**Project Title:** Secret Behind Academic Success

**Team Members:** Joshua Tean / Wilfred Andoh / Yan Shao / Haodong Zhang

**Git Hub Link:** https://github.com/panlovelyda/project-one.git

**Project Description/Outline:**

Data were obtained in a survey of students Maths and Portuguese Language courses in Secondary School. The purpose of the project is to use data analytic tools to reveal the relationship between students’ family socio-economic status, students’ study behaviour, students’ alcohol consumption and students’ grade. In conclusion, we will find which factor affect students’ grade significantly.

**Research Questions to Answer:**

* Basic Data Summary - analysis by gender/age/school/socio-economic status - YS
  + The data set has a nearly equal gender distribution. There are 370 students in this survey, 52.70% of them are girls and 47.30% of them are boys.
  + Based on this dataset, Males performed better in Math while females performed better in Portuguese.
  + From the bar chart, we could conclude that the average age of students who consume alcohol was 16.58 years old, while the median age is 17. I think most students were surprisingly young. If you don’t know, the legal age for alcohol consumption in Portugal is 18.
* How does students’ alcohol consumption impact a student's grade? - WA
  + As it is shown in the pie charts, during the weekdays, nearly 70 % of the students consumed quite a small amount of alcohol, on the contrary, the percentage of students who drank very little on the weekend was around 38%.
  + Based on our data set, we can see that weekday alcohol consumption is way less hence scoring a higher percentage as per, which therefore implies students drink less over weekdays and thus have a higher chance to study and make better grades. As can be seen in their G3 math and G3 Portuguese
  + Unlike weekends, students tend to drink more, and thus have less time in studies, but it would appear the more social interaction, their Portuguese grades are higher. But there is a complete opposite with their math grades, the more alcohol they drink, the less time they have to study hence affecting their math grades.
  + We can conclude that, the ratio of students who drank over weekdays was far less compared to those over the weekend. But notwithstanding there was a significant jump in the number of persons who drank over the weekend.
* How does students’ study behaviour impact a student's grade? – HZ
  + Students Study time has positive impact on Students’ Grade. Although the grade is not overwhelmingly higher for those who study longer, the boxplot shows that with more time spending on study, 75% of the student with same study time have better results
  + Through the scatter plot we can see no significant difference between absences rate and final grade. However, slightly we can see that for Portuguese, the less absent rate lead better final Grade but cannot see that trend in Math.
  + After comparing the average grade of other factors, we can see that extra educational support from school or family, extra paid classes and extracurricular Activities don’t contribute to the final grade.
  + For all students’ behaviours, study time is the key for academic success which also aligns with the old saying that no pain no gain. More time you spend on study, better results you will achieve.
* How does students’ family socio-economic status impact a student's grade? – HZ / JT
  + Students' absences are not affected by where students are living as we can see from the count of bar charts.
  + From where students living and their relative distance to school, we can see that those who live closer to school have slightly advantage on academic results as they probably can spend more on study.
  + From family relationship linear regression figure, we can say that most students have a relatively good family environment which helps student to get a better academic result in Portuguese and Math
  + In general, the family relationship of students’ family socio-economic status has slight positive impact on students’ grade.
* Which factor has the biggest impact on students’ grade – JT
  + Overall, we can see that Students’ daily and weekly alcohol consumption, family relationship, students’ study time, travel time from home to school and classes’ absent rate have impact on Students’ grade.
  + For Portuguese’s Language classes, daily alcohol consumption has most negative impact on academic results while study time has most positive impact.
* Null Hypothesis:
  + Study time has impact on Final Grade. - T test rejects the hypothesis for Math and is true for Portuguese
  + Alcohol consumption have impact on Final Grade. - T test rejects the hypothesis for Math and is true for Portuguese

**Datasets to Be Used:**

<https://www.kaggle.com/datasets/uciml/student-alcohol-consumption>

**Rough Breakdown of Tasks:**

* Use Pandas to clean and format your dataset or datasets. - JT
* Create a Jupyter notebook describing the data exploration and clean-up process. - ALL
* Create a Jupyter notebook illustrating the final data analysis. - JT
* Use Matplotlib to create 6 to 8 visualisations of your data (ideally, at least 2 visualisations per “question” that you ask your data). - ALL
* Save PNG images of your visualisations to distribute to the class and instructional team—and for inclusion in your presentation. - ALL
* Create a write-up summarising your major findings. This should include a heading for each “question” that you asked your data as well as a short description of your findings and any relevant plots. – ALL
* Presentation – Making Power-point on Tuesday / Practice Wednesday & Thursday - ALL