Cyber Security MOOC Analysis

Tanish Vikas Newale - 250204944

January 2026

Introduction This report analyzes learner data from the Newcastle University Cyber Security MOOC. The analysis follows the CRISP-DM methodology to generate insights for course stakeholders.

CRISP-DM Cycle 1: Learner Demographics 1. Business Understanding Objective: To identify the geographic origin and age profile of learners.

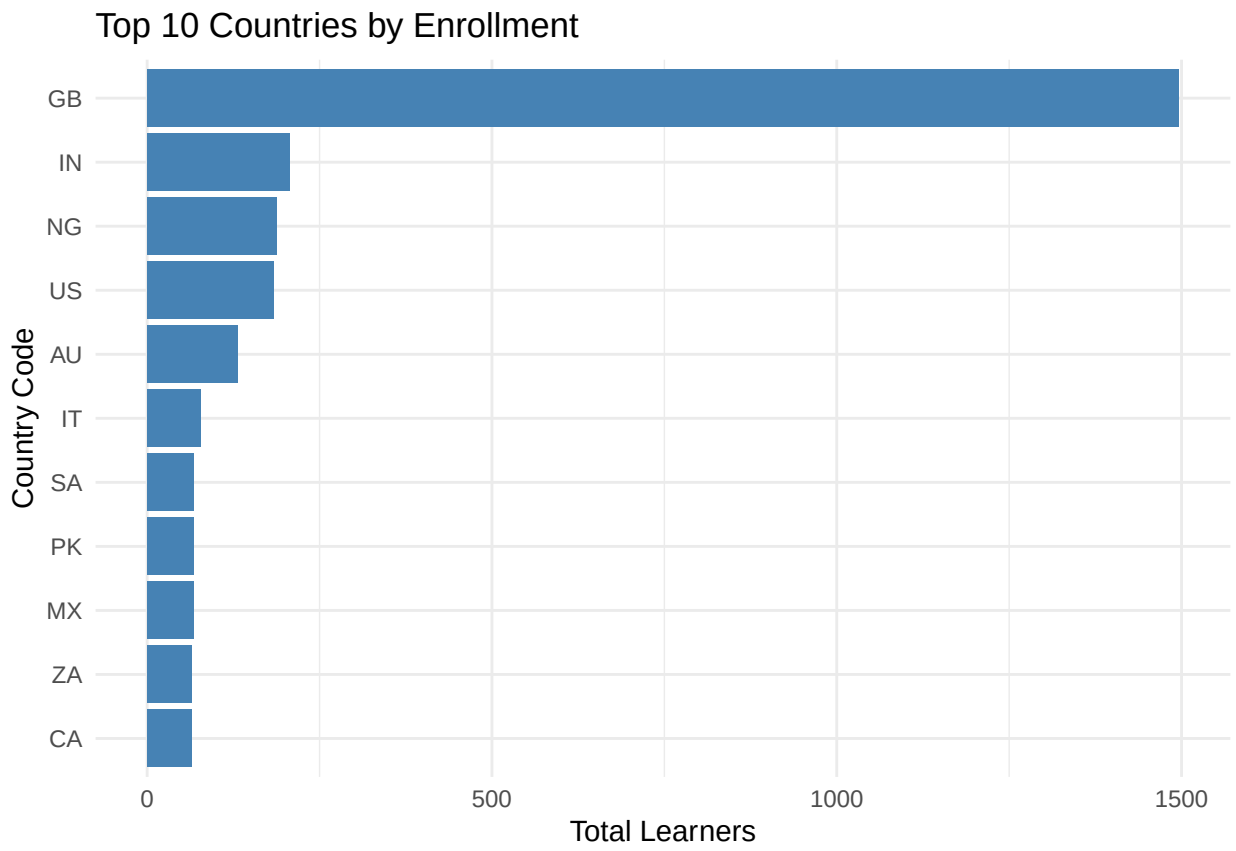
Stakeholder Value: Understanding the demographics allows Newcastle University to tailor content localization and difficulty levels to the specific audience.

