EXPLORING THE IMPACT OF LEARNERS’ DEMOGRAPHIC CHARACTERISTICS ON COURSE COMPLETION AND DROPOUT IN MASSIVE OPEN ONLINE COURSES

Vinu Sherimon 1, Leena Francis 2, Sherimon P.C 3, Disha Devassy 4, Walid Aboraya 5

1 Department of IT, University of Technology and Applied Sciences, Muscat, Sultanate of Oman
2 Department of Science, Indian School Al Seeb, Muscat, Sultanate of Oman
3 Faculty of Computer Studies, Arab Open University, Muscat, Sultanate of Oman
4 Faculty of Computer Studies, Arab Open University, Riyadh, Kingdom of Saudi Arabia
5 Faculty of Education Studies, Arab Open University, Muscat, Sultanate of Oman

Received 10 November 2021
Accepted 19 December 2021
Published 31 January 2022

ABSTRACT

MOOCs (Massive Open Online Courses) have the potential to change education by offering high-quality online courses in a variety of disciplines. However, just a few studies have investigated the impact of MOOC students’ demographics on their completion rates. In this research, we investigate the impact of demographic features of learners in the completion and dropping rate of MOOCs. The data from a survey is used in this study to determine which learner demographic features may have an impact on MOOC completion and dropout. The data was analyzed using Chi-square test. The findings demonstrate that four factors, including gender, marital status, age, and educational level, have no impact on degree of MOOC completion. It was also discovered that marital status, age, and educational level, have no effect on MOOC dropout. However, we discovered a statistically significant link between gender and MOOC dropout ($\chi^2 = 6.347, df = 1, p = 0.012$). These results can be considered in future instructional initiatives.

Keywords: MOOC Completion, MOOC Dropout, Demographic, Online Learning, Lifelong Learning, Online Education

1. INTRODUCTION

Human-technology interaction is a significant driver for bringing about new developments that improve people’s lives and address societal issues. For more than a century, Online learning in education has constantly evolved—both in practice and in definition. Due to the rapid growth of technology in the service sector, the developers targeted the education sector too. Out of these initiatives, Massive open online courses (MOOCs) are a new and emerging style of online learning that provides content such as brief video lectures, digital readings, interactive assignments, discussion forums, and quizzes, and it gained significant popularity among students and educators Liyanagunawardena et al. (2014). Massive open online courses (MOOCs) have risen in popularity, and many universities have put a lot of effort into creating, publicizing, and offering them Onah et al. (2014). Because of the great interest in a variety of MOOCs, providers now have a platform to advertise higher education possibilities more broadly. There have been debates not only about the potentials but also about the
challenges of MOOC as in many academic pursuits.

MOOCs are defined by a mix of four fundamental characteristics: Massive, Open, Online and Course Bates (2014). Massive means it caters to a gigantic number of students for an enormous number of courses from various disciplines. It can provide infinite scalability in terms of possibility for technologically aided tools for learning. Open means it does not have any prerequisites other than access to the internet through vivid devices. It provides the relevant content to all without any conditions. Online means it offers a complete online program, however some of the universities accept a blended mode of course. Students in a cluster format can choose some of the modules such as quiz, case studies physically too. Course term encompasses the whole curriculum content of innumerable subjects which enables anyone to earn a certificate or a badge on the completion.

Pedagogy, quality, completion rates, sustainability, and credit distribution in MOOCs are all key concerns for higher education institutions, according to a white paper by Yuan and Powell (2013). Digital literacy skills, individual differences in motivation, and self-regulation are some of the most essential learner characteristics in the framework of MOOC-based learning. As per the research Christensen et al. (2013), the majority of students enrolled in a Coursera course are from industrialized countries and are young, well educated, and employed. Also, it was observed that males take MOOCs in greater numbers than females, especially in poorer countries Christensen et al. (2013). The extent to which gender progresses the association between region and MOOC learner engagement was investigated in the study conducted by Gameel and Wilkins (2019). This research reports that in three regions, male learners showed significantly higher levels of MOOC engagement than females. There was no substantial difference between male and female learners in the other two regions. So, it may be concluded that the gender disparity depends on the country of the learner.

