ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

ITSC

DEPARTMENT OF SOFTWARE ENGINEERING

Course: Big Data
Lab Report – 1

PREPARED BY:

- **Fruit Teklu** **ATR/7656/07**
- **Kidus Mamuye** **ATR/6157/07**
- **Yared Taddese** **ATR/0593/07**
- **Yehualashet Abebe** **ATR/4554/07**

ADVISORS: Alazar Alemayehu

Dec, 2018

ABSTRACT

R is a programming language and software environment for *statistical analysis*, *graphics representation* and *reporting*. R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team. The main features of R:

- R is a well-developed, simple and effective programming language which includes conditionals, loops, user defined recursive functions and input and output facilities.
- R has an effective data handling and storage facility,
- R provides a suite of operators for calculations on arrays, lists, vectors and matrices.
- R provides a large, coherent and integrated collection of tools for data analysis.
- R provides graphical facilities for data analysis and display either directly at the computer or printing at the papers.

R has a built-in package named "**RMySQL**" which provides native connectivity between with MySql database.

INTRODUCTION

We know that SQL has limited numerical and statistical features, For example, it has no least squares fitting procedures, and to find quantiles requires a sophisticated query.
Not only are basic statistical functions missing from SQL, but in many cases the numerical algorithms used in the basic aggregate functions are not implemented to safeguard numerical accuracy. Also, the wide range of data types may have drawbacks when it comes to performing arithmetic calculations across a row, as some of the conversions from one numeric type to another may produce unexpected truncation and rounding. For these reasons, it is good to use some other tools or programs to do that part and R is the number 1 from the list. The DBI package in R provides a uniform, client side interface to different database management systems, such as MySQL, PostgreSQL, and Oracle. For This assignment we use RMYSQL package, which extends the DBI package to provide a MySQL driver and the detailed inner workings for the generic functions to connect, disconnect, and submit and track queries. The RMYSQL package uses client-side software provided by the database vendor to manage the connection, send queries, and fetch results. The R code the user writes to establish a MySQL driver, connect to a MySQL database, and request results is the same code for all SQL-standard database managers.

Problem Description

Perform CRUD (Create, Read, Update and Delete) Operation using R to the database.