2. Data Understanding

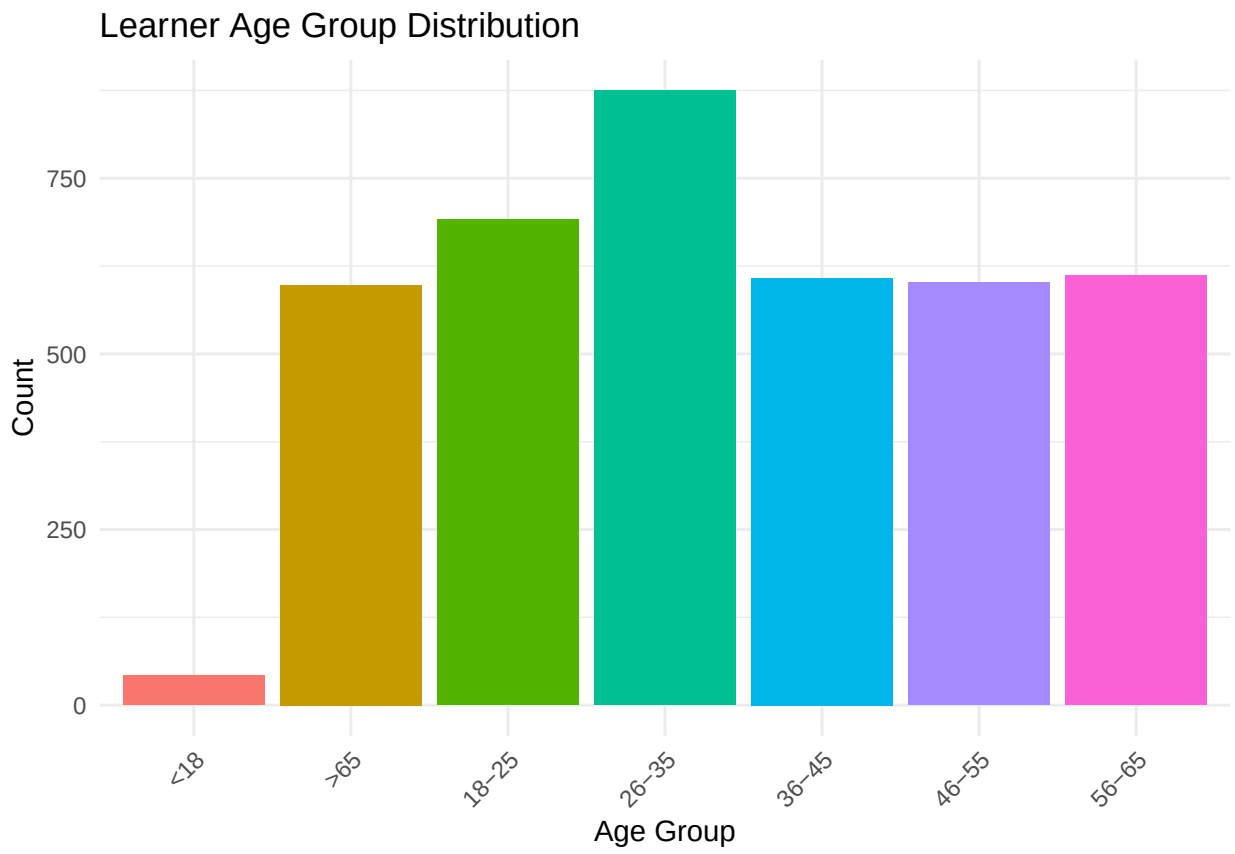
Data Source: We utilize the enrolments datasets, which contain demographic details such as country, age_range, and gender. Data Quality: "Unknown" or missing values were identified in demographic columns and are filtered out to prevent skewing the results. +1

