Part 1: Writing Data to an Excel File
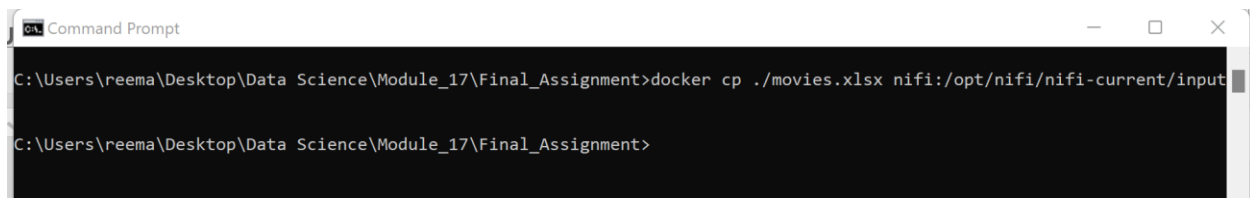
1. Provide a screenshot to show that you successfully created the two folders: input and output.
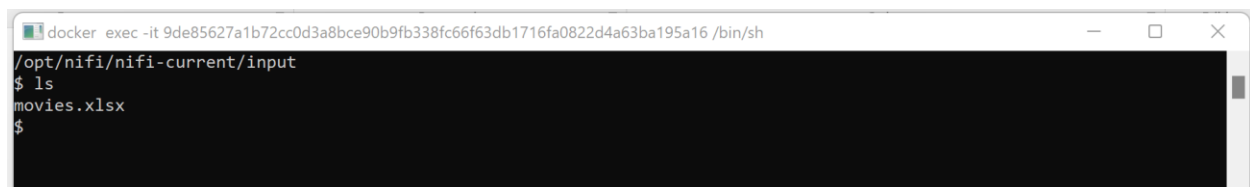

2. Provide a screenshot to show that you successfully copied the movies.xlsx file in the input folder.

3. Provide a screenshot to show that you successfully created the Assignment17 process group.
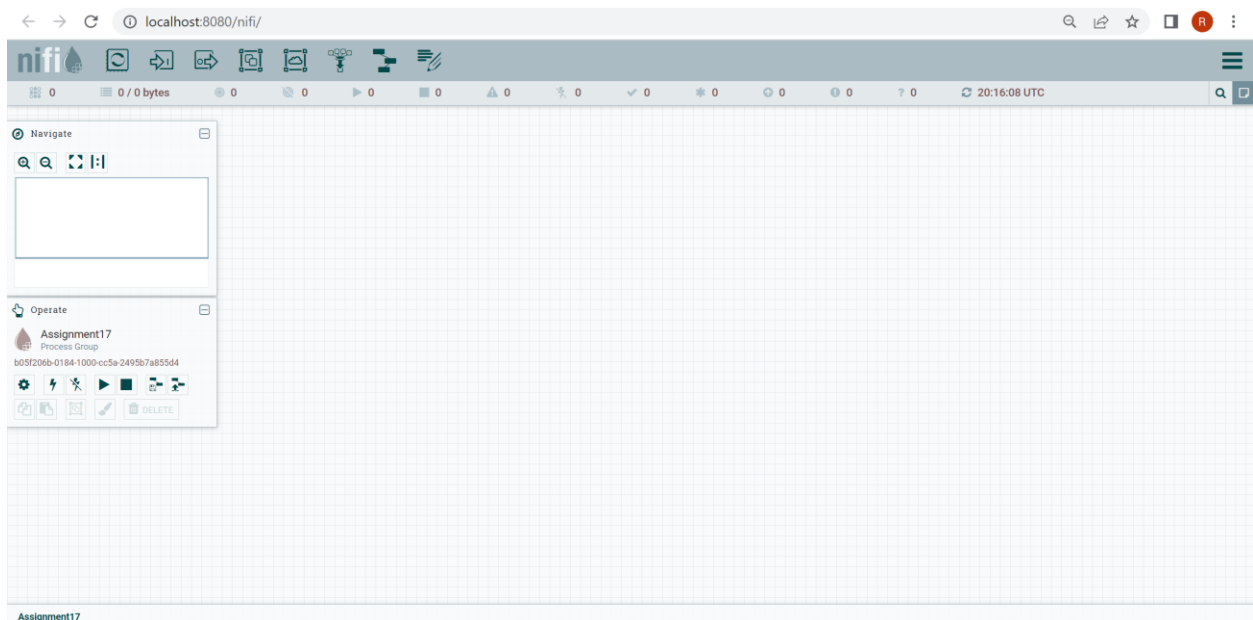
4. Provide a screenshot to show that you correctly configured the *properties* for the GetFile *processor*.
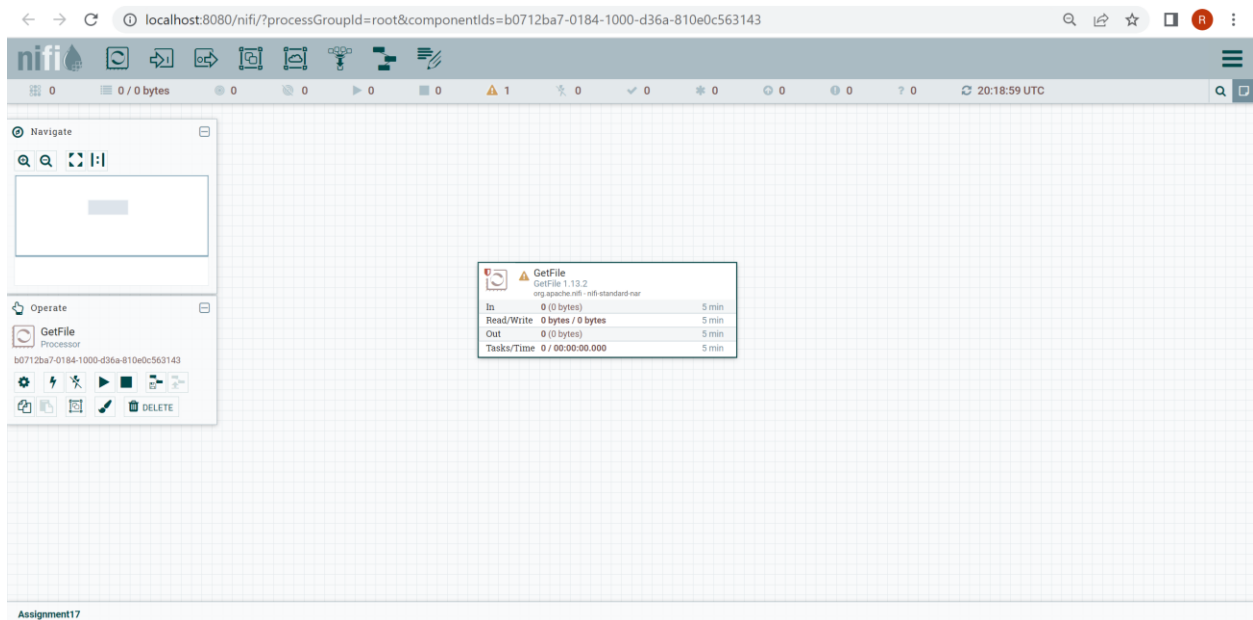
Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Scheduling Strategy ❓

