

Data Analytics for Student Profiling and Academic Counselling

Jizhi Li

Singapore University of Social Sciences
Singapore

Wai Ping Low

Singapore University of Social Sciences
Singapore

Lokesh Bheema Thiagarajan

Singapore University of Social Sciences
Singapore

Lu Chang Peh

Singapore University of Social Sciences
Singapore

Data analytics can be used by universities and schools to have a deeper understanding of student and learning data. By leveraging on data analytics and dashboards, universities and schools can become more proactive in profiling students and anticipating their needs, personalizing approaches to supporting students in academic distress and optimizing the allocation of university's resources to efficiently and effectively counsel these students. In this paper, we outline the analytics framework that can be used on student data to derive insights, to readily observe and predict the students' academic progression and performance, to characterise the academic risk of the student, and to identify the at-risk students at an early stage. With the early alert system in place, these students can then be counselled and rendered student support to be lifted out of the at-risk zone.

Keywords: data analytics, dashboards, profiling, prediction, early alert, student support

Introduction

Data analytics is shaking up every sector in the world by harnessing and extracting meaningful insights from data ([Henebery, 2019](#)). Every time students interact with the university, they leave a digital footprint. This footprint, together with the student's background supplied during admission, is valuable data that can be effectively used for modelling/predicting student behaviour performance ([Lester et al., 2017](#)). Learning has become challenging in this current day environment filled with distractions ([Garrison, 2010](#)). It becomes even more challenging in the case where the students are working adults pursuing their learning and juggling their work and family commitments ([Smith, 2017](#)). Some of these students could potentially be at risk of discontinuing their learning or under-performing academically, thereby leading to increased attrition rates. Timely intervention and counselling provide an effective learning system, which could possibly help in reducing the attrition rate ([Payne, 1973](#)). An effective dashboard with data visualisation and analytics is helpful in quickly identifying the at-risk students and the reasons for their risk levels ([Klerkx et al., 2017](#)). Furthermore, these dashboards would provide simple interface to the end-users, viz., the academic counsellors and faculty, to use and reach out to the at-risk students. This way, the limited resources, such as funding, faculty time and students' time could be effectively utilised to transform the at-risk students and provide them with a conducive learning system ([Beach, 2013](#)).

Background Information

The School of Science and Technology (SST) is one of the five schools in the Singapore University of Social Sciences (SUSS). Currently, the School admits only part-time students in two admission exercises each year, one in January and the other in July. The majority of applicants are diploma holders from local² polytechnics and the rest are from junior colleges, statutory board academies and private education service providers. SUSS does not impose a strict Cumulative Grade Point Average (CGPA) cut-off unlike some of the other local autonomous universities. Most of SUSS's students are adult-learners and have to juggle many commitments at any one time (e.g. family, work and study). Thus, the attrition rate of students is significantly high. The objective of this project is to implement student intervention in a timely manner to enhance student experience and to improve students' chances of completing their studies. This has to be a concerted effort from a number of departments in the university, such as the Teaching and Learning Centre (TLC), Business Intelligence & Analytics (BI&A) department and the School.

² "Local" in this paper refers to Singapore.

Unique and Salient Features of SUSS and SST

The mandate of SUSS and SST is to provide high-quality life-long learning opportunities, adhering to the University and School's mission "to provide lifelong education, equipping learners to serve society". One of the admission criteria is that the applicants are required to have either a full-time job or at least two years of working experience. SST offers nine undergraduate programmes that are organised in three clusters, namely the Engineering, Built Environment and Information and Communication Technology (ICT) clusters. These programmes are:

- Engineering cluster: Aerospace, Biomedical and Electronics engineering programmes
- Built Environment cluster: Building and Project Management, Facilities and Events Management, and Human Factors
- ICT cluster: Mathematics, ICT and Digital Media

In addition, the School offers a Master by Research and an Industrial PhD programme. This project will focus on only the part-time undergraduate degree programmes.

