This is a brief structural analysis of the first SQL query in part ‘Find correlation between weather and complaints’.

https://google.qwiklabs.com/focuses/609?parent=catalog#step9

**SELECT**

**descriptor,**

**sum(complaint\_count) as total\_complaint\_count,**

**count(temperature) as data\_count,**

**ROUND(corr(temperature, avg\_count),3) AS corr\_count,**

**ROUND(corr(temperature, avg\_pct\_count),3) AS corr\_pct**

**From (**

**SELECT**

**avg(pct\_count) as avg\_pct\_count,**

**avg(day\_count) as avg\_count,**

**sum(day\_count) as complaint\_count,**

**descriptor,**

**temperature**

**FROM (**

**SELECT**

**DATE(timestamp) AS date,**

**temperature**

**FROM**

**demos.nyc\_weather) a**

**JOIN (**

**SELECT x.date, descriptor, day\_count, day\_count / all\_calls\_count as pct\_count**

**FROM**

**(SELECT**

**DATE(created\_date) AS date,**

**concat(complaint\_type, ": ", descriptor) as descriptor,**

**COUNT(\*) AS day\_count**

**FROM**

**`bigquery-public-data.new\_york.311\_service\_requests`**

**GROUP BY**

**date,**

**descriptor) x**

**JOIN (**

**SELECT**

**DATE(timestamp) AS date,**

**COUNT(\*) AS all\_calls\_count**

**FROM `<YOUR-PROJECT-NUMBER>.demos.nyc\_weather`**

**GROUP BY date**

**) y**

**ON x.date=y.date**

**) b**

**ON**

**a.date = b.date**

**GROUP BY**

**descriptor,**

**temperature**

**)**

**GROUP BY descriptor**

**HAVING**

**total\_complaint\_count > 5000 AND**

**ABS(corr\_pct) > 0.5 AND**

**data\_count > 5**

**ORDER BY**

**ABS(corr\_pct) DESC**

**Simplified structure:**

**SELECT …**

**FROM (SELECT …**

**FROM a**

**JOIN (SELECT …**

**FROM x**

**JOIN y**

**ON …) b**

**GROUP BY …)**

**GROUP BY …**

**HAVING …**

**ODER BY … DESC**