Stephen Re – Assignment 19

**Part 1: PySpark**

1. Provide a screenshot to show that you correctly pulled the *image* and that the docker-compose.yaml file is present.  
   A black background with white text

   Description automatically generated
2. Provide a screenshot of your Docker Desktop to show that you correctly pulled the *containers.*     
   A screenshot of a computer

   Description automatically generated
3. Provide a screenshot to show that you successfully copied the departuredelays.csv file to the bitnami\_spark\_1 *container*.  
   
4. Provide a screenshot to show that you successfully opened PySpark.  
   A computer screen shot of a program

   Description automatically generated
5. Provide a screenshot to show that you successfully started a PySpark session.  
   
6. Provide a screenshot to show that you successfully defined the spark PySpark session.  
   A screen shot of a computer program

   Description automatically generated
7. Provide a screenshot to show that you successfully defined the assignment19\_3\_data variable.  
   
8. Provide a screenshot to show that you successfully defined the df*dataframe* that contains all of the entries in the departuredelays.csv file.  
   A screen shot of a computer code

   Description automatically generated
9. Provide a screenshot to show that you successfully created a view of the assignment19\_3\_table *dataframe*.  
   
10. Provide a screenshot to show that you selected the correct entries from your data. Your data should display the first 15 flights from PHL to DFW that had a delay of greater than 150 minutes.  
    A screen shot of a computer

    Description automatically generated
11. Provide a screenshot to show that you selected the correct entries from your data. Your data should display the first 10 flights that have a distance of less than 200 miles and the resulting table should contain all of the columns in the original dataset.  
    A screenshot of a computer screen

    Description automatically generated
12. Provide a screenshot to show that you selected the correct entries from your data. Your data should display the first 10 flights that have a distance greater than 600 miles and the resulting table should contain all of the columns in the original dataset.  
    A screenshot of a computer screen

    Description automatically generated

**Part 2: Airflow**

1. Provide a screenshot of your Terminal window response to show that you correctly pulled the Airflow file.  
   A screen shot of a computer

   Description automatically generated
2. Provide a screenshot to show the changed example value ( AIRFLOW\_\_CORE\_\_LOAD\_EXAMPLES set to false).  
   A screen shot of a computer

   Description automatically generated
3. Provide a screenshot to show that the Airflow Docker*containers* are running.  
   A screenshot of a computer

   Description automatically generated
4. Provide a screenshot of your browser window to show that you have successfully logged in to Airflow.  
   A screenshot of a computer

   Description automatically generated
5. Provide a screenshot to show that you created the module19\_assignment.py file.  
   A screenshot of a computer

   Description automatically generated
6. Provide a screenshot to show that you correctly imported the required*libraries*.  
   A screen shot of a computer program

   Description automatically generated
7. Provide a screenshot to show that you set up your DAG correctly, including your last name and email address.  
   A screen shot of a computer program

   Description automatically generated
8. Provide a screenshot to show that you defined the square() *function* correctly.  
   A computer screen shot of a program

   Description automatically generated
9. Provide a screenshot to show that you correctly defined the DAG object.  
   A computer screen shot of a computer code

   Description automatically generated
10. Provide a screenshot to show that you defined the DAG*Task* correctly.  
    A screen shot of a computer program

    Description automatically generated
11. Provide a screenshot of the Airflow UI to show that your DAG is configured correctly.  
    A screenshot of a computer

    Description automatically generated
12. Provide a screenshot of the log to show that the DAG ran successfully.  
    A screenshot of a computer

    Description automatically generated