CIS 4349

Assignment # 1

(50 points)

Due Date: Beginning of class on Tuesday, February 26, 2019

***Introduction:***

You are to design a data warehouse schema for a national taxi company. In your design, you should assume that the company gathers business data from all cities in the US. Using the data warehouse, the company will be able to develop a BI system to determine how successful their taxi service business is. Additionally, the BI solution can potentially help increase the profit on the basis of knowledge (intelligence) generated by the system.

Historically, the company has stored its daily transactional data from each trip of taxi service in a SQL Server database. The data includes driver information, total charge, total mileages, date/time, and the location of each service (trip) provided.

The taxis pick up passengers on different streets, which can be located in different cities in the USA. The system needs to provide daily, monthly and yearly reports on revenues of the company; revenues and total mileages based on streets, cities, and states; and the analysis of relationships, if any, between state, city, driver, month of service and generated revenues.

***Planning Documents:***

None available. You will be responsible for creating:

1. An Excel worksheet to document your data warehouse design (similar to the planning document we used in the class demo).
2. A SQL Script to create a database named **TaxiServiceDB-XXXX**, where XXXX represents your NetID. If your NetID includes hyphens or underscore, please remove themk when using your NetID in database names e.g. a\_h1234 would become ah1234. The database will be created on the university database server accessible at CISSQL.MATRIX.TXSTATE.EDU\CIS4349SP19S251.

**Please note that any reference to “XXXX” in these specifications refers to your NetID, without hyphens or underscores – so if the NetID is a\_h123, it would become ah123.**

***Requirements***

1. Review the OLTP database design presented later in these specifications. Perform the steps we discussed in class to analyze the OLTP database design and identify the tables/data that should be included in your OLAP Data Warehouse database.
2. Next, determine what design pattern you will use (Star, Snowflake, or hybrid of the two) to design your data warehouse. You will identify the measures and dimensions and then determine the design of the Fact table and all Dimension Tables. Finally, present this chosen design in database diagram format for the Data Warehouse you would design for the Taxi Service.
3. Develop an Excel workbook to act as your planning document for the duration of this project. Use the format we have used in class. Complete this workbook to document the data warehouse design you developed in step 2 above. Note: You will be expected to accomplish a similar task on examination.

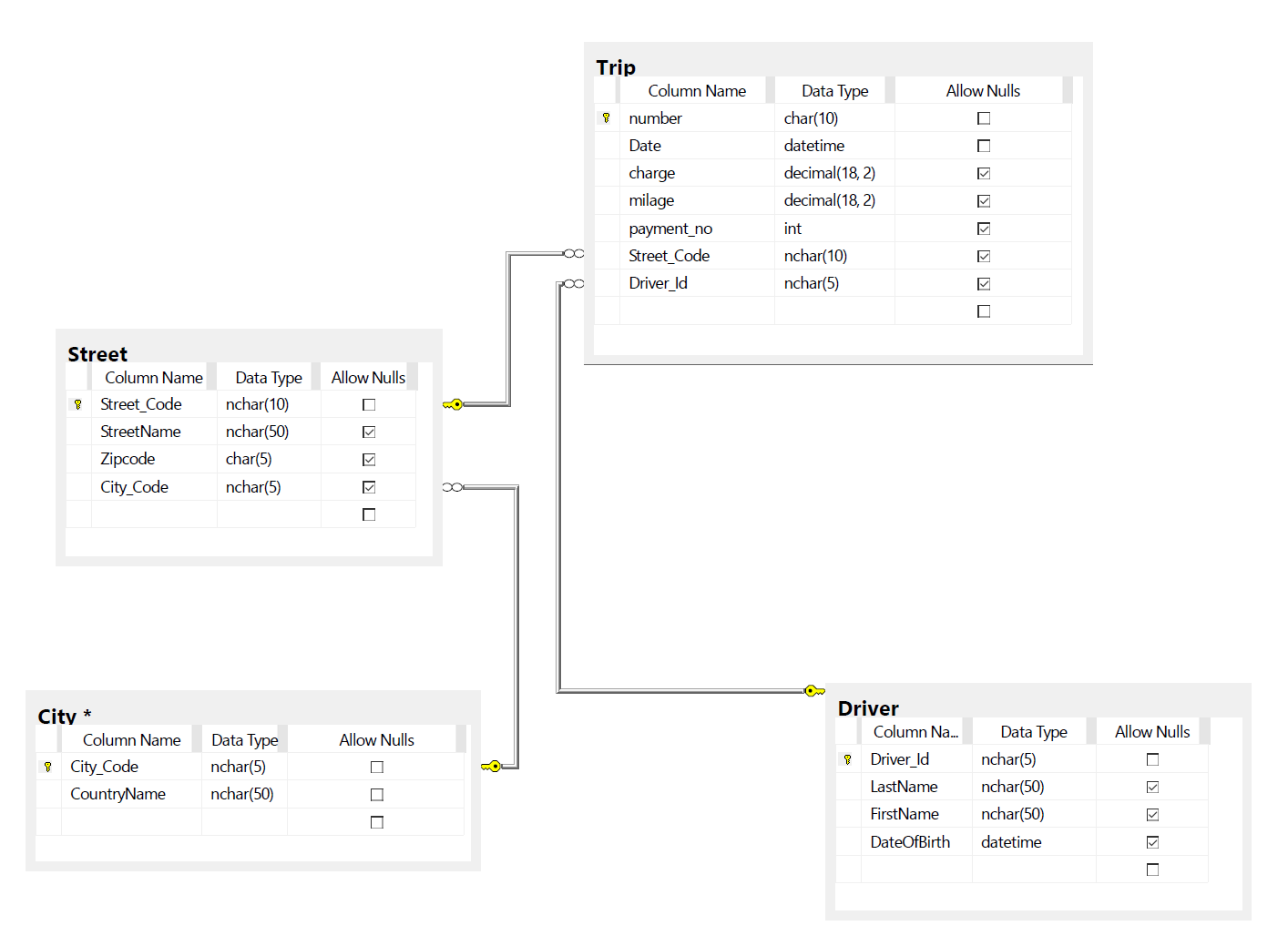
Naming Conventions:

* 1. Fact Table shall be named as Fact*TableName*, where TableName is a meaningful name. There is no length restriction so make the name really meaningful.
  2. Dimension Tables shall be named as Dim*TableName,* again TableName is a meaningful name representative of the dimension it represents.
  3. Each column name shall be a meaningful name to represent the attribute it names. For example, an appropriate name to represent a city field may be CityName rather than just City. Again, there is no length restriction so DO NOT ABBREVIATE or USE SHORT NAMES. Make the names meaningful to represent the field data.

1. The Excel planning document should show the data source, data items to be extracted, source data type, destination, destination data item names, destination data type and details of any conversion to be performed as you move the data from the source (OLTP database) to destination (OLAP Warehouse). This step will help you write the ETL SQL script in the next assignment.
2. Based on the OLTP design presented, develop a SQL script to implement and populate the ServiceDB-XXXX OLTP database on the university database server. The script itself should be named “ServiceDB-XXXX.SQL” and saved in the “Scripts” folder on your USB drive. Just write/type the script. Please do not deploy the database yet. You will do this in the next assignment. Save this script as “ServiceDB-XXXX.sql” in the “Scripts” folder

***Submission Requirements:***

1. The assignment is to be completed individually by each student. This is a requirement. Collaborative submissions and work will be treated as academic dishonesty and handled accordingly.
2. Your complete and working BI Solution project should be available in the “E:\XXXX\Assignment “ folder on a USB flash drive. Your USB drive should have the following artifacts: (XXXX is your NetID without any hyphens or underscores).
   1. A BI Solution created in VS 2017 and named “TaxiService-XXXX”
   2. A BI Project named “TaxiService-Assignment1-XXXX”
   3. This project should contain a folder named “SolutionDocuments”. This folder should include:
      1. Your completed Excel planning document named “TaxiServicePlanningDocument-XXXX”
      2. The SQL script you developed to create the ServiceDB-XXXX database.
3. Submit the USB drive with your BI solution as well as the following printouts arranged per sequence below:
   1. A cover page with your name, project name, and submission date.
   2. Printout of the OLAP diagram you developed based on the data warehouse design.
   3. Printout of the Excel worksheet documenting the source/destination of data based on the design presented in (b).
   4. Printout of the SQL script that will be used to deploy the ServiceDB-XXXX database.
4. Place the USB drive and your printouts on the instructor’s desk before the start of class on the due date. Once the instructor starts class, the assignment is considered late and will not be accepted for grading.



***Figure 1. Design of OLTP Database ServiceDB***