_May**. Select **Data » Extract Data**.
- 7 Select **Fields**. Click **Extract Fields**. The **Selected Fields** dialog box appears.
- 8 In the **Available Fields** list, click the **Start** column header so that all of the fields are listed in sequential

order starting with **CARDNUM**. This order is exactly the same as the order that you used for **Trans1\_May**.

- 9 Click **Add All**. Click **OK**.
- 10 Enter `Trans_May_All` in the **To** field.
- 11 On the **More** tab, select **Append To Existing File**.
- 12 Click **OK**.

*ACL* opens **Trans\_May\_All**. You can now work with this table in *ACL* as you would any other table.

## Verify the new table

Next, you use the **Verify** and **Count** commands to ensure that the two tables have been appended correctly.

### ► Verify the new table

- 1 Select **Data » Verify**.
- 2 Click **Verify Fields**. Click **Add All**. Click **OK**.
- 3 On the **Output** tab, ensure that **Screen** is selected. Click **OK**.

You should get the result: **0 data validity errors detected**.

### ► Compare the Count in each table

- 1 In the **Trans\_May\_All** table, select **Analyze » Count Records**. Click **OK**. You should have 200 records.
- 2 Count the records in the **Trans1\_May** table.  
You should have 86.
- 3 Count the records in the **Trans2\_May** table. You should have 114.

The extract-and-append procedure appears to have been successful.

## Check the combined table for duplicates

Always check for data-entry errors when you combine two tables. Duplicate entries can indicate that someone entered the same transaction twice in the original Excel worksheets.

Duplicate transactions can also indicate fraud. Employees can split credit card transactions to avoid scrutiny. Instead of a \$2000 purchase, the employee might have two separate transactions of \$1000, for example. Alternatively, merchants might charge twice for the same product or service.

Use the Duplicates command to look for transactions of identical amounts.

### ► Check the **Trans\_May\_All** table for duplicates

- 1 Open the **Trans\_May\_All** table.
- 2 Select **Analyze » Look for Duplicates**. The **Duplicates** dialog box appears.
- 3 Select **AMOUNT** in the **Duplicates On** list.
- 4 Click **List Fields**. Click **Add All**. Click **OK**.
- 5 On the **Output** tab, ensure that **Screen** is selected. Click **OK**.
- 6 View the results. You are checking for duplicate **AMOUNT** values.

You should have a transaction for a computer software store and one for Southwest in the same amount. You should also have two transactions at the Hilton one day apart in the same amount.

What do you think is the possible significance of these duplicate amounts?

## Exercises

- 1 Run the Profile command on the **Trans\_May\_All** table. Note the highest and lowest transaction values.
- 2 Check for duplicates in the **CUSTNO** field. Classify the table on the same field. Are the results significant?
- 3 Export a table to XML that lists the employee name, card number, and total dollar value of associated transactions per-card in May. View the file in Internet Explorer.

## ■ Isolate and correct validity errors

In the last section, you verified the data that you imported. However, verifying the data is different from validating it. For example, you determined that the **CARDNUM** field consisted only of characters, so that the data is ASCII text. However, you still must determine whether the card numbers are valid, which means that each entry consists of 16 digits and no other characters.

In this section, you create:

- A filter to display only the valid entries.
- A filter that displays only the invalid entries.

If some entries are invalid, you can either correct the problems, or filter out the records that contain those entries.

### ● Tasks


- Determine whether all the card numbers have exactly 16 digits.
- Make a 16-digit credit card field.
- Extract repaired card number records to a new file.
- Correct the new list of employees.

## Check the format of card numbers

You need to determine which card numbers in the table are in an invalid format. To do this, you apply a filter to the **CARDNUM** field that uses a nested function that consists of **ALLTRIM( )** and **MAP( )**:

- **ALLTRIM( )** removes all trailing blanks from each card number entry.
- **MAP( )** evaluates whether each card number, trimmed of trailing blanks, has exactly 16 digits.

### ► Apply a filter to show invalid card numbers

- 1 With the **Trans\_May\_All** table active, click the **Edit View Filter**  button.

The **Edit View Filter** dialog box appears.

