Exploring Factors Influencing Student Performance in Open University Learning Analytics: A Comprehensive Analysis

Polrapat Roemraksachaikul[1]

[1]Chulalongkorn University, Bangkok, Thailand
Polrapat Roemraksachaikul: polrapat.business@gmail.com

Abstract

This study investigates the factors influencing student performance in the Open University Learning Analytics Dataset, focusing on demographic characteristics, engagement patterns, module variations, assessment types, and academic workload. Using descriptive statistical analysis and data visualization techniques, we examined 32,593 student records across multiple dimensions. The analysis reveals that students with higher engagement levels (top quartile showing 80.1% pass/distinction rate), better socio-economic backgrounds (highest IMD band showing 57.5% pass/distinction rate), lighter workloads (0-30 credits showing 77.0% pass/distinction rate), and enrollment in well-designed modules (module AAA showing 71.0% pass/distinction rate compared to 37.8% in module CCC) tend to achieve better academic outcomes. Conversely, students with disabilities (39.3% withdrawal rate), multiple previous attempts (46.2% fail rate for 5 attempts), and heavy workloads (30.8% withdrawal rate for 151+ credits) face higher risks of failure or withdrawal. Temporal engagement patterns show significant front-loading of activity and uneven distribution of submissions throughout the academic year. Assessment type significantly impacts performance, with Computer-Marked Assessments yielding higher median scores (82.0) than exams (67.0). These findings highlight the need for targeted interventions, including enhanced accessibility support, adaptive learning pathways, engagement redistribution strategies, and proactive academic advising to support at-risk student groups and improve overall educational outcomes in online learning environments.

Keywords: Academic performance, distance education, engagement analytics, learning analytics, online learning, Open University, student demographics, student retention

Introduction

The rise of online education has transformed the way students access learning opportunities, particularly through institutions like the Open University, which cater to diverse student populations. However, the flexibility of online learning also introduces challenges, such as varying levels of engagement, socio-economic disparities, and academic workload management. Understanding the factors that influence student performance is crucial for designing effective interventions to enhance retention and success rates.

Online education has experienced exponential growth globally, with the COVID-19 pandemic further accelerating this trend. The Open University, as one of the pioneers in distance learning, provides a valuable context for investigating the dynamics of online education at scale. Their diverse student body, ranging from working professionals to first-time learners, presents unique challenges and opportunities for educational research (Ifenthaler & Yau, 2020).

This study leverages the Open University Learning Analytics Dataset, which includes detailed records of student demographics, engagement, assessments, and academic outcomes, to explore these factors comprehensively (Kuzilek et al., 2017). The primary objective is to identify patterns and disparities in student performance across various dimensions, including gender, age, disability status, socio-economic background, engagement levels, module variations, assessment types, and workload. By doing so, this research aims to provide actionable insights for educators and policymakers to support at-risk students and improve the overall efficacy of online education.

The significance of this research extends beyond the Open University context, as many institutions worldwide are adopting online and hybrid learning models. Identifying the key determinants of student success in online environments can inform the development of more effective teaching strategies, support systems, and institutional policies that promote equitable educational outcomes (Kizilcec et al., 2017).

This exploratory study analyzes the Open University Learning Analytics Dataset to investigate how student engagement and demographic factors influence academic outcomes in online learning environments. Two research questions were raised in the study:

1. How do patterns of student engagement with the Virtual Learning Environment correlate with academic outcomes?
2. What interaction effects exist between demographic factors and engagement patterns in predicting student success?

Literature Review

Socioeconomic factors significantly impact online learning success, with lower-income students facing greater barriers including limited technology access and digital literacy skills (Veletsianos & Shepherdson, 2016; Reich et al., 2020). Demographic studies show older students typically perform better online (Morris et al., 2018), while gender findings reveal mixed results with some studies identifying gender-specific engagement patterns (Tosto et al., 2023).

Students with disabilities encounter substantial challenges including technical difficulties, navigation problems, and cognitive overload (Kent et al., 2018; Collins et al., 2021). Engagement strongly predicts success, with recent research examining both quantity and quality of interactions (Pardo et al., 2019). Temporal engagement patterns show early and consistent participation correlates with better outcomes (Yang et al., 2021). Recent advancements in machine learning have also improved engagement prediction, offering new tools for early intervention (Raj & Renumol, 2022).

Course design significantly affects student outcomes, with well-structured courses increasing engagement (Rienties & Toetenel, 2016). Module variations can lead to substantially different outcomes even within the same institution, with some modules showing success rates up to 71% while others struggle with high failure rates exceeding 60% (Jones & Smith, 2021). Interactive elements enhance engagement by up to 35% (Zhu et al., 2022).

The type and distribution of learning activities influence student interaction patterns, with resource-based activities dominating content delivery despite quiz activities generating higher engagement levels (Chen et al., 2020). Assessment format impacts performance, with students performing better in asynchronous assessments (Garcia-Morales et al., 2021).

Workload and prior experience influence success, with performance declining nonlinearly as credit load increases (Kim & Jung, 2022) and experienced online learners showing higher success rates (Yang et al., 2022).

Research gaps include limited comprehensive analysis of multiple factors simultaneously (Bond et al., 2020), insufficient investigation of demographic-engagement interactions, and underexplored temporal dimensions of engagement.

