Case:

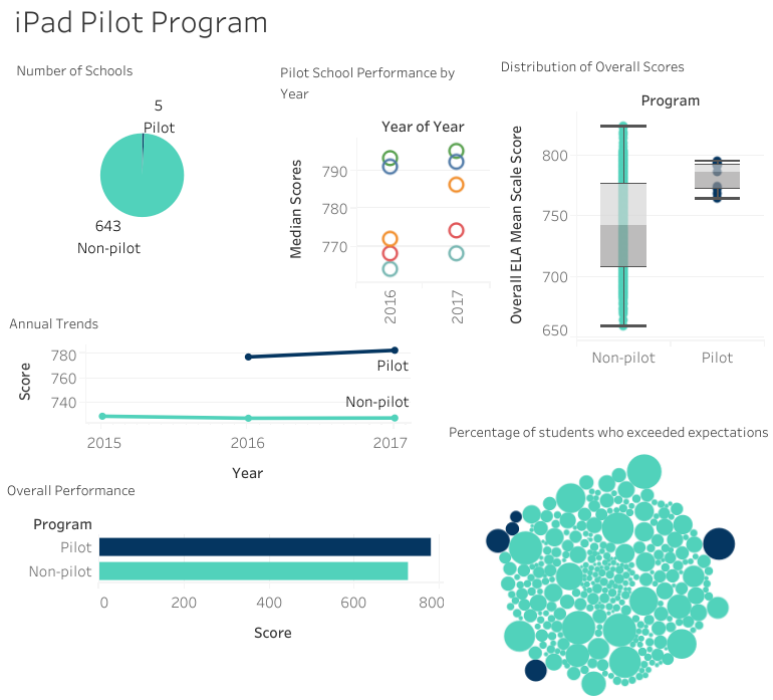
Imagine you are placed on a project at a large school district that has been focused on strategies to improve student outcomes on end of year state tests in English Language Arts (ELA). Before the 2015-2016 school year, the superintendent in collaboration with the school principals, selected 5 schools to participate and pilot a new iPad-based learning system (schools 610534, 609880, 610066, 610082, 610182).

All 5th graders in the 5 schools received a school-issued iPad. During the school day, students spent 1 hour per day completing an ELA program on the iPad. The ELA program had levels of difficulty ranging from 1 - 10, with each level having activities on reading comprehension, writing, and vocabulary. Students needed to demonstrate their mastery of the ELA standards on each level to proceed to a higher and more difficult level.

The superintendent is dedicated to improving student performance across all schools and better-utilizing technology as a learning tool. After conducting the pilot, she wants to expand the iPad program to all 5th graders. As adopting the program across the whole district will be a significant investment, she wants to make sure that it is the right place to prioritize funds before moving forward.

As an initial first step, the superintendent has asked you to look at available State Test Score Data and provide data points that will influence and/or determine immediate next steps. These next steps could vary from further data analysis, to engaging stakeholders on the results of the pilot, to formulating a formal recommendation on whether to roll out the pilot to the rest of the district.

Tableau Dashboard:



Process Notes:

Data set-up and sanitization

- I decided to use excel to sanitize the Data as opposed to R, Python or other software as the cleaning process was very excel-specific (merged cells as opposed to null values or duplicates, for example). It would have been time-consuming and unwieldy to use unsanitized data in other softwares.
- I removed blank rows with a filter and delete function, and deleted blank columns.
- I unmerged the years column values with a custom macro I'd coded earlier (there are now easier ways around this that also work).
- I first checked for duplicates. This involved identifying what is the unique value for our purposes in this dataset. T creating a checker column with concatenates of the columns containing School ID, Year and Test Name as these are the unique combinations to result from this dataset, ie, each school should have one test score per year per grade-level test.
- Then I created a Pilot school column that identified whether a school was one of the five included in the program, using a combination of IFERROR and VLOOKUP
  =IFERROR(VLOOKUP([SchoolIDCell],[PilotSchoolsRange],1,0),"No")
- This can also be done natively in R with the following function: transform(data, type = ifelse(`School ID` %in% c(610066, 610082, 610182, 610534, 609880), "Pilot", "Non-Pilot"))

Key Insights and Recommendations:

- The iPad Pilot program seems to have generally favourable results for the majority of participating students. The pilot program was limited to 5 schools, and expansion of the program to other schools in the program is recommended.
- However, given that the range of scores is narrowerin the pilot schools as opposed to the non-pilot schools, it's unclear whether the program is having a counterproductive effect on those students who, under other circumstances, would have exceeded expectations.
- While more granular information on the subscores of the test seem to have been collected, they are incomplete in their inclusion within this dataset. Access to the subscore would greatly improve the scope for recommendations.