Despite the widespread excitement for MOOC courses and platforms and their rapid growth, there has been increasing alarm about the number of MOOC dropouts. To some extent, these research studies have investigated the demographics and motives of learners who are taking a course Ali Alshehri (2021), Bayeck (2016), Gameel and Wilkins (2019), Goldberg et al. (2015), Guo and Reinecke (2014), Jiang et al. (2018), Morris et al. (2015), Shrader et al. (2016). Only a few research have been conducted on the influence of demographics of MOOC students and their completion/ dropout rates. In this article, the researchers aim to explore the impact of demographic factors such as age, gender, educational level and marital status on course completion and course dropout. As a result, the following questions will be addressed by this study.

RQ1: How do age, gender, educational level, and marital status affect a learner’s decision to complete a MOOC?

RQ2: To what extent do factors such as age, educational attainment, gender, and marital status influence a learner’s decision to withdraw a MOOC?

2. OBJECTIVES

1) To analyse the relationship between age, educational level, gender, and marital status and MOOC course completion.

2) To analyse the relationship between age, educational level, gender, and marital status and dropping of MOOC.
3. RESEARCH HYPOTHESES

The following hypothesis are framed to test the objectives. H₁ through H₄ is used to test objective 1 and H₄ through H₈ is used to test objective 2.

H₁: There exists a significant relationship between age and MOOC course completion.

H₂: There exists a significant relationship between gender and MOOC course completion.

H₃: There exists a significant relationship between educational level and MOOC course completion.

H₄: There exists a significant relationship between marital status and MOOC course completion.

H₅: There exists a significant relationship between age and dropping of MOOC.

H₆: There exists a significant relationship between gender and dropping of MOOC.

H₇: There exists a significant relationship between educational level and dropping of MOOC.

H₈: There exists a significant relationship between marital status and dropping of MOOC.

4. LITERATURE REVIEW

4.1. IMPACT OF DEMOGRAPHIC DATA ON MOOC COURSE COMPLETION

This study Alshehri et al. (2021) estimates students' purchase behaviour and the revenues of massive open online course based on a learner's activity clickstream and demographic data acquired from futurelearn.com's MOOC platform. This study analysed data from 23 runs split among 5 MOOC courses on 4 different topic areas, all given by the University of Warwick through Future Learn. The research used multiple machine learning techniques to predict course purchase ability, including Random Forest, Gradient Boosting, AdaBoost, and XGBoost. By including learner demographics, such as gender, age group, level of education, and nationality, greater accuracies of 0.83 to 0.95 were obtained, which had a significant impact on the model performance.

This article Panagiotakopoulos et al. (2021) summarizes the findings of a study conducted during a Massive Open Online Courses (MOOCs) for smart city professionals, in which demographic and personal data, as well as a limited number of contextual information between learners and the MOOC, were analysed to determine the factors influencing the decision to start or not start the MOOC. Based merely on the number of first-day logins in the MOOC and a few demographic data like the learner's present profession and the amount of study hours he or she can contribute on a weekly basis, the prediction achieved an accuracy of 85 percent.

The amount to which various parameters influence the completion of a Massive Open Online Course is investigated in this research (MOOC) Zhang et al. (2019). A pre-course survey was used to collect data on those characteristics, such as gender, age, education level, proficiency in English, and motivation for course enrolment. The results reveal that individual over the age of 50 have a better chance of finishing the MOOC. MOOC completion rates are also higher when the MOOC provides experiences that complement students' present academic backgrounds and when
the MOOC is offered by a reputable educational institution. Guo and Reinecke (2014) discovered that age was strongly connected with the quantity of coverage, which was then positively correlated with grade in a study of four edX MOOCs with data from 140,546 students. Breslow et al. (2013) discovered no link between age and MOOC completion, and no link between gender and MOOC accomplishment, but barely a minor relationship between highest degree attained and MOOC achievement.

4.2. IMPACT OF DEMOGRAPHIC DATA ON DROPPING OF MOOC

This investigation and analysis of student dropout behaviour in online learning with the identification of reasons and understanding of their impact is the focus of this research Rawat et al. (2021). This study revealed that a student’s likelihood of dropping out of a MOOC is largely determined by both course-related and student-behavioural characteristics. Many of the dropout students were between the ages of 19 and 25. The relationship between students’ pre-existing misconceptions and their retention in a MOOC setting is investigated in this study Chen et al. (2020). The research findings include a positive impact of age on course survival whereas student gender and educational level did not affect the course survival.

