CSE 511 – Project Phase 1

Group15

Group Members

|  |  |  |
| --- | --- | --- |
| Aniruddha Mondal | 1219300407 | amondal8@asu.edu |
| Ravi Maddi | 1222610472 | rmaddi1@asu.edu |
| Sagar Subbaiah Kuppanda Cariappa | 1219783773 | skuppand@asu.edu |
| Vishrut Jha | 1221914707 | vkjha@asu.edu |

The goal of this project is to extract data and perform spatial queries using user-defined functions. We use Scala to implement these functions.

**Problem Description:** There are two user-defined functions which we must develop ST\_Contains and ST\_Within. We use these functions to run four spatial queries -

1. Range Query – This query is used to find all the points P located within a rectangle R.
2. Range Join Query – This is used to find all possible point, rectangle pairs such that the point is within the rectangle. We utilize points P and rectangles R.
3. Distance Query – We have a fixed-point P and a distance D. Here we need to find all possible points in the radius of D from P.
4. Distance Join Query – Here we have two sets of points P1 and P2, and a given distance D. This query is used to find a set of points (p1, p2) where p1 is part of P1 and p2 is part of P2 such that the distance between p1 and p2 is less than or equal to D.

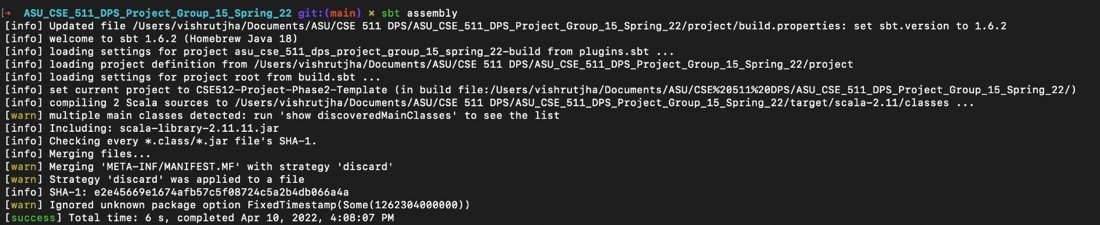
**ST\_Contains:** This function takes two input parameters. One input parameter is a point, and another is the diagonal ends of a rectangle. It checks if the given point is within the rectangle or not.

**ST\_Within:** There are 3 arguments passed to this function. Two points p1 and p2, and a distance D. This function is used to check if the distance between the points p1 and p2 is less than or equal to D.

**Testing our code:**

**Below are the steps to run the code and test its functionality:**

1. After having all the prerequisites installed properly and verifying the installation of spark running the command “spark-shell”, we will foloow the below steps for executing the project.
2. We will open command prompt and navigate to the path where the project has been kept.
3. We will then run the command “sbt assembly” as shown in the screenshot:

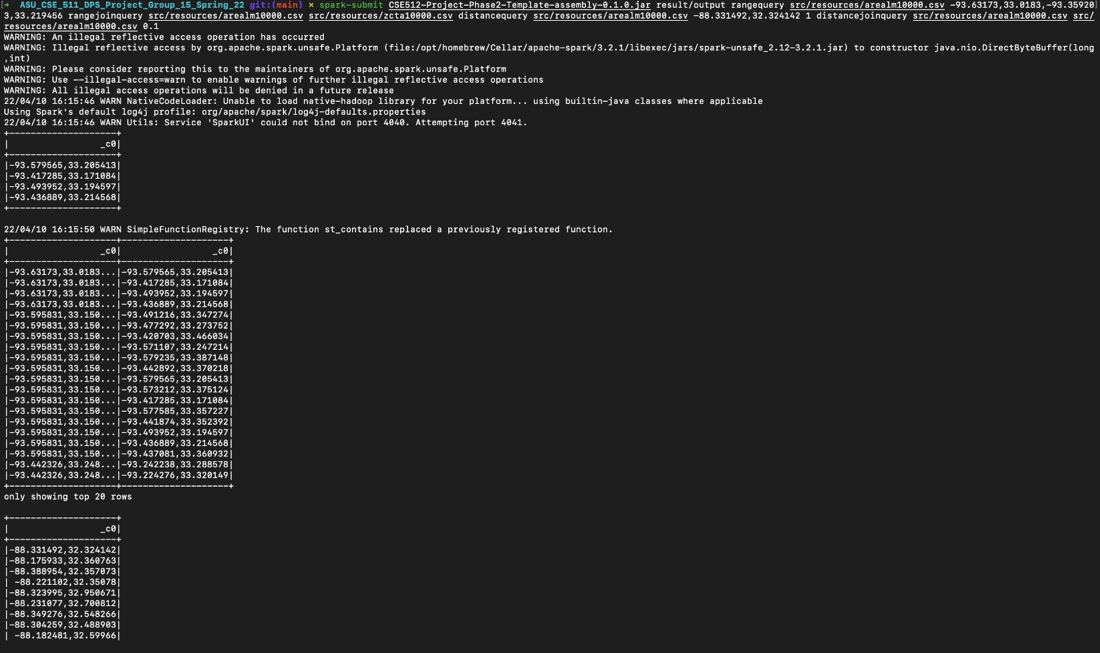


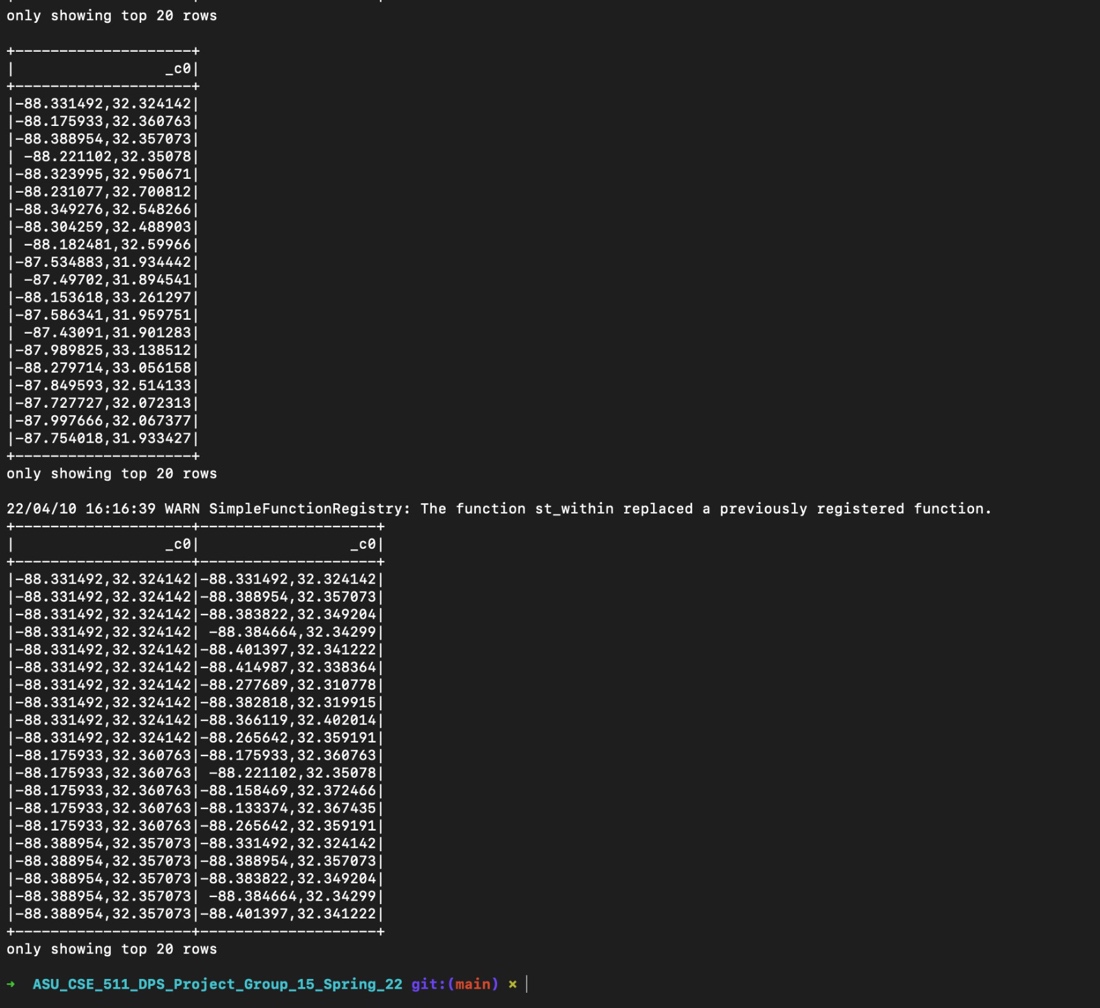
On successfully running the command we will have the target folder. Inside the target folder the jar file gets created.

1. We will now copy the jar file and paste it in the root folder.
2. After this we will run the below command to fire the execution:

spark-submit CSE512-Project-Phase2-Template-assembly-0.1.0.jar result/output rangequery src/resources/arealm10000.csv -93.63173,33.0183,-93.359203,33.219456 rangejoinquery src/resources/arealm10000.csv src/resources/zcta10000.csv distancequery src/resources/arealm10000.csv -88.331492,32.324142 1 distancejoinquery src/resources/arealm10000.csv src/resources/arealm10000.csv 0.1

**Below are the screenshots of the run:**





1. On running the above command we will have our target folder where we will be getting the 4 output sub folders. Below is the screenshot for the same:

