

Databases for Small Business

Essentials of Database Management,
Data Analysis, and Staff Training
for Entrepreneurs and Professionals

—
Anna Manning

Apress[®]

www.it-ebooks.info

DATABASES FOR SMALL BUSINESS

ESSENTIALS OF DATABASE MANAGEMENT,
DATA ANALYSIS, AND STAFF TRAINING FOR
ENTREPRENEURS AND PROFESSIONALS

Anna Manning

Apress®

Databases for Small Business: Essentials of Database Management, Data Analysis, and Staff Training for Entrepreneurs and Professionals

Copyright © 2015 by **Anna Manning**

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

ISBN-13 (pbk): 978-1-4842-0278-4

ISBN-13 (electronic): 978-1-4842-0277-7

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director: Welmoed Spahr

Acquisitions Editor: Robert Hutchinson

Developmental Editor: Douglas Pundick

Technical Reviewer: Richard Roiger

Editorial Board: Steve Anglin, Louise Corrigan, James DeWolf, Jonathan Gennick,

Robert Hutchinson, Celestin Suresh John, Michelle Lowman, James Markham,

Susan McDermott, Matthew Moodie, Jeffrey Pepper, Douglas Pundick, Ben Renow-Clarke,

Gwenan Spearing

Coordinating Editor: Rita Fernando

Copy Editor: Kezia Endsley

Compositor: SPi Global

Indexer: SPi Global

Distributed to the book trade worldwide by Springer Science+Business Media New York, 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springer.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail rights@apress.com, or visit www.apress.com.

Apress and friends of ED books may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Special Bulk Sales–eBook Licensing web page at www.apress.com/bulk-sales.

Any source code or other supplementary materials referenced by the author in this text is available to readers at www.apress.com. For detailed information about how to locate your book's source code, go to www.apress.com/source-code/.

Apress Business: The Unbiased Source of Business Information

Apress business books provide essential information and practical advice, each written for practitioners by recognized experts. Busy managers and professionals in all areas of the business world—and at all levels of technical sophistication—look to our books for the actionable ideas and tools they need to solve problems, update and enhance their professional skills, make their work lives easier, and capitalize on opportunity.

Whatever the topic on the business spectrum—entrepreneurship, finance, sales, marketing, management, regulation, information technology, among others—Apress has been praised for providing the objective information and unbiased advice you need to excel in your daily work life. Our authors have no axes to grind; they understand they have one job only—to deliver up-to-date, accurate information simply, concisely, and with deep insight that addresses the real needs of our readers.

It is increasingly hard to find information—whether in the news media, on the Internet, and now all too often in books—that is even-handed and has your best interests at heart. We therefore hope that you enjoy this book, which has been carefully crafted to meet our standards of quality and unbiased coverage.

We are always interested in your feedback or ideas for new titles. Perhaps you'd even like to write a book yourself. Whatever the case, reach out to us at editorial@apress.com and an editor will respond swiftly. Incidentally, at the back of this book, you will find a list of useful related titles. Please visit us at www.apress.com to sign up for newsletters and discounts on future purchases.

The Apress Business Team

This book is dedicated in loving memory of my wonderful parents, Peter and Pam, who could not have been more supportive of me.

Contents

About the Author	ix
About the Technical Reviewer	xi
Acknowledgments	xiii
Introduction	xv
Chapter 1: How Data Can Benefit Your Small Business	1
Chapter 2: A Small Engineering Company: Case Study	17
Chapter 3: A Small Law Firm: Case Study	27
Chapter 4: A Small Nonprofit: Case Study	39
Chapter 5: Aligning Your Database With the Goals of Your Small Business	51
Chapter 6: Choosing Between Spreadsheets and Databases	67
Chapter 7: Designing Your Small Business Database	79
Chapter 8: Data Protection, Security, and Privacy Policy	123
Chapter 9: Collecting Your Data	131
Chapter 10: Cleansing Your Data	145
Chapter 11: Maintaining Your Database	167
Chapter 12: Searching Your Database	173
Chapter 13: Analyzing Your Data	197
Chapter 14: Reporting Your Small Business Data Analysis	217
Chapter 15: Acting on Reports	229
Chapter 16: Acting on Outside Requests	241
Chapter 17: Archiving Your Database	253
Chapter 18: Storing Your Database in the Cloud	265
Chapter 19: Searching Online Databases	271
Chapter 20: Training Your Staff	281
Appendix	291
Index	313

