Cyber Security Data Analysis Report

Roshan Pandey

11/11/2021

Introduction

Online learning platforms are getting popularity day by day and hence, generating tons of data. This data can be used to help the course or the platform to grow and gain more traction by improving the course even further. So, in this report we will try to answer questions like, which device learners use more often to access the content, from which region more learners are taking up the course, over the various iterations of the course, number of enrolled learners are increasing or decreasing.

Objective

There are mainly two objectives:

1. Target Right Audience.
2. Course Improvement.

Target right audience:

> We want to know from which **location** mostly people are enrolling for the course, what are their employment background and status, their age range etc. This will allow the course provider to target right audience.

Course Improvement:

> Accessing the content should be devices(**Desktop, Mobile Phone, Tablet**) friendly based on which device most learners are using.

About the data

Data consists of 7 iterations of the course, Cyber Security: Safety at Home, Online, in Life.

1. "Cyber-security-1_archetype-survey-responses"- file consists of responses of learners for archetype surveys for 1st iteration of the course.
2. "Cyber-security-2_enrolments"- file consists of Enrollment details of the learner for 2nd iteration of the course.
3. "Cyber-security-3_leaving-survey-responses"- file consists of responses of learners for course leaving survey for 3rd iteration.

4. “Cyber-security-4_question-response”- file consists of responses of learners for assessment questions for 4th iteration of the course.
5. “Cyber-security-5_step-activity”- file consists of step wise first visit and last completion time of learners for 5th iteration of the course.
6. “Cyber-security-6_team-members”- file consists of the team and user role of people for the 6th iteration of the course.
7. “Cyber-security-6_video-stats”- file consists of data related to the length of the video, no. of views, no. of downloads, device viewed on, part of world viewed from etc. for 6th iteration of the course.
8. “Cyber-security-7_weekly-sentiment-survey-responses”- file consists of responses of learners about their sentiments for the week’s learning.

Data Processing

We are considering data into two different formats, mentioned bellow: 1. Combined data: Combining data of same file genres from all the iterations. For example, enrollment data from each iteration is been combined row wise (rbind)

How is course performing over different iterations?