Timer driven ⌄

Concurrent Tasks ❓

1

Run Schedule ❓

15 sec

Execution ❓

All nodes ⌄

CANCEL   APPLY

---

localhost:8080/nifi/?processGroupId=root&componentIds=b0712ba7-0184-1000-d36a-810e0c563143

Navigate

Operate

GetFile
Processor
b0712ba7-0184-1000-d36a-810e0c563143

🗑 DELETE

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Required field ➕

Property	Value
Input Directory ❓	/opt/nifi/nifi-current/input
File Filter ❓	movies.xlsx
Path Filter ❓	No value set
Batch Size ❓	10
Keep Source File ❓	false
Recurse Subdirectories ❓	true
Polling Interval ❓	0 sec
Ignore Hidden Files ❓	true
Minimum File Age ❓	0 sec
Maximum File Age ❓	No value set
Minimum File Size ❓	0 B
Maximum File Size ❓	No value set

CANCEL   APPLY

Assignment17

5. Provide a screenshot to show that you correctly configured the *properties* for the ConvertExcelToCSVProcessor *processor*.

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Scheduling Strategy ❓

Timer driven ▼

Concurrent Tasks ❓

1

Run Schedule ❓

15 sec

Execution ❓

All nodes ▼

CANCEL   APPLY

Configure Processor

⚠️ **Invalid**

| SETTINGS | SCHEDULING | **PROPERTIES** | COMMENTS |

Required field   ➕

Property		Value	
Sheets to Extract	❓	Assignment	
Number of Rows to Skip	❓	**0**	
Columns To Skip	❓	No value set	
Format Cell Values	❓	**false**	
CSV Format	❓	**Custom Format**	
Value Separator	❓	**,**	
Include Header Line	❓	**true**	
Quote Character	❓	**"**	
Escape Character	❓	****	
Comment Marker	❓	No value set	
Null String	❓	No value set	
Trim Fields	❓	**true**	
Quote Mode	❓	Do Not Quote Values	

CANCEL APPLY

6. Provide a screenshot to show that you correctly configured the *properties* for the PutFile *processor*.

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Name

PutFile   ☑ Enabled

Id
b1090219-0184-1000-f4e1-5b0b0b901649

Type
PutFile 1.13.2

Bundle
org.apache.nifi - nifi-standard-nar

Penalty Duration ❓

30 sec

Yield Duration ❓

1 sec

Bulletin Level ❓

WARN ⌄

Automatically Terminate Relationships ❓
☑ failure
Files that could not be written to the output directory for some reason are transferred to this relationship
☑ success
Files that have been successfully written to the output directory are transferred to this relationship

CANCEL   APPLY

Configure Processor

⚠ **Invalid**

SETTINGS	SCHEDULING	PROPERTIES	COMMENTS

Scheduling Strategy ❓

Timer driven ⌄

Run Duration ❓

0ms 25ms 50ms 100ms 250ms 500ms 1s 2s

Lower latency — Higher throughput

Concurrent Tasks ❓

1

Run Schedule ❓

15 sec

Execution ❓

All nodes ⌄

CANCEL   APPLY

← → C ⓘ localhost:8080/nifi/?processGroupId=root&componentIds=b0efa7b9-0184-1000-cdcd-9b5d69fe48e8

Configure Processor

⚠ Invalid

SETTINGS	SCHEDULING	PROPERTIES	COMMENTS

Required field ＋

Property	Value
Directory	/opt/nifi/nifi-current/output
Conflict Resolution Strategy	fail
Create Missing Directories	true
Maximum File Count	No value set
Last Modified Time	No value set
Permissions	No value set
Owner	No value set
Group	No value set