For This lab we will create three table Author ,Category and Book table

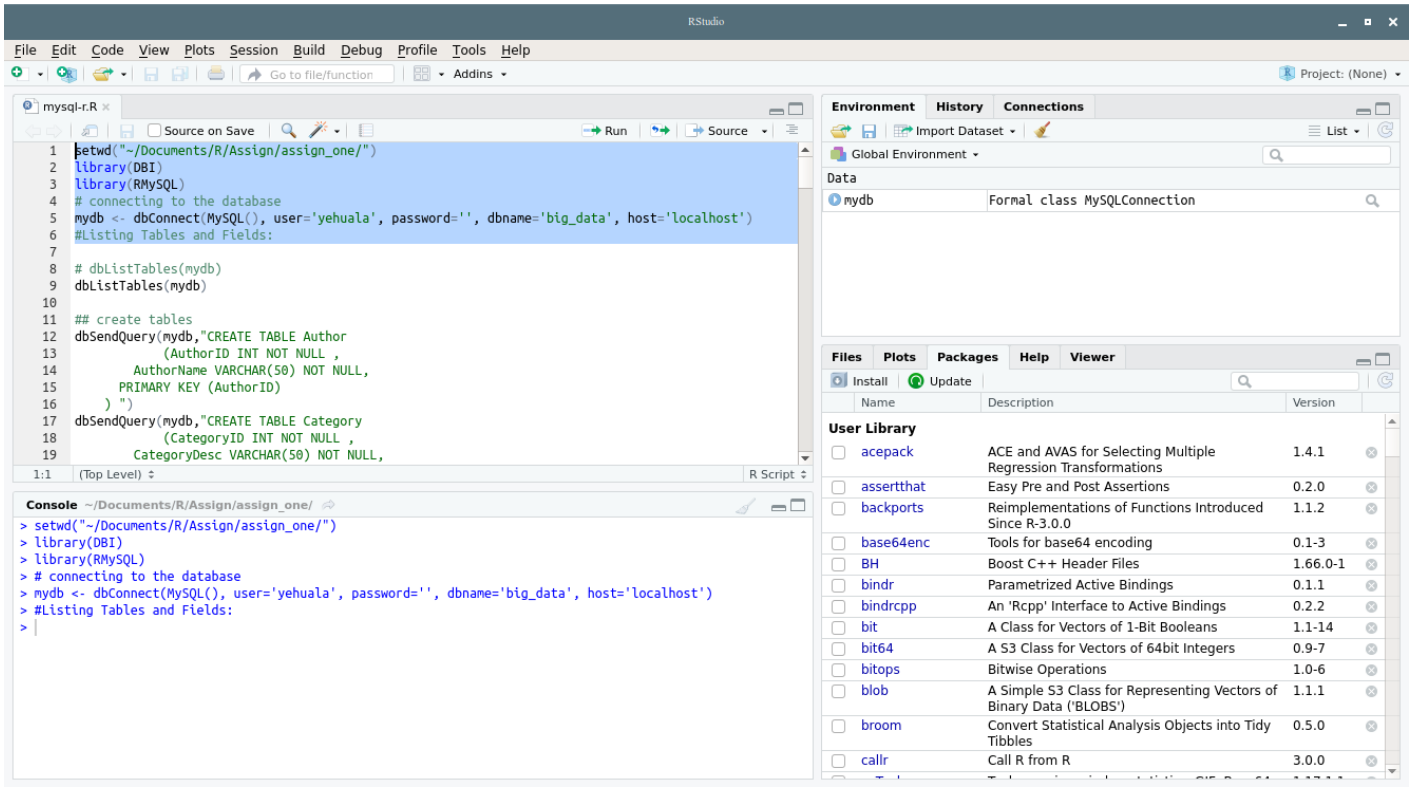