In order to graduate, a student has to attain a minimum CGPA of 2.0 (out of 5.0) upon completion of the credit units (CUs) required for graduation, which are 130 CUs and 170 CUs for a basic degree and honours degree, respectively. When a student's CGPA drops below 2.0 for the first time, the student will receive an Academic Warning notification. If subsequently the student still scores CGPA below 2.0, then the student is given an Academic Termination and removed from the programme. In SST, the at-risk students were identified as the students with CGPA below 2.3. The identified at-risk students were given academic counselling by the Head of Programme (HOP)³ through emails, phone calls and/or face-to-face meetings. Through interviews with more than 200 of these at-risk students, it emerged that their weak academic performance was due to poor time management and lack of administrative awareness, viz., academic progression, course requirements, available options, etc. These at-risk students were generally in the first three semesters of study and they contribute significantly to the attrition rate.

Peer tutoring has been shown to reduce the student attrition rate (Bryer, 2012). To further assist these at-risk students, SST set up a peer mentoring support network in 2018. This network was designed to help new students settle into university life. Senior students or recent graduates take the role of peer mentors and are matched with freshmen from their undergraduate programme for the first semester. The mentors provide signposting and provide survival tips to the freshmen. SST also has a peer tutoring scheme where academically strong seniors provide additional academic support to students in selected courses.

The current interventions are either pre-emptive, where peer mentors provide skills/information to guide the freshmen, or retroactive where students receive academic counselling after failing to attain the minimum CGPA. There are students who are not freshmen nor at-risk, but whose deteriorating CGPA may result in a lower class of Honours classification for them. These students are outside the current intervention radar and hence do not receive any academic counselling or support unless they are proactive in seeking help. This project studies the feasibility of using an early alert system to detect at-risk students across the entire CGPA range from 0 to 5.0, so that timely intervention and support could be rendered.

Business Intelligence & Analytics (BI&A)

The Business Intelligence & Analytics (BI&A) department was set up to provide information for data-driven and evidence-based decision making and planning in SUSS. This covers not only the digital warehousing of critical data, but also the twin functions of reporting and analytics, and these are achieved through training and project collaborations.

Teaching and Learning Centre (TLC)

The Teaching and Learning Centre (TLC) aims to promote excellence in teaching, supports the learning needs of students, and strengthens ties among faculty to foster a vibrant academic community. This study focused on the two student-facing functions of the centre, namely: meeting the learning needs of students through a range of

³ HOP is the full-time academic faculty who oversees the undergraduate degree programme.

support measures such as resources of study skills and workshops and to share teaching tools, good practices and guidelines, and building a repository of teaching and learning resources.

Predictive Modelling for Academic Performance

In order to generate early alert signals for students who may potentially under-perform academically, the predictive model uses Semester Grade Point Average (SGPA) as the target and predicts how a student is projected to academically fare in each semester. Input variables used for prediction include students' demographics, work background, prior education and their academic performance in SUSS. Academic data is updated at the start of each semester while the other student's particulars were mostly collected upon admission to the university. Derived variables were computed using the collected variables to provide insights on students' learning pattern in and across semesters.

Table 1: Input variables for the model and dashboard

Demographic & Work Background (12)	Collected	Age at year of intake Residency status Gender Marital status Race Present employment status Industry Current designation level Gross annual salary Company sponsorship status Applicability of degree to work environment Relevant industry experience
	Collected	Ordinary level (O-level) English grade O-level Mathematics grade Final education category based on qualification Final education awarding institution Final education study mode Degree/polytechnic CGPA
Prior Education (8)	Derived	Whether Diploma is relevant Years since last study
	Collected	Discipline CUs exempted Restart status Transfer-of-Programme (TOP) status Minimum CUs required Previous semester's CGPA
SUSS Academic (17)	Derived	Years into degree Deferment status Ratio of CUs taken to Min CUs (indicates level of programme completion) Ratio of CUs withdrawn to CUs taken (relative proportion of courses withdrawn) Ratio of CUs taken for university core courses (UCOR) to CUs taken (relative proportion of UCOR courses taken) Ratio of no. of exam (EQP) courses taken to total no. of courses taken (relative proportion of courses with exam component) Total CUs withdrawn Total CUs completed Total CUs taken for UCOR Total no. of EQP courses taken No. of active semesters

A total of 37 variables, as listed in Table 1, were used as inputs to the models. The model selection mechanism is such that the model with the best prediction performance is retained and the rest are sieved out. To minimise over-fitting and to ensure that the model is generalizable, this study used the Construction, Validation & Testing (CVT)

method for model building and selection. In this method, a construction dataset is used for model training and testing dataset is used for model testing. The final model selected is then applied onto a deployment dataset for scoring.

