Module 3: Normalization

Lab: Normalizing Data

Scenario

The retail company that you work for uses a spreadsheet to track customer orders, but as the number of orders increases, the data has become more difficult to maintain. Management have decided to switch to a relational database. You have been charged with designing a new database schema for the orders data. You have obtained a reduced copy of the current orders spreadsheet that contains a small representative sample of the data; you will use this to help you to design the schema.

Objectives

After completing this lab, you will have:

- Normalized a data set to first normal form.
- Normalized a data set to second normal form.
- Normalized a data set to third normal form.
- Denormalized a data set.

Lab Setup

Estimated Time: 60 minutes

Ensure that the 10985C-MIA-DC and 10985C-MIA-SQL virtual machines are both running, and then log on to 10985C-MIA-SQL as **ADVENTUREWORKS\Student** with the password **Pa$$w0rd**.
Estimated Time: 60 minutes

Exercise 1: Normalizing to First Normal Form

Scenario

Having obtained the sample data set, you will analyze it and normalize it to first normal form.

The main tasks for this exercise are as follows:

1. Review the Data Set
2. Normalize the Data Set to First Normal Form

Task 1: Review the Data Set

- Review the data on the **Raw Data** tab in **D:\Labfiles\Lab03\Starter\Normalize to 1NF.xlsx** and note any aspects that cause it to violate first normal form.

Task 2: Normalize the Data Set to First Normal Form

1. Decompose the table in the **Raw Data** worksheet to create a new version of the data that is in first normal form. Document your solution in the **Normalize to 1NF** worksheet.
2. On the **1NF Solution** worksheet, review the suggested solution.

3. Close **Normalize to 1NF.xlsx**, saving any changes if prompted.

> **Results**: In this exercise, you normalized a data set to first normal form.

Exercise 2: Normalizing to Second Normal Form

Scenario

Having normalized the data set to first normal form, you will now normalize it to second normal form.

The main tasks for this exercise are as follows:

1. Review the Data Set in First Normal Form
2. Normalize the Data Set to Second Normal Form

Task 1: Review the Data Set in First Normal Form

- In the **D:\Labfiles\Lab03\Starter** folder, open **Normalize to 2NF.xlsx**, and then on the **Normalize to 2NF** tab, review the table to identify any features of the data that might cause the table to violate second normal form.

Task 2: Normalize the Data Set to Second Normal Form

1. Decompose the table in the **Normalize to 2NF** worksheet to create a new version of the data that is in second normal form. Document your solution in the free space in the **Normalize to 2NF** worksheet.
2. On the **2NF Solution** worksheet, review the suggested solution.
3. Close **Normalize to 2NF.xlsx**, saving any changes if prompted.

> **Results**: In this exercise, you normalized the data set to second normal form.

Exercise 3: Normalizing to Third Normal Form

Scenario

Having normalized the data to second normal form, you will now normalize it to third normal form.

The main tasks for this exercise are as follows:

1. Review the Data Set in Second Normal Form
2. Normalize the Data Set to Third Normal Form

Task 1: Review the Data Set in Second Normal Form

- In the **D:\Labfiles\Lab03\Starter** folder, open **Normalize to 3NF.xlsx**, and then on the **Normalize to 3NF** tab, review the table to identify any features of the data that might cause the table to violate third normal form.

Task 2: Normalize the Data Set to Third Normal Form

1. Decompose the table in the **Normalize to 3NF** worksheet to create a new version of the data that is in third normal form. Document your solution in the free space in the **Normalize to 3NF** worksheet.
2. On the **3NF Solution** worksheet, review the suggested solution.
3. Close **Normalize to 3NF.xlsx**, saving any changes if prompted.

> **Results**: In this exercise, you normalized the data set to third normal form.

Exercise 4: Denormalizing Data

Scenario

Having normalized the data set to third normal form, you will now denormalize a table. During your research, you identified the types of queries that will be issued against the database. These include a query that returns the total sales amount for each order. This query runs frequently, so you want to investigate ways to ensure that the response times for it are as fast as possible. To do this, you will denormalize the Orders table.

The main tasks for this exercise are as follows:

1. Review the Data Set in Third Normal Form
2. Denormalize the Orders Table

Task 1: Review the Data Set in Third Normal Form

- In the **D:\Labfiles\Lab03\Starter** folder, open **Denormalize.xlsx**, and then on the **Denormalize** tab, review the tables.

Task 2: Denormalize the Orders Table

1. Denormalize the **Orders** table to meet the requirements of the scenario for this exercise.
2. On the **Denormalized Solution** tab, review the suggested solution.
3. Close **Denormalize.xlsx**, saving any changes if prompted.

> **Results**: In this exercise, you denormalized a table.

Lab Review

Question: In the suggested solution, the Orders table has a surrogate key as its primary key. Why do you think this key was chosen instead of a composite candidate key, such as the OrderDate and Customer columns?

Question: In the suggested solution, in the Customers table, are there any other columns that you could break down even further? If so, why might you do this?

©2016 Microsoft Corporation. All rights reserved.

The text in this document is available under the Creative Commons Attribution 3.0 License, additional terms may apply. All other content contained in this document (including, without limitation, trademarks, logos, images, etc.) are **not** included within the Creative Commons license grant. This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

This document is provided "as-is." Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it. Some examples are for illustration only and are fictitious. No real association is intended or inferred. Microsoft makes no warranties, express or implied, with respect to the information provided here.