The location of this multiple datasets comes under the California SAT, <https://data.world/education/california-sat-reports>

I found the dataset in the dataset sources provided by the professor privately in the field of education. The data can be downloaded for free and it was available in the website of “data-world” which is famous for free datasets. As this dataset is free and the website is legal, downloading of this dataset doesn’t require any legal permissions. The project is regarding multiple datasets which contains California SAT scores from 1998 to 2016. I am aiming to combine the datasets from 2010 to 2016 and their number of rows are 2431, 2473, 2595, 2293, 2332 and 2335 rows. The sizes ranged from 500KB to 600KB. The common columns the datasets contained are County number, District number, School number, School name, grade 12th marks, mean score, average score in math and verbal. Multiple datasets talk about the number of students increased over the years for exams and how the scores have increased over the years. The data combined won’t have more than 2600 rows as that is the highest number of rows and the maximum the combined dataset occupied is 3-4MB for approximation. The dataset contains two types of dataset.

Hypothesis: I am assuming that the mean score has increased by atleast 10 percent for every years. Additionally, I am also assuming that the number of takers are increasing by 15 percent.

Question: The percentage increase per school for exam takers over the years. Additionally, which county presented the largest change in the increase of takers and mean scores and are they same?