Image 1: Connect to the Database

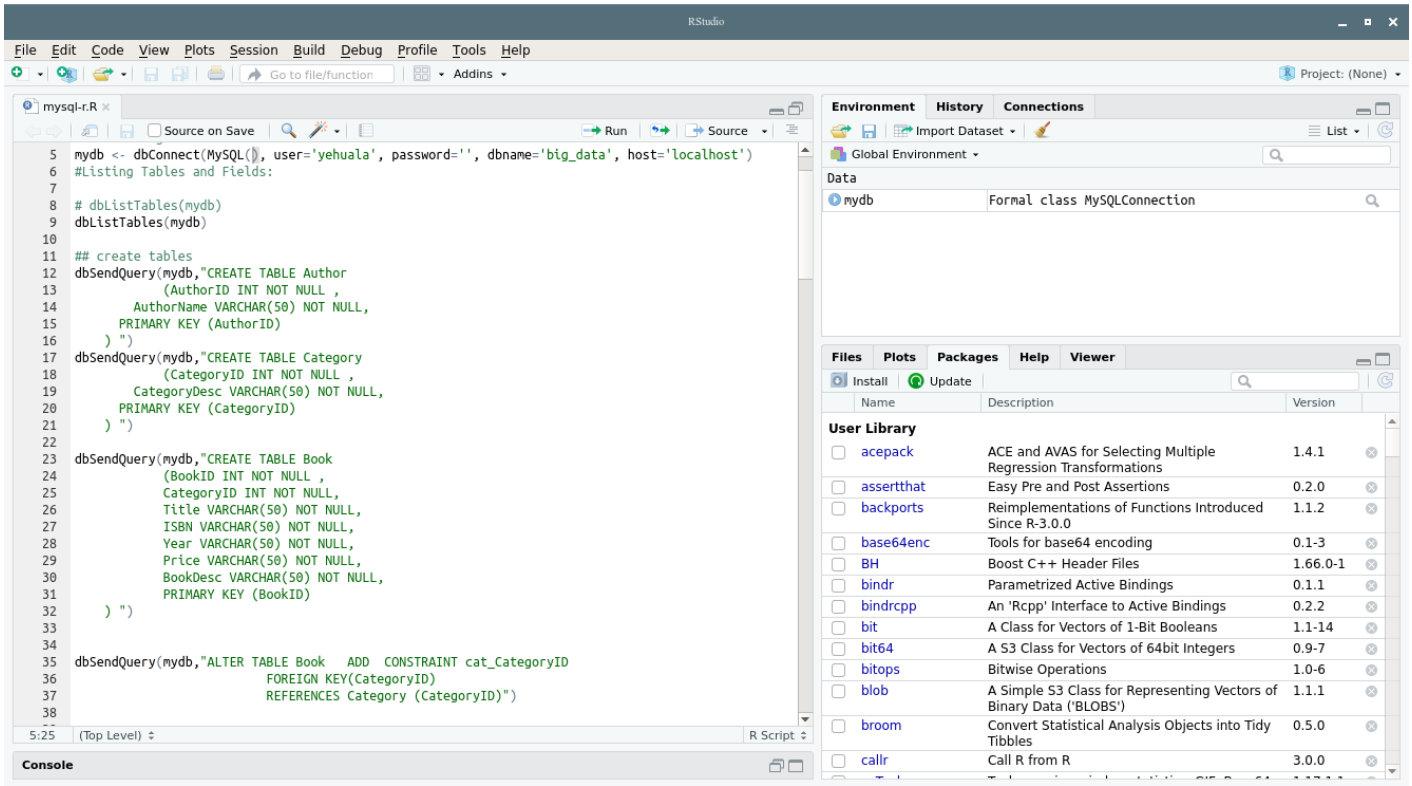

Image 2: Create Table using Normal query