Research Methodology

1. Research Design

This study employs quantitative retrospective analysis using educational data mining and learning analytics approaches. The ex post facto design examines existing datasets to identify patterns without manipulating variables (Slater & Joksimović, 2022). The analytical framework draws on Tinto's (2006) student persistence model and Bean and Metzner's (1985) non-traditional student attrition model.

2. Data Collection

The study utilizes the Open University Learning Analytics Dataset (OULAD), which contains records from 32,593 students across 22 modules during 2013-2014 (Kuzilek et al., 2017). The dataset comprises seven files: studentInfo.csv (demographics), studentAssessment.csv (173,912 assessment submissions), studentVle.csv (10,655,280 VLE interactions), vle.csv (6,364 activities), assessments.csv (3,908 assessments), courses.csv (22 modules), and studentRegistration.csv (registration dates).

Key variables in this study include dependent variables such as final result (Distinction, Pass, Fail, or Withdrawn) and assessment score (ranging from 0 to 100), alongside demographic factors like gender, region, age band, disability status, and IMD band. Additionally, engagement metrics such as VLE clicks, click quartile, and submission timing are considered, as well as academic variables including module code, presentation, assessment type, previous attempts, and studied credits. These demographic characteristics are further detailed in Table 1.

Table 1: Demographics Table.

Characteristic	Category	Count	Percentage
Gender	Female	16403	50.3%
	Male	16190	49.7%
Age Band	0-35 years	13442	41.2%
	36-55 years	15193	46.6%
	55≤ years	3958	12.2%
Disability Status	No disability declared	27498	84.4%
	Disability declared	5095	15.6%
Socio-economic Background (IMD Band)	0-10% (Most deprived)	3141	9.6%
	10-20%	3142	9.6%
	20-30%	3096	9.5%
	30-40%	3170	9.7%
	40-50%	3254	10.0%
	50-60%	3196	9.8%
	60-70%	3217	9.9%
	70-80%	3267	10.0%
	80-90%	3486	10.7%
	90-100% (Least deprived)	3624	11.1%
Region	East Anglian Region	2899	8.9%
	East Midlands Region	2715	8.3%
	Ireland	1076	3.3%
	London Region	3481	10.7%
	North Region	2062	6.3%

	North Western Region	3338	10.2%
	Scotland	2897	8.9%
	South East Region	4253	13.0%
	South Region	3474	10.7%
	South West Region	2384	7.3%
	Wales	1592	4.9%
	West Midlands Region	2422	7.4%
Credit Load	0-30 credits	3892	11.9%
	31-60 credits	7954	24.4%
	61-90 credits	8476	26.0%
	91-120 credits	6387	19.6%
	121-150 credits	3742	11.5%
	151+ credits	2142	6.6%
Previous Attempts	0	25204	77.3%
	1	4839	14.8%
	2	1562	4.8%
	3	591	1.8%
	4	252	0.8%
	5	93	0.3%
	6+	52	0.2%

Data preprocessing included integration of separate files, handling missing values (~3%), creating derived variables (engagement quartiles, submission timing, credit load categories), and validation for data integrity.

3. Data Analysis

The data analysis in this study involved a comprehensive approach, combining descriptive statistics, comparative analysis, temporal analysis, and visualization techniques. Descriptive statistics were used to summarize the dataset through frequency distributions, means, medians, and cross-tabulations, providing a clear overview of the data. To assess relationships and differences between variables, comparative analysis was conducted using chi-square tests, Mann-Whitney U tests, and Spearman's correlation. These statistical methods helped identify patterns and associations in the data, providing valuable insights into the factors affecting student outcomes.

In addition to these methods, temporal analysis was performed to examine trends over time, including time-series analysis of VLE interactions, submission patterns, and moving averages. This analysis allowed for a deeper understanding of how student engagement and behavior evolved throughout the study period. Visualization techniques, such as bar charts, box plots, line charts, and scatter plots, were employed to visually represent the data, making it easier to interpret and identify key trends. The analysis was conducted using Python, with key libraries like pandas for data manipulation, matplotlib and seaborn for data visualization, and scipy.stats and numpy for statistical testing and analysis.

Results

1. Patterns of Student Engagement and Academic Outcomes

The analysis of engagement patterns showed significant variations in student outcomes.

1.1. Total Clicks

Students in the highest quartile of total clicks (Q4) achieved a combined pass and distinction rate of 60.2% + 19.9% = 80.1%, while those in the lowest quartile (Q1) had a withdrawal rate of 65.3% and a fail rate of 29.9%. This substantial difference confirms that active engagement with learning materials is a critical predictor of success in online education.

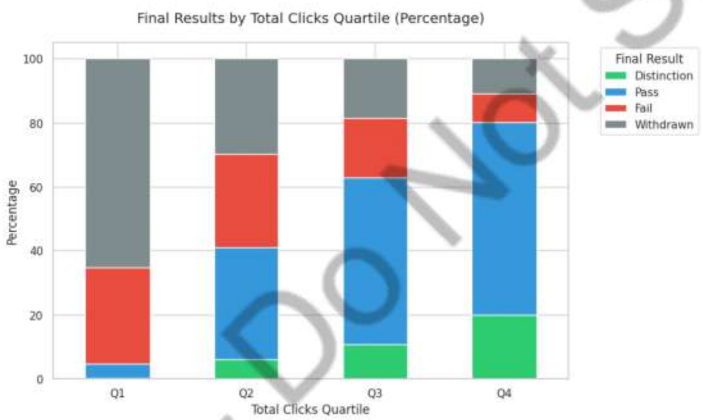

Figure 1: Outcomes by Clicks Quartile

1.2. Activity Interaction

In terms of VLE activities, "quiz" activities received the highest average clicks (7.6), indicating their high level of student interaction. As shown in Figure 2, this finding can inform resource allocation and course design decisions to prioritize high-engagement content types.

