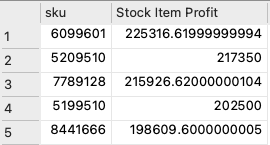
Updates: (Oct 8th-Oct 13th)

* Created a GitHub repository accessible to four team members, a TA, and Professor Diego
* Installed pgAdmin, established a connection to the MLDS Postgres server, and conducted an initial exploration of the pgAdmin interface
* Established the tables in PostgreSQL (deptinfo, skstinfo, strinfo) according to the database schema and provided access to all team members (Note: we faced issue importing the other two tables because of ambiguous column label and mismatch between the number of columns in the schema and the actual dataset.)
* Summary Statistics:

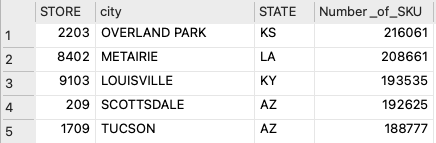
1. Mean, Min, Max, SD of “COST”, “RETAIL” in table skstinfo

Summary Statistics:

* Found a lot of problems with the data, including incomplete foreign keys and poorly defined and duplicated columns in trnsact.
* In particular, the last (undeclared) column of each table has columns of 0 and 1.
* Total 787669 distinct SKU inside the skuinfo table.
* Shown the Top 5 Most Profitable Stock Items, sum(retail) - sum(cost)



* Figured the Top 5 busiest store in the US (based on counting the number of stock item)



Next Steps: (Oct 16th-Oct 21st)

* Understand, clean, and import the other two tables (skuinfo, trnsact) into PostgreSQL
* Explore data and the table relationship