CANCEL   APPLY

Assignment17

---

7. Provide a screenshot to show that you successfully connected all the *processors* with the correct relationships.

Configure Connection

DETAILS	SETTINGS

From Processor
GetFile
GetFile

To Processor
ConvertExcelToCSVProcessor
ConvertExcelToCSVProcessor

Within Group
Assignment17

Within Group
Assignment17

For Relationships
☑ success

CANCEL   APPLY

Create Connection

DETAILS   SETTINGS

From Processor
ConvertExcelToCSVProcessor
ConvertExcelToCSVProcessor

Within Group
Assignment17

For Relationships
☑ failure
☑ original
☑ success
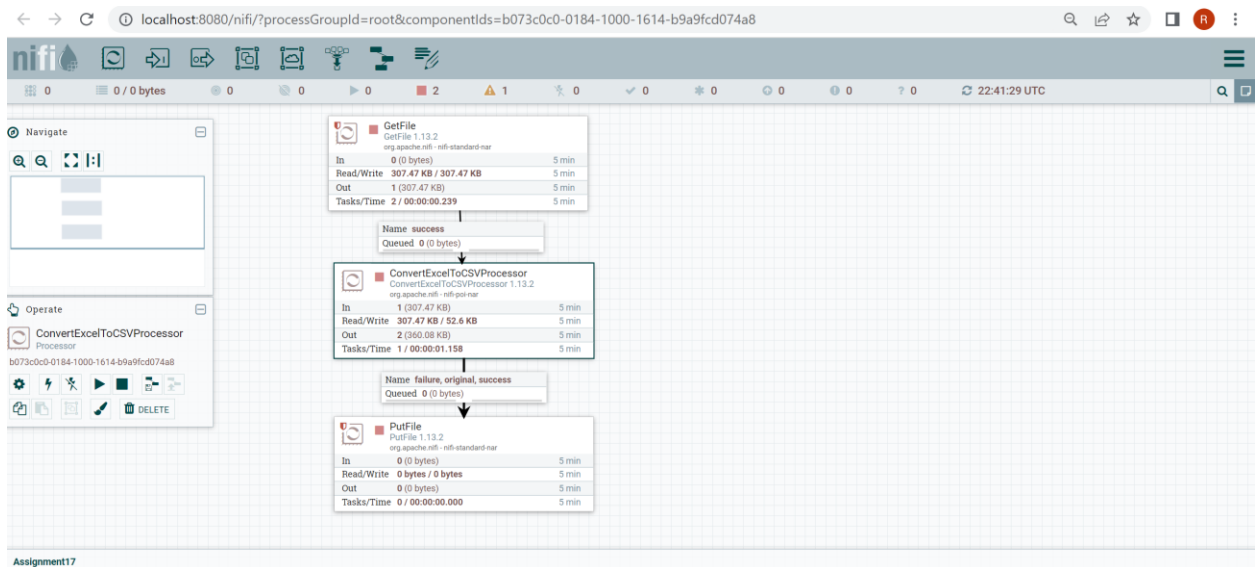
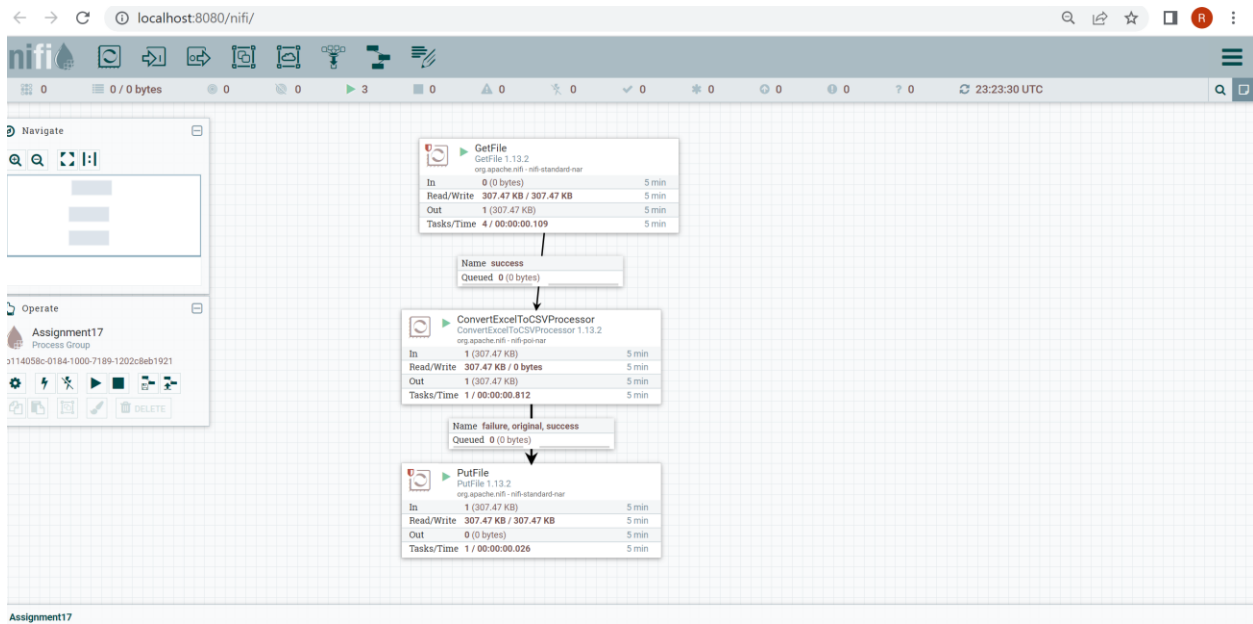To Processor
PutFile
PutFile

