Exercise 01.03: Working with Partitions

In this exercise, you will:

- Create a new table that allows querying videos by title and year using a composite partition key

Background

Your peers need to query videos based on 'title' and 'added_year'. The new columns for this table are:

Column Name	Data Type
title	text
added_year	int
added_date	timestamp
description	text
user_id	uuid
video_id	timeuuid

Steps

1. Use a text editor to view the file */home/ubuntu/labwork/partitions/videos_by_title_year.csv* file.

2. At prompt, navigate to the directory */home/ubuntu/labwork/partitions/*.

3. Start `cqlsh`.

4. Create a second table in the *killrvideo* keyspace called 'videos_by_title_year.'

    a. Use the column names and data types as shown in the structure above.

    b. Confirm users can query this table on both 'title' and 'added_year' by combining them into a composite partition key.

```
cqlsh> USE killrvideo;
cqlsh:killrvideo> CREATE TABLE videos_by_title_year (
title TEXT,
added_year INT,
added_date TIMESTAMP,
description TEXT,
```

```
user_id UUID,
video_id TIMEUUID,
PRIMARY KEY ((title, added_year))
);
```

5. Load the data from the *videos_by_title_year.csv* file using the COPY command.

```
cqlsh> COPY videos_by_title_year (title, added_year, added_date,
description, user_id, video_id) FROM 'videos_by_title_year.csv' WITH
HEADER=true;
Using 1 child processes

Starting copy of killrvideo.videos_by_title_year with columns [title,
added_year, added_date, description, user_id, video_id].
Processed: 430 rows; Rate:     858 rows/s; Avg. rate:    1225 rows/s
430 rows imported from 1 files in 0.351 seconds (0 skipped).
```

6. Use the command SELECT to COUNT(*) the number of imported rows. Confirm this amount matches the rows the COPY command reported as imported.

```
cqlsh> SELECT COUNT(*) FROM killrvideo.videos_by_title_year;
 count
-------
   430

(1 rows)
```

7. Try running queries on the 'videos_by_title_year' table to query on a specific 'title' and 'added_year'.

    a. Below are some example 'title' and 'added_year' entries. How would you construct a `cqlsh` query to pull up information on each video title?

```
SELECT * FROM videos_by_title_year WHERE title = 'Introduction To
Apache Cassandra' AND added_year = 2014;
SELECT * FROM videos_by_title_year WHERE title = 'Sleepy Grumpy Cat'
AND added_year = 2015;
SELECT * FROM videos_by_title_year WHERE title = 'Grumpy Cat: Slow
Motion' AND added_year = 2015;
SELECT * FROM videos_by_title_year WHERE title = 'AzureDev' AND
added_year = 2015;
```

Example entries:

title	added_year
Introduction to Apache Cassandra	2014
Sleepy Grumpy Cat	2015
Introduction to Azure Search	2015
AzureDev	2015

8. What error is returned when you try to query on just 'title' or just 'added_year'? Why?
9. Exit `cqlsh`.

END OF EXERCISE