- 2 In the **Expression** box, enter the following expression all on one line:

```
MAP(ALLTRIM(CARDNUM),
"9999999999999999")
```

The string of sixteen nines indicates that the filter determines whether each **CARDNUM** value consists of sixteen digits.

- 3 Click **OK**. *ACL* applies the filter.
- 4 Count the number of records. As you can see, 195/200 records remain in the view, which means that 5 records do not meet the criteria.

- 5 In the **Filter** text box, add **NOT** to the beginning of the expression:

```
NOT MAP(ALLTRIM(CARDNUM),
"9999999999999999")
```

Press **ENTER**. **NOT** causes *ACL* to display the invalid entries.

6 Count the records. You should have 5/200 visible.



The invalid card numbers contain either extraneous dashes or spaces. In this case, you can see all of the invalid entries at a glance, but in a large project, you might not be able to do so.

## Create a 16-digit card number field

Now you compare these card numbers to another list of card numbers that do not contain dashes and spaces. To do this, create a computed field in **Trans\_May\_All** in which all of the card numbers are sixteen digits long with no spaces or dashes.

The field that you create must also contain all of the valid card numbers without changing them. Invalid card numbers must be corrected without altering the digits.

### ► Create a computed field with corrected card numbers

- 1 In the **Trans\_May\_All** table, click the **Remove Filter**  button to remove the filter. The view now displays all of the records.
- 1 Select **Edit » Table Layout**. The **Table Layout** window appears.
- 2 Click the **Add a New Expression**  button.
- 3 Enter `CCN_Corrected` in the **Name** field.
- 4 Click the **f(x)** button. The **Expression Builder** appears. The expression that you create will populate the **CCN\_Corrected** field.
- 5 Enter the following expression:
 

```
ALLTRIM(INCLUDE(CARDNUM,"0123456789"))
```
- 6 Click **OK**. The **Edit Fields/Expressions** tab appears again.
- 7 Click the green checkmark.
- 8 Close the **Table Layout** window.

You can now add the new field to the view.

► Add the `CCN_Corrected` computed field to the view

- 1 Right-click the `CARDNUM` header. Select **Add Columns**. The **Add Columns** dialog box appears.
- 2 Double-click `CCN_Corrected`. Click **OK**.
- 3 Remove the `CARDNUM` column.


Because you now have the corrected card number field, you no longer need to see the `CARDNUM` field.

Now you apply a filter to ensure that the `CCN_Corrected` field contains only digits.

- 4 In the **Filter** text box, enter:

```
MAP(CCN_Corrected, "9999999999999999")
```

Press **ENTER**.

- 5 Count the number of records. All 200 are visible.
- 6 Click the **Remove Filter**  button to remove the filter.

Your next step is to extract this information to a new table.

## Extract the corrected records to a new table

Now that you have created a field that contains only valid card numbers, your next step is to extract the records to a new table. You then use this new table for further analysis.

► **Extract the corrected records to a new table**

- 1 Select **Data » Extract Data**.
- 2 Select **Fields**. Click the **Extract Fields** button. The **Selected Fields** dialog box appears.
- 3 Click **Add All**. Because `CARDNUM` contains invalid entries, you do not want to include that field in your new table. Remove `CARDNUM` from the **Selected Fields** list. Click **OK**.
- 4 In the **To** text box, enter the file name `Trans_May_All_CCN_Corrected`. Click **OK**.

*ACL* opens **Trans\_May\_All\_CCN\_Corrected**.



## Export the combined transaction table to Excel

You have created and checked a table that combines the two transaction tables that represent May 2003 credit card transactions. You might want to have this data in Excel format as well.

### ► Export the `Trans_May_All_CCN_Corrected` table to Excel

- 1 With the `Trans_May_All_CCN_Corrected` table active, select **Data » Export to Other Application**.
- 2 Select **Fields**.
- 3 Select the fields you want to export in the order you want them to appear. The fields will appear in the Excel worksheet in this order, from left to right.
- 4 From the **Export As** drop-down list, select **Excel**.
- 5 In the **To** field enter the filename `Trans_May_All_Corrected`. Click **OK**.