Since the past semester's data is used for model construction which will then be deployed on the current data, it is necessary to test if the model is stable over time. Therefore, in this study the recent one semester data is used as the testing dataset. If the model performance is consistent for both the construction and testing dataset, it indicates that data is relatively homogenous at different points-in-time and that the model built with historical data is applicable on current and future data. The five-step CVT procedure is as follows:

1. Conduct five-fold cross-validation using the construction dataset
2. Use the full construction dataset to construct a model and test it on the full testing dataset
3. Repeat steps (1) and (2) for various different models
4. Select the model with highest R^2 and lowest Mean Absolute Error (MAE) as the best model
5. Apply the selected model on the deployment dataset to obtain prediction scores

In this method, step (1) validates the model performance in a robust way while step (2) tests the out-of-sample performance of the model on unseen data. Ideally, a good model should be stable in its performance for both the validation and testing steps.

The procedure was applied on various models such as Linear Regression, k Nearest Neighbour (kNN), Neural Network and Decision Tree. The model that produced the best and also consistent results was considered as the optimal model for deployment. Scoring will be done using the new data in the up-coming semester to predict students' SGPA and compare it with the actual SGPA of the latest semester.

These predictions will be integrated and deployed into the Dashboard for Predictive Modelling of Student's Academic Performance, acting as additional information and providing early signs of potentially under-performing students so that targeted support can be provided.

Dashboard Construction

There are three overlapping dashboards created for the School, namely: Application Data (Figure 1), Student Profile (Figure 2) and Predictive Modelling of Student's Academic Performance (Figure 3). In addition, the dashboard consists of three layers of information; each successive layer allows the user to "zoom in" for additional details, views and perspectives. This is rather similar to how [Eckerson \(2010\)](#) categorised them – bottom layer as detailed reporting view (individual students), middle layer as multidimensional view (where one can explore or "slice and dice" the data) and top layer as summarised using a graphical view. The data are extracted from the data warehouse, which collects SUSS data, sourced from applications such as the University's Student Information Management System (SIMS) that allows users to generate information and insights for decision making. The data warehouse is currently developed to house the complete SUSS student data. It will also provide point-in-time snapshots to support users in their data needs.

The first two dashboards are built from factual data, in which the Dashboard for Students' Application Data provides an overview on all applicants of the School, their demographics and awarding institution, and the Student Profile Dashboard presents background information of admitted students, giving the School a comprehensive view of the current student population. On the other hand, the Dashboard for Predictive Modelling contains both factual and prediction data, allowing for extrapolation of students' future performance. These can be used together with students' historical performance data to understand association between past and future performance so as to design customised coaching.

Dashboard for Students' Application Data

Most student applicants are from the five local polytechnics. The secondary pipeline is junior colleges. Finally, there is a small group of applicants from statutory board academies and private education providers. The dashboard will enable the School and HOPs to visualise the number of applicants, places offered and actual enrolment of students, demographics of the applicants, prior education, programme they have applied, funding status etc. For instance, from the charts/graphs in the dashboard, the Dean and HOPs can determine if there is a healthy ratio of application-to-offer and offer-to-enrolment student numbers.



Figure 1: Dashboard on Student Application Data

Table 2 lists a few observations obtained from the dashboard and actions proposed or taken to address these issues.