About the Author



Anna Manning is a founding director of an engineering startup company, having worked for eight years as a data scientist at Manchester University. A specialist in data mining and data protection, she was one of the original designers of data mining based software for protecting confidential data now used by national statistics agencies worldwide. She has also served as a data analyst for a nonprofit organization for a number of years. Dr. Manning's papers in the field of data mining have been published in various research journals. She has a PhD in data mining and a master's degree in informatics from the University of Manchester, an honors degree in law from the University of Law, and a BA honors in mathematics from Trinity College, Dublin. Anna enjoys cycling, debating, and going to the

theater. She is a parish councilor in the village where she lives, near Chester in northwest England.

About the Technical Reviewer



Richard Roiger is a professor emeritus at Minnesota State University, Mankato, where he taught and performed research in the Computer & Information Science department for 27 years. Dr. Roiger holds a PhD in computer & information sciences from the University of Minnesota. After retirement, he continues to serve as an adjunct faculty member, teaching courses in data mining, artificial intelligence, and research methods. Richard is a textbook writer and has published several journal articles. Dr. Roiger is a board member of the Retired Education Association of Minnesota, where he serves as their financial advisor. Richard enjoys interacting

with his grandchildren, traveling with his wife, writing, and pursuing his musical talents.

Acknowledgments

A huge thank you to my husband, Howard, for helping me with the content of the book and providing me with endless encouragement (and cups of tea). The wheelbarrow design in the Smart Wheelbarrows Inc. case study came from Howard. He produced the drawings at the beginning of the chapter and explained which components would be necessary to build the wheelbarrow. He proofread most of the chapters before they were submitted, providing me with invaluable feedback.

I am extremely grateful to my technical reviewer, Richard Roiger, for agreeing to be part of this project, particularly as it was with short notice. We met through a mutual colleague at a data mining conference in Houston, Texas, in 2005 and later when I was visiting Minnesota State University in 2006. His contribution has helped enormously with the quality of the work.

Thank you to Joe Monaghan for agreeing to let me base the Connecting South Side case study on his ideas. Joe does a fantastic amount of work in Liverpool in the UK to help people who are struggling.

Many thanks to my good friend Lisa Marie Allen from California who was able to help me with the American English that was used in this book. Being British, American terminology didn't always come easily to me. She also advised me on the location of the legal case study that's set in California.

Thank you to the team at Apress for everything that they have done. Thank you to Robert Hutchinson for his initial enthusiasm when I approached Apress about this book and for helping me develop the ideas into something marketable. I faced a rather tough time while writing the book—my father died in June 2015 after several months of illness—and I appreciated the support from Rita Fernando who was very patient during this time.

Introduction

Most people who run a small business are familiar with word processing and spreadsheets. However, relatively few are familiar with databases, even though the software is easily available and often free. I have written this book with the aim of addressing this knowledge gap.

This book is written with complete database beginners in mind, with an assumption that you have experience of spreadsheets. The book shows you how to create a database from scratch, all the way through to analyzing the data and presenting it in reports. The aim is that you can build the databases presented in the book and use them as a test suite to experiment on.

Four case studies are considered throughout the book. The aim of these case studies is to provide a good variety of small businesses. The examples are:

- A small online business selling greetings cards
- A small engineering business
- A small legal firm
- A small nonprofit

Even if your business is, for example, a legal firm, it is still worth reading the other examples as well. The important point to remember is that databases are not difficult to learn. If you are familiar with spreadsheets, it is only a small step to using databases.

Relational databases (the databases addressed in this book) have been with us for over 40 years. The fundamental criteria for their design and use have not changed enormously, despite the massive changes in technology that have occurred during this time. At present, databases are following a number of new trends—for example, moving away from the desktop and into the cloud. However, wherever your database is stored and however the instructions are given, you will need to understand the fundamentals of database design in order to be effective.

Database design is part of a process of data handling and processing. Before you can enter data into the database, the data needs to be collected and cleansed. Once it is stored, you will need to know how to take advantage of it by analyzing it and assessing whether your business is meeting its goals. You will need to be able to write clear and meaningful reports based on your analysis.