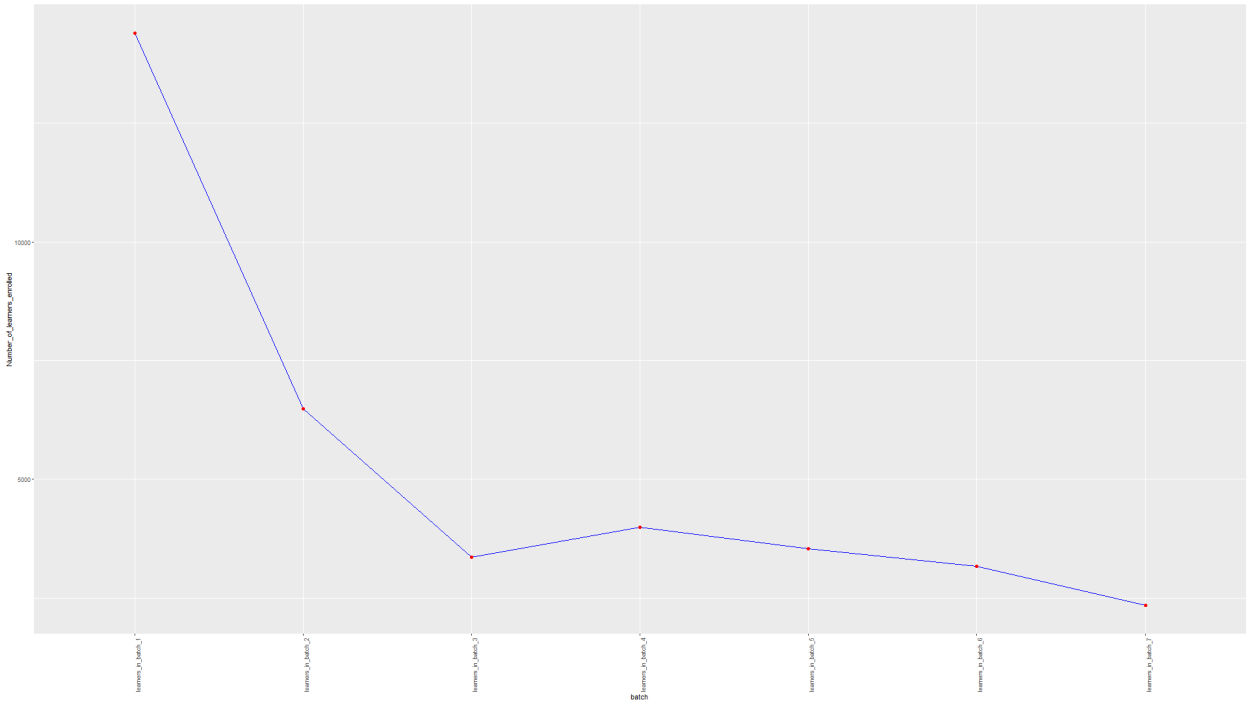

Figure 1: Continent from where most learners are enrolling

From which region most learners are enrolling?

From figure 1, we can observe most learners are coming from Europe

From fig 2, we can see that mostly learners are enrolled from the UK, India, and USA and there are considerable number of people from regions like Australia, Saudi Arabia, Nigeria, Mexico, and Russia. So, based on number of people from various locations advertisement can be optimized. For example, rather than

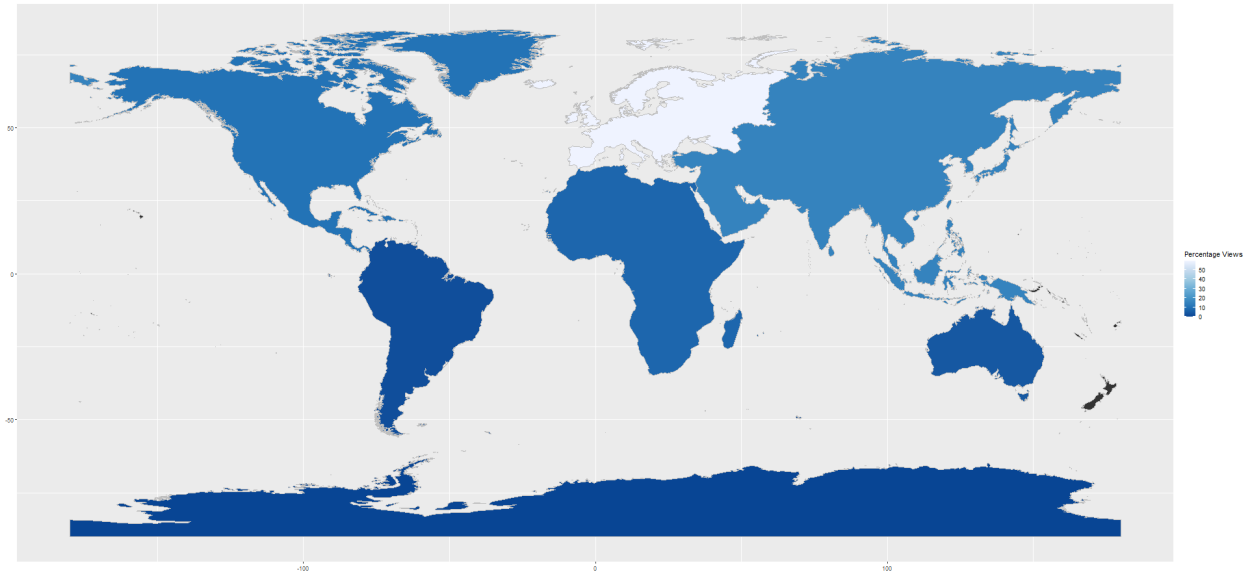

Figure 2: Continent from where most learners are enrolling

advertising the course or the platform with same intensity in all the regions, importance can be given to regions like the UK and India. This way cost of investment in advertising can be reduced and profits can be increased and that cost can also be used for reinvesting in the course or platform to improve it even further.

Which type of people are enrolling for the course?

Which device is being mostly used to access the content?

