

Explain the function of each option in the context of import operation.

5. The following is a list of command options for Sqoop export operation:

- `connect jdbc:mysql`
- `//< database_endpoint >/<database name>`
- `table <table_name>`
- `username`
- `password`
- `export-dir`

Explain the function of each option in the context of export operation.

## Practice Exercises

The following exercises use data files stored at the Student Resources website. Go to the text website at <http://www.prospectpressvt.com/textbooks/asllani-big-data/>, scroll to the horizontal red menu bar, and click on “Student Resources.” You’ll find the data files organized by chapter. Please download the files to your computer before using them.

**Exercise 7-1** For this exercise, you will use the following files: `products.txt`, `orders.txt`, `order_details.txt`, `categories.txt`, and `customers.txt`.

1. Create a database using a MySQL command like `create table <table_name>` or `insert into table <table_name> values`.
2. Upload all the tables from MySQL to HDFS using the Sqoop `import` command.
3. Using Pig, perform the following operations (some of which may take several Pig Latin commands):
  - a. Load data from `products` and `categories` into Pig.
  - b. Find the number of products for each category and upload the results back to HDFS.
  - c. Load data from `orders` and `order_details` into Pig.
  - d. Find the total amount for each order and upload the results back to HDFS.
4. Place the results from b and d above into the MySQL database using the Sqoop `export` command.
5. Perform a `select all` to verify that the results were exported correctly.

**Exercise 7-2** Using the same files, perform the same operations as in Exercise 1 but using Hive instead of Pig. Specifically:

1. Create a database using MySQL (if you have not created the database before).
2. Upload all the tables from MySQL to HDFS (if you have not uploaded them before) using the Sqoop `import` command.

6. The following is a list of command options for exporting a table directly into Hive:

- `connect jdbc:mysql`
- `//<database_endpoint>/<database name>`
- `table <table_name>`
- `username <user name>`
- `password`
- `hive-import`

Explain the function of each option in the context of importing from or exporting to Hive.

3. Using Hive, perform the following operations:
  - a. Load data from `products` and `categories` into Hive.
  - b. Find the number of products for each category and upload the results back to HDFS.
  - c. Load data from `orders` and `order_details` into Hive.
  - d. Find the total amount for each order and upload the results back to HDFS.
4. Place the results from b and d back into the MySQL database using the Sqoop `export` command.
5. Perform a `select all` to verify that the results were exported correctly.

**Exercise 7-3:** For this exercise, you will use the following two files: `CustomerDetails.txt` and `LatePayments.txt`. Perform the following operations:

1. Log in to MySQL, create the database, and insert the records into the two tables of this database.
2. Verify that both tables are created correctly and contain records.
3. Load the data from these two tables into HDFS using the Sqoop `import` command.
4. Using Hive, create a similar database and create and populate both tables.
5. Perform the following:
  - a. Display customers (ID, name, year of birth, and zip code) from the state of New York.
  - b. Display the total number of customers who have late payments for each zip code.
  - c. Display the customers who have any late transactions and the date of each late transaction for the customers born before 1985.
  - d. Display the cities and the number of customers who do not have a credit card for each city, sorted by the number of customers.
  - e. Display the states and the number of customers who have late payments and use mobile banking in each state.