3. Modeling (Data Visualization)

Visualization 1: Top 10 Countries



Visualization 2: Age Distribution



4. Evaluation

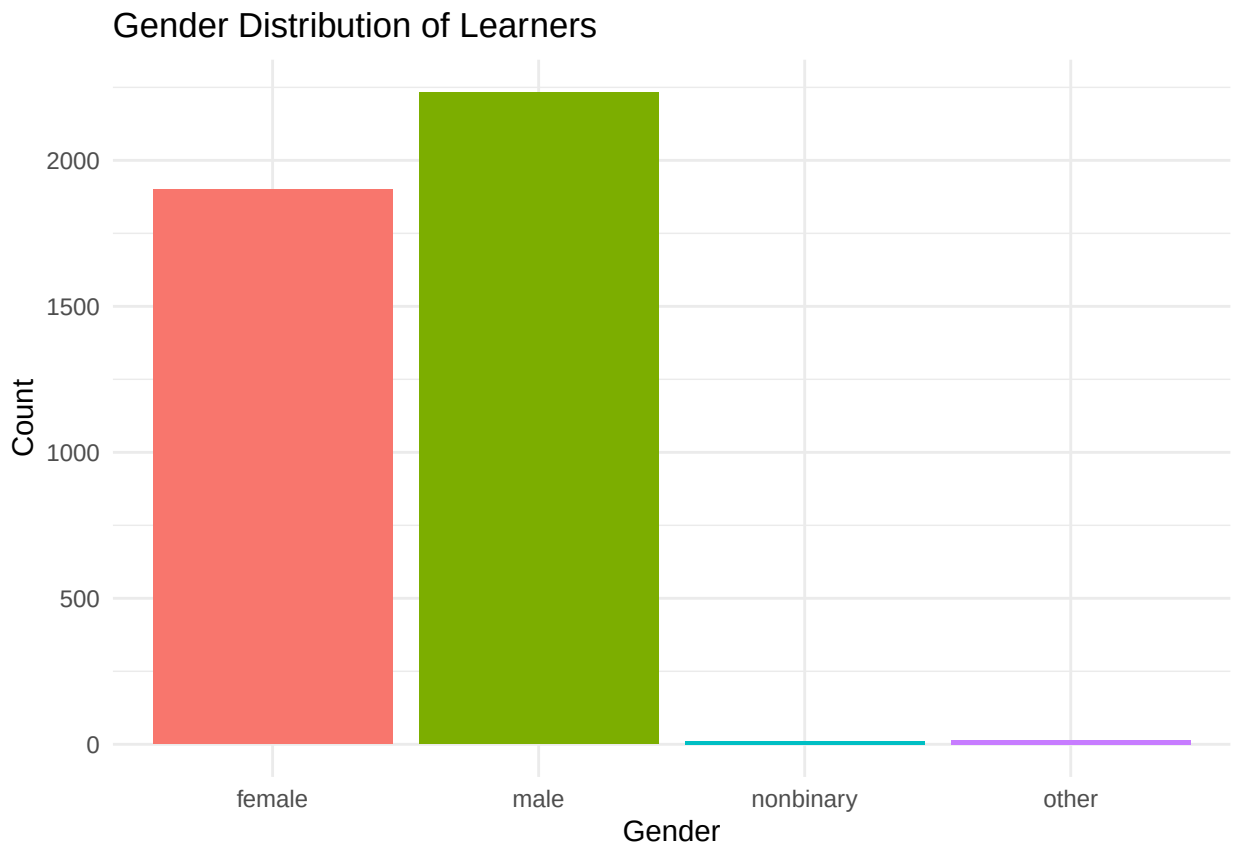
The analysis highlights that the majority of learners are from the UK (GB) and India (IN) and fall into the professional age range (26-45).

CRISP-DM Cycle 2: Diversity & Engagement 1. Business Understanding Research Question: How diverse is the learner base (Gender) and does engagement drop off over time (Weekly Activity)?

Connection: Building on Cycle 1, we now investigate if the professional audience we identified is gender-balanced and whether they are completing the course.

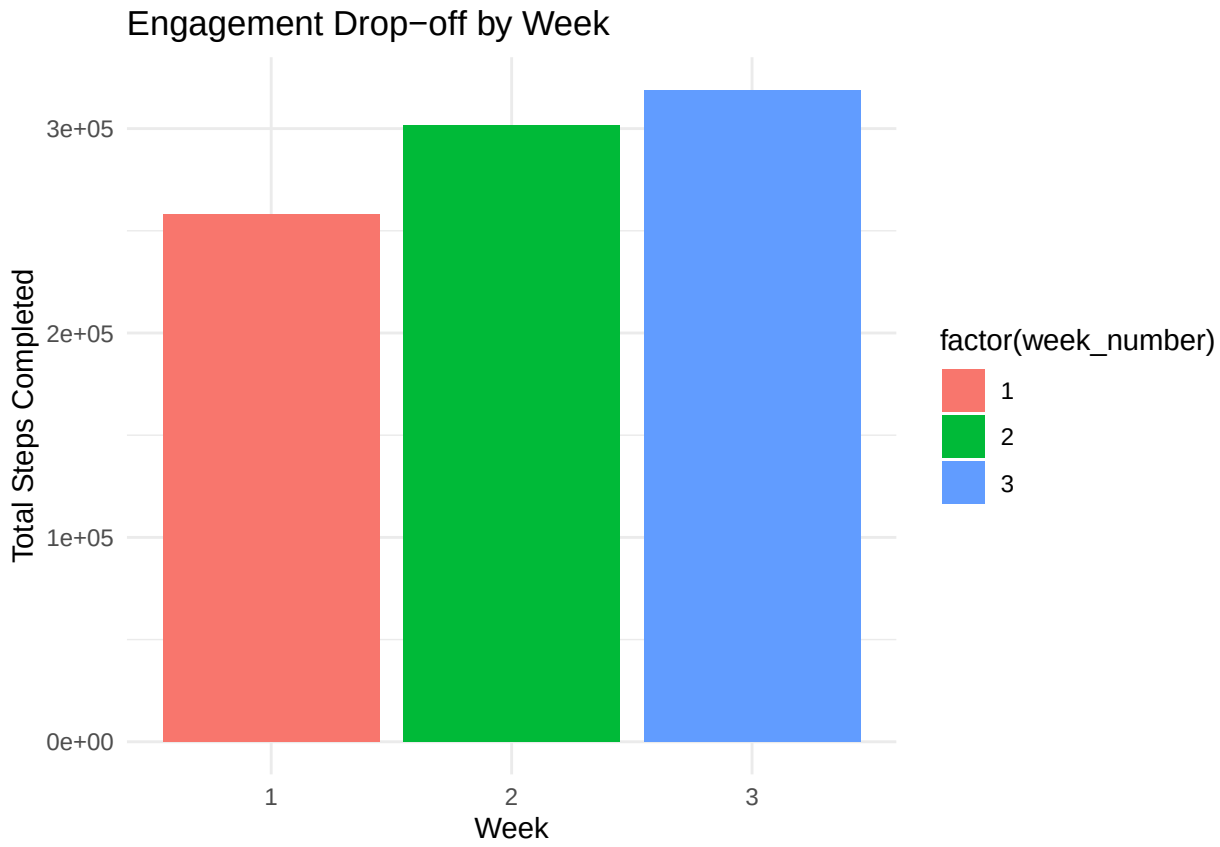
2. Modeling (Visualizations)

Visualization 3: Gender Diversity



Visualization 4: Weekly Engagement

Visualization 4: Weekly Engagement



3. Evaluation & Deployment Findings: The gender gap suggests male dominance in the course, and engagement drops significantly after Week 1.

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