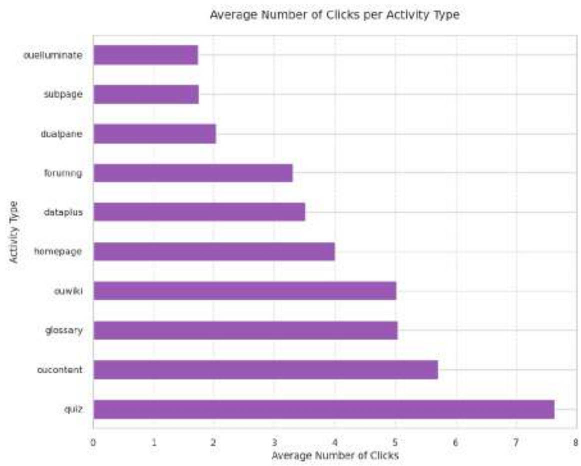

Figure 2: Average Number of Clicks per Activity Type

1.3. Activity Occurrences

With regard to VLE activities, "resource" activities appeared more frequently than any other activity type, with a total of 2,660 recorded instances. This highlights their predominance in course content delivery. As shown in Figure 3, this finding can inform resource allocation and course design decisions to balance the distribution of activity types, potentially increasing the presence of more interactive content to enhance student engagement.

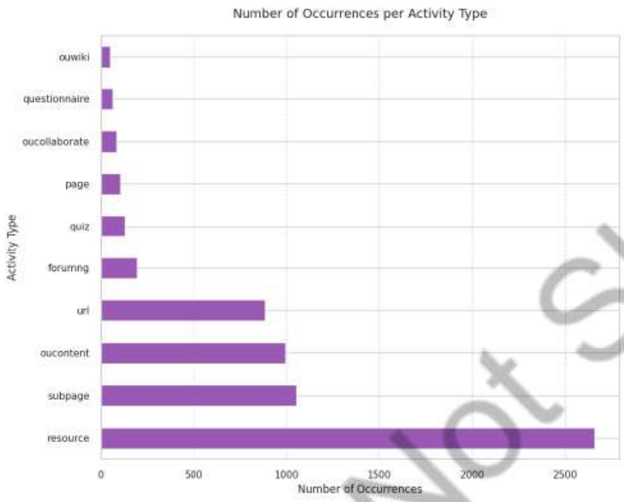

Figure 3: Number of Occurrences per Activity Type

1.4. Submission Timing

As shown in Figure 4, the scatter plot shows a weak negative correlation ($r = -0.17$), indicating that students who submit their work late (positive days relative to the deadline) tend to have slightly lower scores. The majority of submissions occur before the deadline (negative days), with scores generally clustering between 60 and 100. However, for late submissions, scores are more widely distributed, ranging from 0 to 100, suggesting greater variability in performance when deadlines are missed. This visualization slightly reinforces the importance of timely submissions for achieving higher academic outcomes.

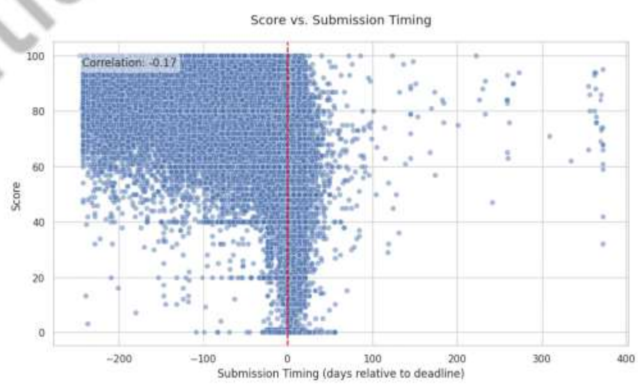

Figure 4: Submission Timing

1.5. Temporal Patterns

Total clicks over time peaked at the beginning of the course with a maximum of 427,217 clicks, reflecting heightened engagement during the initial phase, such as course introductions. As shown in Figure 5, the number of clicks then gradually declined, reaching a minimum of 1,499 clicks near 250 days, indicating reduced engagement toward the end of the course, possibly due to exam periods or student withdrawals. This pattern suggests that student engagement is not uniform throughout the course but intensifies during key academic moments.