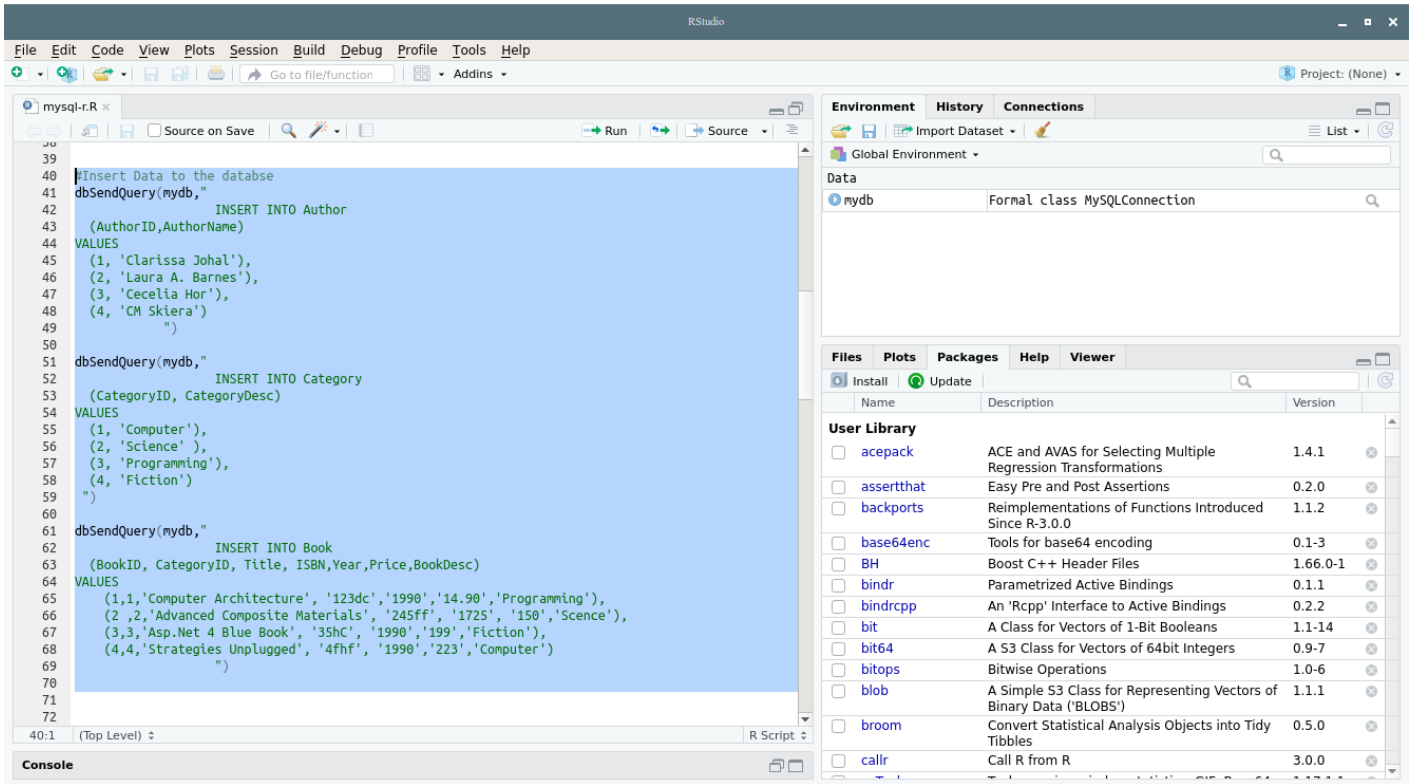

Image 3: Insert Data to the database

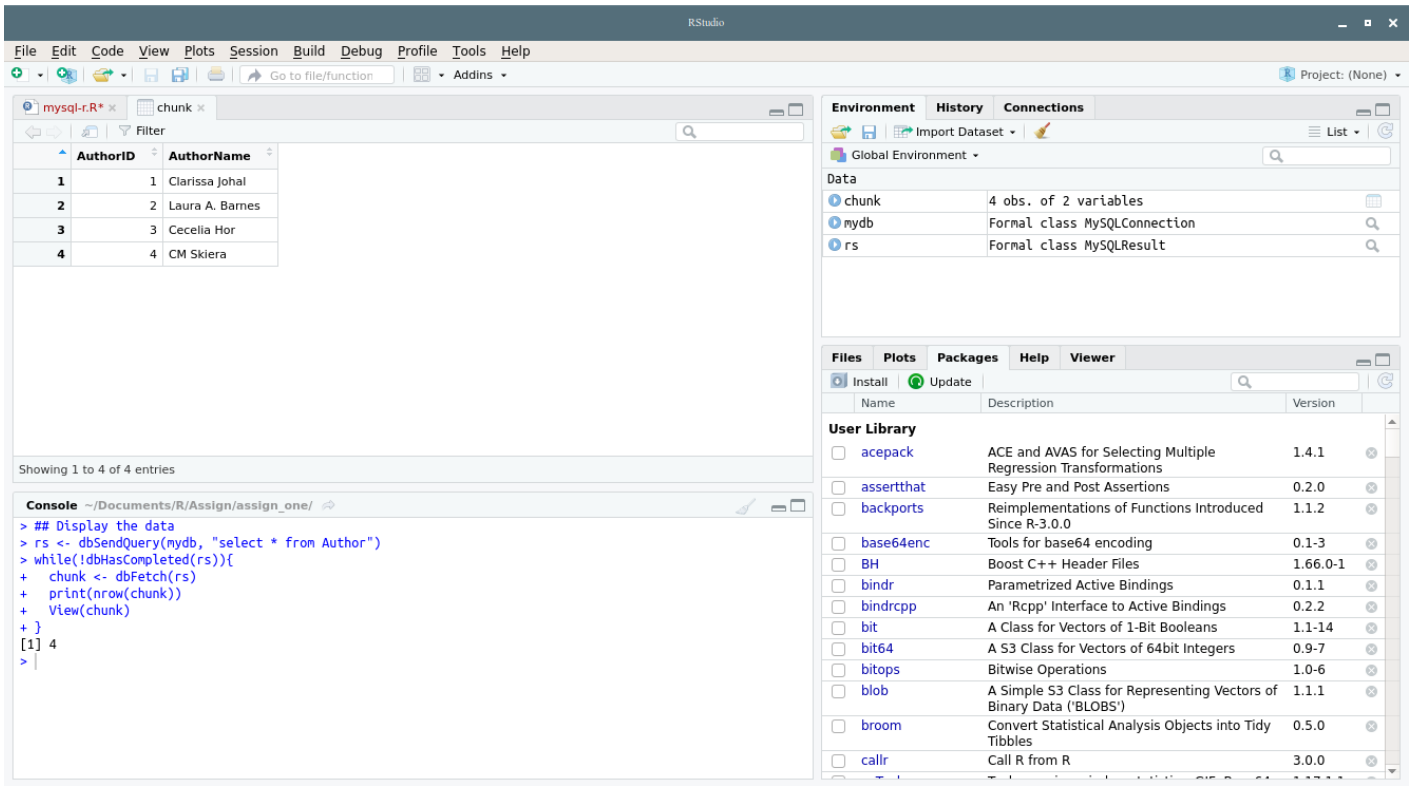