Table 2: Observations from the Application Dashboard and Actions Taken

Observations from the dashboard w.r.t. Applicants' Data	Actions proposed/taken to address the issues w.r.t. Applicants' Data
Gender imbalance – there are significantly more male than female applicants	To market the programmes with female student ambassadors to overturn the impression that the Science, Technology, Engineering and Mathematic (STEM) industries are dangerous, dirty, demanding and dominated by men, to one that is professional, progressive and productive that women can have an important role to play
Age profile of applicants	Targeted marketing efforts focusing on Generation Y, Generation Z and the millennials through social media since they are more IT-savvy
Education background and awarding institutions	The School can collaborate with Institutes of Higher Learning (IHLs) that are major pipelines to the Schools' programme, in articulation pathways such as Through-train and Earn-and-Learn programmes

Student Profile Dashboard

The Student Profile Dashboard presents a big picture of all students in SST. Consolidated information of all students including those who have graduated, terminated and withdrawn from their programmes provide a holistic view of all past and present students, allowing the dashboard user to have an overview of the students'

demographics and prior education and employment status. This dashboard also charts the students' academic progression, i.e., each students' CGPA in the latest semester.

From the dashboard in Figure 2, it can be observed, for instance, that there are 2,777 active students in the School across the nine disciplines/programmes. The dashboard allows the Dean and HOPs to monitor the number of admissions into each of the programmes, as well as to filter and focus on specific groups of students such as Continuous Education and Training (CET) students, Transfer of Programme (TOP) students, Restart students, and students within a certain CGPA range. With lifelong learning and CET becoming more commonplace, the School has to cater to demand-driven education and braiding workplace and in-class learning to scaffold and impart sector/profession specific, generic and transferable skillsets that matter (Deegan and Martin, 2018). Table 3 lists the observations and actions taken by observing the student profile dashboard.