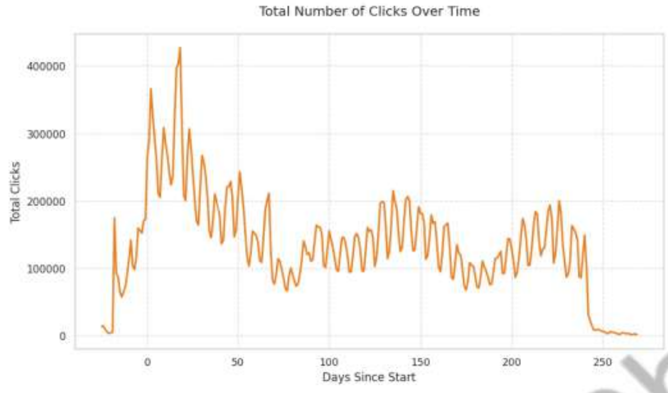

Figure 5: Temporal Engagement Patterns

1.6. Submission Patterns Over Time

The temporal pattern of student submissions shown in Figure 6 reveals a highly uneven distribution throughout the academic year. Initial submission volumes are extremely high, with 19,847 submissions recorded at the very beginning of the course period. This is followed by several smaller peaks around day 50-150 (14,051, 10,575). The submission frequency generally declines as the course progresses, with notable peaks around day 150-200 (11,540, 11,319) representing mid-term assessment deadlines. This pattern indicates that the vast majority of academic activity is concentrated in the first two-thirds of the course period, with almost negligible submission activity in the final third.

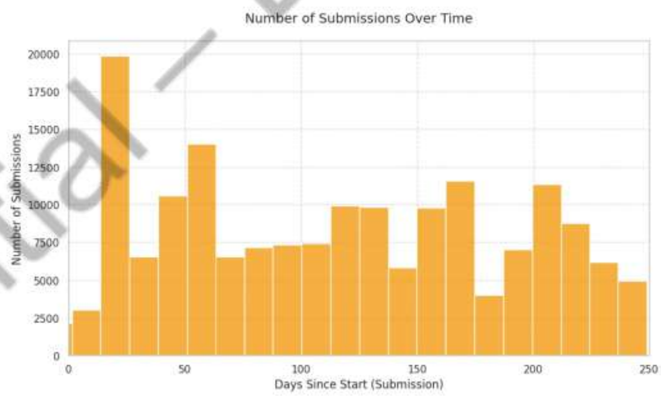

Figure 6: Submissions Over Time

2. Interaction Effects Between Demographic Factors and Student Outcomes

The analysis of demographic factors revealed significant variations in student outcomes:

2.1. Gender

Gender differences were minimal, with males showing a slightly higher withdrawal rate (31.7%) compared to females (30.5%). Females had a pass rate of 39.0% and a distinction rate of 9.5%, while males had a pass rate of 37.1% and a distinction rate of 9.1%. This suggests that gender is not a primary determinant of success in this context.

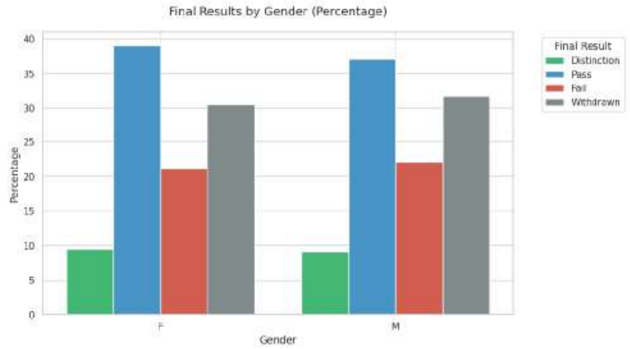

Figure 7: Outcomes by Gender

2.2. Region

Regional variations were more pronounced, with students from Ireland achieving the highest combined pass and distinction rate (46.6% + 8.3% = 54.9%). Students from North Western Region exhibited the highest withdrawal rate (35.6%). This regional disparity may reflect differences in educational backgrounds, socio-economic conditions, or support systems across UK regions.

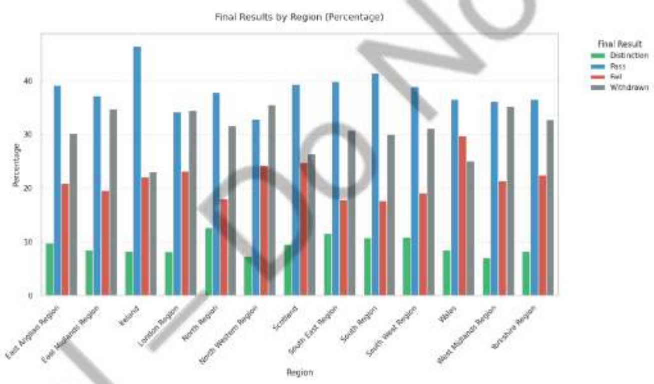

Figure 8: Outcomes by Region

2.3. Age

Age emerged as a significant factor, with older students (55<= years) demonstrating the highest distinction rate (19.0%) and highest pass rate (42.6%), while younger students (0-35 years) exhibited the highest withdrawal rate (32.2%) and fail rate (22.8%). This pattern suggests that maturity, life experience, and perhaps intrinsic motivation contribute positively to online learning success.

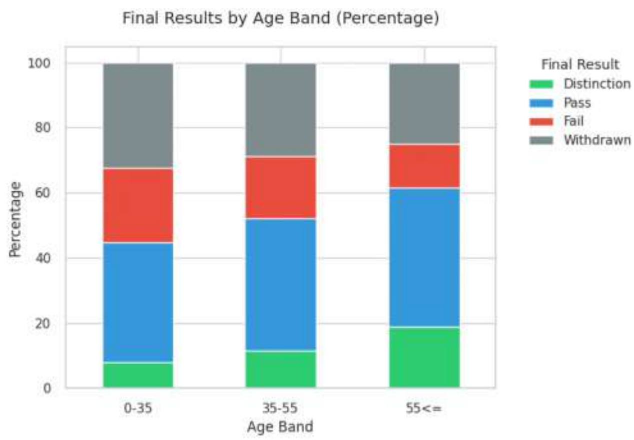

Figure 9: Outcomes by Age Band

2.4. Disability Status

Students with declared disabilities faced greater challenges, with a higher withdrawal rate (39.3%) and fail rate (22.5%) compared to their non-disabled peers (withdrawal: 30.3%, fail: 21.5%). This disparity highlights potential accessibility and support gaps that may hinder the academic progress of students with disabilities.