From Figure 4 we can see that most of the times course is being accessed on desktop and hence, the learning platform and content can be optimized further more for desktop first and then focus can be shifted to other devices like mobile and tablets etc.

Figure 5 shows although mobile phones are being significantly less used as compared to desktops, they are increasing with iterations. So, platform and course content optimization for mobile phone should also be given equal importance.

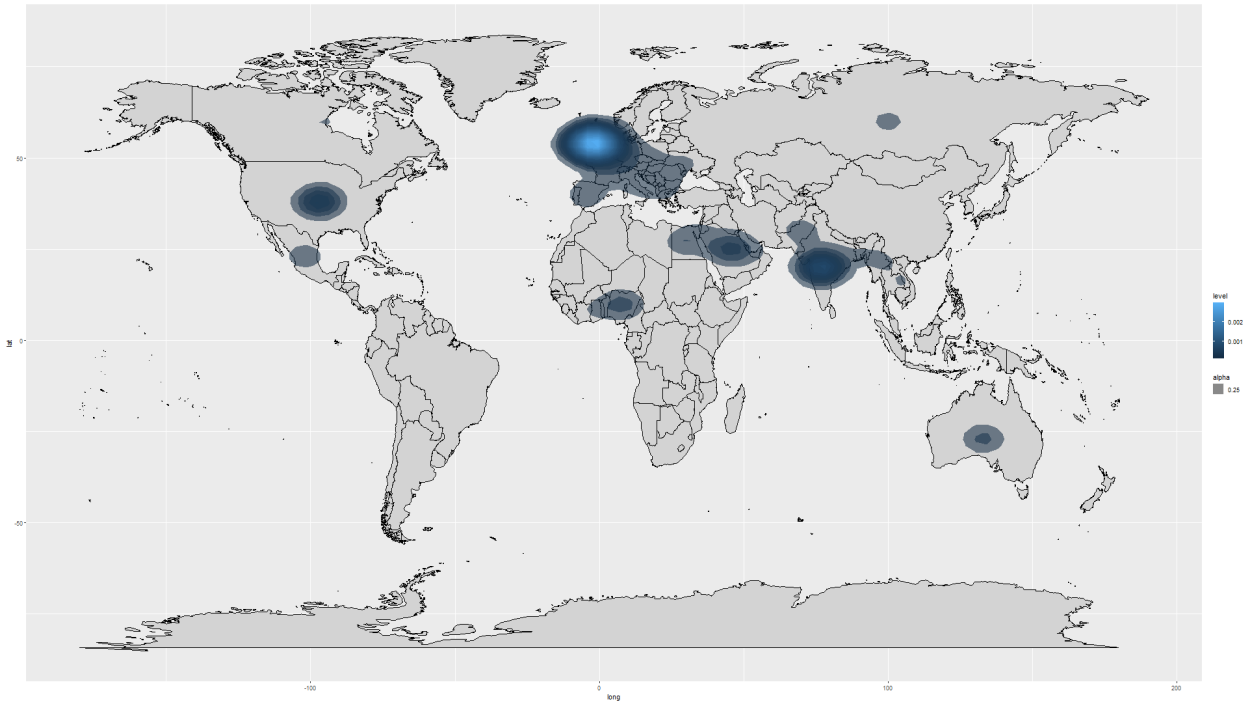

Figure 3: Most widely used device to access the content

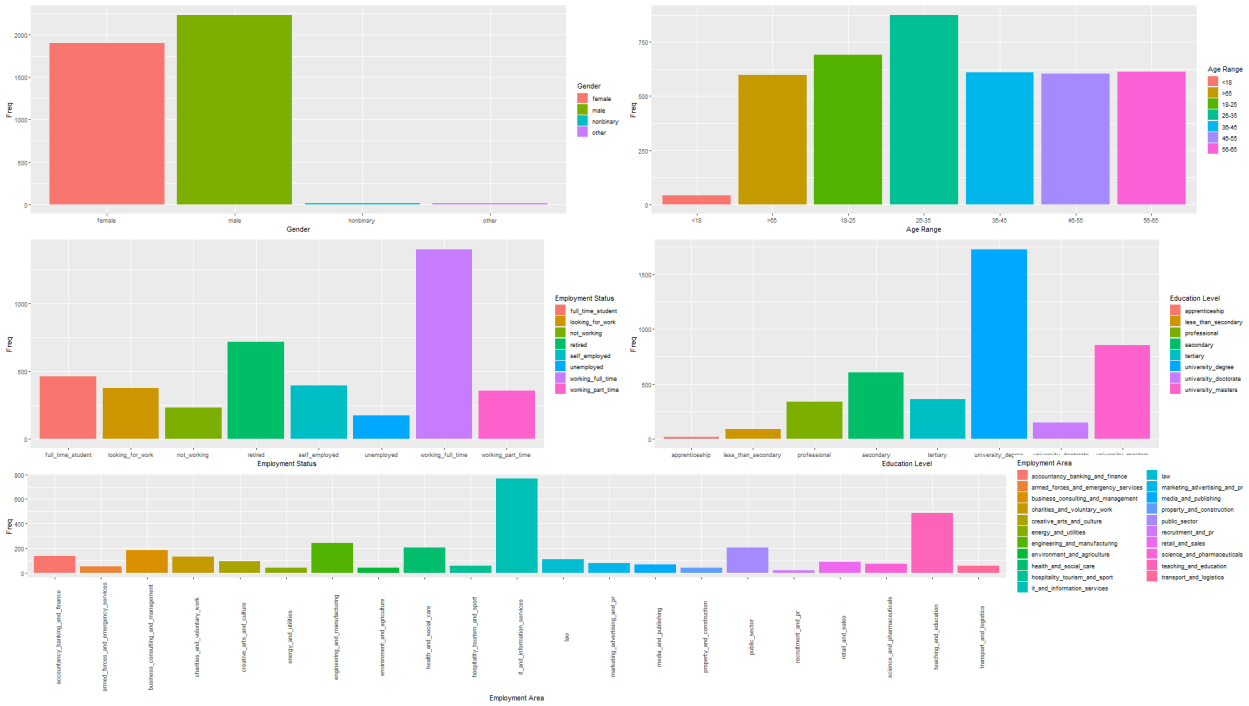

Figure 4: Which type of people are enrolling for the course?

