DATA QA CASE STUDY

Instructructions -

Analyze the data using SQL, Python, or other scripting languages, and run automation to generate the aggregation below from the raw data supplied. Compare the output already supplied and mark if the aggregations based on the source data are correct. While submitting the case study back, retain all the scripts as is and also provide feedback on the data issues, inconsistencies, and notes on transformations made/required with any other observations.

* Total Districts
* Total unique schools under each district
* Total students under each district
* Session status/subject summary rolled up to School and District hierarchy.
* Same Tutor continuity rate against unique students per subject rolled up to District in %’s(Tutor matchup-rate)
* School with the Highest average exit ticket score in Math and English separately
* Average exit ticket scores rolled up to the district.
* Mastery Rate %
* District with Highest/Lowest Tutor/Session rating