**Hive Lab – Aparna Pavithran (axp161730)**

**Question 1:**

CREATE TABLE business (businessid STRING, fullAddress STRING, categories STRING) ROW FORMAT DELIMITED FIELDS TERMINATED BY '^';

CREATE TABLE review (reviewid STRING, userid STRING, businessid STRING, stars FLOAT) ROW FORMAT DELIMITED FIELDS TERMINATED BY '^';

CREATE TABLE user (userid STRING, name STRING, url STRING) ROW FORMAT DELIMITED FIELDS TERMINATED BY '^';

LOAD DATA INPATH "/yelp/business/business.csv" OVERWRITE INTO TABLE business;

LOAD DATA INPATH "/yelp/review/review.csv" OVERWRITE INTO TABLE review;

LOAD DATA INPATH "/yelp/user/user.csv" OVERWRITE INTO TABLE user;

/\*to run in Data Bricks

LOAD DATA INPATH "/FileStore/tables/yck4y8os1478572995296/business.csv" OVERWRITE INTO TABLE business;

LOAD DATA INPATH "/FileStore/tables/yck4y8os1478572995296/review.csv" OVERWRITE INTO TABLE review;

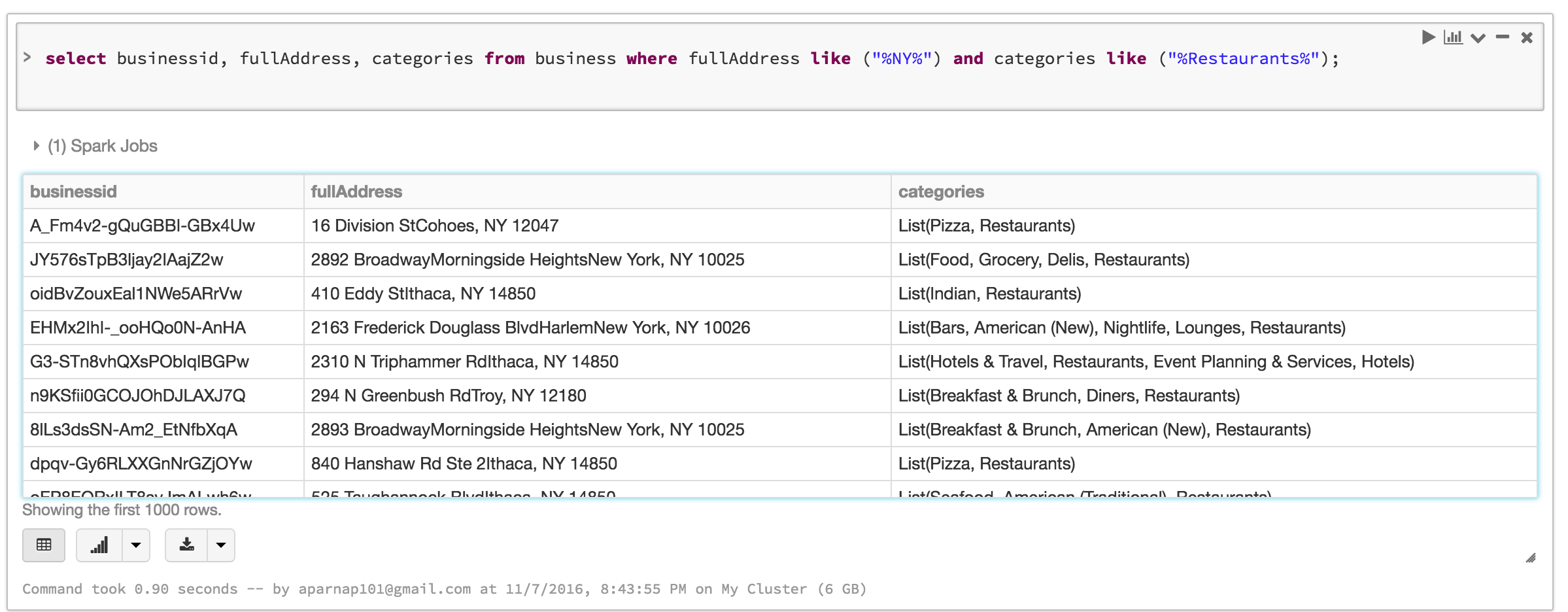
LOAD DATA INPATH "/FileStore/tables/yck4y8os1478572995296/user.csv" OVERWRITE INTO TABLE user;

