

# EXPLORATORY DATA ANALYSIS OF THE SURVEY FOR FUTURE LEARN'S CYBER SECURITY COURSE.

---

SHUBHAM RAJESH

210462188

# Data Description:

Future Learn provided the following data for the analysis of the increase in the dropout rate for their Cyber Security course.

In this analysis only leaving survey response data is considered.

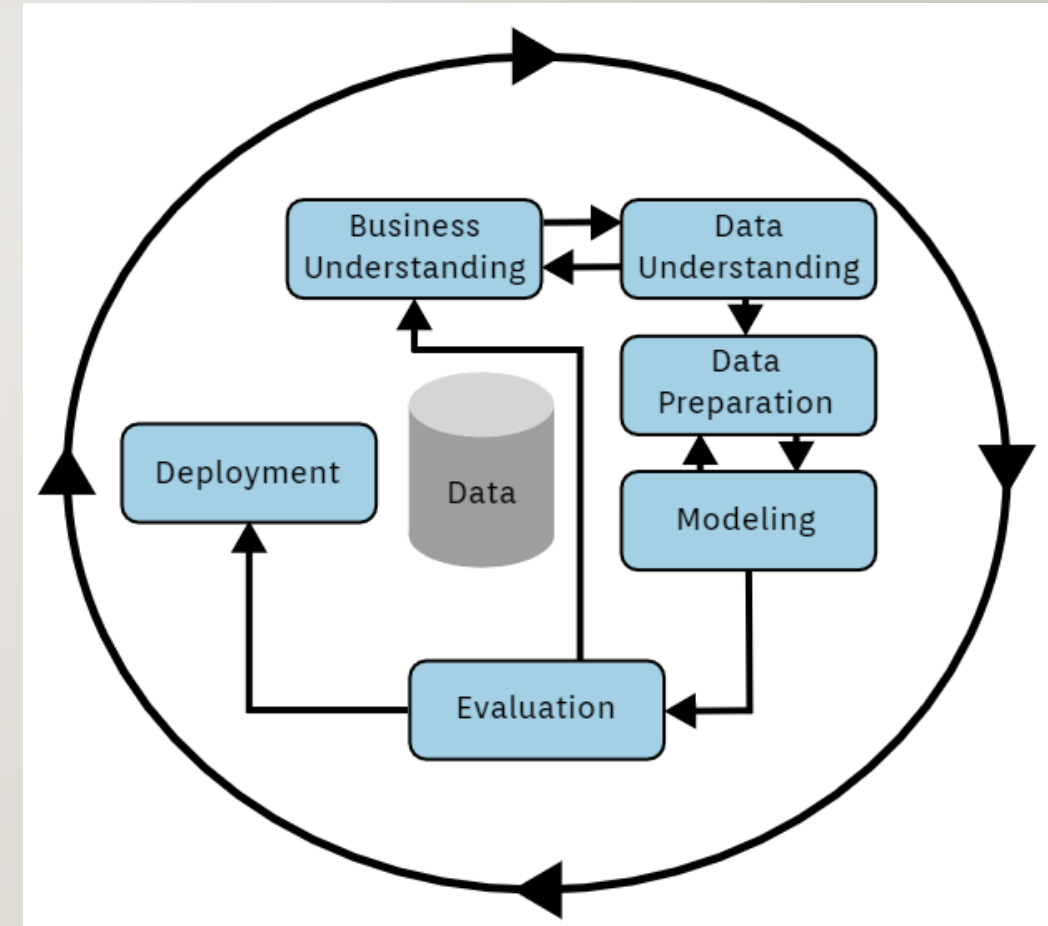
id	learner_id	left_at	leaving_reason	last_completed_step_at	last_completed_step	last_completed_week_number	last_completed_step_number
2831	5cbc8827-23de-415c-aa41-b3f2c9f28c0f	2017-11-27 17:51:03 UTC	The course wasn't what I expected	2017-11-20 18:01:16	1.16	1	16
5012	f5cb04e7-e09a-4cce-8f28-4a04b203f27a	2017-11-29 18:02:17 UTC	The course won't help me reach my goals	2017-11-28 21:01:15	1.15	1	15
5459	5b4373c7-72cf-4675-8af6-c5589497d392	2017-11-30 01:17:18 UTC	The course wasn't what I expected	2017-11-19 22:02:11	1.1	1	1
7493	02fbf753-9bbe-4236-a0a6-a5d25f3d65f8	2017-12-01 13:33:27 UTC	The course wasn't what I expected	2017-12-01 13:31:32	3.2	3	20
13663	470171ac-7fc9-47bc-bfca-819682ae75ca	2017-12-05 00:22:18 UTC	The course required more time than I realised	2017-11-15 02:47:13	1.3	1	3
14154	a9d385ab-5785-43d4-a126-bd7f070b932c	2017-12-05 09:29:09 UTC	The course required more time than I realised	2017-11-15 12:31:19	1.19	1	19
15137	206f1ef2-408a-4ede-a3d2-edbceea5d0c9	2017-12-05 19:35:56 UTC	The course wasn't what I expected	2017-12-05 19:34:32	3.2	3	20
16110	9a4b15f1-e7df-41f8-a586-97fff9db0486	2017-12-06 12:55:08 UTC	The course won't help me reach my goals	2017-11-19 12:40:12	1.2	1	2
16419	94a0c34f-44c4-4729-8b3e-1eeb19caf8ab	2017-12-06 15:52:09 UTC	The course wasn't what I expected	2017-12-06 15:51:15	1.5	1	5
16967	de6f99e0-4f64-4325-88f5-abb9196be8b3	2017-12-07 00:14:06 UTC	The course won't help me reach my goals	2017-12-07 00:13:18	3.18	3	18

# CRISP-DM Methodology

This report examines a Massive Open Online Course (MOOC) on Cyber Security for late 2017 and 2018 using the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology and R programming.

Stages of CRISP-DM:

1. **Business Understanding:** Identify business objectives, plan an approach.
2. **Data Understanding:** Explore the data and check its quality.
3. **Data Preparation:** Preparing data for fitting in the model.
4. **Modeling:** Model Building and Assessment.
5. **Evaluation:** Check if the model fulfils the business objective.
6. **Deployment:** Plan for monitoring and maintenance, make a final report.



# Business Problems:

- What were the most and least common reasons for students dropping out of the Cyber Security course? Week wise analysis for the most common leaving reason?
- Which week did the most and least participants leave the Cyber Security course? Also, what was the most last common step completed by students before dropping out of the course?





# Pie chart to compare most and least common leaving reasons for students to have left the course

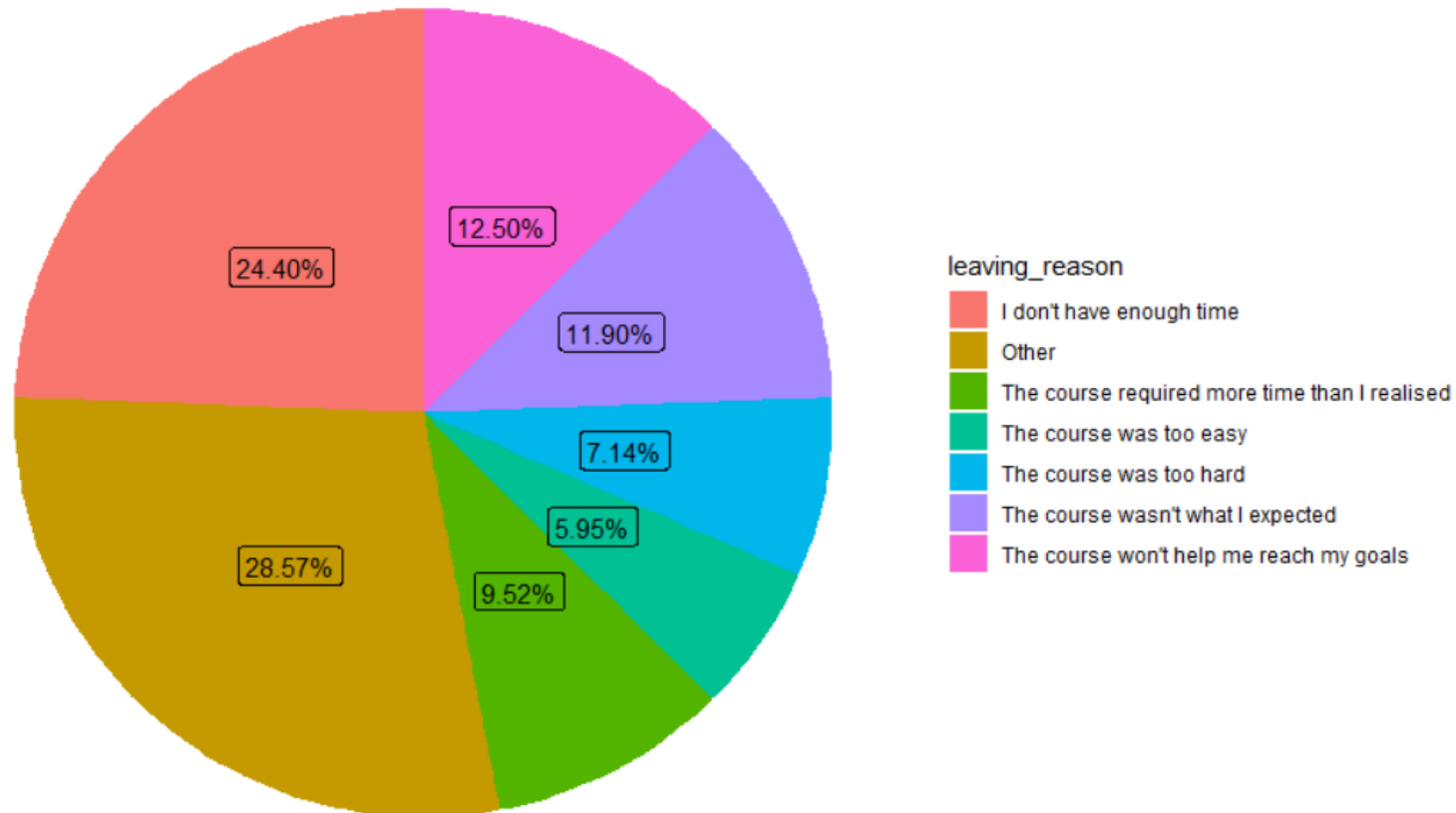


Figure: I

# Week-wise analysis of Leaving reason: 1<sup>st</sup> Week

Bar plot for identifying which was the most common Leaving Reason for students that left the course in 1st Week

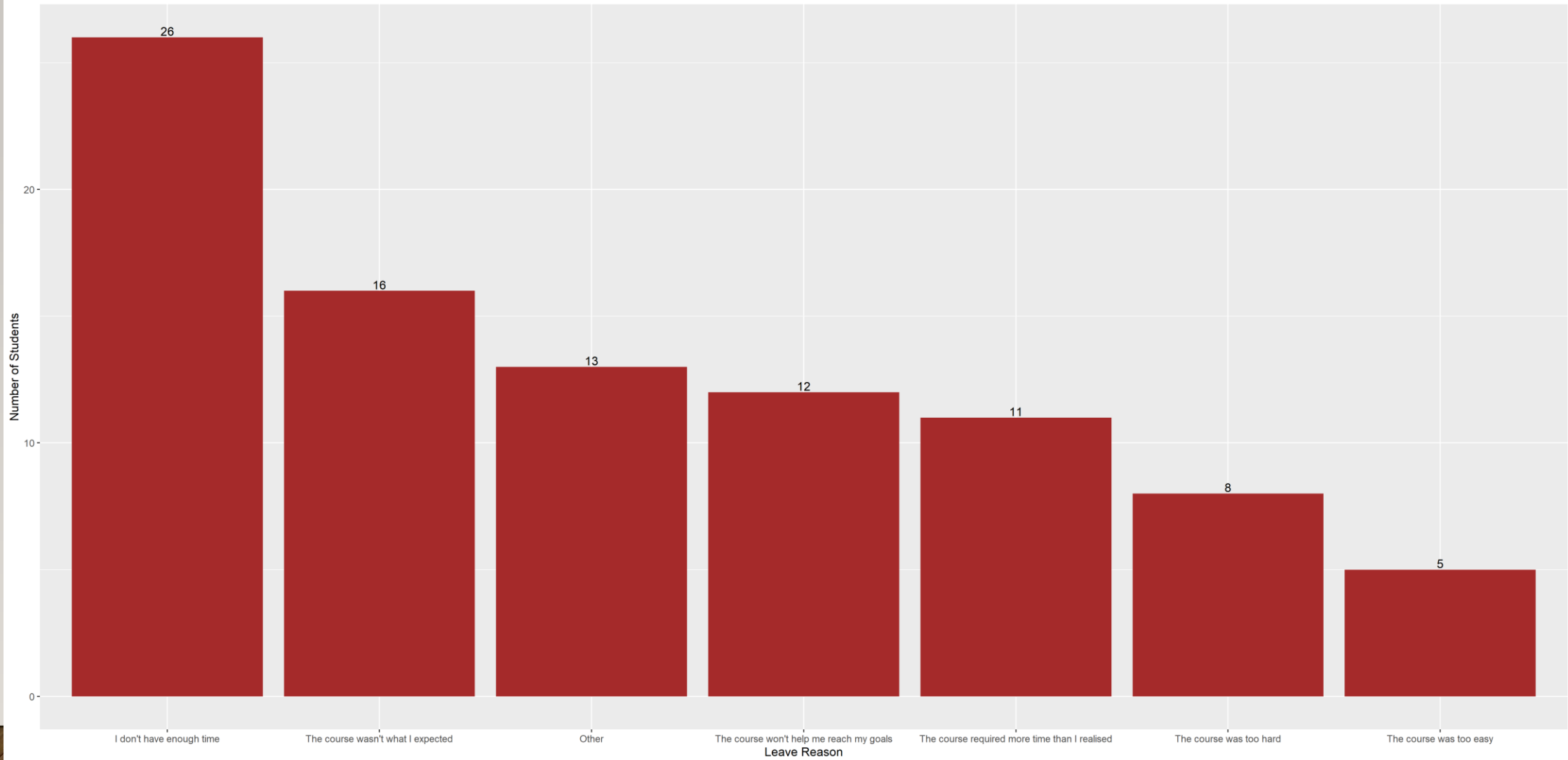


Figure: 2

# Week-wise analysis of Leaving reason: 2nd Week

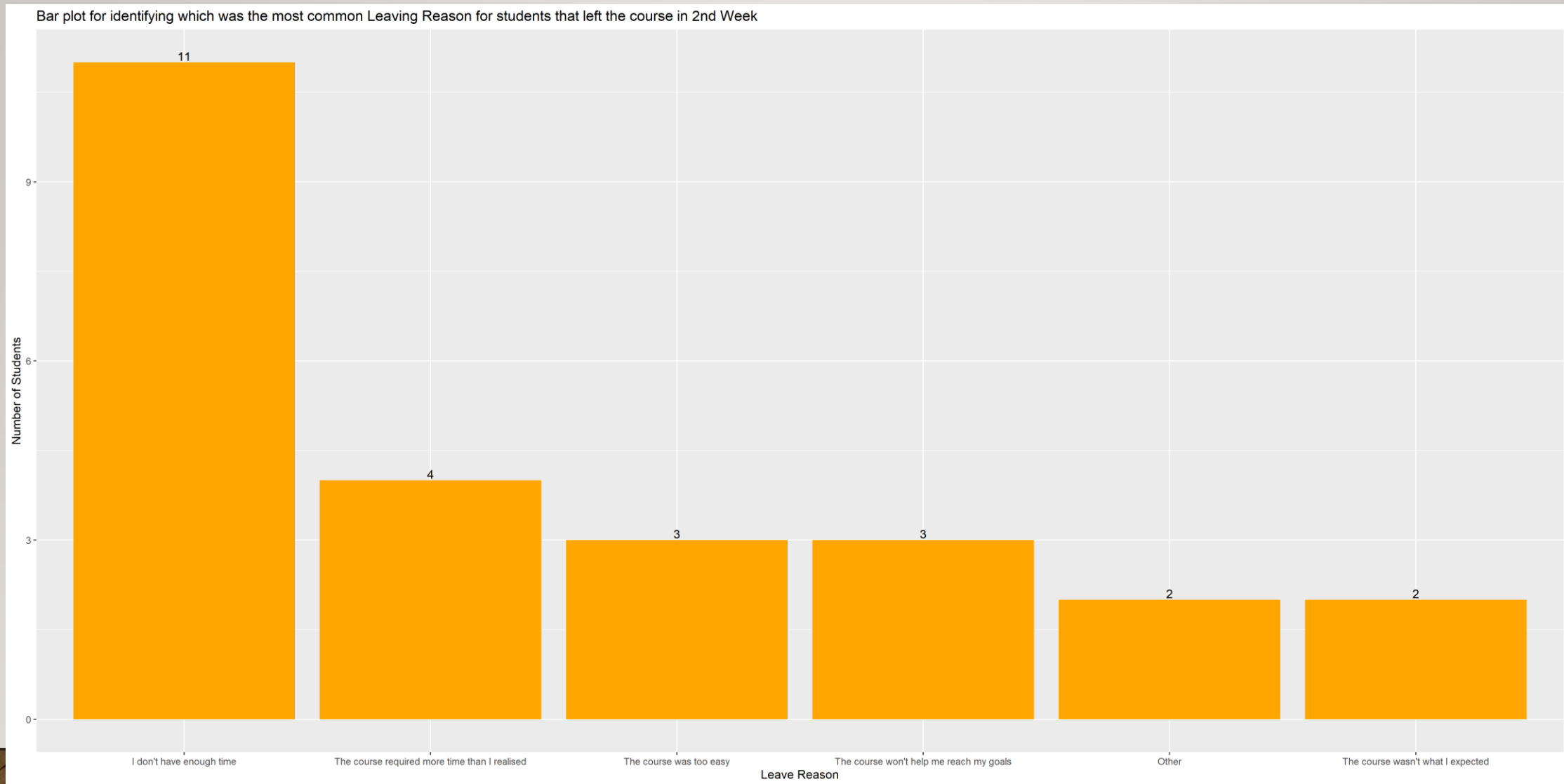


Figure: 3

# Week-wise analysis of Leaving reason: 3rd Week

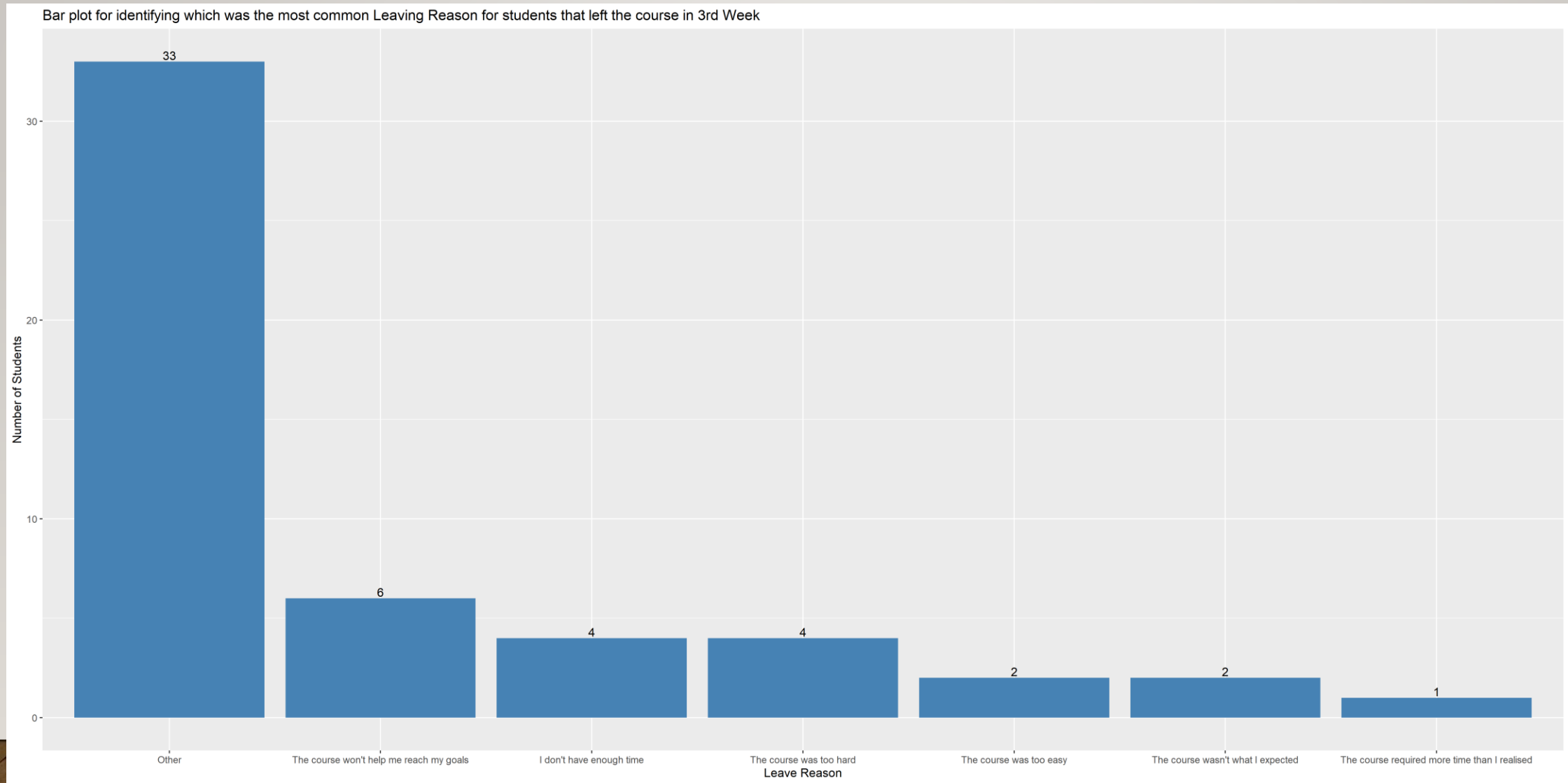


Figure: 4



