

Cyber Security MOOC Analysis

Tanish Vikas Newale - 250204944

January 2026

1. Introduction

Massive Open Online Courses (MOOCs) have become a strategic tool for universities to expand global reach, enhance brand visibility, and support lifelong learning. In technical domains such as Cyber Security, MOOCs also play a critical role in addressing global skill shortages by providing accessible, flexible learning opportunities to diverse learner groups.

This report presents an in-depth analysis of learner data from the Newcastle University Cyber Security MOOC. The primary aim is to extract actionable insights that can support academic planners, marketing teams, and course designers in improving learner targeting, inclusivity, and retention. The analysis is structured using the CRISP-DM (Cross Industry Standard Process for Data Mining) framework, ensuring a systematic and industry-recognized approach to data-driven decision-making.

The report goes beyond visual outputs by interpreting patterns, contextualising findings within higher education strategy, and translating analytical results into practical recommendations for stakeholders.

2. Methodological Framework: CRISP-DM

CRISP-DM provides a structured six-phase lifecycle for data analytics projects: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. For clarity and relevance, this report focuses on two complete analytical cycles:

Cycle 1: Learner Demographics

Cycle 2: Diversity and Engagement

Each cycle begins with a clearly defined business question and concludes with evaluative insights and strategic implications.

CRISP-DM Cycle 1: Learner Demographics 2.1 Business Understanding

Objective: To identify the geographical distribution and age profile of learners enrolled in the Cyber Security MOOC.

Strategic Importance: Understanding who the learners are and where they come from is essential for:

Designing culturally and professionally relevant course content

Selecting appropriate examples, case studies, and assessment styles

Informing international marketing and recruitment strategies

Aligning course difficulty with learner maturity and professional experience

For a globally offered MOOC, demographic insights directly influence both academic quality and commercial sustainability.

2.2 Data Understanding

Data Sources: The analysis draws on enrolment datasets generated during the MOOC registration process. Key variables include:

Country of residence

Age or age range

Gender

Data Quality Considerations: Initial exploration revealed the presence of missing, blank, or “Unknown” values within demographic fields. These entries were systematically filtered out to ensure that:

Aggregate trends are not distorted

Visualisations accurately reflect active and identifiable learners

Conclusions remain statistically and conceptually valid

This step ensures analytical integrity before any modelling or interpretation is performed.

2.3 Modelling and Visual Analysis

Visualization 1: Top 10 Countries by Enrollment

The country-level analysis highlights a strong international footprint for the course. The highest concentration of learners originates from:

United Kingdom (GB)

India (IN)

Other countries such as the United States, Nigeria, and Australia also feature prominently, indicating interest from both developed and emerging economies.

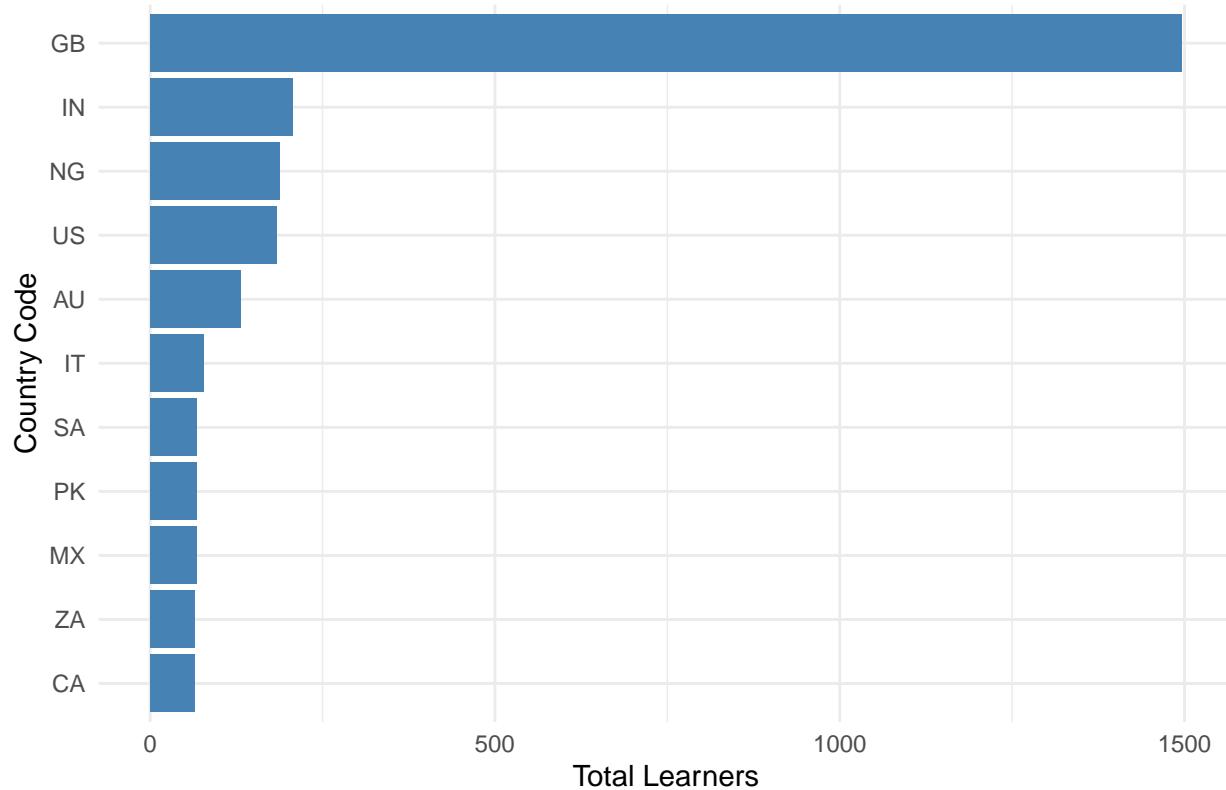
Interpretation:

