Quiz 5

This quiz will give you practice on group by statements and aggregate functions when grouping by a single column. For this quiz please use the data from here:

`bigquery-public-data.new_york_trees.tree_census_2015`

Read over the schema page as it will explain what each column in the dataset means.

1.

When counting the number of trees by health, which group within health has the least amount of trees?

Which has the most? Note: Null can be an answer.

2.

For each type of tree (use the spc_common column), calculate the average diameter (use the tree_dbh column).

How many types of trees have an average diameter strictly greater than 10?

3.

Of all the trees that have a health status of "Poor" and a tree diameter greater than 10 (tree_dbh > 10), how many are Damaged and how many are Not Damaged (as measured by the sidewalk column)?

4.

Consider the trees for which user_type is "TreesCount Staff" or "NYC Parks Staff", and spc_common is not "London planetree", and curb_loc is "OffsetFromCurb". For the trees that meet these conditions, find the maximum tree_dbh for each of the different categories/groups in the guards column.

5.

What is the maximum, minimum, and average tree_dbh across the entire data set?

HINT: When using aggregate functions, you do not have to use a group by. You can just use aggregate functions in a simple select statement from the table. It will work as long as you are only returning aggregated columns and not ordering by any column.

For example (although this is rather meaningless) it will still execute:

```
SELECT

  SUM(tree_dbh)

FROM

  `bigquery-public-data.new_york_trees.tree_census_2015`
```