In this research Feng et al. (2019), a systematic investigation was conducted regarding the dropout problem in MOOCs using a dataset from Xuedong, one of China’s largest MOOCs. The finding of the research states that young people are more likely to drop out, so age is considered as a significant factor. Another important factor was gender, with female users being more likely to withdraw science courses and male users being more likely to drop non-science courses. Also, educational level plays a role in the decision to drop MOOC courses, as higher dropout rates were from master’s and doctorate degree holders.

5. RESEARCH METHODOLOGY

5.1. RESEARCH DESIGN AND SAMPLING

The study used a descriptive survey as its research design. This design is deemed appropriate since it allows the researcher to gather information using standard processes based on highly organized research instrument(s) and well-defined study themes and variables. The data was collected from students, faculty, and others from different institutions in Sultanate of Oman during the period May 2021–June 2021. A total of 79 people took part in the survey. Convenience sampling method, one of the non-probabilistic sampling techniques was adopted in this research.

5.2. RESEARCH INSTRUMENT, DATA COLLECTION AND DATA ANALYSIS

To obtain the needed information from the participants, a self-developed questionnaire was used. There were two sections to the questionnaire (Part I and Part II). Part I contained questions to collect the personal details of the respondents, whereas Part II contained questions related to the reasons of enrolment of MOOC, experience of doing MOOC, preferences of choosing MOOC, strengths, and limitations of MOOC etc. Google Forms was used to design the questionnaire. After seeking the approval from the ethics committee of the institutions, the questionnaire was distributed through email.

Descriptive statistics like frequency and percentages, as well as inferential statistics like Chi-square ($\chi^2$) were used to analyse the completed surveys. To
analyse the demographic factors, descriptive statistics such as frequency and percentages were utilized, while inferential statistics such as Chi-square ($\chi^2$) were used to test the hypotheses at the 0.05 level.

6. RESULTS

The data was examined in two stages. The data was analysed using descriptive analysis (frequency and percentage) and inferential statistics.

6.1. ANALYSIS OF DEMOGRAPHIC DATA

Table 1 gives a quick overview of the participants' individual characteristics.

<table>
<thead>
<tr>
<th>Table 1 Demographic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>36-45</td>
</tr>
<tr>
<td>46-55</td>
</tr>
<tr>
<td>26-35</td>
</tr>
<tr>
<td>&gt;55</td>
</tr>
<tr>
<td>16-25</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Marital status</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Unmarried</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Educational level</td>
</tr>
<tr>
<td>PhD</td>
</tr>
<tr>
<td>Graduate</td>
</tr>
<tr>
<td>Undergraduate</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Figure 1 show the gender wise sample distribution. Female participants constitute (54.4%) and male participants (45.6%). These findings support earlier studies Bayeck (2016) and Morris et al. (2015) showing that female students make up the majority of MOOC participants.
Figure 2 presents the breakdown of the age and gender of the sample. It can be seen from the data of Figure 2 that most of the participants fall in the range of 36-45 years as reported in the research Morris et al. (2015).

The pie chart in Figure 3 illustrates the gender-wise marital status of the participants. It is evident from the data that female married and male married report the highest percentage with 39% compared to the other groups.

Figure 3 Distribution of sample by gender and marital status

Figure 4 presents the educational level of the survey participants. Out of the 79 students, 54.4% had a Graduate degree; 39.2% had a Ph.D., while 5.1% had an Undergraduate and 1.3% had other degrees.

Figure 4 Distribution of sample by educational level
6.2. ANALYSIS OF RESEARCH QUESTIONS
Chi-square (χ²) test was used to test the hypothesis.

6.2.1. RQ1: TO WHAT EXTENT AGE, GENDER, EDUCATIONAL LEVEL, AND MARITAL STATUS INFLUENCE A LEARNER TO COMPLETE A MOOC

H₁: There exists a significant relationship between age and MOOC course completion.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit χ² value</th>
<th>Calc χ² value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>79</td>
<td>4</td>
<td>0.05</td>
<td>9.488</td>
<td>3.404</td>
<td>0.492</td>
<td>Reject H₁</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 3.404, with the p-value of 0.492 which is greater than 0.05. The chi-squared critical value at df = 4 is 9.488. We reject H₁ because the test statistic is smaller than the critical value. So, there is NO sufficient evidence to prove that there exists a significant relationship between age and MOOC course completion.