High UK participation suggests effective domestic outreach and strong institutional brand recognition

Significant enrolment from India and Nigeria reflects demand for cyber security skills in fast-growing digital economies

The diversity of countries indicates the MOOC’s success as a global learning product

Top 10 Countries by Enrollment



Visualization 2: Age Distribution of Learners

The age distribution analysis shows that the majority of learners fall within the 26–45 age range, commonly associated with early to mid-career professionals.

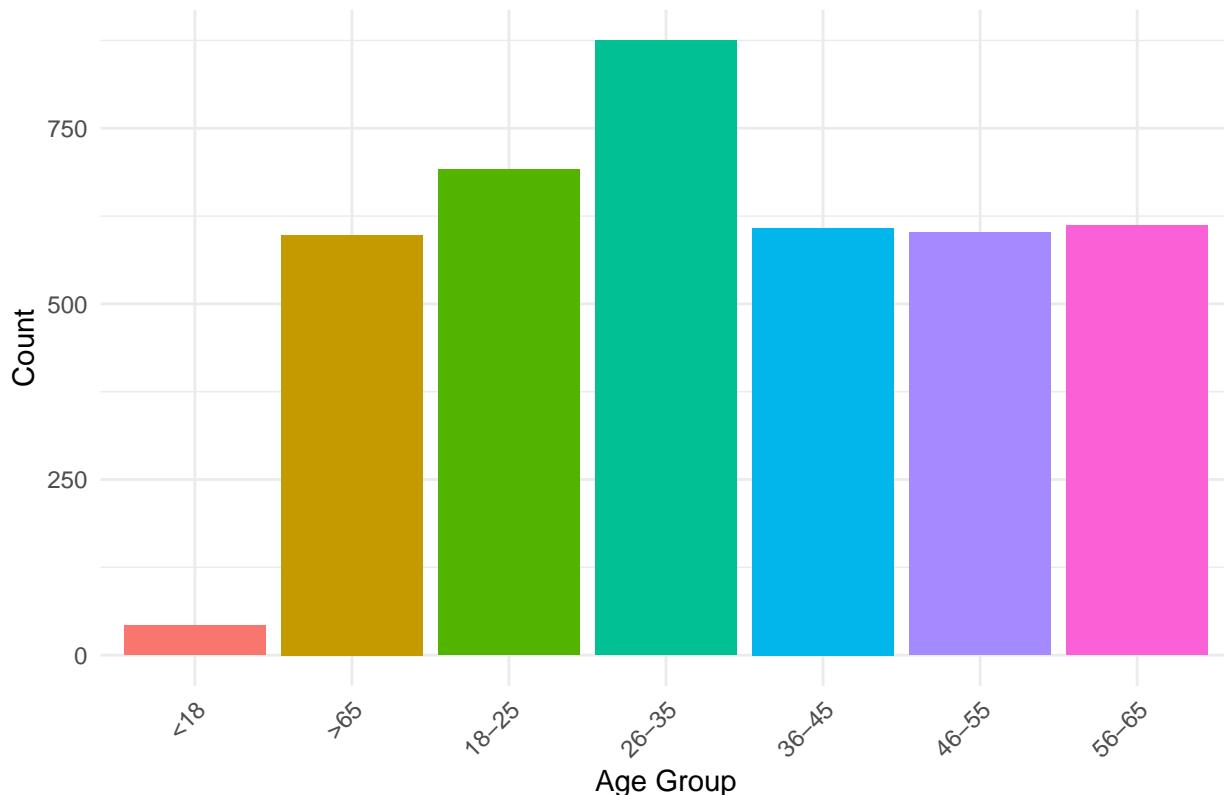
Interpretation:

Learners are likely motivated by career progression, reskilling, or role transition

The relatively lower participation from under-18 and over-65 groups suggests that the course primarily appeals to the active workforce

Course pacing, assessment complexity, and applied content are well aligned with professional learners rather than purely academic audiences

Learner Age Group Distribution



2.4 Evaluation (Cycle 1)

Cycle 1 confirms that the Cyber Security MOOC predominantly serves a global, professional, and career-oriented audience. The geographic and age patterns validate the course's positioning as a practical and industry-relevant programme rather than an introductory academic course.