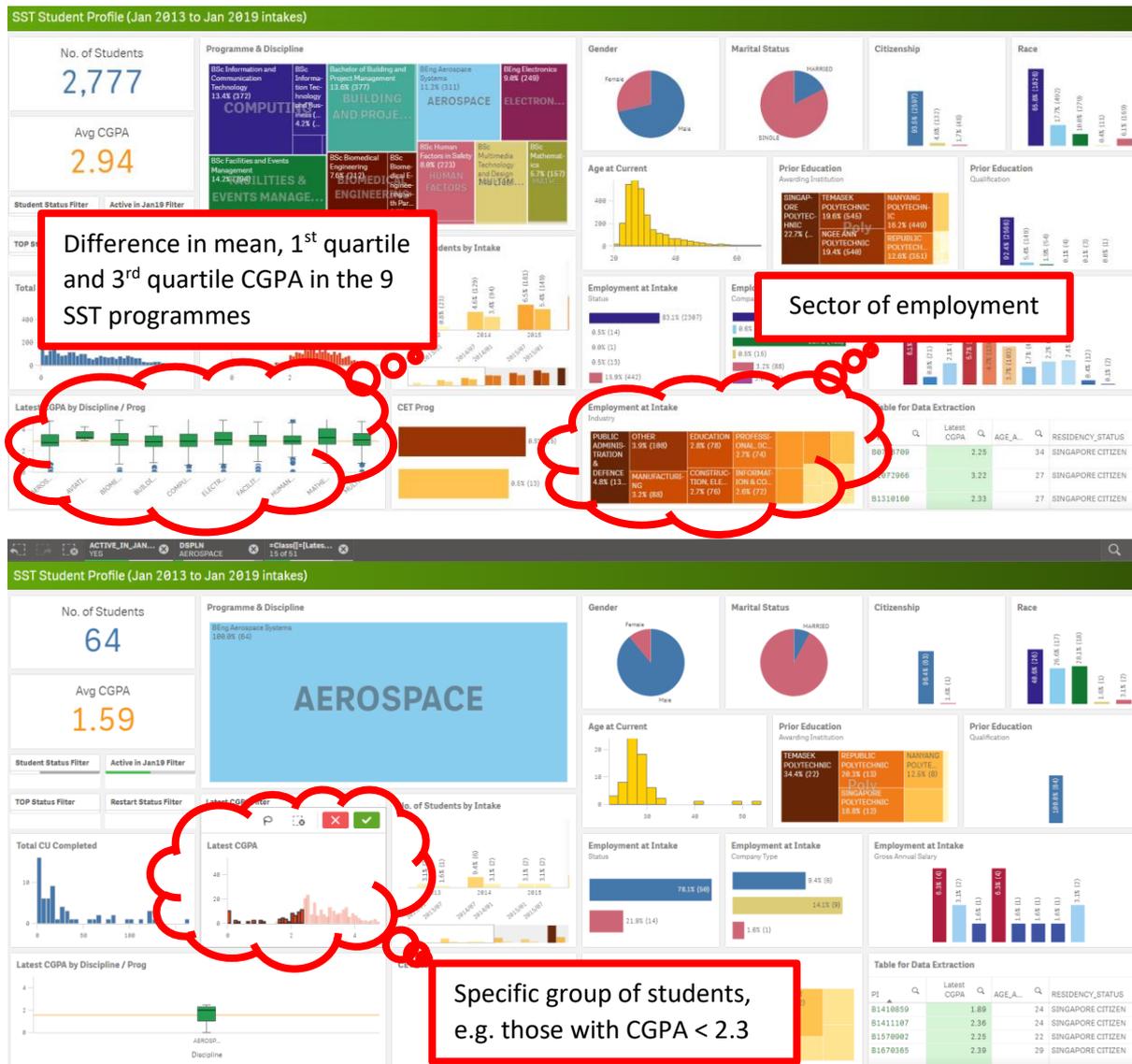


Figure 2: Dashboard on Student Profile

Table 3: Observations from the Student Profile Dashboard and Actions Taken

Observations from the dashboard w.r.t. Student Profile	Actions proposed/taken to address the issues w.r.t. Student Profile
Some programmes have a mean, 1 st quartile and 3 rd quartile CGPA that are significantly/comparatively higher than the other programmes	To review if the courses and assessments are lenient compared to other programmes.
The major sectors of employment can be identified at the school and programme levels	Outreach and marketing efforts can be more targeted; Training can be customized to integrate in-demand industry skills
Specific groups of students, for instance CET students, TOP students, Restart students, or students with a specific range of CGPA can be identified	These students can be identified and tracked for their academic performance, and intervention measures can be implemented

Dashboard for Predictive Modelling of Student’s Academic Performance

For the model results to be readily usable by faculty members and staff, the results/predictions are integrated into the dashboard to provide early alert signals for potentially at-risk students. Charts and visualisations are able to compare students’ predicted score against their actual latest CGPA/SGPA, and categorise such predicted change in CGPA/SGPA by whether there will be an improvement, deterioration or relatively constant trend (Figure 3).