You can also select the path for the file. If you enter just the filename, the new file will be saved to the last folder that you used to store your *ACL* data.

## Create an error-free employee record table

For May, you have been given a slightly different list of employees to use: `Employees_May_with_CCN.xls`. The table includes only those employees who have credit cards with Metaphor.

### ► Import the new table of employees

- 1 Import `Employees_May_with_CCN.xls` into the **May 2003** folder and save it as `Employees_May_with_CCN`.
- 2 Scroll through this new table.

You can see that the bottom record has “End of File” in the **First\_Name** field, and that all the other field entries for this record are empty. Clearly, this is not an employee record, and it should be filtered.

### ► Filter invalid employee records

- 1 With the `Employees_May_with_CCN` table active, right-click the **End of File** entry.
- 2 Select **Quick Filter » Not Equal**.

The invalid record is now filtered out.

If you extract the valid data to another table, then you do not have to worry about preserving the filter on the original table.

### ► Extract the valid records to a new table

- 1 Select **Data » Extract Data**.
- 2 Select **Fields** button. Click **Extract Fields**. Click **Add All**. Click **OK**.
- 3 In the **To** text box, enter the filename `Employees_May_Corrected`. Click **OK**.

## Exercises

- 1 Suppose you wanted all card numbers to have the form: xxxx-xxxx-xxxx-xxxx. What strategy would you follow to create a computed field like that?
- 2 If you had the information available, checking the Social Security Number (SSN) for each employee would help determine whether each record was unique. How would you check for duplicate Social Security Numbers? How would you ensure that all Social Security Numbers were in the format of one long string of digits: xxxxxxxx?
- 3 Devise a strategy for analyzing your May transactions that closely mirrors the strategy you followed for the April transactions. Find the unacceptable transactions for this set of transactions.
- 4 For your own projects, think of various approaches that improve the quality of your data before you import it into *ACL*. Think of ways to test your data after you import it into *ACL*.

## ■ Ideas for further analysis

In this tutorial, you have worked through several examples of how to perform analyses using *ACL*. What other analyses could you perform with the data?

Consider these questions. You can likely come up with numerous ideas that are applicable to your own work.

- Do the transactions have unacceptable Merchant Category Codes?
- Do the Merchant Category Codes that are used match the cardholder's responsibilities?
- Does the cardholder conduct repeated transactions at regular intervals with the same merchant?
- Does the cardholder have several transactions within a brief time period (say 72 hours) with the same merchant?
- Does the cardholder have transactions occurring on weekends or holidays?
- Do an unusual number of transactions have rounded values, such as \$1000, \$200, or \$450?
- Are there credits charged back to a card number, and what are the explanations for these credits?
- Does the cardholder regularly come close to or exceed his or her credit limit?
- Has there been a sudden increase in the cardholder's purchases?
- Is the cardholder responsible for charging other people's expenses?
- Which five or ten employees spent the most? To which customers did they charge their expenses? Is

there any relation between spending a great deal and charging unacceptable transactions?

- Determine the total spending by department on a monthly basis. Use the **DeptCode** field in the **Employees** table.
- If an employee has charged air fare, what other travel-related expenses has he or she charged? For example, if an employee charges air fare, he or she is likely to charge a hotel room and transportation. Is this the case? What are some possible explanations for the patterns that appear?
- If the codes for caterers and bakeries were unacceptable, how much would money have been saved?

## ■ Chapter review

In this chapter you accomplished the following tasks:

- Imported an Excel workbook with two worksheets.
- Combined information from two tables into one table.
- Corrected a data field that included invalid data.
- Extracted the new transactions table to Excel.
- Corrected a table by filtering out an invalid record and extracting the results to a new table.
- Created a table that combines all of the CC transactions into one table.