After reading this book and following the examples, you will be in the position to design a simple database of your own with tables, relationships, and simple queries over those tables. You will be able to analyze the data and write reports based on the results.

Even if you choose not to create a database yourself, after reading this book you will be an “intelligent customer.” You will better understand potential database designs presented to you and will be able to decide whether they are being reasonably priced or, whether, as is often the case, they are an extremely simple design behind a fancy interface.

How Data Can Benefit Your Small Business

Collecting and analyzing data is important to your small business because it can improve efficiency and profitability. Data can provide a record of what has been going on—who your customers are, what their demographic is, and what they've bought. It can enable you to find trends—for example, your customers' favorite products. There are hidden patterns in data that are important to your business, such as groups of products that sell together. Data can provide an archive that can be searched.

When those in small businesses find out what an enormous difference their data can make to them they are generally fascinated to find out more. The fact is, for small businesses, data is collected about many aspects of their operations already. All businesses collect details about income and expenditure in order to satisfy tax requirements. Many businesses collect the names and addresses of their customers so that they can contact them.

The aim of this book is to demonstrate that the data you are collecting has benefits far beyond its initial purpose. The book will explain the value of your data and show you how to take full advantage of it. The book is designed to be an introduction to databases and a signpost to further references, should these be necessary.

As a starting point, I am assuming that you have a basic familiarity with packages such as word processing and spreadsheets, but are completely inexperienced with databases. Further, I am aware that most small businesses don't have only one person working exclusively on their data and that most of their staff has several roles in the company. I also assume that your small business is constrained by a small budget.

You may have been using your data ineffectively for some time without realizing its full potential. If this is the case you will be delighted to have the improved insights and data analysis presented in this book. And, if you are simply storing your data without using it for anything beyond what's absolutely necessary, you will be pleasantly surprised to realize that the first and most important steps of learning the techniques for processing and analyzing your data are well within your reach.

Your business could be at one of a number of stages. You could be starting out and be keen to making sure you are fully abreast of the latest technology. You may have been operating for some time and wish to encourage further growth. Or your business may be experiencing challenging times and is looking for help to move forward. Whatever your circumstances, you will find that your data can make a significant difference to the profitability of your business.

What Is Meant by Data and Databases

The following sections describe the meaning of data and spreadsheets. Databases are introduced by using spreadsheets as a starting point.

Data

Data can take many forms, including numbers, text, images, hyperlinks, and sound. Data can be about any subject at all, or be about more than one subject.

Spreadsheets

As mentioned, I am assuming that you are familiar with spreadsheets, some common examples being Microsoft Excel, LibreOffice Calc, and Google Sheets. I will use a spreadsheet as a starting point for describing databases.

Spreadsheets and databases are both used to store and manage data. A very simplistic way of expressing the difference is to regard a spreadsheet as a large sophisticated programmable calculator and a database as a highly efficient electronic filing system that makes data quickly available for look-up and analysis. The main differences between spreadsheets and databases are how they store and manipulate data and often the amount of data involved.

A spreadsheet stores data items (e.g., numbers, text, and hyperlinks) in cells, with multiple cells represented in a system of rows and columns. Values in one cell can be related to values in other cells with the relationship defined by a formula.

Two of the most serious problems with spreadsheets are as follows:

- *Storing repeated data:* The same data may be stored in multiple spreadsheets. If you need to make changes to the data, the same change is likely required in every location where the data is stored, thus increasing the risk of errors.
- *Finding data:* Finding an item of data in a spreadsheet can involve scanning across numerous columns.

HYPOTHETICAL EXAMPLE OF A SMALL BUSINESS: CARDS FOR EVERYONE INC.

This book uses three case studies for the purposes of illustration: a small engineering firm, a small law firm, and a small nonprofit. They are introduced in the next three chapters. In addition, you'll read about a hypothetical example that depicts a small online business called Cards for Everyone Inc., which sells a range of cards online.

Cards for Everyone Inc. buys cards from a range of suppliers and displays images of them on its web site for sale. The web site was designed by a contractor so that employees of the company can update it themselves. The business has three employees: Pat, Zeph, and Leona.