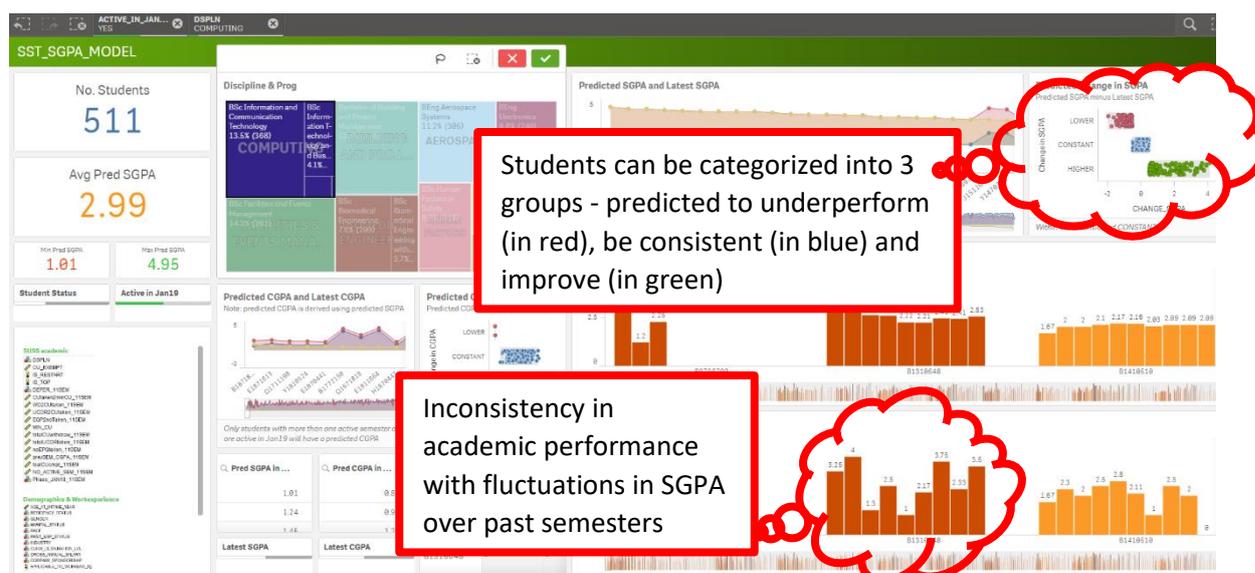


Figure 3: Dashboard on Predictive Modelling of Student’s Academic Performance

Depending on the cases, customised advisory or encouragement notes can be sent to the respective students. To relate academic performance to learning pattern, the dashboard also displays students’ historical grades, course withdrawal and failure information to provide insights on the potential factors or reasons for the predicted academic under-performance.

Comparison of predicted score across groups such as restart students versus non-restart students, or across different disciplines and programmes is able to provide a bigger picture on how these factors are related to academic performance (Figure 4). The filters allow the dashboard users to zoom in to a selected subgroup of the student population, for example, students who are at the verge of getting Academic Warning or those who are close to getting a First Class Honours degree award.

