

Simple SQL - 9/11

Announcements

- **IMPORTANT:** Do not leak AWS account credentials! This means you **SHOULD NOT** share over the internet nor commit to GitHub. Share your credentials physically with teammates or securely in Stache.

Reading Quiz

- Q1: B.
- Q2: B.
- Q3: A.
- Q4: C. `DISTINCT` is the keyword that is used to remove duplicates from the results of a `SELECT` statement.
- Q5: A. The array datatype is part of Postgres but not part of the SQL standard. This is an example of an extension.

Entity Relationship Diagrams

Each table represents an entity and each row in the table describes the attributes of the entity. The connections between the tables represent the relationships between entities. Note the crows feet and bar at the end of the connections. These represent the cardinality of the relationships.

Athena

Presto is a SQL query engine which allows you to query all sorts of data sources including delimited files of all sizes. Athena is just the open source Presto software packaged as a cloud service.

Simple SQL

```
SELECT * from aisles;
```

This selects all the records along with all its attributes from the `aisles` table.

```
SELECT * from departments LIMIT 15;
```

This selects only 15 records from the `departments` table.

Instacart

It is our job to be able to find out which tables relate to which. We need to determine different things from manually inspecting the dataset or prior knowledge of the dataset. This can include finding primary keys, foreign keys, datatypes, and relationship cardinality.

Practice Problem 1

The Order_Products table requires a compound primary key of order_id and product_id since it does not have any other unique key using just a single column.

Next Class

Try to get Athena up and running by next class in order to troubleshoot during office hours or after class. You will need it for lab time next Monday.