In this mini-lesson you will set up a pipeline to perform ETL to read data from a MongoDB database, transform the data, and save the contents to a JSON file for further processing.

For this mini-lesson you will use Docker *containers*. First, create a Docker network called NifiNetwork by running the command below in a Terminal window:

| docker network create NifiNetwork |
| --- |

Next, create a Docker *container* running NiFi and associate it to the NifiNetwork that you created:

| docker run --name nificontainer -p 8080:8080 --network NifiNetwork -d apache/nifi:1.13.2 |
| --- |

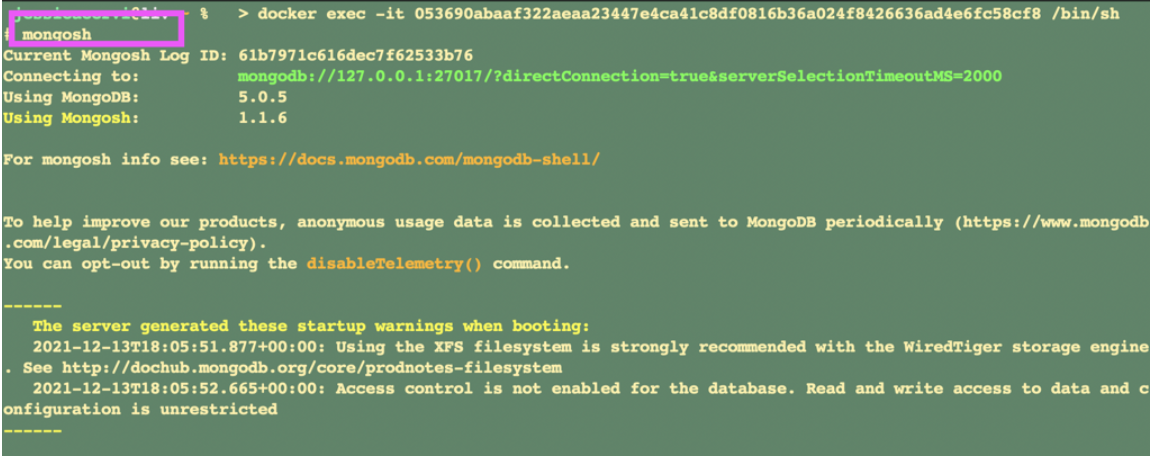
Following the steps in [Mini-Lesson 17.2](https://classroom.emeritus.org/courses/10605/pages/mini-lesson-17-dot-2-using-nifi-to-extract-data-from-an-excel-file-45-00), open the NiFi *container* bash window and create two folders named input and output inside of the /opt/nifi/nifi-current/ folder.

In addition to the NiFi Docker *container* that you created*,* you will need to create another *container* for the MongoDB database. Make sure that the MongoDB database *container* is associated with the NifiNetwork network along with the NiFi Docker *container*. Run the following Docker command to create the MongoDB *container*:

| docker run -p 27017:27017 --name some-mongo --network NifiNetwork -d mongo |
| --- |

From the Docker desktop, select the CLI bash for the MongoDB *container*. From the bash prompt, enter mongosh and select Enter. You are now inside the MongoDB shell.

You should see the following output:

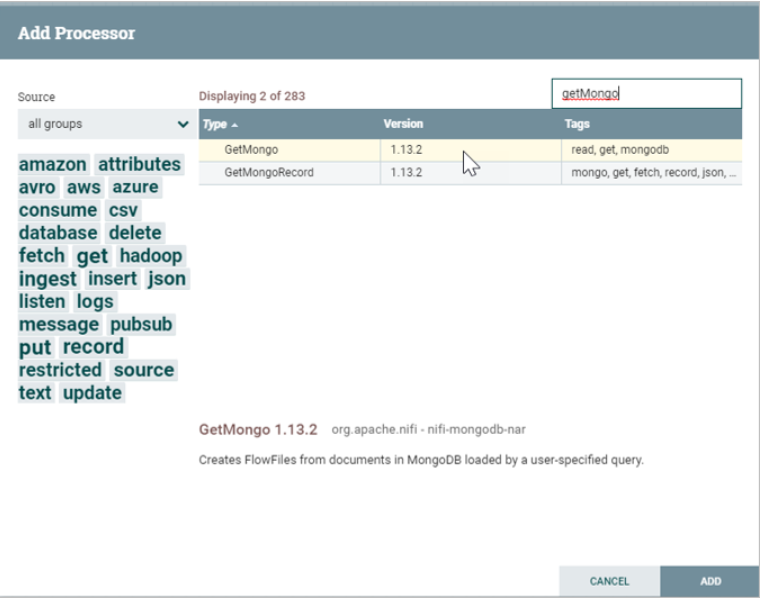
Copy and paste the command below into the mongosh shell to populate the MongoDB database:

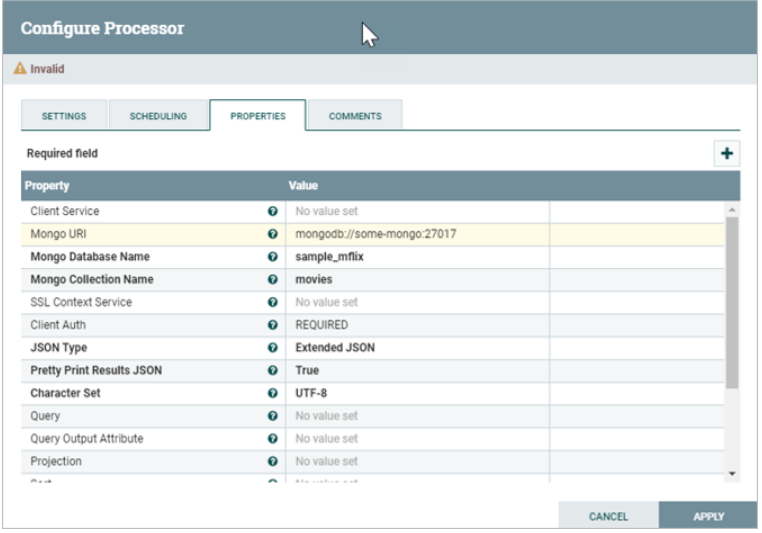
| use sample\_mflix  db.movies.insertMany([  {  title: "Jurassic World: Fallen Kingdom",  genres: [ "Action", "Sci-Fi" ],  runtime: 130,  rated: "PG-13",  year: 2018,  directors: [ "J. A. Bayona" ],  cast: [ "Chris Pratt", "Bryce Dallas Howard", "Rafe Spall" ],  type: "movie"  },  {  title: "Tag",  genres: [ "Comedy", "Action" ],  runtime: 105,  rated: "R",  year: 2018,  directors: [ "Jeff Tomsic" ],  cast: [ "Annabelle Wallis", "Jeremy Renner", "Jon Hamm" ],  type: "movie"  } ]) |
| --- |

You should see the following output:

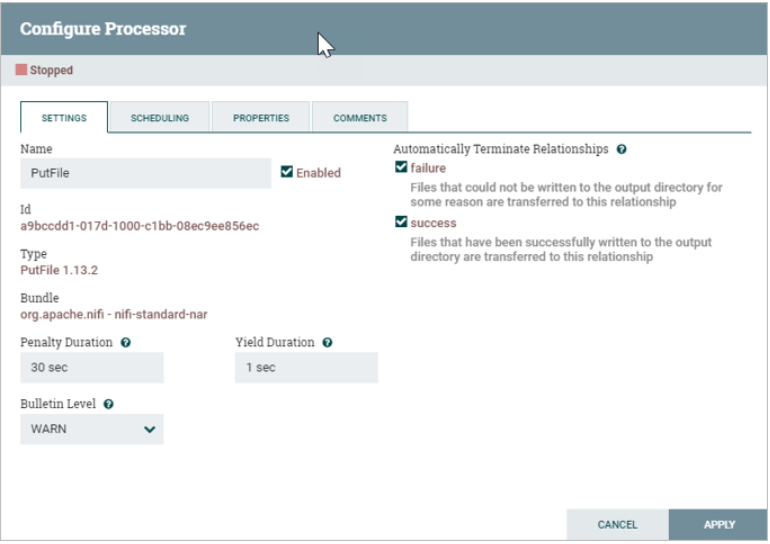
To perform ETL on a MongoDB database using NiFi, follow the steps below:

1. Following the steps in [Mini-Lesson 17.2](https://classroom.emeritus.org/courses/10605/pages/mini-lesson-17-dot-2-using-nifi-to-extract-data-from-an-excel-file-45-00) and create a new process group titled Mini-Lesson 17.4.
2. Add a new *processor* titled GetMongo.