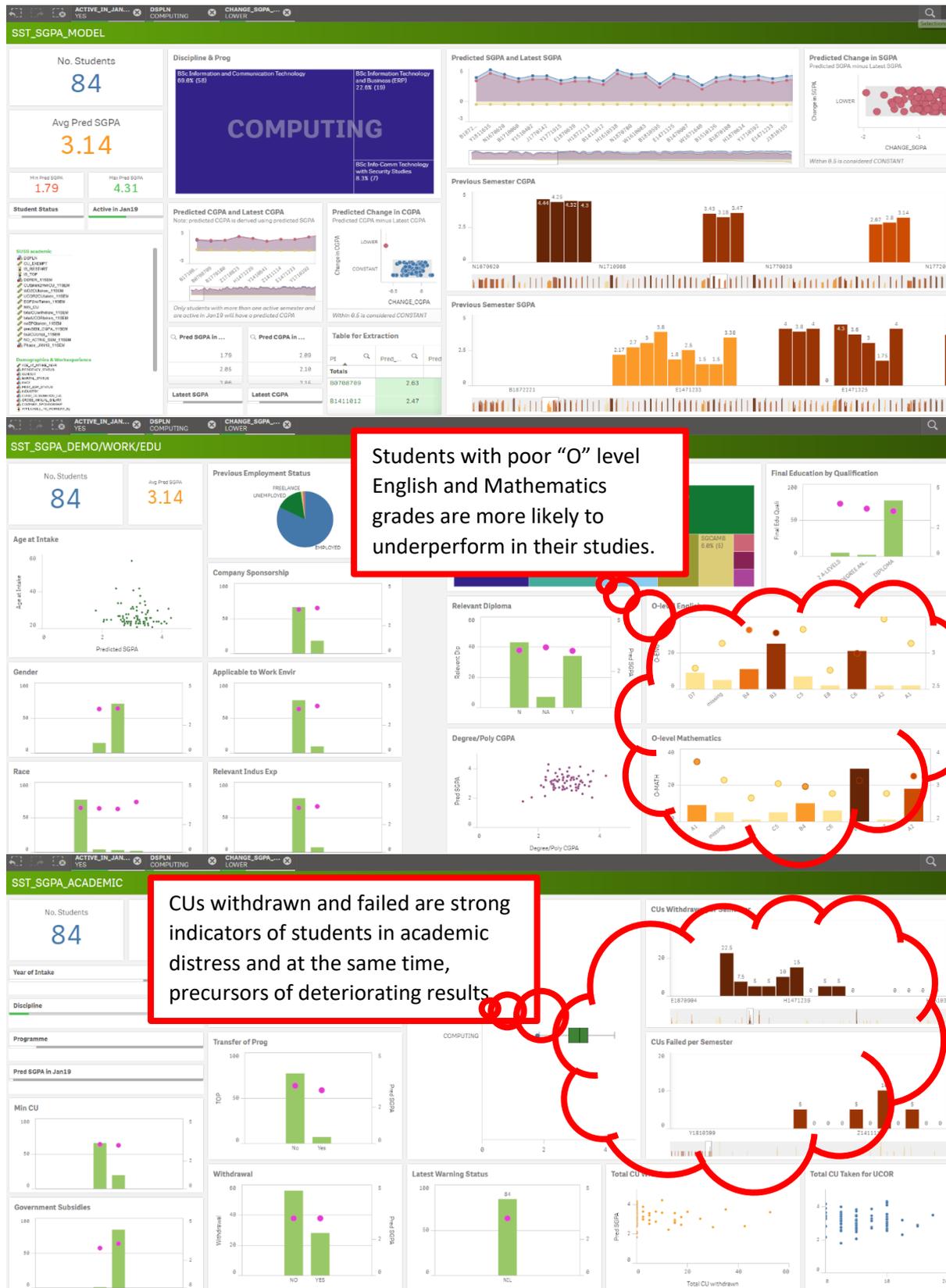


Figure 4: Dashboard on Relating Factors to Student's Academic Performance

By integrating the predicted SGPA into a dashboard, it translates raw data into visually appealing and easily accessible forms of graphical representations, allowing faculty members to generalise insights for timely and targeted support and early intervention. Compared to a static report or data file, the dashboard's dynamic and interactive features allow users to focus on a particular group of students, study their profile, historical academic performance and understand what could possibly contribute to their under-performance. By focusing on a subset of the data, customised support plan can be designed to address the specific case more effectively.

These the predictive modelling dashboard and Table 4 lists the observations from this dashboard and the corresponding actions proposed/taken.

Table 4: Observations from the Predictive Dashboard and Actions Taken

Observations from the dashboard w.r.t. Predictive Modelling	Actions proposed/taken to address the issues w.r.t. Predictive Modelling
Students with poorer English and Mathematics grades from their O-levels are more likely to fare badly in their degree studies	To partner with TLC to offer bridging and supplementary courses. On top of that, the School has initiated a student buddy and student mentor scheme. These initiatives will hopefully, get these students up to speed
To identify students who have recently dropped out, or are projected to drop out of a particular degree classification	To send advisory/encouragement notes to offer the students a friendly nudge, and to recommend them to attend courses customised by TLC
To identify students with inconsistent academic performance	To send advisory/encouragement notes to offer the students a friendly nudge, and to recommend them to attend courses customised by TLC

Conclusion

There are many benefits to having the dashboards, which enable users to access, interact and analyse up-to-date information to facilitate data-driven planning and timely decision-making. Anecdotal feedback/evidence has shown that students who have received counselling or participated in peer mentoring schemes have benefitted in terms of performing better in their studies and reporting a higher satisfaction level during their candidature. In addition, a better understanding of the students' in terms of their academic background, learning needs and learning preferences can aid the School to customise learning, enabling more students to succeed ([Christensen et al., 2010](#)).

Nonetheless, the dashboard is not without its limitations due to a multitude of factors, such as old data, non-reliability of data, over-fitting etc. One approach to circumvent these issues is to establish data governance structures, and to ensure data stewardship and cross-institutional agreement on data definitions (Wolf et al., 2016). Moreover, researchers have cautioned that dashboards and automated messages can be a double-edged sword and backfire to detrimentally affect students' mental health and stress levels, in addition to altering their behaviour in the wrong way (Straumsheim, 2017).

In fact, dashboards can convey snapshots of important measures, but they are poor at providing the nuance and context that effective data-driven decision making demands. Shapiro (2017) coined these drawbacks the importance trap, context trap and causality trap. We have to be cognizant that every dashboard, particularly for predictive modelling, is built on a set of priorities and assumptions about what is important, and it is essential that it is customised to the School's needs and with a nuanced and contextual understanding of the School and its environment. The fallacy of analytics is that it is often misconstrued as representing some sort of unbiased and dispassionate truth. To avoid equating "empirical" and "quantitative" with "objective", the user has to exercise judgement and discretion to reconcile and validate the information and interpretation from the dashboard.