Within Group
Assignment17

CANCEL   ADD

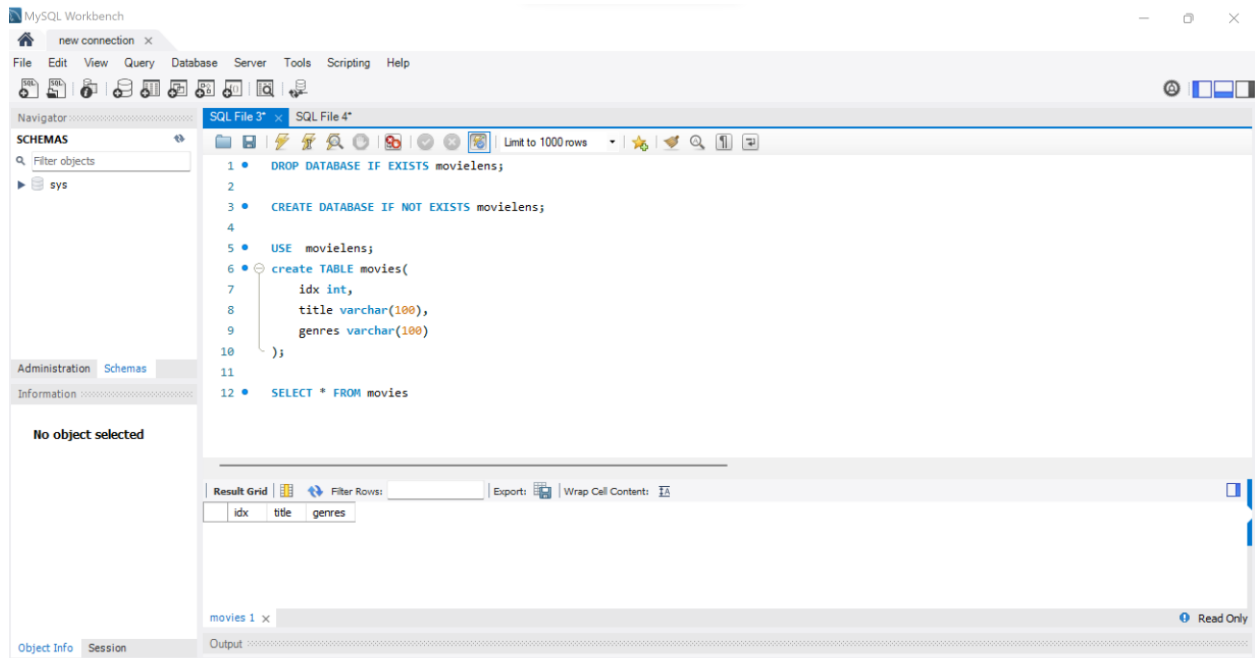8. Provide a screenshot to show that all the *processors* are running (as indicated by a green arrow).

9. Provide a screenshot to show that the movies_Assignment.csv file has been created.
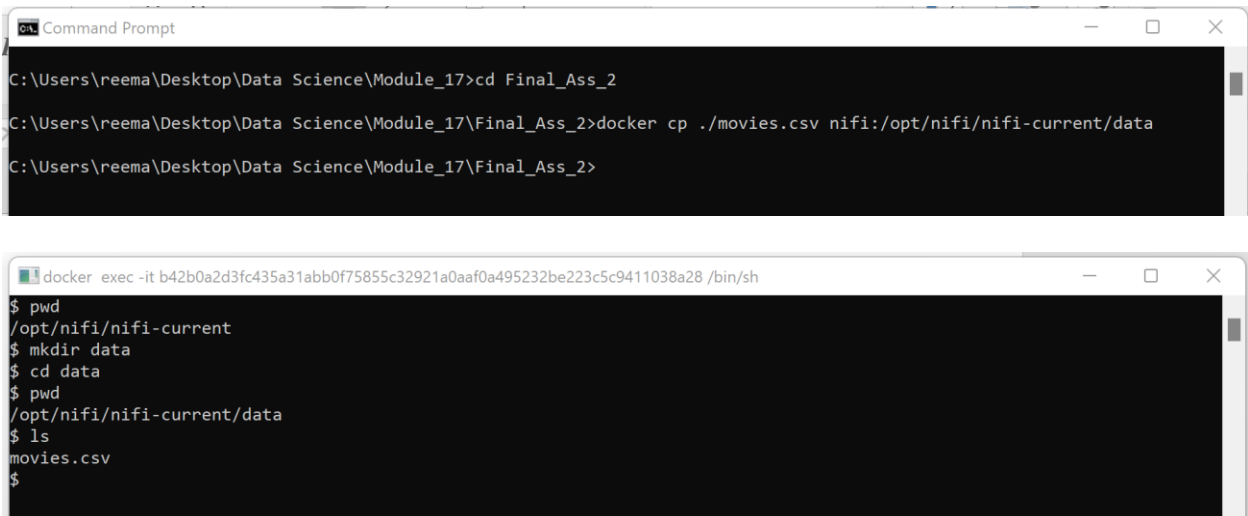