H₂: There exists a significant relationship between gender and MOOC course completion.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit χ² value</th>
<th>Calc χ² value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>79</td>
<td>1</td>
<td>0.05</td>
<td>3.841</td>
<td>0.641</td>
<td>0.423</td>
<td>Reject H₂</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 0.641, with the p-value of 0.423 which is greater than 0.05. The chi-squared critical value at df = 1 is 3.841. Because the test statistic is less than the critical value, we reject H₂. So, there is NO sufficient evidence to prove that there exists a significant relationship between gender and MOOC course completion.

H₃: There exists a significant relationship between educational level and MOOC course completion.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit χ² value</th>
<th>Calc χ² value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational level</td>
<td>79</td>
<td>3</td>
<td>0.05</td>
<td>7.815</td>
<td>2.375</td>
<td>0.498</td>
<td>Reject H₃</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 2.375, with the p-value of 0.498 which is greater than 0.05. The chi-squared critical value at df = 3 is 7.815. As the test statistic is less than the critical value, we reject H₃. So, there is NO sufficient evidence to prove that there exists a significant relationship between educational level and MOOC course completion.
H4: There exists a significant relationship between marital status and MOOC course completion.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit $\chi^2$ value</th>
<th>Calc $\chi^2$ value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>79</td>
<td>1</td>
<td>0.05</td>
<td>3.841</td>
<td>0.111</td>
<td>0.739</td>
<td>Reject H4</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 0.111, with the p-value of 0.739 which is greater than 0.05. The chi-squared critical value at df = 1 is 3.841. As the test statistic is less than the critical value, we reject H4. So, there is NO sufficient evidence to prove that there exists a significant relationship between marital status and MOOC course completion.

6.2.2. RQ2: TO WHAT EXTENT AGE, GENDER, EDUCATIONAL LEVEL, AND MARITAL STATUS INFLUENCE A LEARNER TO DROP A MOOC

H5: There exists a significant relationship between age and dropping of MOOC.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit $\chi^2$ value</th>
<th>Calc $\chi^2$ value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>79</td>
<td>4</td>
<td>0.05</td>
<td>9.488</td>
<td>7.457</td>
<td>0.114</td>
<td>Reject H5</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 7.457, with the p-value of 0.114 which is greater than 0.05. The chi-squared critical value at df = 4 is 9.488. As the test statistic is less than the critical value, we reject H5. So, there is NO sufficient evidence to prove that there exists a significant association between age and dropping of courses.

H6: There exists a significant relationship between gender and dropping of MOOC.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit $\chi^2$ value</th>
<th>Calc $\chi^2$ value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>79</td>
<td>1</td>
<td>0.05</td>
<td>3.841</td>
<td>6.347</td>
<td>0.012</td>
<td>Accept H6</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 6.347, with the p-value of 0.012 which is less than 0.05. The chi-squared critical value at df = 1 is 3.841. As the test statistic is greater than the critical value, we accept H6. So, there is exists a significant association between gender and dropping of courses.

H7: There exists a significant relationship between educational level and dropping of MOOC.
Table 8 Chi-square analysis to determine the relationship between educational level and dropping of MOOC

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit $\chi^2$ value</th>
<th>Calc $\chi^2$ value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational level</td>
<td>79</td>
<td>3</td>
<td>0.05</td>
<td>7.815</td>
<td>6.684</td>
<td>0.083</td>
<td>Reject $H_7$</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 6.684, with the p-value of 0.083 which is greater than 0.05. The chi-squared critical value at df = 3 is 7.815. As the test statistic is less than the critical value, we reject $H_7$. So, there is NO sufficient evidence to prove that there exists a significant association between educational level and dropping of courses.

$H_8$: There exists a significant relationship between marital status and dropping of MOOC.

Table 9 Chi-square analysis to determine the relationship between marital status and dropping of MOOC

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>df</th>
<th>Significance level</th>
<th>Crit $\chi^2$ value</th>
<th>Calc $\chi^2$ value</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>79</td>
<td>1</td>
<td>0.05</td>
<td>3.841</td>
<td>0.598</td>
<td>0.439</td>
<td>Reject $H_8$</td>
</tr>
</tbody>
</table>

The chi-squared test statistic is calculated as 0.598, with the p-value of 0.439 which is greater than 0.05. The chi-squared critical value at df = 1 is 3.841. As the test statistic is less than the critical value, we reject $H_8$. So, there is NO sufficient evidence to prove that there exists a significant association between marital status and dropping of courses.