The role of the Dean, HOPs or any programme administrator is unique and daunting in many ways. A well-designed dashboard will support the Dean and HOPs to draw on and understand information from a plurality of systems and processes (Wolf et al., 2016).

The School is unable to provide statistical data at this nascent stage of the project because it has been in a state of flux, converting from a private to public, autonomous university, and in the midst of doing so, intentionally

heightening the academic standards of the programmes. In this case, the historical data will not suffice. Moving forward, and with the School and programmes at steady state, this can be tracked and studied in depth.

Last but not least, BI&A will be improving and refining the entire framework of learning analytics with more aspects of student data in order to enhance the reliability and precision of the predictive modeling for students' academic performance. By incorporating users' feedback and intervention that has been deployed to help students, the performance of such a student support system can also be tracked, evaluated and improved.

References

- Beach, C. (2013). *At-Risk Students: Transforming Student Behavior*. R&L Education.
- Bryer, J. (2012). Peer Tutoring Program for Academic Success of Returning Nursing Students. *The Journal of the New York State Nurses' Association*, 43(1), 20–22.
- Christensen, C., Johnson, C. W., & Horn, M. B. (2010). *Disrupting Class, Expanded Edition: How Disruptive Innovation Will Change the Way the World Learns*. McGraw Hill Professional.
- Deegan, J. & Martin, N. (2018). *Demand Driven Education – Merging Work and Learning to Develop the Human Skills That Matter*. Pearson.
- Eckerson, W. W. (2010). *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*. John Wiley & Sons.
- Garrison, T. (2010). Teaching and Learning in The Age of Distractions. *Oceanography*, Vol. 23, pp. 180–181. <https://doi.org/10.5670/oceanog.2010.36>
- Henebery, B. (2019). *How Data Analytics is Shaking Up the Education Sector*. <https://www.theeducatoronline.com/k12/technology/e-learning/how-data-analytics-is-shaking-up-the-education-sector/259565>
- Klerkx, J., Leuven, K. U., Belgium, Verbert, K., Duval, E., Leuven, K. U., ... Belgium. (2017). *Learning Analytics Dashboards*. Handbook of Learning Analytics, pp. 143–150. <https://doi.org/10.18608/hla17.012>
- Lester, J., Klein, C., Rangwala, H., & Johri, A. (2017). *Learning Analytics in Higher Education*. ASHE Higher Education Report, Vol. 43, pp. 9–135. <https://doi.org/10.1002/aehe.20121>
- Payne, J. (1973). *Counselling and academic achievement*. British Journal of Guidance & Counselling, Vol. 1, pp. 19–25. <https://doi.org/10.1080/03069887308259348>
- Smith, N. (2017). *Research on Problems Faced by Part-Time University Students*. <https://classroom.synonym.com/research-problems-faced-parttime-university-students-17471.html>
- Shapiro, J (2017). *3 Ways Data Dashboards Can Mislead You*. <https://hbr.org/2017/01/3-ways-data-dashboards-can-mislead-you>
- Straumsheim, C. (2017). *Study Finds that Giving Students Timely Feedback about Their Performance in a Class can Benefit Students with Lower Grades*. <https://www.insidehighered.com/news/2017/02/02/study-explores-effect-data-dashboards-student-performance>
- Singapore University of Social Sciences. (n.d.). School of Science and Technology | Singapore University of Social Sciences. Retrieved July 21, 2019, <https://www.suss.edu.sg/about-suss/schools/sst>
- Wolf, M., Taimuty, M., Patel, M & Meteyer J. (2016). *The Dean's Information Challenge: From Data to Dashboard*. <https://er.educause.edu/articles/2016/11/the-deans-information-challenge-from-data-to-dashboard>

Please cite as: Li, J. Z., Low, W. P., Bheema Thiagarajan, L. & Peh, L. C. (2019). Data Analytics for Student Profiling and Academic Counselling. In Y. W. Chew, K. M. Chan, and A. Alphonso (Eds.), *Personalised Learning. Diverse Goals. One Heart. ASCILITE 2019 Singapore* (pp. 196-205).