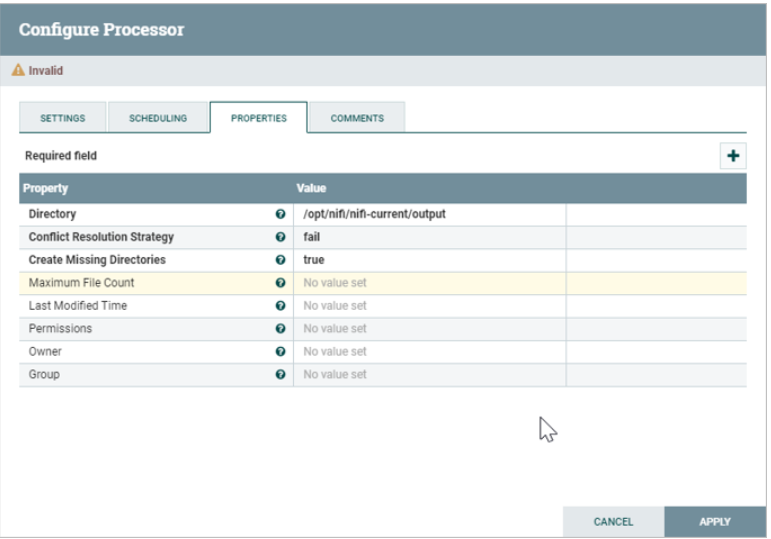


3. Right-click on the GetMongo *processor* and select Configure. In the PROPERTIES tab, set the Mongo URI field equal to mongodb://some-mongo:27017. Note that some-mongo is the name of the MongoDB *container* used in this example. If you used another title, then you must provide that title instead of some-mongo. Set the Mongo Database Name field equal to sample\_mflix and the Mongo Collection Name field equal to movies.

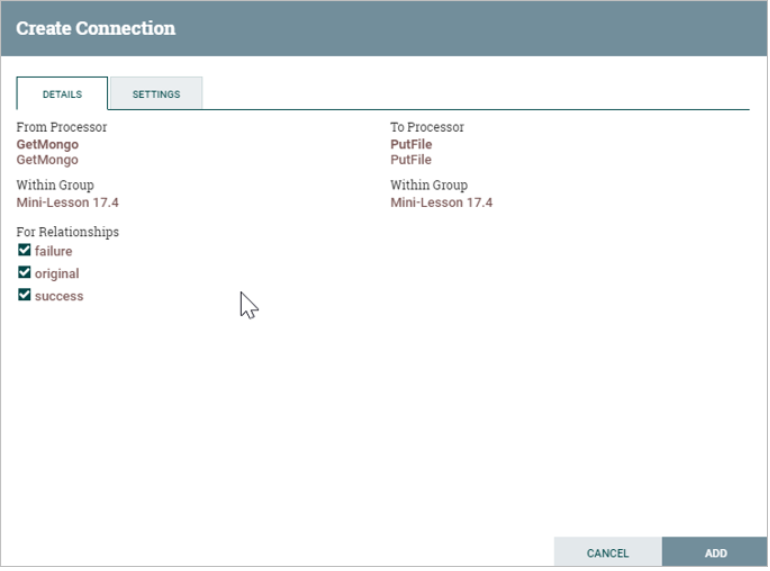
4. Add a new *processor* titled PutFile.

5. Right-click on the PutFile *processor* and select Configure. In the SETTINGS tab, select failure and success under the Automatically Terminate Relationships tab.