These insights establish a foundation for deeper analysis into learner inclusivity and behavioural engagement.

CRISP-DM Cycle 2: Diversity and Engagement 3.1 Business Understanding

Research Questions:

How diverse is the learner population in terms of gender?

Does learner engagement decline as the course progresses?

Strategic Rationale: While enrolment numbers indicate reach, sustained engagement and inclusivity determine long-term success. Gender balance is a key concern in cyber security education, a field traditionally dominated by males. Similarly, engagement patterns help identify structural weaknesses in course design.

3.2 Modelling and Visual Analysis

Visualization 3: Gender Distribution of Learners

The gender distribution reveals a clear imbalance, with male learners significantly outnumbering female and non-binary participants.

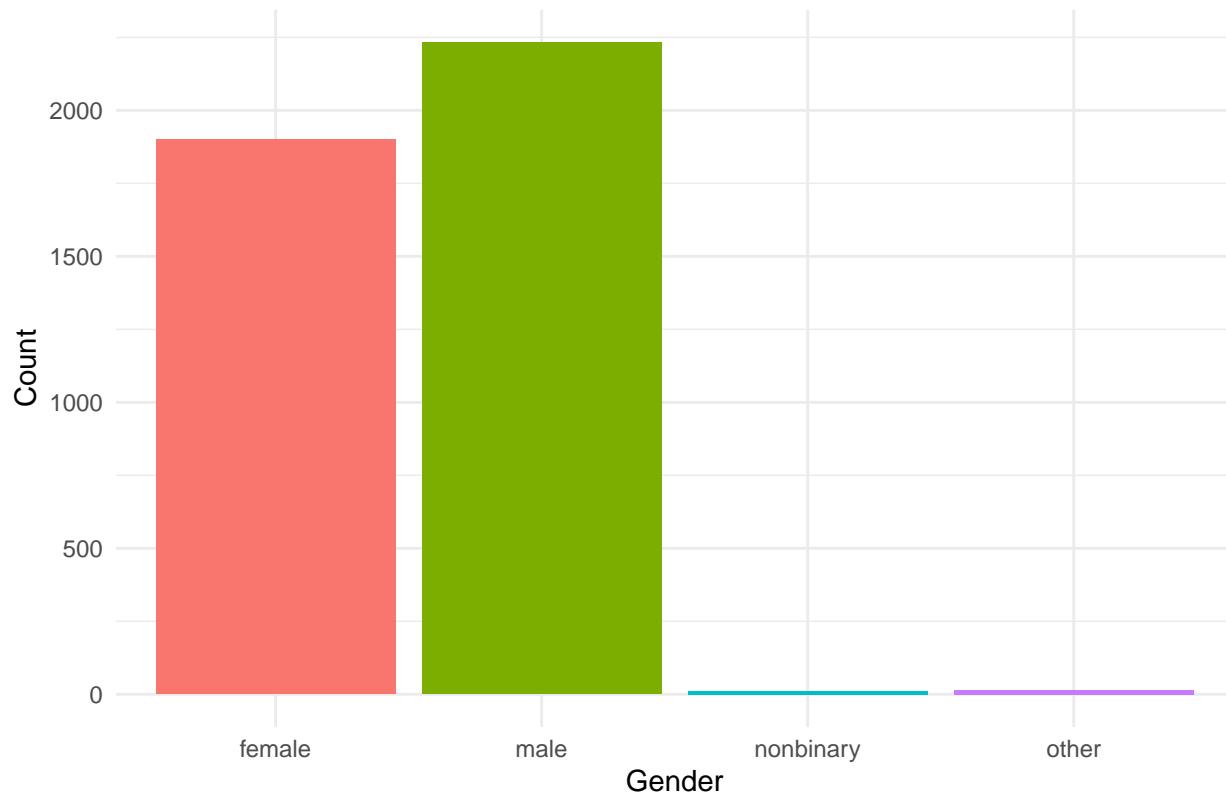
Interpretation:

The findings mirror broader industry trends within cyber security and STEM fields

Structural barriers, lack of targeted outreach, and perception issues may discourage female participation

Without intervention, such imbalance risks reinforcing existing inequalities in the cyber security workforce