**A**

- Absolute value
  - Profile command results **12**
- Acceptable\_Codes.mdb
  - data file **5**
- Access. *See* Microsoft Access data
- Account number categories
  - Metaphor Corporation **4**
- Allowable expenses
  - Metaphor Corporation **4**
- ALLTRIM() function
  - using **43, 45, 52, 54, 57, 59, 70**
- Analysis strategy
  - planning **7**
- Analytical strategy
  - planning **6**
- Append option
  - Extract command **67**
- Applications
  - exporting to other **35**
- Ascending
  - Quick Sort order **15**
- ASCII
  - data type **28**
- ASCII data
  - formatting in expressions **36**
- Audit trail
  - leaving **53**
  - log tab **15**

**B**

- Backup copy

- ACL project **26**
- BETWEEN() function
  - using **43**
- Blank entries
  - removing **55**
- Blanks
  - removing leading and trailing **45, 70**
- Bonuses
  - finding total **15**

**C**

- Calculated fields. *See* Computed fields
- Categories
  - data type **29**
  - expense, Metaphor Corporation **4**
- Cells
  - blank **55**
- Character data
  - applications **28**
  - formatting in expressions **36**
  - numerals in **24**
- Characteristics of data
  - checking **27**
- Checkmark
  - green **17**
- Classify command
  - using **46**
- Code categories
  - Metaphor Corporation **5**
- Codes
  - acceptable **5**
  - common **51**
  - comparing **54, 57**

- Column titles
  - changing **25**
  - and field names **25**
- Column width
  - changing in view **17**
- Columns
  - adding to view **52, 54, 58, 72**
  - changing sort order in view **10, 15**
  - removing from view **55**
- Combining data
  - methods **67**
- Command
  - defined **15**
- Command line
  - using **66**
- Command results
  - viewing details **13**
- Commands
  - Classify **46**
  - Count **11, 16**
  - DISPLAY **66**
  - Duplicates **13, 69**
  - Export **35, 73**
  - Extract **56, 58**
  - Join **50**
  - Profile **12**
  - Relations **50, 54, 58, 60**
  - running on single fields **11**
  - selecting fields to include **13**
  - Statistics **11**
  - Summarize **42**
  - Total **15, 34**
  - Verify **68**
- Comma-separated-value files

## Index

- importing 25
- Company\_Departments.txt
  - data file 5
- Computed fields
  - adding to view 57
  - creating 17, 52, 54
- Conditions
  - multiple 38
- Count command
  - using 11
- COUNT field
  - Summarize command 42
- Credit card policies
  - Metaphor Corporation 4
- Credit cards
  - finding expiry date 34
- Credit\_Cards\_Metaphor.xls
  - data file 5
- Credit\_Cards\_Metaphor.xls worksheet
  - importing 23
- CSV files
  - importing 25
- Cycle of analysis
  - typical 6

## D

- Data
  - appending 67
  - combining 50
  - errors
    - eliminating 7
  - exporting 35
  - filtering 72
  - format information 7
  - importing through the wizard 22
  - integrity testing 13, 27
  - isolating specified portions 43
  - modifying with computed fields 28

- samples for testing 7
- sources 6
- validity errors, finding 30
- verifying type 30
- viewing in original format 6

- Data Definition Wizard
  - using 22
- Data definitions
  - modifying 28
- Data files
  - opening 10
  - tutorial 5
- Data format
  - checking 27
- Data type
  - changing 24, 35
  - checking 51, 53, 66
  - verifying 27
- Data type categories
  - Table Layout window 29
- Data-entry errors
  - finding 69
- Date data
  - Statistics command 11
- Dates
  - formatting in expressions 16
  - selecting in Expression Builder 34
- Decimals
  - rounding 17
- Default view
  - displaying 17
- Delimited files
  - importing 23
- Descending
  - Quick Sort order 15
- Digits
  - using as character data 24
- DISPLAY command
  - using 66

- Documentation
  - project 30
- Double entries
  - finding 13
- Drill-down filter
  - command results 13
- Duplicate records
  - finding 13
- Duplicates command
  - using 13, 69

## E

- Edit Field Properties screen
  - using 24
- Edit view filter
  - dialog box 16, 17, 34
- Employee numbers
  - finding duplicates 13
- Employees.csv
  - data file 5
- Empty cells
  - removing 55
- End of file record
  - deleting 73
- Entries
  - duplicate 13, 55, 69
  - invalid, displaying 70
- Errors
  - data validity 30
  - data-entry 7, 69
  - testing for 65
  - validity 68
- Excel. *See* Microsoft Excel data
- Expenses
  - allowable
    - Metaphor Corporation 4
  - isolating 43, 45
  - summarizing by code 42