Pat, the manager, has a number of roles, including buying the supplies and marketing the business. Pat must keep the web site up to date as stock levels change.

Zeph and Leona are assistants who process the orders by locating cards in the stockroom and putting them into envelopes for distribution. When new stock arrives, they place it in the storeroom and pass the details on to Pat so that she can update the web site.

A database addresses both of these issues. Repeated data items are minimized by splitting the data into tables (made up of rows and columns) so that, ideally, each data item is stored only once. Very efficient data retrieval is made possible by relationships defined between the tables: they link the data together and make retrieval efficient. Chapter 6 explains the pros and cons of spreadsheets and databases in more detail.

The next section gives an example of a database used by Cards for Everyone Inc.

A Database Used by Cards for Everyone Inc.

Typical data for a given customer in Cards for Everyone Inc. is their name, address, telephone number, and e-mail address. Table 1-1 shows an example database table containing this data for three customers. Each customer is allocated a row in the table, also known as a *record*. Every record is divided into five columns, one for each part of the record (i.e., Customer ID, Name, ZIP Code, etc.). Every customer record has a unique customer identification number, called the Customer ID. No two customers can have the same Customer ID, thus avoiding duplication.

Table 1-1. Basic Customer Database Table

Customer ID	First Name	Last Name	ZIP Code	Phone	E-Mail
1	Lisa	Garcia	MI 48823	517-xxx-xxxx	Lgarcia@hotmail.com
2	John	Williams	OR 97062	503-xxx-xxxx	Jwilliams@gmail.com
3	Steve	Jones	FL 33901	239-xxx-xxxx	Sjones@aol.com

Further database tables at Cards for Everyone Inc. contain details of products, suppliers, and invoices. As with the customers, each product, each supplier, and each invoice is given a unique identification number in their respective tables. Identification numbers can appear in more than one table in order to help construct relationships. At this point, these tables are for illustration purposes only; Chapter 7 contains more detail about their meanings and construction. Table 1-2 shows a table of five products for Cards for Everyone Inc., with the Product ID in column 1. Notice that the table has a column for the identification number of the supplier in column 2.

Table 1-2. Basic Product Database Table

Product ID	Supplier ID	Product Name	Category	Price (\$)	Quantity Available
11	100	Cats	Birthday	2.00	5
20	200	Roses	Thankyou	3.00	8
23	200	Boats	Birthday	2.50	10
42	300	Hearts	Valentines	4.50	9
61	300	Rabbits	New Baby	5.00	11

A list of hypothetical suppliers is given in Table 1-3.

Table 1-3. Basic Suppliers Database Table

Supplier ID	Supplier Name	ZIP Code	Telephone	E-Mail
100	Special Occasions	IA 52241	319-xxx-xxxx	admin@specialoccasions.com
200	Old Favorites	CA 92591	503-xxx-xxxx	office@oldfavorites.com
300	Handmade Cards	FL 33351	954-xxx-xxxx	enquiries@handmadecards.com

Five invoice records are shown in Table 1-4 with the corresponding Customer IDs in column 2.

Table 1-4. Basic Invoices Database Table

Invoice No.	Customer ID	Date	Amount (\$)
1001	1	11/4/14	...
1002	2	11/4/14	...
1003	3	11/4/14	...
1004	2	12/14/14	...
1005	1	12/18/14	...

At present these four tables are little more than four separate spreadsheets of data, as shown in Figure 1-1.

CUSTOMERS TABLE			
Customer ID	Name	...	
1	Lisa Garcia	...	
2	John Williams	...	
3	Steve Jones	...	
...	

INVOICES TABLE			
Invoice Number	Customer ID	Date	...
1001	1	11/04/14	...
1002	2	11/04/14	...
1003	3	11/04/14	...
...

PRODUCTS TABLE			
Product ID	Supplier ID	Product Name	...
11	100	Cats	...
20	200	Roses	...
23	200	Boats	...
...

SUPPLIERS TABLE		
Supplier ID	Supplier Name	...
100	Special Occasions	...
200	Old Favorites	...
300	Handmade Cards	...
...

Figure 1-1. Separate Database Tables for Cards for Everyone Inc.