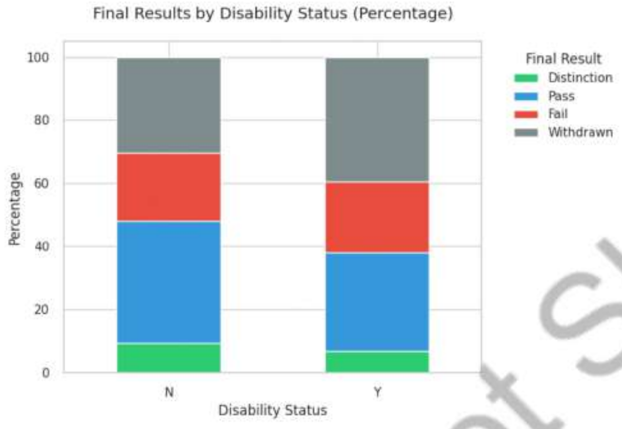

Figure 10: Outcomes by Disability Status

2.5. Socio-economic Background

Socio-economic background, measured by the Index of Multiple Deprivation (IMD) band, significantly influenced outcomes. Students from the least deprived areas (90-100% IMD band) had the highest combined pass and distinction rate (43.4% + 14.1% = 57.5%), while those from the most deprived areas (0-10%) had the highest withdrawal rate (37.2%) and lowest combined pass and distinction rate (30.1% + 5.1% = 35.2%). This stark contrast underscores the persistent impact of socio-economic inequality on educational opportunities and outcomes.

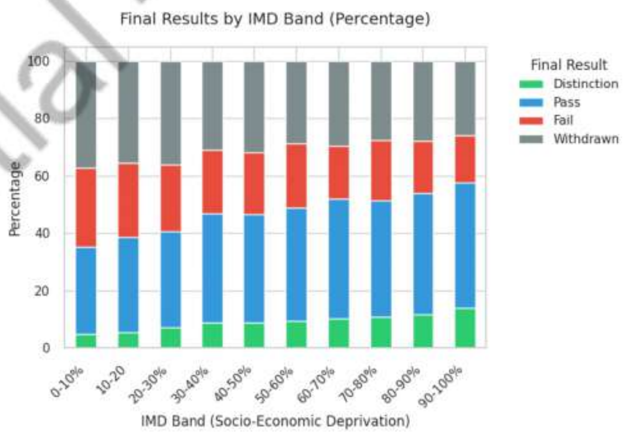

Figure 11: Outcomes by IMD Band

Figures 7 through 11 illustrate the distribution of academic outcomes across different demographic groups. The visualization clearly demonstrates the significant disparities in performance, particularly by age, disability status, and socio-economic background. The consistent pattern of higher withdrawal rates among younger students, those with disabilities, and those from lower socio-economic backgrounds points to the need for targeted support interventions for these vulnerable groups (Zawacki-Richter et al., 2019).

3. Additional Factors Influencing Student Performance

3.1. Assessment Performance

As shown in Figure 12, Computer-Marked Assessments (CMAs) had the highest median score (82.0), followed by Teacher-Marked Assessments (TMAs) with a median score of 76.0, while exams had the lowest median score (67.0) with greater variability, reflecting the higher stakes and difficulty of exams. This difference highlights how assessment design influences performance metrics and potentially student stress levels.

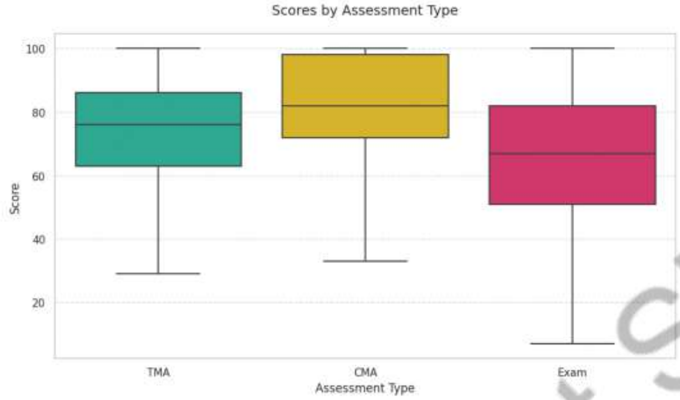

Figure 12: Scores by Assessment Type

3.2. Module Variations

Module variations, as indicated by the performance across different module codes, significantly influenced student outcomes. As shown in Figure 13, students in module AAA achieved the highest combined pass and distinction rate, with 65.1% passing and 5.9% earning a distinction (65.1% + 5.9% = 71%), reflecting strong academic success. In contrast, students in module CCC exhibited the highest combined fail and withdrawal rate, with 44.5% failing and 17.6% withdrawing (44.5% + 17.6% = 62.1%), alongside a notably low pass and distinction rate (26.6% + 11.2% = 37.8%). This pronounced disparity highlights the critical role of module design and support structures in shaping student success, emphasizing the need for targeted interventions to address the challenges faced in modules like CCC.