# Bar-plot to compare in which week did the most and least participants left the Cyber Security course

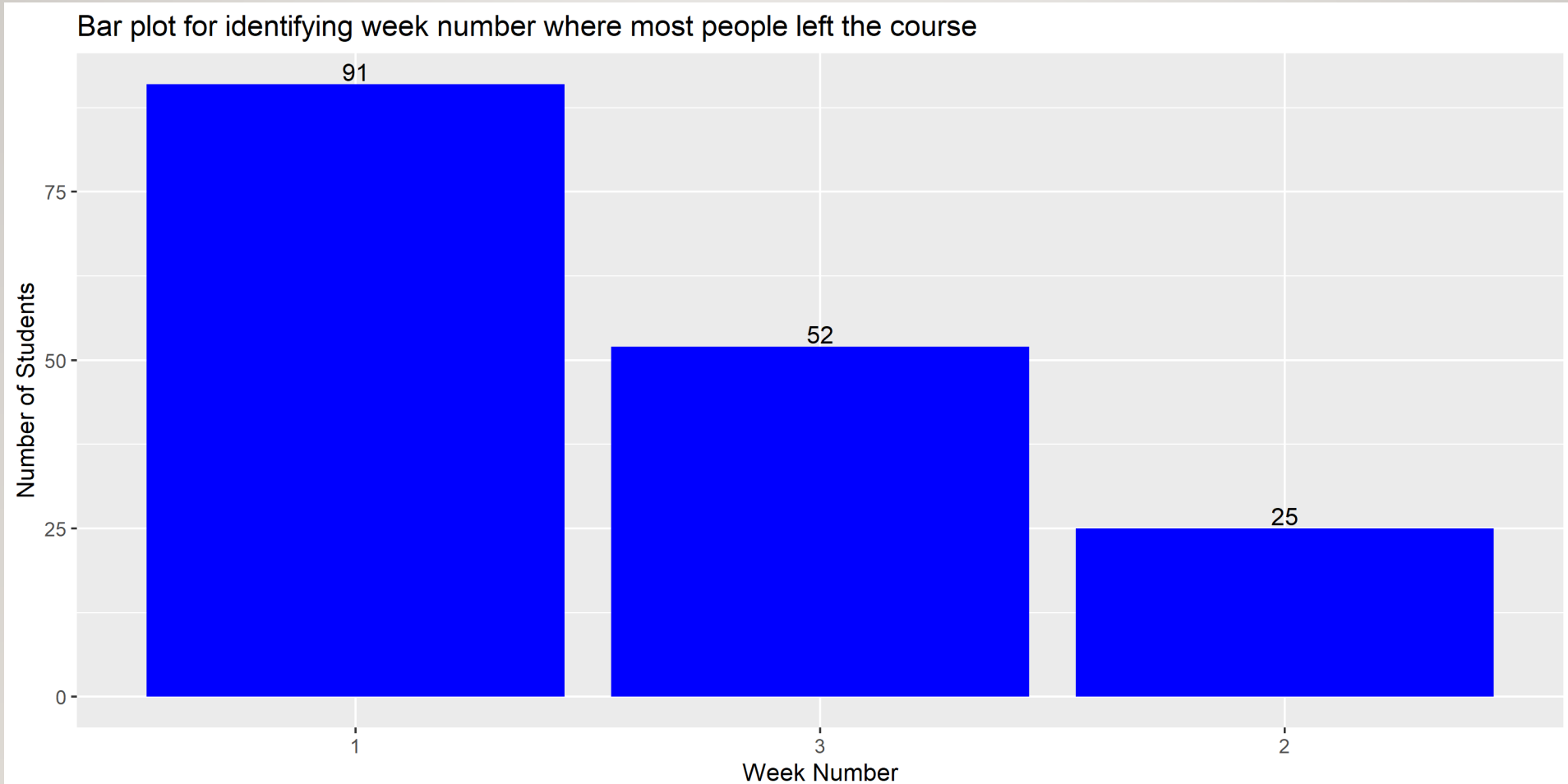


Figure: 4

# Bar-plot to analyze the most common Last step completed for students that left the course

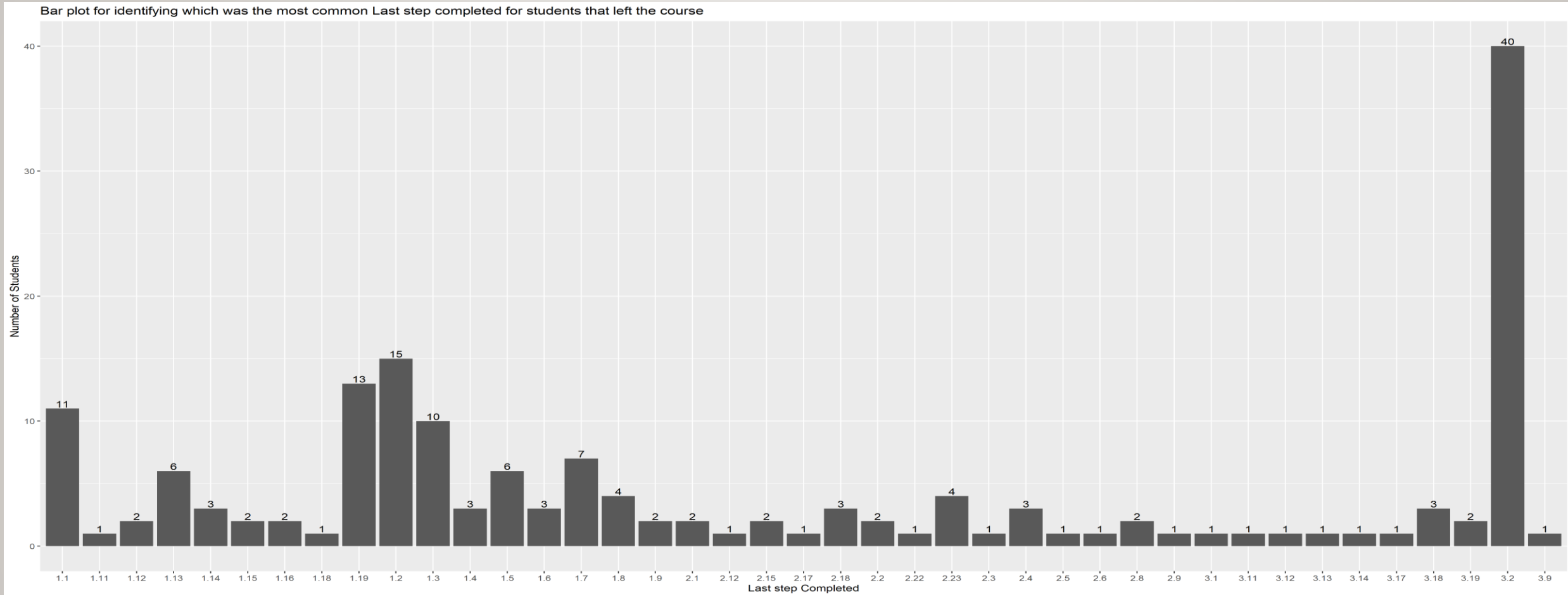


Figure: 5

# Conclusion

- Majority of students chose "Other", and the second most popular answer was "I don't have enough time" so it means students didn't feel they should invest their time in this course.
- Also, by observing the last completed steps analysis, it can be speculated that most people completed the last step (i.e. Step 3.2) but didn't complete the course, so the "Other" leaving reason can't be neglected.
- Therefore, the major reason for the increase in the dropping out rate in the third week can be concluded that learners found the quiz hard and could not clear it.
- And for the first week, it can be concluded that most students who left the course completed Step 1.2 (i.e. Why are you here?) So, most people understood the course outcomes are irrelevant to what they wanted and then decided not to continue with the course.



Thank-you!