## Index

Expiry dates  
    finding 34

Export  
    using Excel format 73

Export command  
    using 35, 73

Expression  
    defined 16

Expression Builder  
    *See also* Edit view filter  
    functions list 35  
    opening 52  
    using 17

Expressions  
    checking syntax 36  
    creating and saving 17  
    formatting  
        character data 36  
        multiple conditions in 38

Extract command  
    append option 67  
    using 44, 56, 58

**F**

Field length  
    changing 66  
    checking 51, 53, 59

Field names  
    and column titles 25  
    importing from data file 24

Field order  
    changing in the view 10  
    selecting for export 73

Field separators  
    specifying 24

Field totals  
    finding 34

Field type. *See* Data type

Field width  
    adjusting in view 17

Fields  
    adding to view 52, 54, 58  
    blank entries in 55  
    changing order 60  
    combining from multiple tables 50  
    computed 28, 52, 54  
    editing 28  
    eliminating invalid characters 70  
    harmonizing 51, 53, 59  
    key 51, 66  
    ordering sequentially 67  
    removing from view 55  
    sort order 15, 37  
    verifying content 27

Files  
    importing 23  
    Metaphor data 5

Filters  
    applying multiple 45  
    creating 16, 34, 72  
    removing 13, 17  
    saving 37  
    setting multiple conditions 38

FIND( ) function  
    using 35

Folders  
    creating 64

Format  
    data  
        checking 27  
        data file 6  
        date expressions 16  
        expression  
            verifying 36  
        invalid 70

Formats  
    export 35

Functions  
    ALLTRIM( ) 43, 45, 57, 70  
    BETWEEN( ) 43  
    FIND( ) 35  
    INCLUDE( ) 71  
    MAP( ) 70  
    MATCH( ) 45  
    nesting 70  
    SUBSTRING( ) 52, 54, 57  
    using 35

## G

Green checkmark  
    saving changes 17

## H

Harmonization  
    key fields 51, 53, 59

Header rows  
    using as field names 24

Headings  
    and column order 10  
    and sort order 15

History  
    project 30

HTML  
    report format 61

Hyperlink  
    command results 13

## I

Identical  
    entries, finding 69  
    records, finding 13

Import  
    CSV files 25

## Index

- Excel worksheets 23, 65
- Microsoft Access data 25
  - using the wizard 22
- INCLUDE() function
  - using 71
- Intactness of data
  - testing 13
- Integrity of data
  - verifying 27
- Invalid characters
  - eliminating 70
- Invalid formats
  - finding 70
- Invalid records
  - filtering out 73
- Invoices
  - duplicate 69

**J**

- Join command
  - compare with Relations 50
  - using 52

**K**

- Key fields
  - preparing 51, 53, 59

**L**

- Length of field
  - checking 51, 53, 59, 66
- Link
  - See Hyperlink, command results
- Log tab
  - viewing command entries 15
- Logical expressions
  - filtering data 16

**M**

- MAP() function
  - using 70
- MATCH() function
  - isolating data 45
- Maximum value
  - Profile command results 12
- MDB files
  - importing 25
- Merchant Category Code list
  - about 5
- Metaphor Corporation
  - about 3
- Methodology
  - documenting 30
- Microsoft Access data
  - importing 25
- Microsoft Excel
  - data, importing 23, 65
  - export format 73
- Minimum value
  - Profile command results 12
- Multiple tables
  - working with 67

**N**

- Navigator pane
  - about 10
- Nested functions
  - using 70
- Notes tab
  - displaying 11
  - using 30
- Number of records
  - counting 11
- Numerals
  - using as character data 24

- Numeric data
  - alpha characters in 30
  - Profile command 11, 12
  - Statistics command 11

**O**

- Objectives
  - tutorial 2
- Operators
  - using in filters 38, 45
- OR statement
  - MATCH() function 45
- Order
  - ascending or descending 15, 37
  - columns in the view 10
- Order of fields
  - arranging in view 67
- Output fields
  - selecting 13
- Overview of data
  - obtaining 11
- Overview tab
  - about 10
  - organizing 64