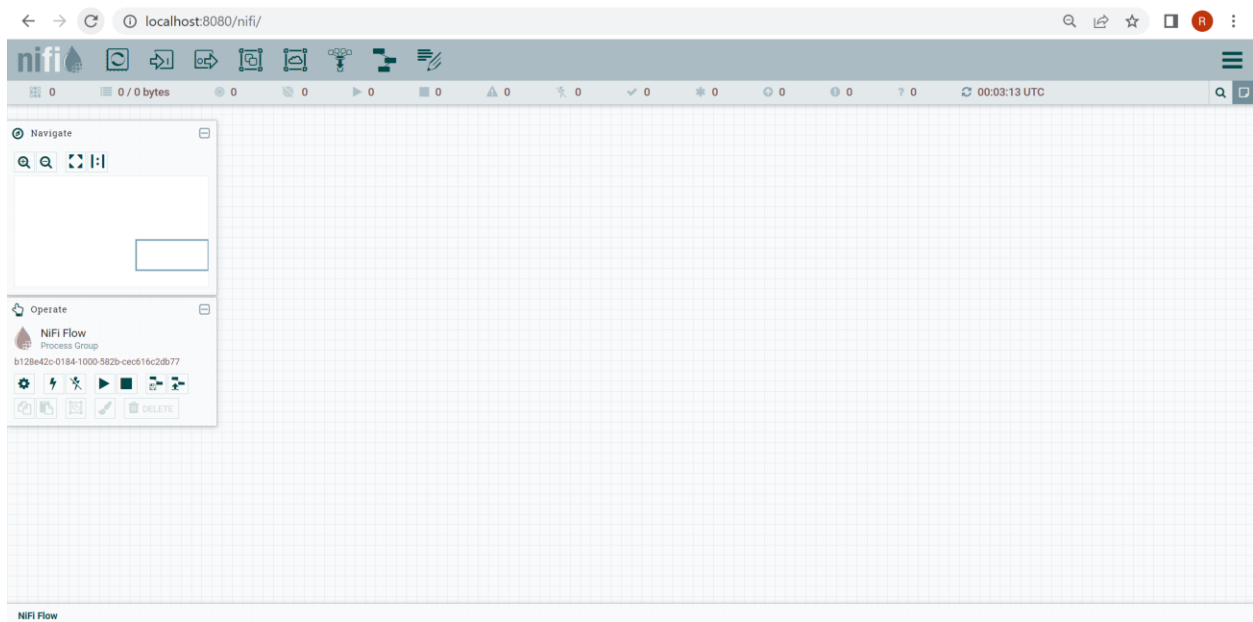**Part 2: Writing Data to an SQL Database**

1. Provide a screenshot of your MySQL Workbench to show that you have successfully initialized an empty movies table in the movielens database.
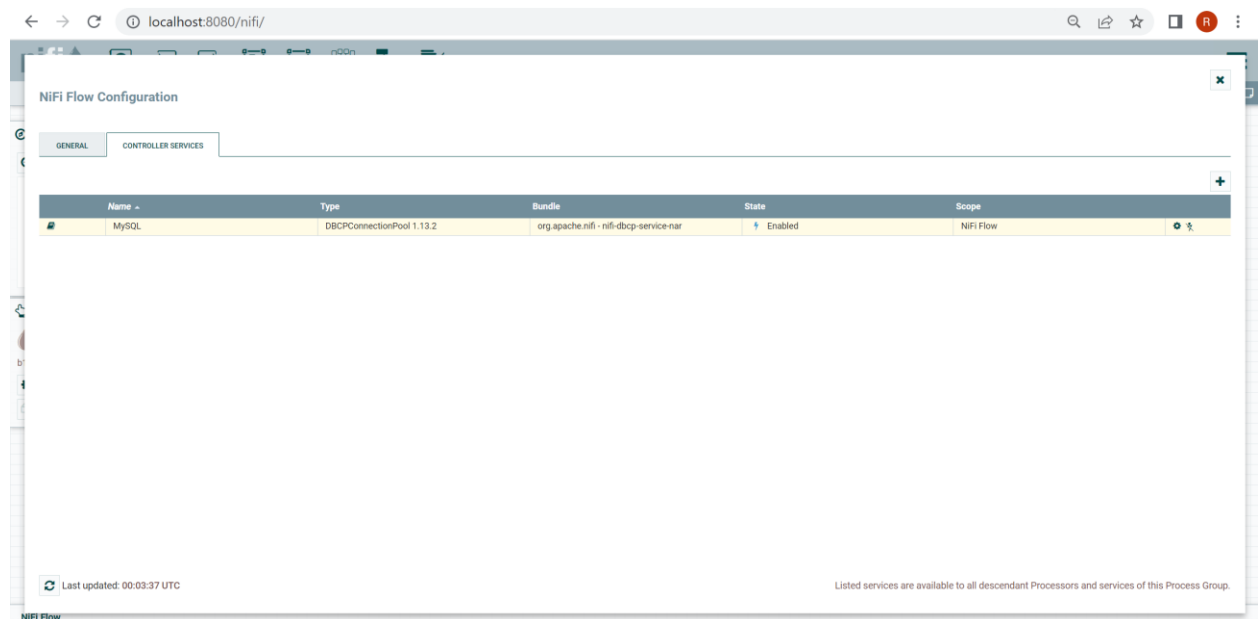
2. Provide a screenshot to show that the movies.csv file is now on the NiFi *server*.

3. Provide a screenshot to show that you successfully opened the NiFi UI.

4. Provide a screenshot to show that you successfully created and enabled the MySQL *controller* service.

5. Provide a screenshot of the *controller* screen to show that the three *controller* services (*reader*, *writer*, and MySQL) are enabled.

Controller Service Details

SETTINGS | **PROPERTIES** | COMMENTS

Required field

Property		Value	
Database Connection URL	❓	**jdbc:mysql://myfinalsql:3306**	
Database Driver Class Name	❓	**com.mysql.jdbc.Driver**	
Database Driver Location(s)	❓	/opt/nifi/drivers/mysql-connector-j-8.0.31.jar	
Kerberos Credentials Service	❓	No value set	
Kerberos Principal	❓	No value set	
Kerberos Password	❓	No value set	
Database User	❓	root	
Password	❓	Sensitive value set	
Max Wait Time	❓	**500 millis**	
Max Total Connections	❓	**8**	
Validation query	❓	No value set	
Minimum Idle Connections	❓	0	
Max Idle Connections	❓	8	
Max Connection Lifetime	❓	-1	