7. DISCUSSION

Four of the variables tested, namely age, gender, educational level, and marital status, had no significant relationship with degree of MOOC course completion. Regarding the relationship between age and course completion, our results are contradictory to the study of Guo and Reinecke (2014) where the researchers found that the elder group of learners (> 50 years old) achieve more MOOC coverage than the youngest group (< 20 years old) and the research of Impney and Formanek (2021) which looks into course enrolment, learner behaviour and learner demographics, before and after the COVID-19 pandemic declaration by World Health Organization (WHO), where a substantial increase was observed in a MOOC enrolment of population aged 18 to 30 compared to pre-COVID19 pandemic.

We did not find any statistical relationship between gender and the course completion. Our research findings are consistent with the findings of Ali Alshehri (2021) research reported from Saudi Arabia that, there is no correlation between gender and MOOC completion. Our results are also consistent with the research Shrader et al. (2016) which found that both male and female participants have an equal chance of being enrolled. The research also reports that although females and males may register in different courses at varying rates, once registered in a course, their chances of course involvement at various levels are identical. The findings of Jiang et al. (2018) also show that MOOC was equally likely to be completed by males and females. But the results are contrary to the findings reported by Jiang et al.
Exploring the impact of learners' demographic characteristics on course completion and dropout in Massive Open Online Courses

(2018) which found that females had a 12 percent lower likelihood of enrolling in at least one MOOC than males.

While exploring the relationships between educational level and course completion, our results contradict with the results of Impey and Formanek (2021) research, that in comparison to individuals with more advanced degrees, the percentage of learners with a high school diploma, some college education, or a bachelor’s degree has increased after the declaration of COVID-19 as pandemic by WHO. The research reported by Goldberg et al. (2015) also states that MOOCs can attract students from a variety of educational backgrounds. According to their findings, individuals with a high completion rate and level of engagement came from lower levels of education, demonstrating that exceptional online learners do not require a university degree. Our results show that to be successful online learners, they do not need a specific degree.

According to our research results, three out of the four characteristics studied, namely age, educational level, and marital status, had no significant link with the degree of MOOC dropouts, whereas we found statistically significant relationship between gender and MOOC dropout. Our findings of no relationship between age and the MOOC dropout contradicts the findings of Feng et al. (2019). As per the research findings of Feng et al. (2019), learners in the younger age are more likely to drop the course. Similarly, the above research also presents a relationship between educational level and course dropping. This contradicts our findings because we didn't find statistical evidence to prove the relationship between educational level and course dropping. However, our findings were consistent with those of Feng et al. (2019) when it came to the association between gender and course dropping, with female users being more likely to drop science courses and male users being more likely to quit non-science courses.

We didn’t come across any solid research trying to investigate the relationship between marital status and the course completion/ dropouts. As per our research findings, we failed to prove any statistically signification relationship between marital status and course completion, and between marital status and course dropout.

There are certain limitations to this research. Firstly, the small sample size; the data was gathered from a small number of people. Secondly, we didn't focus on specific MOOC providers or nature of courses. The information was gathered from learners who enrolled in at least one online course, regardless of the course domain or provider.

8. CONCLUSION

Massive Open Online Courses (MOOCs) have become a popular way for self-directed learners to gain new skills and knowledge. This work contributes to the MOOC literature by reporting the impact of demographics of MOOC learners on course completion and dropout. Four of the characteristics investigated, including age, gender, educational level, and marital status, had no significant link with the degree of MOOC course completion, according to our findings. However, we discovered a statistically significant association between gender and MOOC dropout, but we were unable to demonstrate the influence of age, educational level, or marital status on MOOC dropout. Future research should investigate if and how the patterns of results found in our analysis alter as MOOC providers and course topics change.

These studies’ findings are expected to serve as a starting point for further debates. Policymakers in education and university officials will benefit from this
research. They can utilize the information to create policies and target certain groups of students to provide a positive academic atmosphere and boost their institution's brand image.

REFERENCES


Exploring the impact of learners' demographic characteristics on course completion and dropout in Massive Open Online Courses