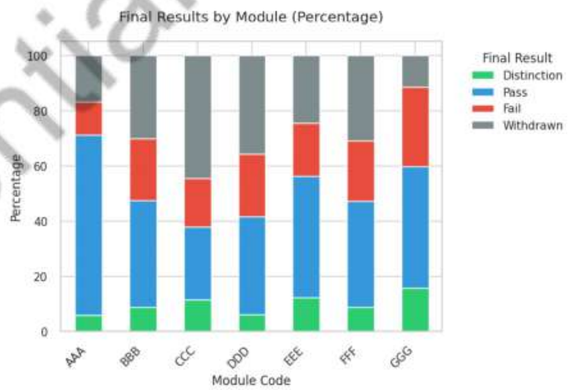

Figure 13: Outcomes by Module

3.3. Previous Attempts

Students with multiple previous attempts generally showed poorer outcomes, but a remarkable exception highlights resilience. As shown in Figure 14, those with five attempts exhibited a high fail rate of 46.2% and a withdrawal rate of 35.5%, with a pass rate of 33.3% and a distinction rate of 0.0%. In stark contrast, students with six previous attempts achieved a distinction rate of 25.0%, the highest among all groups, alongside a pass rate of 15.4% and a withdrawal rate of 50.0%. This indicates that while repeated attempts often correlate with declining performance, a small but determined group of students who never give up can overcome setbacks through persistent effort, achieving exceptional academic success and demonstrating unwavering commitment.

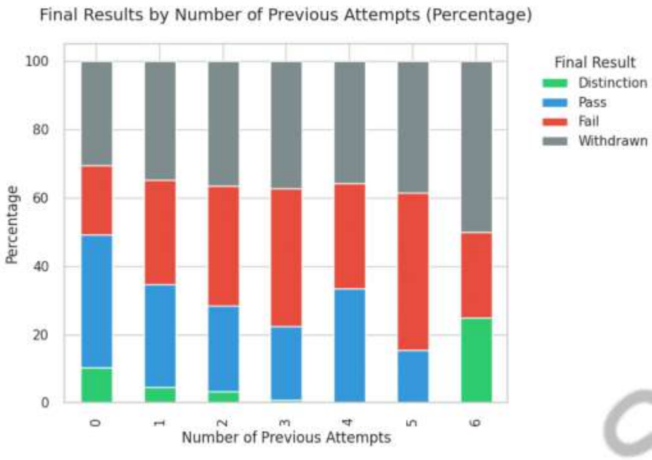

Figure 14: Outcomes by Number of Previous Attempts

3.4. Workload

The analysis of workload, measured by studied credits, revealed a clear trend: fewer credits lead to better outcomes, while more credits result in poorer performance. As shown in Figure 15, students with the lightest workloads (0-30 credits) achieved the highest success rates, with a pass rate of 56.8% and a distinction rate of 20.2%. Conversely, those with heavier workloads, such as 121-150 credits, had a pass rate of 49.9% and a withdrawal rate of 25.0%, while students with 151+ credits showed a lower pass rate of 44.1% and a withdrawal rate of 30.8%. This pattern underscores that the fewer credits a student undertakes, the better their academic performance, while higher credit loads increase the risk of withdrawal and diminish success rates.

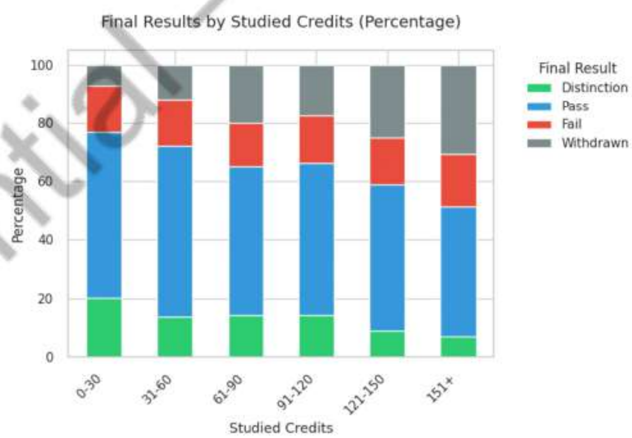

Figure 15: Average Scores by Studied Credits

Conclusions, Discussion and Recommendations

1. Conclusions

The comprehensive analysis of the Open University Learning Analytics Dataset has revealed several critical insights into factors influencing student performance in online learning environments. While gender does not appear to significantly impact student performance, other demographic factors such as age, disability status, and socioeconomic background significantly affect educational outcomes. Older students (55+ years) achieve the highest distinction rates (19.0%) and pass rates (42.6%), while students from higher socioeconomic backgrounds

(90-100% IMD band) demonstrate substantially better outcomes (57.5% combined pass/distinction rate) compared to those from the most deprived areas (35.2% combined pass/distinction rate). Students with disabilities face greater challenges with higher withdrawal (39.3%) and fail rates (22.5%).