Image 4: Display the data for Author Table

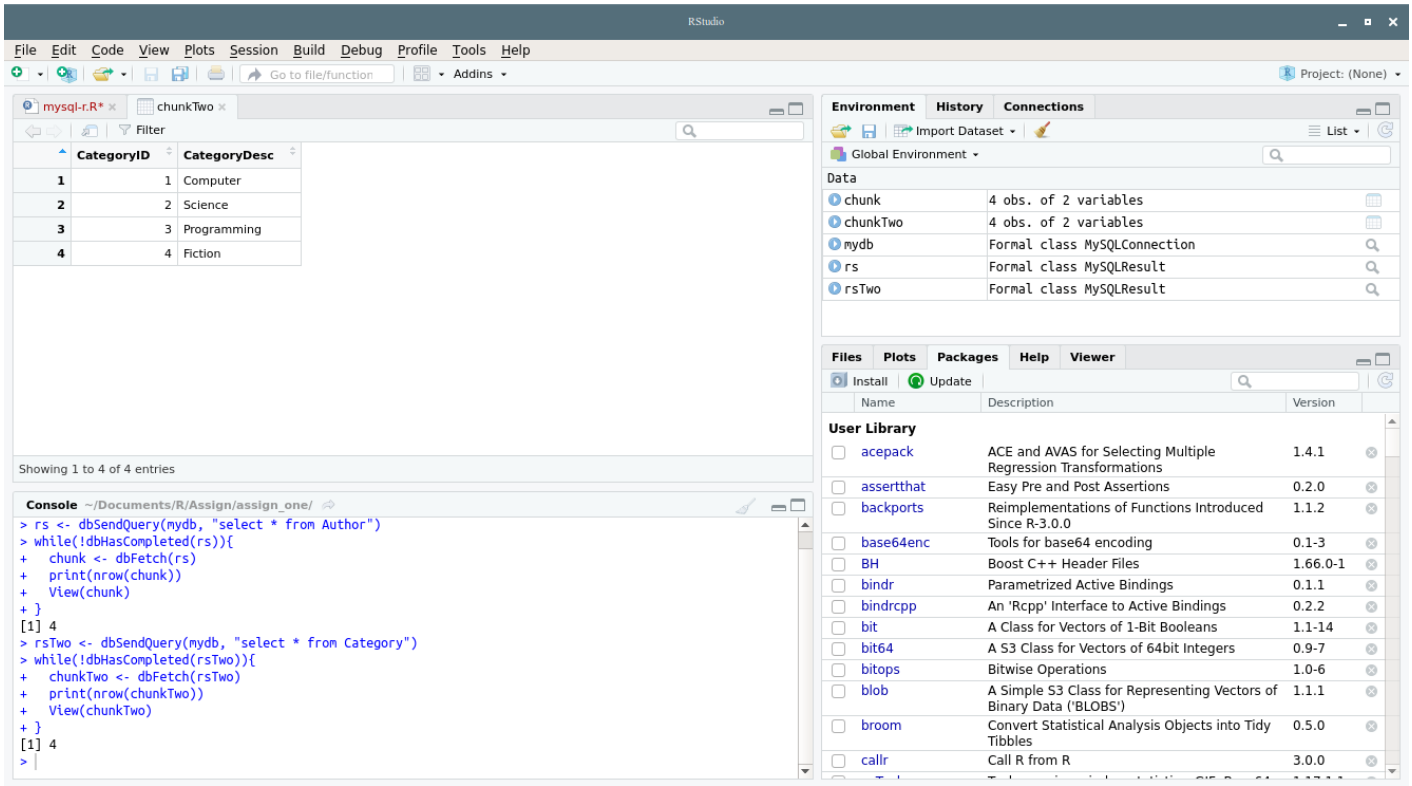

Image 5: Display the data for Category table

