**Docker, Jupiter, DLab- Task 1**

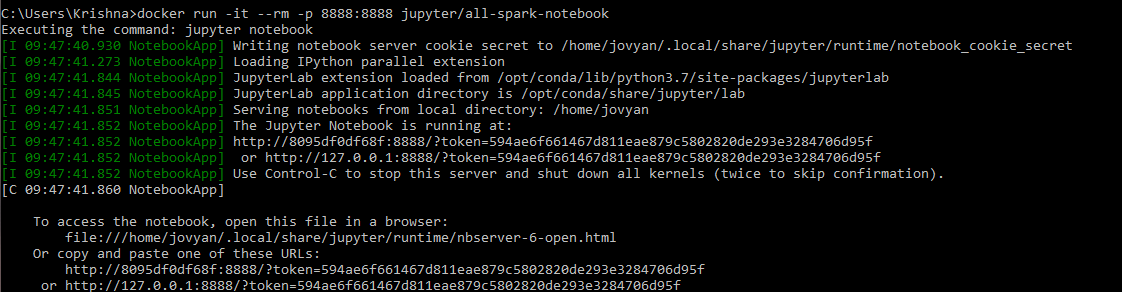
**Docker:**

Docker is a tool designed to make it easier to create, deploy, and run applications by using containers. Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and deploy it as one package.

**Jupyter notebook:**

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

**Starting Docker All Spark Notebook:**



Programs in container are isolated from the outside world. So, Docker gives an option to open port for external connections.

In the above image, I have opened two ports 8888:8888(Inside:outside). If we not mention outside port then Docker will assign random available port.

**Importing packages and defining functions:**

In the below image, we can see that I have write down all the import packages code.

I have also declare my function with it.

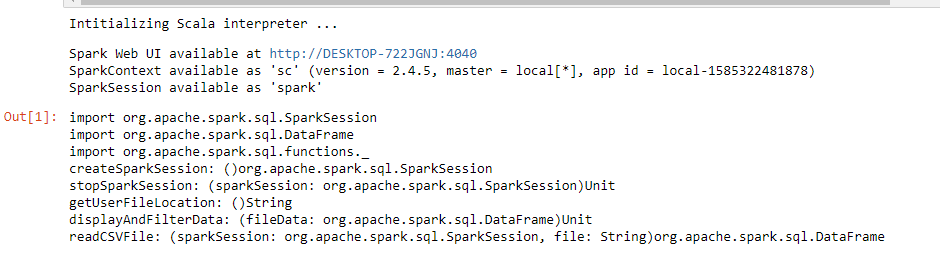
The scope of these imported packages & function is throughout the Jupyter-notebook. So, we don’t need to import or define function again and again.

This increase the reusability of the program.



**Output of above code:**

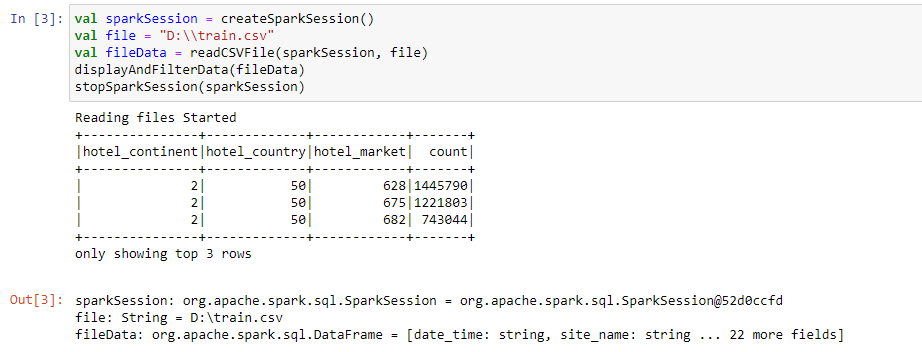
In the below image, we can see that packages are successfully imported. And functions are successfully declared.



**Function Calling:**

In the below image, you can see that I have called the all the previously declared functions.

So, this increase the reusability.



In the above, finally we get the desired output.

**Logs are displayed on the command prompt:**

In the below image, you can see that file is readied by the spark in 31 partition. And the desired output will be generated on the Jupyter notebook.