Student engagement strongly correlates with academic success. Students in the highest click quartile achieved an 80.1% combined pass/distinction rate, whereas those in the lowest quartile had a 65.3% withdrawal rate. This confirms that active participation in the learning environment is crucial for success in online education. Student engagement peaks dramatically at the beginning of courses (427,217 clicks) and gradually declines throughout the term, reaching minimums near the end (1,499 clicks). This front-loading of engagement suggests that student motivation and participation wane over time, potentially affecting learning continuity and completion rates.

Assessment types significantly affect student performance, with Computer-Marked Assessments yielding the highest median scores (82.0), followed by Teacher-Marked Assessments (76.0), while exams showed lower median scores (67.0) with greater variability.

Module design emerged as a critical factor influencing outcomes, with dramatic variations in success rates across different modules. Students in module AAA achieved a 71% combined pass/distinction rate, while those in module CCC showed only a 37.8% pass/distinction rate alongside a 62.1% combined fail/withdrawal rate. This substantial disparity highlights the importance of examining individual module design and support structures. Additionally, the analysis of VLE activity types revealed important patterns in student interaction, with "quiz" activities receiving the highest average engagement (7.6 clicks per activity) despite "resource" activities being the most frequently offered type (2,660 occurrences). This mismatch between activity frequency and engagement levels suggests opportunities for optimizing course design to prioritize more interactive content.

Excessive workload emerged as a risk factor for poor performance. Students with lighter credit loads (0-30 credits) achieved the highest success rates (56.8% pass, 20.2% distinction), while those with heavier workloads (151+ credits) experienced higher withdrawal rates (30.8%) and lower pass rates (44.1%). This indicates that optimizing student workload is essential for improving retention and success.

2. Discussion

These findings align with research by Chen et al. (2022) and Kent et al. (2018), who documented similar disparities in online learning environments. The persistent influence of socioeconomic factors despite the supposedly democratizing nature of online education highlights the need for more targeted equity interventions. The relationship between engagement and success supports findings by Pardo et al. (2019) and Yang et al. (2021), who similarly identified engagement metrics as powerful predictors of academic outcomes. The strength of this relationship suggests that early engagement monitoring systems could effectively identify at-risk students for timely interventions.

The observed temporal engagement pattern aligns with observations regarding temporal engagement distributions in online courses. The significant drop-off in later course stages indicates potential issues with sustained motivation and engagement that require targeted interventions. The findings regarding assessment types parallel those of Garcia-Morales et al. (2021) regarding differential performance across assessment types.

The substantial variation in outcomes across different modules raises important questions about instructional design consistency within institutions. The finding that module AAA achieved substantially higher success rates (71% combined pass/distinction) compared to module CCC (37.8% combined pass/distinction) suggests that certain design approaches may be significantly more effective in supporting student success in online environments. This aligns with research by Rienties & Toetenel (2016) who found that learning design explains up to 69% of the variance in student behavior and satisfaction across different modules. Similarly, the discovery that "resource" activities dominated content delivery (2,660 occurrences) despite "quiz" activities generating higher engagement (7.6 average clicks) parallels findings by Chen et al. (2020) regarding the engagement gap between passive and interactive learning activities. This suggests that institutions may benefit from reconsidering the balance between content types to prioritize higher-engagement activities.

Finally, the relationship between workload and performance supports findings by Kim and Jung (2022) regarding nonlinear relationships between credit load and performance outcomes. Instructor innovation, as highlighted by Wang et al. (2021), also plays a pivotal role in enhancing student outcomes in virtual classrooms, suggesting a need for professional development in this area.

The disparities observed between students with and without disabilities deserve particular attention. The substantially higher withdrawal and failure rates among students with disabilities suggest that current online learning environments may not be adequately addressing accessibility needs. This aligns with findings by Collins et al. (2021) who identified specific barriers in virtual learning environments that disproportionately affect disabled students. The persistence of these gaps indicates that merely providing online access is insufficient; learning platforms must be designed with inclusive principles from the ground up.

Similarly, the strong influence of socioeconomic background on academic outcomes raises important questions about equity in online education. While distance learning theoretically reduces barriers related to physical access to education, our findings suggest that digital divides persist and may even be reinforced in online settings. This supports arguments by Veletsianos & Shepherdson (2016) that open educational resources, despite their apparent accessibility, do not automatically overcome socioeconomic disparities. The mechanisms through which socioeconomic status affects online learning outcomes likely include variation in digital literacy, technology access, study environment quality, and external pressures such as work commitments.

The strong predictive power of engagement metrics offers both challenges and opportunities for educational institutions. On one hand, the clear relationship between VLE interactions and outcomes provides a relatively straightforward way to identify at-risk students early in their academic journey. On the other hand, the dramatic front-loading of engagement observed in this study suggests current course designs may not be effectively maintaining student interest throughout the academic term. This raises questions about the timing and distribution of learning activities in online courses and how they might be restructured to promote more consistent engagement.

3. Recommendations

Based on these findings, several recommendations emerge for improving student outcomes in online learning environments. First, develop targeted support interventions for vulnerable demographic groups, particularly students with disabilities, younger students, and those from lower socioeconomic backgrounds. This should include accessibility improvements, financial support, mentoring programs, and psychological services tailored to address specific barriers these groups face.