\*/

**Question 2:**

select businessid, fullAddress, categories from business where fullAddress like ("%NY%") and categories like ("%Restaurants%");

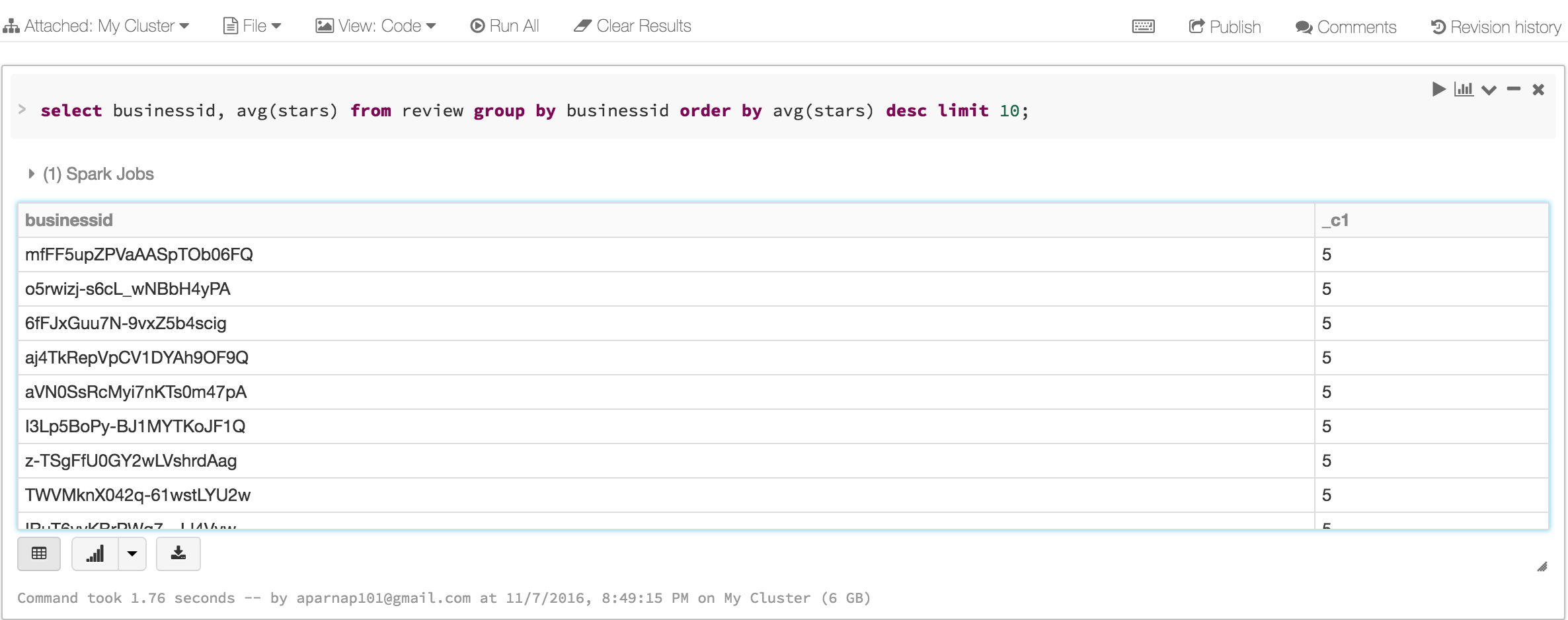
**Result:**



**Question 3:**

select businessid, avg(stars) from review group by businessid order by avg(stars) desc limit 10;

**Result:**



**Question 4:**

create view businessview as

select businessid, substr(fullAddress,(length(fullAddress)-4)) as zipcode from business;

select zipcode, count(businessid) from businessview group by zipcode order by count(businessid) desc limit 10;

**Result:**