Gender Distribution of Learners



Visualization 4: Weekly Engagement and Drop-off

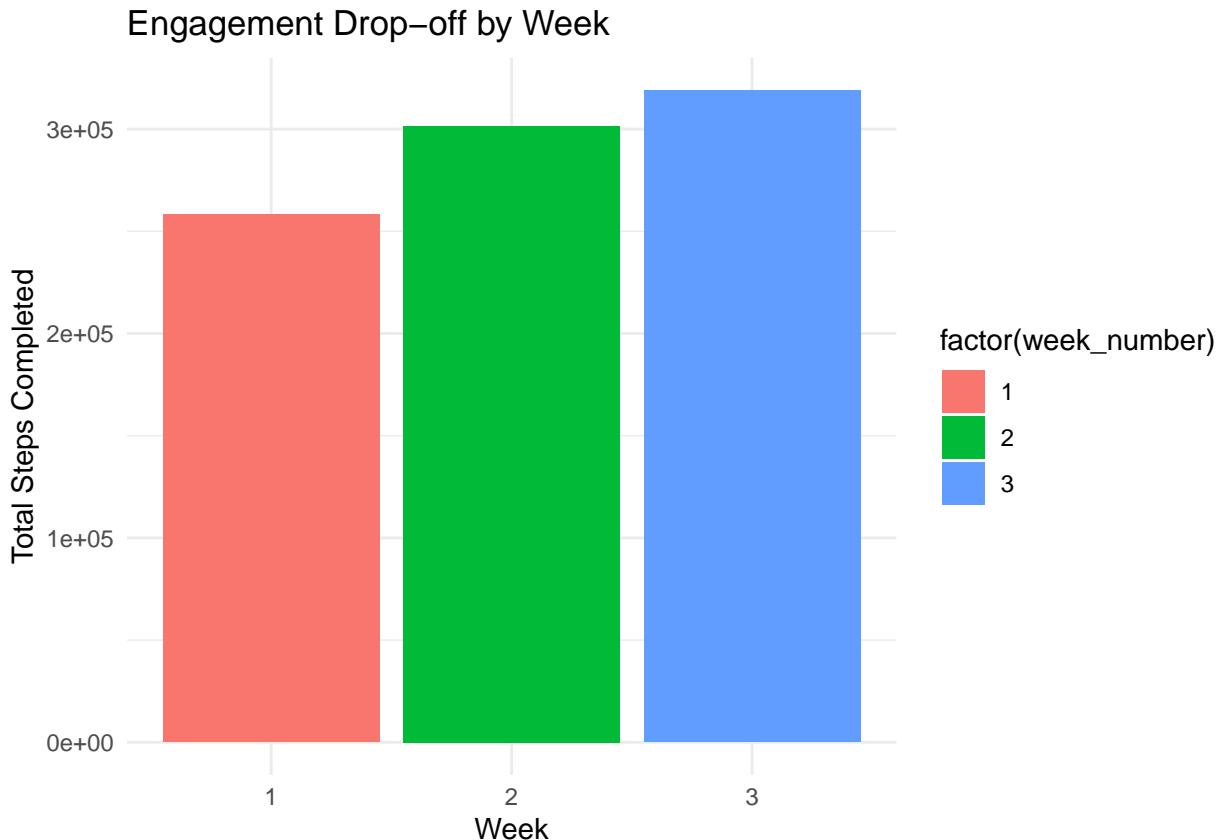
Weekly engagement analysis demonstrates a sharp decline in activity after Week 1, followed by continued reduction in subsequent weeks.

Interpretation:

Initial curiosity and enrolment motivation do not consistently translate into sustained participation

Learners may experience time constraints, cognitive overload, or misalignment between expectations and course demands

Early-course friction is a critical risk point for learner attrition



3.3 Evaluation and Deployment (Cycle 2)

Key Findings:

- The learner base lacks gender diversity
- Engagement drops significantly after the first week, indicating retention challenges
- Strategic Recommendations:
 - Introduce female-focused marketing campaigns and partnerships with women-in-tech organisations
 - Highlight female role models and instructors within course materials
 - Implement early “quick win” activities such as short quizzes, badges, or applied tasks in Week 2
 - Provide clearer onboarding guidance to align learner expectations with course structure
 - These interventions can enhance both inclusivity and completion rates.

4. Conclusion

This analysis demonstrates how structured data analytics can support evidence-based decision-making in higher education. By applying the CRISP-DM framework, the report systematically transforms raw enrolment and activity data into strategic insights.

The Cyber Security MOOC successfully attracts a global, professional audience but faces challenges related to gender diversity and learner retention. Addressing these issues through targeted design and marketing strategies can significantly enhance the course’s educational and societal impact.

Overall, this study highlights the value of combining robust technical analysis with rich theoretical interpretation to inform meaningful organisational action.

Deployment: Newcastle University should target female professionals in marketing campaigns and introduce “quick win” milestones in Week 2 to improve retention.