**P**

- Plain text
  - export format 35
- Policies
  - credit card usage 4, 5
- Profile command
  - using 12
- Project methodology
  - documenting 30
- Project Navigator
  - location 10
- Project planning



## Index

- analysis strategy 7
- Projects, ACL
  - about 10
  - adding data 23
  - creating 22
  - opening 10
  - organizing 64
  - saving 26
  - viewing content 11
- Properties of ACL project
  - displaying 11
- Push-pin button
  - about 66

## Q

- Quick Filter
  - using 55
- Quick Sort
  - about 15
  - using 37
- Quotation marks
  - character data 36
  - date data 16

## R

- Ranges of data
  - finding 43
- Record layout. *See* Data, format information
- Record of commands
  - Log tab 15
- Records
  - counting 11, 68, 72
  - extracting 44, 56, 58
  - invalid
    - filtering out 73
  - sorting 37
- Relations command

- compare with Join 50
  - using 54, 58, 60
- Relations dialog box
  - arranging tables 55
- Reports
  - creating 61
- Results
  - displaying multiple 66
  - recording 53
  - saving 37
- Reverse quotation marks
  - dates 16

## S

- Save
  - automatic 26
- Save the Open Project
  - button 26
- Screen
  - ACL
    - described 10
- Scripts
  - writing 66
- Search for substrings
  - using FIND( ) 36
- Separators
  - field 24
- Single fields
  - selecting 11
- Sort
  - Quick 37
- Soundness of data
  - testing 13
- Sources
  - data 6
- Spaces
  - eliminating 45, 70
  - trimming 43

- Start positions
  - checking 66
- Statistics command
  - using 11
- Strings
  - comparing 70
  - finding 36
- SUBSTR( ). *See* SUBSTRING( ) function
- SUBSTRING( ) function
  - using 52, 54
- Substrings
  - finding 36
- Summarize command
  - using 42
- Summary
  - tutorial objectives 2

## T

- Table Layout window
  - changing data type 35
  - checking data characteristics 27
  - displaying 17
- Table layouts
  - comparing 66
  - editing 28, 65
- Tables
  - adding new columns 52, 54, 58
  - appending 67
  - arranging
    - Relations dialog box 55
  - combining 50
  - Company\_Dept 57
  - creating 23
  - Employees 57
  - empty 53
  - exporting 35, 73
  - extracting records 56, 58
  - joining 52

## Index

- methods of combining 67
- moving to new folder 64
- opening 11
- relating 54, 58, 60
- removing blank entries 55
- saving 26
- Unacceptable\_Codes 23
- Tabs
  - displaying multiple results 66
- Task sheet
  - sample 30
- Test data
  - using to refine strategy 7
- Text
  - export format 35
- Text files
  - importing 23
- Titles of fields
  - importing 24
- Total command
  - using 15, 34
- Total values
  - Profile command results 12
- Totals
  - accumulating per value 46
- Trans\_April worksheet
  - importing 23
- Trans\_April.xls
  - data file 5
- Transaction totals
  - per credit card 46
- Transactions
  - comparing 54
  - identifying unacceptable 53
  - isolating by category 43
  - split 69
  - summarizing by code 42
- Type
  - drop-down list 35
  - Type of data
    - changing 24
    - finding 28
- U**
  - Unacceptable\_Codes.txt file
    - about 5
    - importing 23
  - Use first row as field names
    - option 24
- V**
  - Validate
    - compare with verify 70
  - Validity errors
    - checking 65
  - Values
    - comparing 57
    - finding in multiple tables 50
  - Verify
    - compare with validate 70
  - Verify button
    - checking expressions 36
  - Verify command
    - using 30, 68
  - Views
    - adding new columns 52, 54, 58
    - column sort order 15
    - field order
      - changing 60
    - removing
      - blank entries 55
      - columns 55
- W**
  - Web browser
    - viewing reports 61
  - Width of fields
    - adjusting in view 17
  - Window
    - Table Layout 17, 35
  - Worksheets
    - importing multiple 65
    - selecting for import 23





Data you can trust. Results you can see.