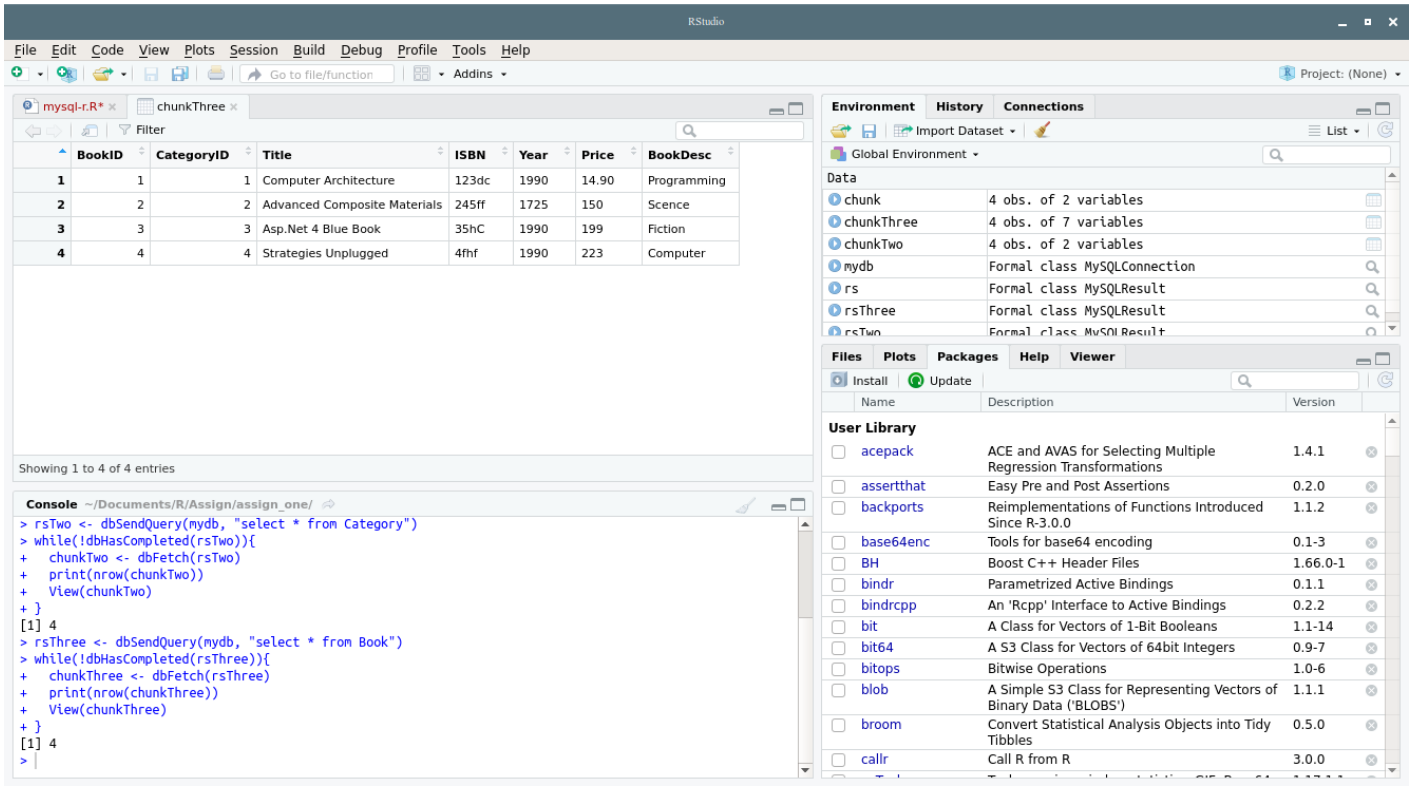

Image 6: Display the data for Book table

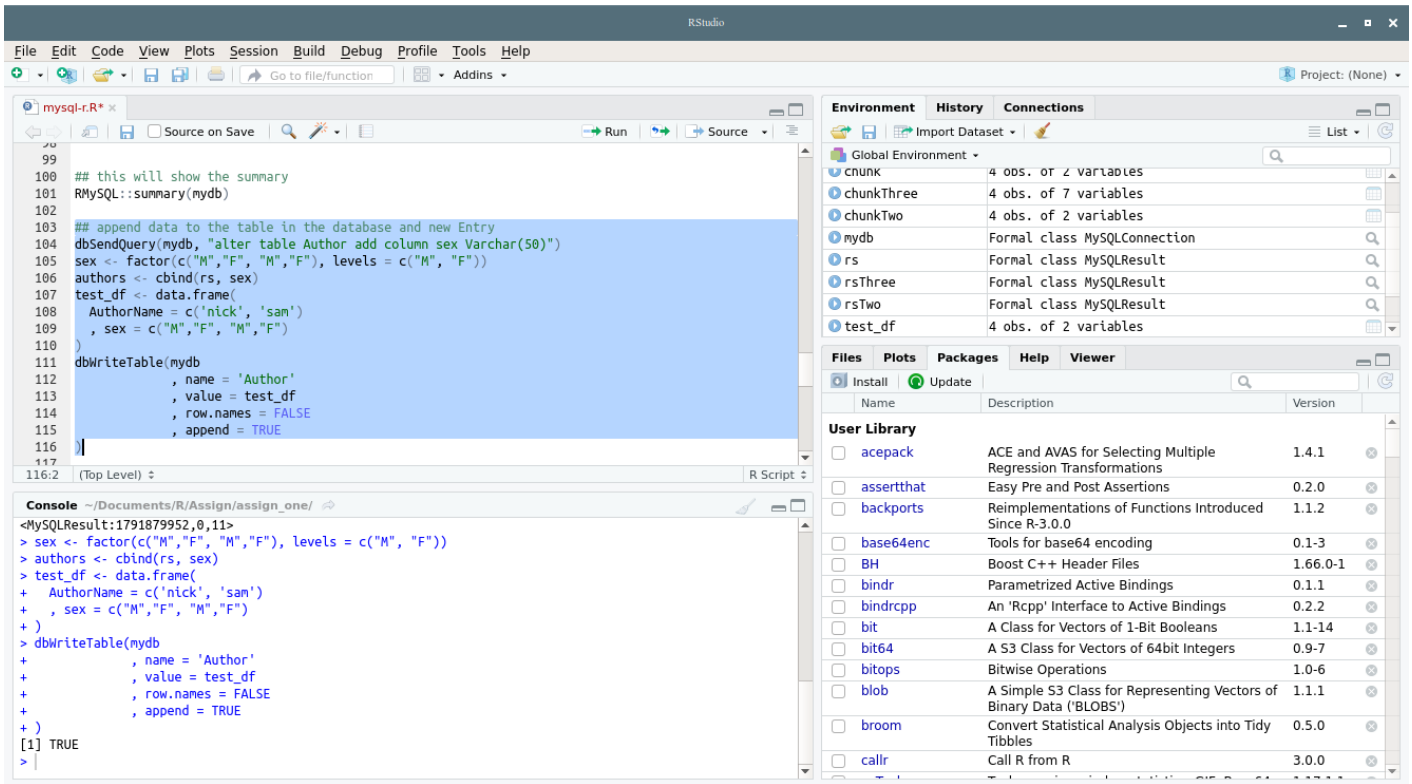