OK

Configure Controller Service

SETTINGS	PROPERTIES	COMMENTS

Required field ➕

Property		Value	
Schema Access Strategy	❓	**Use String Fields From Header**	
CSV Parser	❓	**Apache Commons CSV**	
Date Format	❓	No value set	
Time Format	❓	No value set	
Timestamp Format	❓	No value set	
CSV Format	❓	**Custom Format**	
Value Separator	❓	,	
Record Separator	❓	\n	
Treat First Line as Header	❓	true	
Ignore CSV Header Column Names	❓	false	
Quote Character	❓	"	
Escape Character	❓	\	
Comment Marker	❓	No value set	
Null String	❓	No value set	

CANCEL   APPLY

Configure Controller Service

SETTINGS	**PROPERTIES**	COMMENTS

Required field　　　　　　　　　　　　　　　　　　　　　　　＋

Property		Value
Schema Write Strategy	❓	**Do Not Write Schema**
Schema Cache	❓	No value set
Schema Access Strategy	❓	**Inherit Record Schema**
Date Format	❓	No value set
Time Format	❓	No value set
Timestamp Format	❓	No value set
Pretty Print JSON	❓	false
Suppress Null Values	❓	**Never Suppress**
Output Grouping	❓	Array
Compression Format	❓	none

CANCEL　　APPLY

localhost:8080/nifi/

NiFi Flow Configuration

GENERAL	**CONTROLLER SERVICES**

Name ▲	Type	Bundle	State	Scope	
CSVReader	CSVReader 1.13.2	org.apache.nifi - nifi-record-serialization-services-n...	Enabled	NiFi Flow	⚙ ⚡
JsonRecordSetWriter	JsonRecordSetWriter 1.13.2	org.apache.nifi - nifi-record-serialization-services-n...	Enabled	NiFi Flow	⚙ ⚡
MySQL	DBCPConnectionPool 1.13.2	org.apache.nifi - nifi-dbcp-service-nar	Enabled	NiFi Flow	⚙ ⚡

Last updated: 00:03:37 UTC　　　　　　　Listed services are available to all descendant Processors and services of this Process Group.

NiFi Flow

6. Provide a screenshot of your complete data pipeline, including all five *processors*: GetFile, SplitText, ConvertRecord, ConvertJSONToSQL, and PutSQL.

Configure Processor

⚠ **Invalid**

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Name

get movies file ☑ Enabled

Id
b162db2d-0184-1000-aa3f-06112fdf2c4d

Type
GetFile 1.13.2

Bundle
org.apache.nifi - nifi-standard-nar

Penalty Duration ❓

30 sec

Yield Duration ❓

1 sec

Bulletin Level ❓

WARN ⌄

Automatically Terminate Relationships ❓
☐ success
　　All files are routed to success

CANCEL   APPLY

Configure Processor

⚠ Invalid

SETTINGS	SCHEDULING	**PROPERTIES**	COMMENTS

Required field   ＋

Property		Value	
Input Directory	❓	/opt/nifi/nifi-current/data	
File Filter	❓	[^\.].*	
Path Filter	❓	No value set	
Batch Size	❓	10	
Keep Source File	❓	false	
Recurse Subdirectories	❓	true	
Polling Interval	❓	0 sec	
Ignore Hidden Files	❓	true	
Minimum File Age	❓	0 sec	
Maximum File Age	❓	No value set	
Minimum File Size	❓	0 B	
Maximum File Size	❓	No value set	

CANCEL   APPLY

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Name

SplitText   ☑ Enabled

Id
b1679300-0184-1000-7cfd-44ead8f7f2a9

Type
SplitText 1.13.2

Bundle
org.apache.nifi - nifi-standard-nar

Penalty Duration ❓   **Yield Duration** ❓

30 sec   1 sec

Bulletin Level ❓

WARN   ⌄

Automatically Terminate Relationships ❓

☑ **failure**
If a file cannot be split for some reason, the original file will be routed to this destination and nothing will be routed elsewhere

☑ **original**
The original input file will be routed to this destination when it has been successfully split into 1 or more files

☐ **splits**
The split files will be routed to this destination when an input file is successfully split into 1 or more split files

CANCEL   APPLY

Configure Processor

⚠ **Invalid**

| SETTINGS | SCHEDULING | **PROPERTIES** | COMMENTS |

Required field   ➕

Property		Value	
Line Split Count	❓	1	
Maximum Fragment Size	❓	No value set	
Header Line Count	❓	1	
Header Line Marker Characters	❓	No value set	
Remove Trailing Newlines	❓	true	

CANCEL   APPLY

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | **PROPERTIES** | COMMENTS |

Required field ➕

Property		Value	
Record Reader	❓	CSVReader	→
Record Writer	❓	JsonRecordSetWriter	→
Include Zero Record FlowFiles	❓	true	

CANCEL | APPLY

Configure Processor

⚠ **Invalid**

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Name

ConvertJSONToSQL   ☑ Enabled

Id
b16a85dd-0184-1000-c713-eb7d1b2135da

Type
ConvertJSONToSQL 1.13.2

Bundle
org.apache.nifi - nifi-standard-nar

Penalty Duration ❓

30 sec

Yield Duration ❓

1 sec

Bulletin Level ❓

WARN ⌄

Automatically Terminate Relationships ❓

☑ **failure**

A FlowFile is routed to this relationship if it cannot be converted into a SQL statement. Common causes include invalid JSON content or the JSON content missing a required field (if using an INSERT statement type).