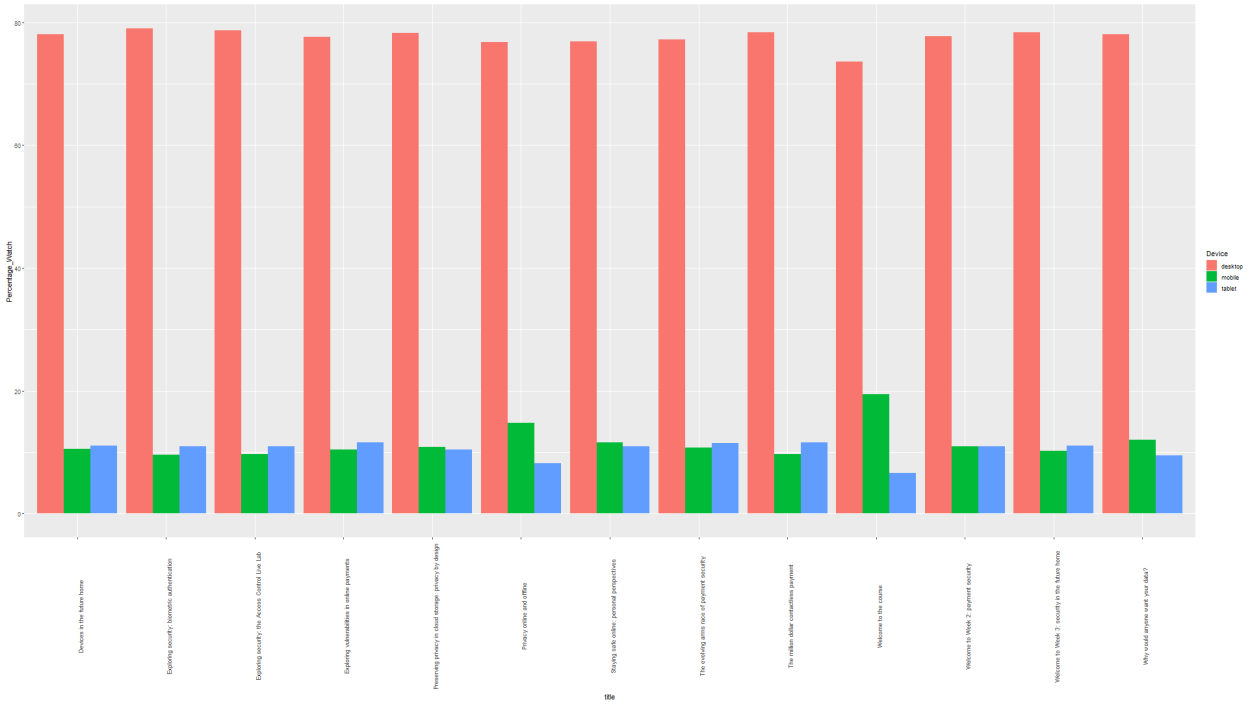

Figure 5: Most widely used device to access the content in each iteration

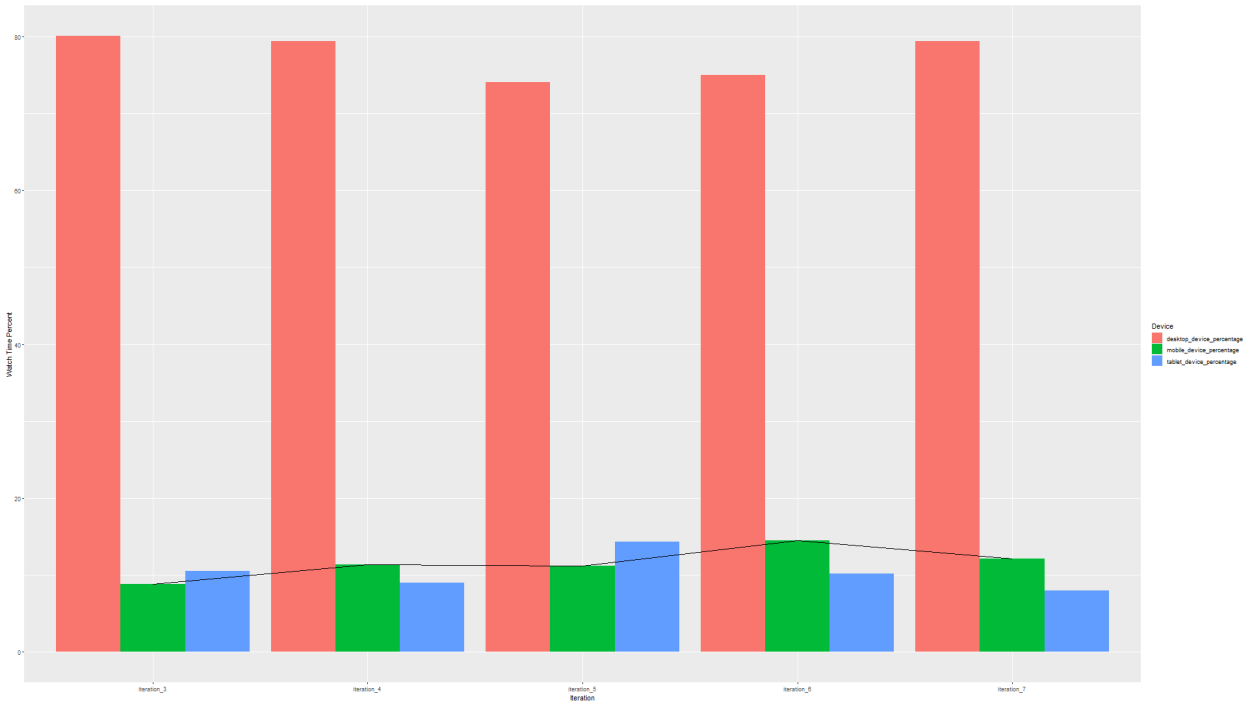

Figure 6: Most widely used device to access the content in each iteration