A database enables these tables, and their data, to be linked together using the unique identifiers of each table, as shown in Figure 1-2. A database structure such as the one shown in Figure 1-2 prevents you from having to store data repeatedly, risking both typing and mismatch errors. Databases are excellent for pulling data together, separating the themes, and enabling efficient storage, updates, and retrieval.

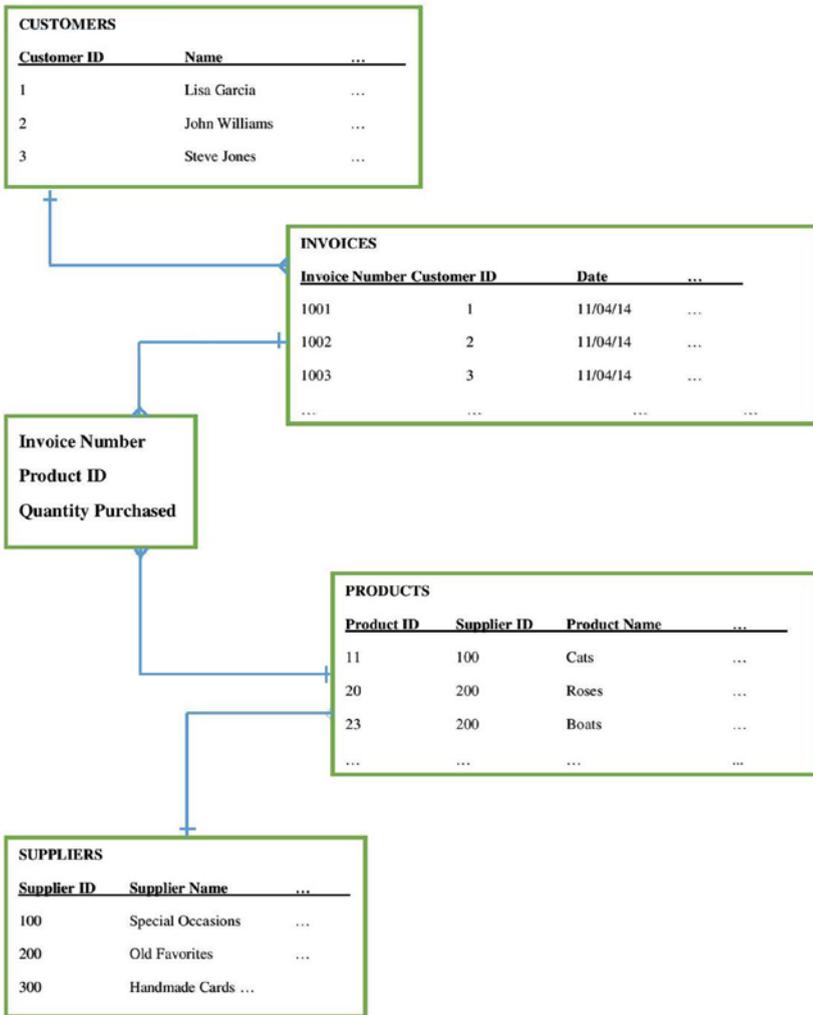


Figure I-2. Linked database tables for Cards for Everyone Inc.

The resulting connected data structure can be viewed in three main ways:

Query: This is a question asked about data in a database, from one or more tables or from other queries. Queries can be used to filter data, to perform calculations, to summarize data, and to perform management tasks. For example, when an order comes in, it would be useful to know if there are enough products in stock and, if not, where to order more from. This involves

querying the Products and Suppliers tables simultaneously. Queries can be highly complex, spanning many tables. Figure 1-3 shows a very simple database query from Cards for Everyone Inc. The query has counted the number of cards of each category in Table 1-2.

The screenshot shows the Microsoft Access interface for a database named 'CARDS FOR EVERYONE INC. DATABASE'. The 'All Access Objects' pane on the left shows a tree view with Tables (CUSTOMERS, INVOICES, PRODUCTS, SUPPLIERS), Queries (COUNT CATEGORY QUERY), Forms (CUSTOMER ENTRY FORM), and Reports (COUNT CATEGORY REPORT). The 'COUNT CATEGORY QUERY' is selected, and its results are displayed in a Datasheet View. The table has two columns: 'Category' and 'CountOfCategory'. The data rows are: Birthday (2), New Baby (1), Thankyou (1), and Valentines (1). The status bar at the bottom indicates 'Record: 1 of 4'.