Image 7: Add new Column to the table and add new entry to that table

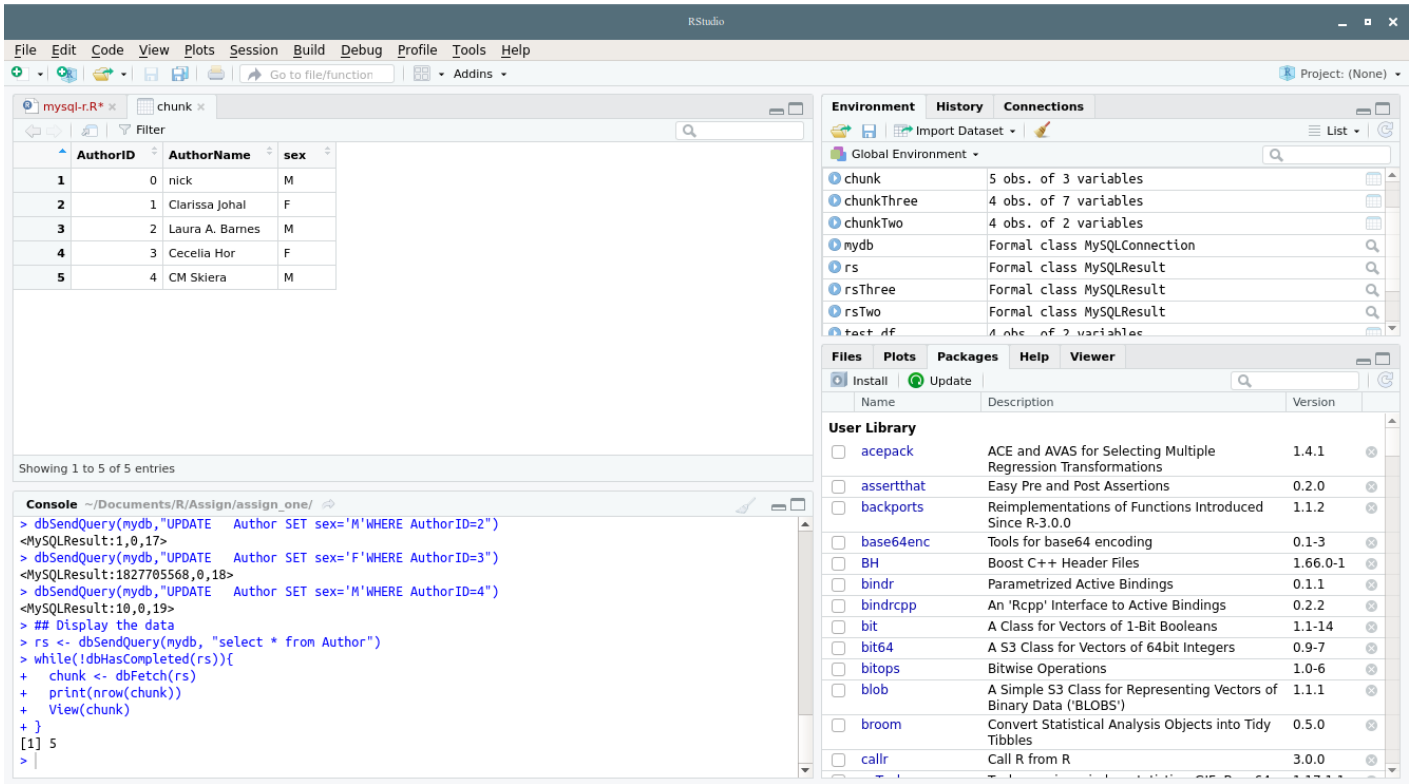

