**Assignment 4 - pandas\_challenge – Kenneth Le**

**Background**

You are the new Chief Data Scientist for your local government area. In this capacity, you'll be helping the school board and mayor make strategic decisions regarding future school budgets and priorities.

As a first task, you've been asked to analyse the area-wide standardised test results. You'll be given access to every student's maths and reading scores, as well as various information on the schools they attend. Your task is to aggregate the data to showcase obvious trends in school performance.

### Instructions

Using Pandas and Jupyter Notebook, create a report that includes the following data. Your report must include a written description of at least two observable trends based on the data.

#### Written Report (15 points)

To receive all points, the written report presents a cohesive written analysis that:

* Summarises the analysis (5 points)
* Draws two correct conclusions or comparisons from the calculations (10 points)

**Anaylsis**

This report aims to analyse the various schools within the local government area in order to assist the schoolboard and mayor to make strategic decisions regarding the future school budgets and priorities. The data provided covers 15 local schools with information in regards to what type of schools there are, current budget allocated, size of the schools as well as further details on the students enrolled at the current schools and their maths and reading scores. We hope to discover trends in the data through our analysis and provide insight to our stakeholders.

For starters, we have analysed the data and created various statistics to assist in our analysis. In total, we have 15 schools in the local area with a total of 39,170 students where the average scores for maths and reading is 70.34 and 69.98 respectively.

If we are to look at the schools with the highest maths and reading scores we see that within the top 5 schools are mostly Independents (4 out of 5) with Griffin High School and Cabrera High School taking the top 2 spots in terms of overall passing score percentages for maths and reading with Griffin achieving a Overall Pass % of 81.34 and Cabrera achieving a pass percentage of 80.79%. On the other side of the spectrum we see that the bottom 5 for Overall passing score are mostly Government schools (4 out of 5 again) with Hernandez High School and Huang High School placing in the bottom 2 with a score of 66.36% and 66.71% respectively. From this we may assume that Independent School are more successful based of the overall passing score percentages.

If we look further into this angle that Independent Schools are more successful that Government schools, we can look at the budget allocation of these schools and see if we should continue funding as is or to allocate more towards a specific type of school.

If we look at the current budget allocation, we can see that the majority of the fund allocation goes into the Government school budget compared to the Independent schools. The lowest budget amount of Government school is $1,763,916.00 (Ford High School), which is higher than the highest budget allocation for Independent school with Wilson High School only receiving $1,319,574.00. Further to this, in our analysis we looked at how much of the budget is allocated per student at each school and found that in government schools the range of per student budget allocation is always in the highest range ($645-680) compared with Independent Schools (<$585). This would indicate that funding and budget allocation goes towards Government schools more so that Independent schools.

If we combine the 2 sets of data and analysis, we can conclude that even though more of the budget is allocated towards Government schools compared to Independent schools, the overall passing percentage of students in Independent schools is higher than those in Government schools. This is a good indication that perhaps the budget allocation should be more favourable towards Independent schools to further support their students to achieve even more overall passing rates in the future. However, this is not a wholly conclusive assessment as perhaps there is always a degree for error in the data. A limitation in this data would perhaps be towards the student numbers in Government schools being much higher than that of Independent schools and with such a higher volume of students, perhaps the scores may have been more skewed towards the lower end with some outstanding student scores being brought down by peers in said schools. If we were to normalize the student numbers in each school we may get a more clearer picture of the overall passing percentages.