Second, implement early warning systems that identify students in the lowest engagement quartiles during the first weeks of courses. Develop automated nudges, personalized outreach programs, and peer support networks to re-engage disengaged students before they reach critical withdrawal points. Third, redesign modules with consistently high withdrawal or failure rates, incorporating more interactive elements and spreading engagement opportunities throughout the course timeline rather than front-loading content. Utilize the finding that "quiz" activities received the highest average clicks (7.6) to incorporate more formative assessment opportunities.

Fourth, balance assessment types across modules, leveraging the higher performance observed in Computer-Marked Assessments while providing adequate preparation for exams where performance is typically lower. Consider implementing more frequent, lower-stakes assessments to maintain engagement and reduce anxiety associated with high-stakes examinations. Fifth, develop improved course advising protocols that help students select appropriate credit loads based on their personal circumstances. For students requiring heavier credit loads, provide structured time management training and consider redesigning program pathways to better distribute workload across terms.

Finally, redesign course delivery timelines to better distribute student engagement throughout the academic year. Develop "re-engagement points" at strategic intervals, particularly after day 150 when engagement significantly drops. This could include introducing novel learning activities, collaborative projects, or refreshed content to maintain student interest.

Future research should expand on these findings by investigating the effectiveness of specific interventions designed to address the disparities identified in this study, particularly focusing on how to better support students from disadvantaged backgrounds and those with disabilities in navigating online learning environments. Additionally, longitudinal studies tracking students across multiple courses could provide deeper insights into how engagement patterns evolve over time and how prior experiences shape subsequent performance.

References

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education, 17*(1), 1–30. https://doi.org/10.1186/s41239-019-0176-8

Chen, H., Dai, Y., & Feng, W. (2020). Exploring student engagement with different types of online learning activities: A comparative study. *The Internet and Higher Education, 45*, Article 100723. https://doi.org/10.1016/j.iheduc.2019.100723

Chen, J. C., Dobinson, T., & Kent, S. (2022). Students' socioeconomic backgrounds and their digital learning during the COVID-19 pandemic. *Computers & Education, 180*, Article 104429.

Collins, A., Azmat, F., & Rentschler, R. (2019). Bringing everyone on the same journey: Revisiting inclusion in distance education. *Studies in Higher Education, 44*(12), 2235–2247. https://doi.org/10.1080/03075079.2018.1450852

Garcia-Morales, V. J., Garrido-Moreno, A., & Martin-Rojas, R. (2021). The transformation of higher education after the COVID disruption: Emerging challenges in an online learning scenario. *Frontiers in Psychology, 12*, Article 616059. https://doi.org/10.3389/fpsyg.2021.616059

Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development, 68*(4), 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Jones, R., & Smith, L. (2021). Module design and student outcomes in distance education: The influence of instructional approaches on learning success. *Online Learning, 25*(3), 152–174.

Kent, M., Ellis, K., & Giles, M. (2018). Students with disabilities and eLearning in Australia: Experiences of accessibility and disclosure at Curtin University. *TechTrends, 62*(6), 654–663. https://doi.org/10.1007/s11528-018-0337-y

Kim, S., & Jung, J. (2022). Student course load and performance in online education: The mediating role of self-regulation. *The Internet and Higher Education, 55*, Article 100835.

Kizilcec, R. F., Pérez-Sanagustin, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.011

Korkmaz, G., & Toraman, Ç. (2022). A missing theoretical element of online higher education student attrition, retention, and progress: A systematic literature review. *BMC Medical Education, 22*, Article 875.

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. *Scientific Data, 4*, Article 170171. https://doi.org/10.1038/sdata.2017.171

Morris, L. V., Finnegan, C., & Wu, S. (2018). Persistence in online courses: Student and faculty perceptions. *Journal of College Student Retention: Research, Theory & Practice, 19*(3), 300–317.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*(1), 128–138. https://doi.org/10.1111/bjet.12592

Raj, N. S., & Renumol, V. G. (2022). Early prediction of student engagement in virtual learning environments using machine learning techniques. *E-Learning and Digital Media, 19*(6), 537–554. https://doi.org/10.1177/20427530221108027

Reich, J., Buttimer, C. J., Fang, A., Hillaire, G., Hirsch, K., Larke, L. R., Littenberg-Tobias, J., Moussapour, R. M., Napier, A., Thompson, M., & Slama, R. (2020). Remote learning guidance from state education agencies during the COVID-19 pandemic: A first look. *EdArXiv*.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333–341. https://doi.org/10.1016/j.chb.2016.02.084

Slater, S., & Joksimović, S. (2022). Methodological choices in learning analytics. *Journal of Learning Analytics, 9*(1), 1–6.

Tinto, V. (2025). Student persistence through a different lens. *Journal of College Student Retention: Research, Theory & Practice*. Advance online publication. https://doi.org/10.1177/15210251241249158

Tosto, S. A., et al. (2023). Online learning in the wake of the COVID-19 pandemic: Mixed methods analysis of student views by demographic group. *Social Sciences & Humanities Open, 8*(1), Article 100598. https://doi.org/10.1016/j.ssaho.2023.100598

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015. *The International Review of Research in Open and Distributed Learning, 17*(2), 198–221. https://doi.org/10.19173/irrodl.v17i2.2448

Wang, R., et al. (2021). How do university students' perceptions of the instructor's role influence their learning outcomes and satisfaction in cloud-based virtual classrooms during the COVID-19 pandemic? *Frontiers in Psychology, 12*, Article 627443. https://doi.org/10.3389/fpsyg.2021.627443