Image 8: Display the new add column

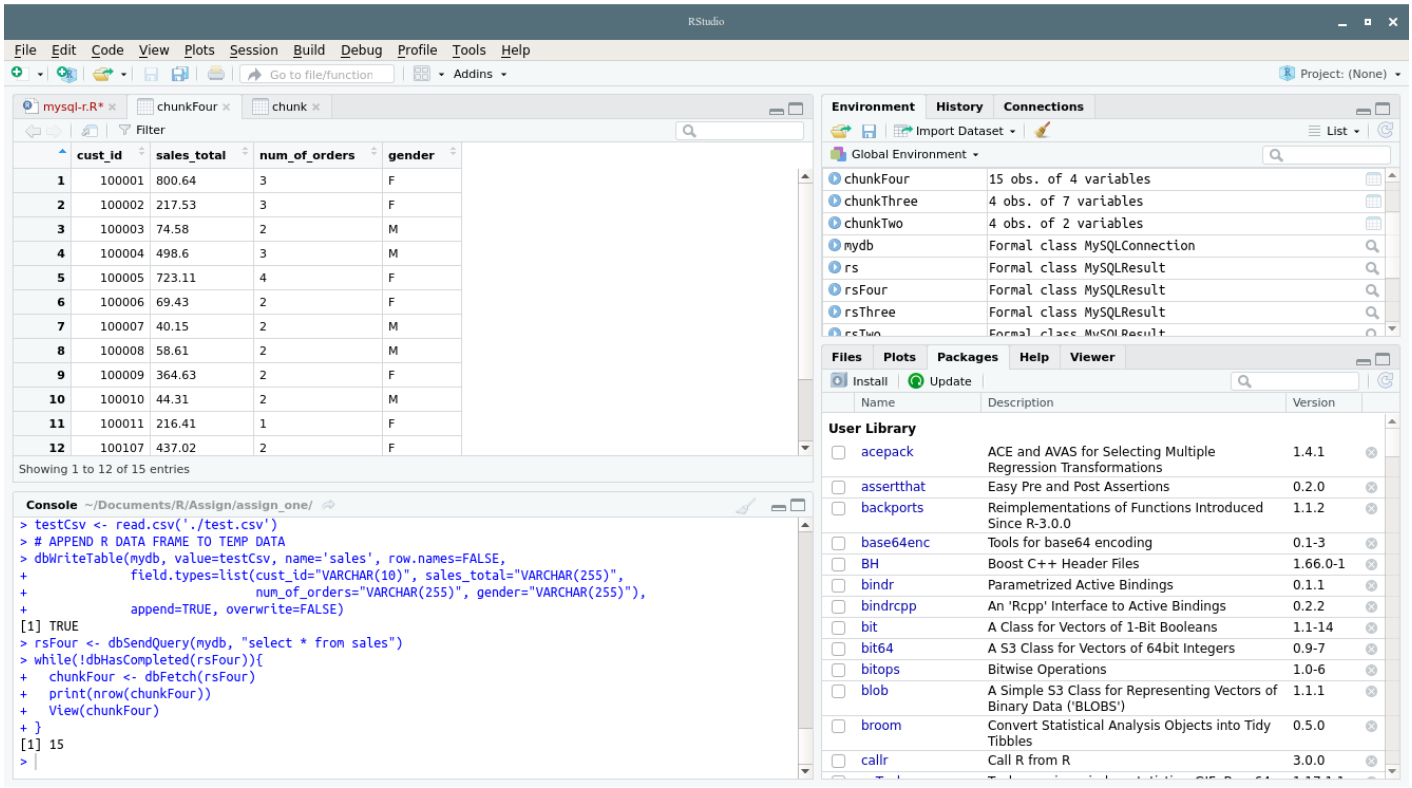

Image 9: Create new table from csv