☑ **original**

When a FlowFile is converted to SQL, the original JSON FlowFile is routed to this relationship

☐ **sql**

A FlowFile is routed to this relationship when its contents have successfully been converted into a SQL statement

CANCEL   APPLY

Configure Processor

⚠ **Invalid**

| SETTINGS | SCHEDULING | **PROPERTIES** | COMMENTS |

Required field   ✚

Property		Value	
JDBC Connection Pool	❓	**MySQL**	→
Statement Type	❓	**INSERT**	
Table Name	❓	**movies**	
Catalog Name	❓	movielens	
Schema Name	❓	No value set	
Translate Field Names	❓	true	
Unmatched Field Behavior	❓	Ignore Unmatched Fields	
Unmatched Column Behavior	❓	Fail on Unmatched Columns	
Update Keys	❓	No value set	
Quote Column Identifiers	❓	false	
Quote Table Identifiers	❓	false	
SQL Parameter Attribute Prefix	❓	**sql**	
Table Schema Cache Size	❓	100	

CANCEL   APPLY

Read/Write 0 bytes / 0 bytes

Configure Processor

⚠ **Invalid**

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Name

PutSQL   ☑ Enabled

Id
b16cea17-0184-1000-b57f-44d4643164c4

Type
PutSQL 1.13.2

Bundle
org.apache.nifi - nifi-standard-nar

Penalty Duration ❓   Yield Duration ❓

30 sec   1 sec

Bulletin Level ❓

WARN   ⌄

Automatically Terminate Relationships ❓

☑ **failure**
A FlowFile is routed to this relationship if the database cannot be updated and retrying the operation will also fail, such as an invalid query or an integrity constraint violation

☐ **retry**
A FlowFile is routed to this relationship if the database cannot be updated but attempting the operation again may succeed

☑ **success**
A FlowFile is routed to this relationship after the database is successfully updated

CANCEL   APPLY

Read/Write 0 bytes / 0 bytes   Read/Write 0 bytes / 0 bytes

Configure Processor

⚠ Invalid

| SETTINGS | SCHEDULING | PROPERTIES | COMMENTS |

Required field ➕

Property		Value	
JDBC Connection Pool	❓	MySQL	→
SQL Statement	❓	No value set	
Support Fragmented Transactions	❓	true	
Database Session AutoCommit	❓	false	
Transaction Timeout	❓	No value set	
Batch Size	❓	100	
Obtain Generated Keys	❓	false	
Rollback On Failure	❓	false	

CANCEL   APPLY

---

localhost:8080/nifi/?processGroupId=root&componentIds=b16cea17-0184-1000-b57f-44d4643164c4

🔲 0   ≡ 0 / 0 bytes   ▶ 0   ■ 0   ► 0   ■ 0   ⚠ 5   ⚡ 0   ✓ 0   ✱ 0   ⊘ 0   ⓘ 0   ? 0   🔄 00:55:15 UTC

Navigate

Operate

PutSQL
Processor
b16cea17-0184-1000-b57f-44d4643164c4

DELETE

⚠ **get movies file**		
GetFile 1.13.2		
org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

⚠ **SplitText**		
SplitText 1.13.2		
org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

⚠ **ConvertRecord**		
ConvertRecord 1.13.2		
org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

⚠ **ConvertJSONToSQL**		
ConvertJSONToSQL 1.13.2		
org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

⚠ **PutSQL**		
PutSQL 1.13.2		
org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