Category	CountOfCategory
Birthday	2
New Baby	1
Thankyou	1
Valentines	1

Figure 1-3. Simple query for Cards for Everyone Inc.

Forms: These are electronic versions of paper forms and a useful means of visualizing the data. They are composed from one or more tables of the database or from one or more queries and can be used to enter, edit, or display data. Forms are a useful method of searching the data through filtering. For example, forms can be used as invoices, drawing together customer and product details, which can be edited and searched (based on the customer, products bought, etc.) Figure 1-4 shows a very simple form for Cards for Everyone Inc., which provides a method for viewing data in the Customers table and a means for entering data. The form is a far more user-friendly way than inputting data directly into the table.

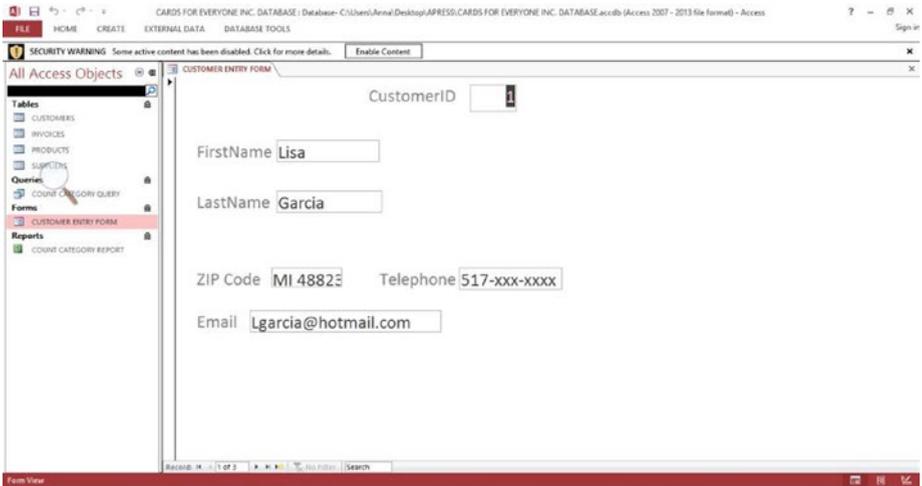


Figure I-4. Sample form for Cards for Everyone Inc.

Reports: They enable a document to be designed from database tables or queries that can be shared easily in electronic form or viewed away from a computer. Reports can be used to summarize which products have been selling well and which customers have been buying them. Figure I-5 shows a report of the query data from Figure I-3.

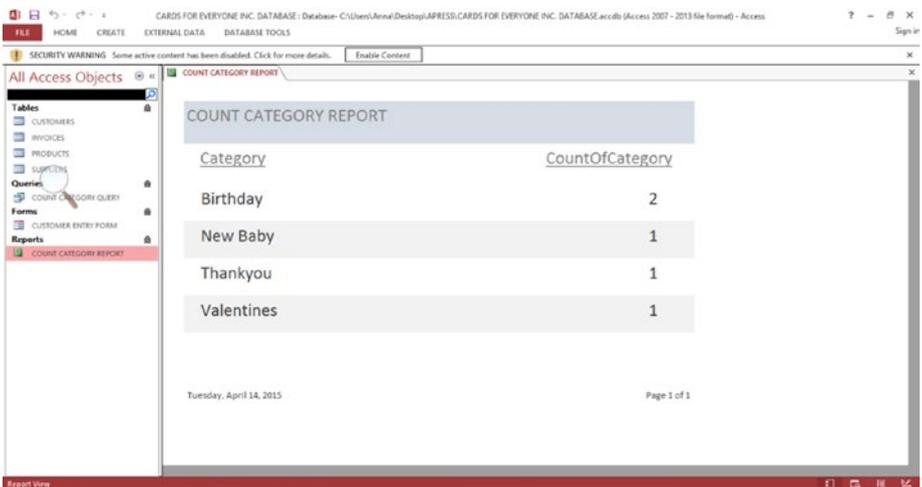


Figure I-5. Sample report for Cards for Everyone Inc.