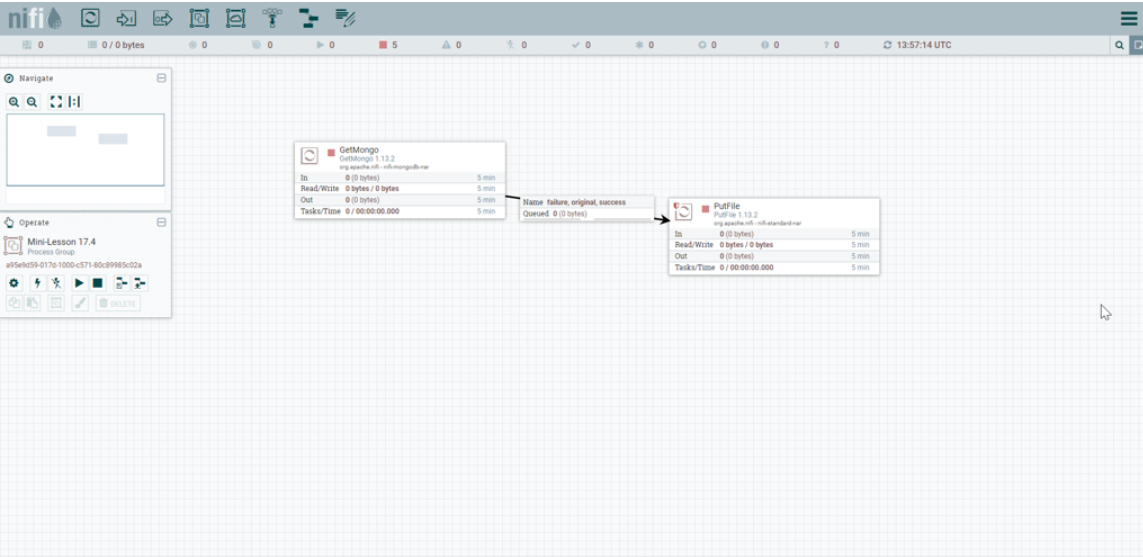
In the PROPERTIES tab, set the Directory field equal to /opt/nifi/nifi-current/output.



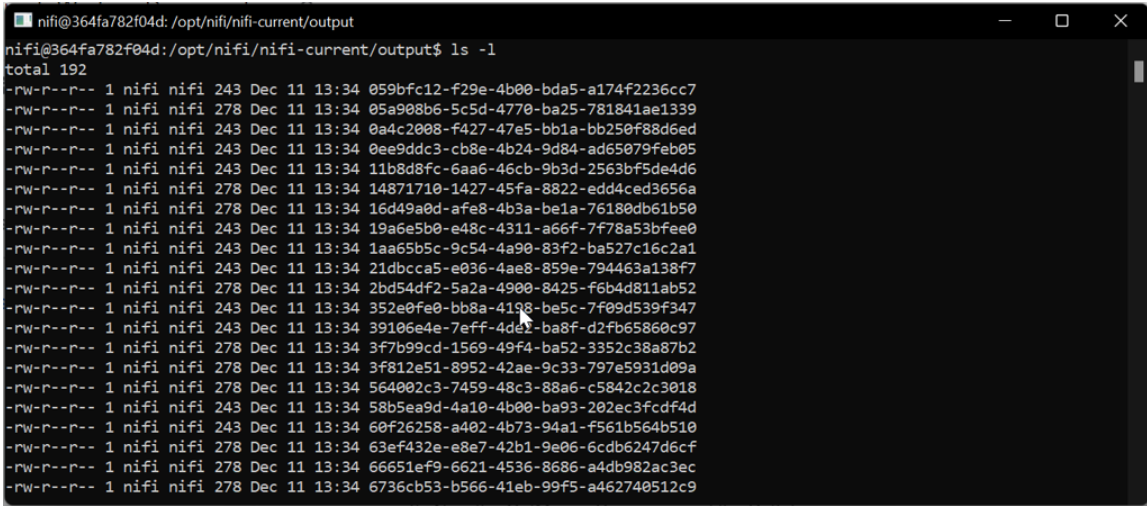
6. Connect the GetMongo and PutFile *processors*. Select failure, original, and success for the relationships.



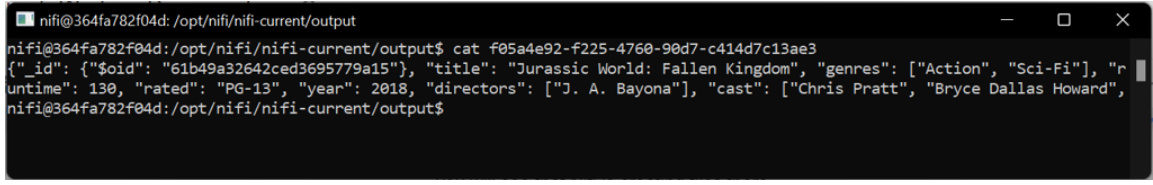
7. Start the GetMongo and PutFile *processors*.



8. Navigate to the /opt/nifi/nifi-current/output *directory* on the NiFi *server* and use the ls command to display the files created by NiFi.



9. Use the cat command to display the contents of one of the files. You will notice that it is a JSON-formatted file:



Now you know how to use NiFi to read data from a MongoDB database to perform ETL.