7. Provide a screenshot of all five *processors* to show that the correct *connectors* have been added between the *processors*.
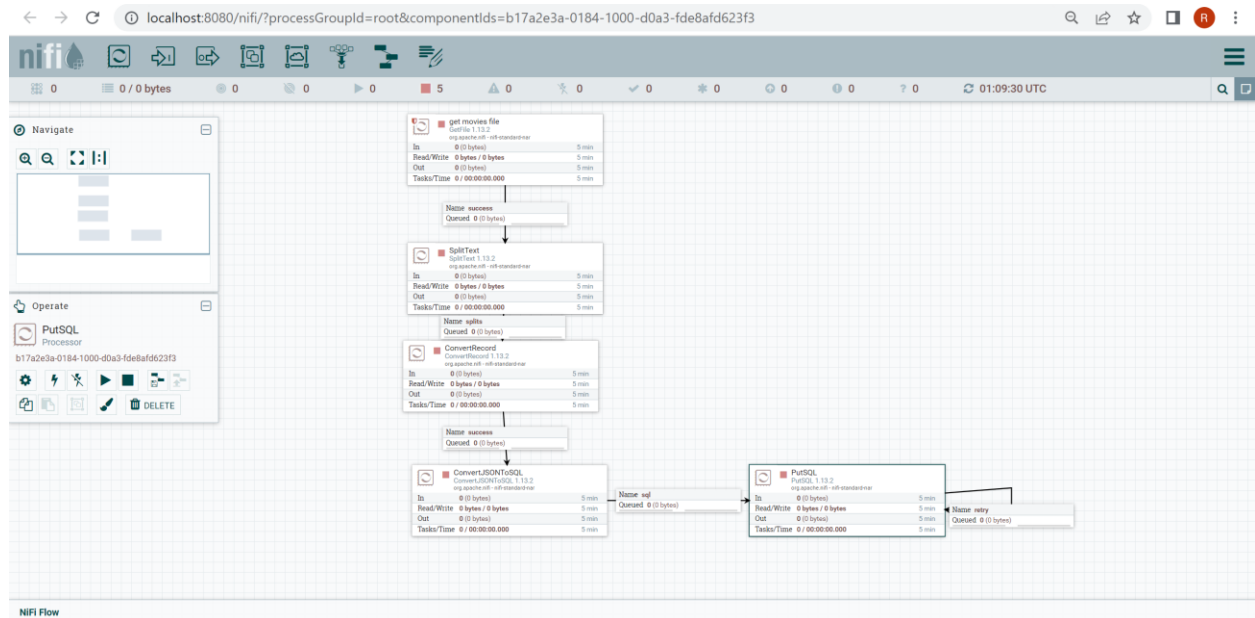
8. Provide a screenshot of your NiFi browser screen to show all five *processors* connected and running.
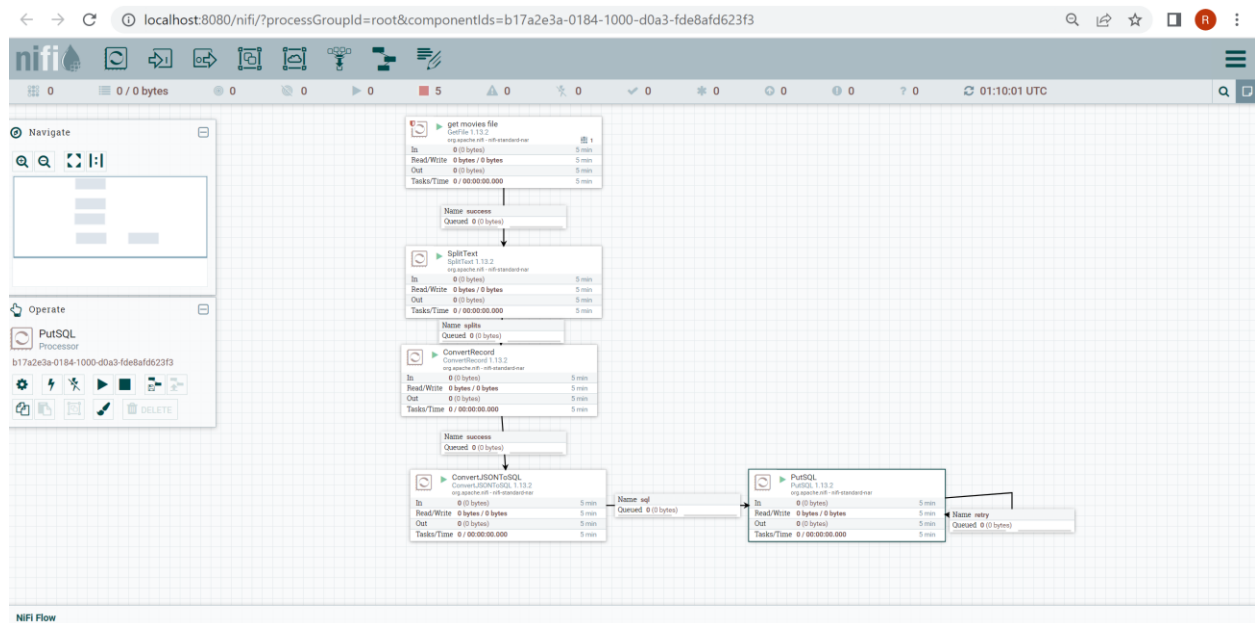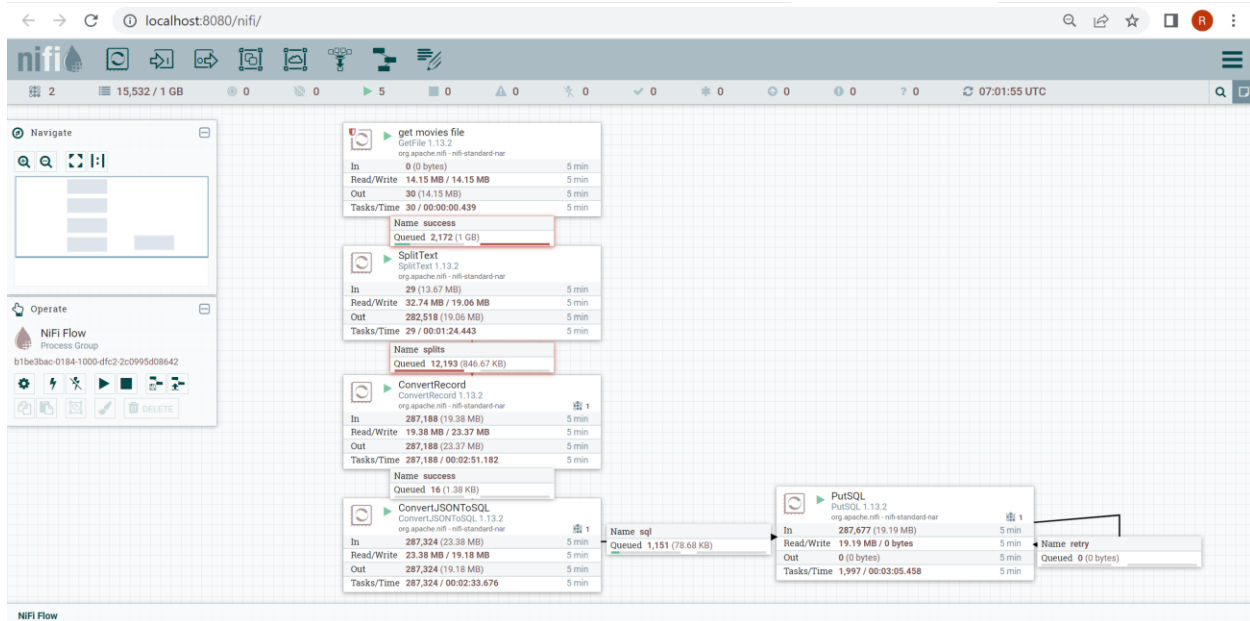
9. Provide a screenshot of the result of this *query* to show that the movies table in the movieslens database is now saturated with data.