Word Separator Test

Experiment 2, Experimentation & Evaluation 2021

Filippo Piloni & Alessandro Zanzi

Abstract

This experiment aimed to see how using separators between words affects how easily we read programming code. Fifteen participants of different age, sex and coding background took part in a test where their reaction time in reading words written in camel case, kebab case, or without a separator has been measured. The results showed a clear and meaningful difference in code readability when a separator was used compared to when there wasn't one. Interestingly, the type of separator (camel case or kebab case) didn't seem to have an higher impact in terms of the readability of the code.

1. Introduction

Effective communication through written code is crucial in software development, and the readability of source code significantly impacts the efficiency of understanding and maintaining software systems. One aspect that has been explored in natural language reading studies is the use of explicit separators between words, which has been shown to improve reading speed by up to 20%. The question that arises is whether this finding holds true for programming code as well. In particular, does the use of a specific separator influence the speed of reading composed identifiers in code?

The motivation behind this experiment lies in the potential impact on code readability and developer productivity. In software development, developers often encounter composed identifiers, such as variable names or function names, that consist of more than one word. The common conventions for such identifiers include camelCase (e.g., readingThisText) and kebab-case (e.g., reading-this-text). Understanding whether these conventions affect code reading speed can provide valuable insights into best practices for code style and contribute to the ongoing discussions within the software development community.

To investigate the impact of different separators on code reading speed, our experiment will focus on three conditions: camelCase, kebab-case, and no separator. Participants will be presented with code snippets containing composed identifiers written in each of these styles. The experiment will be conducted in a controlled environment where participants will read and comprehend the provided code snippets. We will measure the time taken by participants to read and understand the code in each condition. By analyzing the reading speeds across the three conditions, we aim to draw conclusions regarding the effectiveness of separators in enhancing code readability. This experiment holds the potential to inform coding conventions and practices, contributing to the ongoing efforts to improve code quality and developer experience.

Hypotheses:

Null Hypothesis: there will not be a significant difference in the response time across different word formatting styles

2. Method

2.1 Variables

Independent variable	Levels
Separator between the words Number of words	Camel case, kebab case, no separator 2, 3

Dependent variable	Measurement Scale
Time required to click the correct word	Microseconds

Control variable	Fixed Value
Words to choose	Software-related words

Blocking variable	Levels
Computer abilities Language proficiency Prior exposure to formatting style	Low, Medium, High Low, Medium, High No exposure, some exposure, high exposure

2.2 Design

Type of Study (check one):

Observational Study	**Quasi-Experiment**	**X Experiment**

Number of Factors (check one):

Single-Factor Design	**X Multi-Factor Design**	Other

Between vs. Within (check one):

Between Group Design (independent measures)	**X Within Subject Design** (repeated measures)	Other

The study incorporates multiple independent variables, forming a multi-factor design. This approach allows us to explore the interactions and combined effects of these variables, providing a more comprehensive understanding of the phenomena under investigation.

For the experiment, all the possible combinations of independent variables have been taken into consideration for all the control variable. Each separator (camel case, kebab case, and no separator) is tested with both two and three words.

The experiment has been conducted using the within subject design each participant serves as their control, experiencing all levels of the independent variable. In this case, each participant reads code snippets in all three conditions (camelCase, kebab-case, and no separator).

2.3 Participants

The experiment involved fifteen participants, all recruited through communication channels like messages and emails within and outside a university environment. The participants exhibited a diverse set of characteristics, encompassing various factors such as age, sex, occupation, field of study, education level, usage patterns, and engagement with coding at home.

Out of the fifteen participants, 10 were male, while the remaining 5 were female. The majority of the participants, specifically 73% (11 individuals), were currently engaged in studying computer science, reflecting a focus on individuals with a background in this field, with an high-school diploma as highest degree.

The age range of the participants varied, with individuals ranging from 17 to 38 years old. Occupations included both study and work, with participants pursuing education at different levels, from high school to a Ph.D. level. Fields of study covered a spectrum, with a notable concentration in computer science.

Education levels among the participants included bachelor's, master's, and Ph.D. degrees, indicating a diverse educational background.

Additionally, participants' coding habits at home were explored, with some indicating an habit to code at home (10 of them). This diversity in coding habits provides insights into the participants' extracurricular engagement with the subject matter outside formal educational settings.

2.4 Apparatus and Materials

- Computer: model Dell Precision 5550
- Web Application: Next.js 14
- Libreoffice: version 6.4.7.2 40 (Used to save the datas)
- Jasp: version 0.18.1 (Used for statistics and graphs)

2.5 Procedure

In conducting the experiment, the participants have been engaged through an online survey platform. The experimental setup involved presenting participants with a series of 30 questions, each displaying a word in either camel case, kebab case, or without a separator. The main objective for each participant was to identify the correct word from a set of three distractors.

The participants accessed the survey through a designated website. Upon initiation, each question prompted the display of a word formatted in one of the specified cases. The participants were required to carefully examine the word and select the correct one from the three alternatives provided. The survey platform recorded both the time taken by the participants to respond to each question.

Participants were instructed to approach each question independently, ensuring that their selections were based on accurate identification rather than random guessing. The order of the presentation of the questions was randomized to minimize potential order effects.

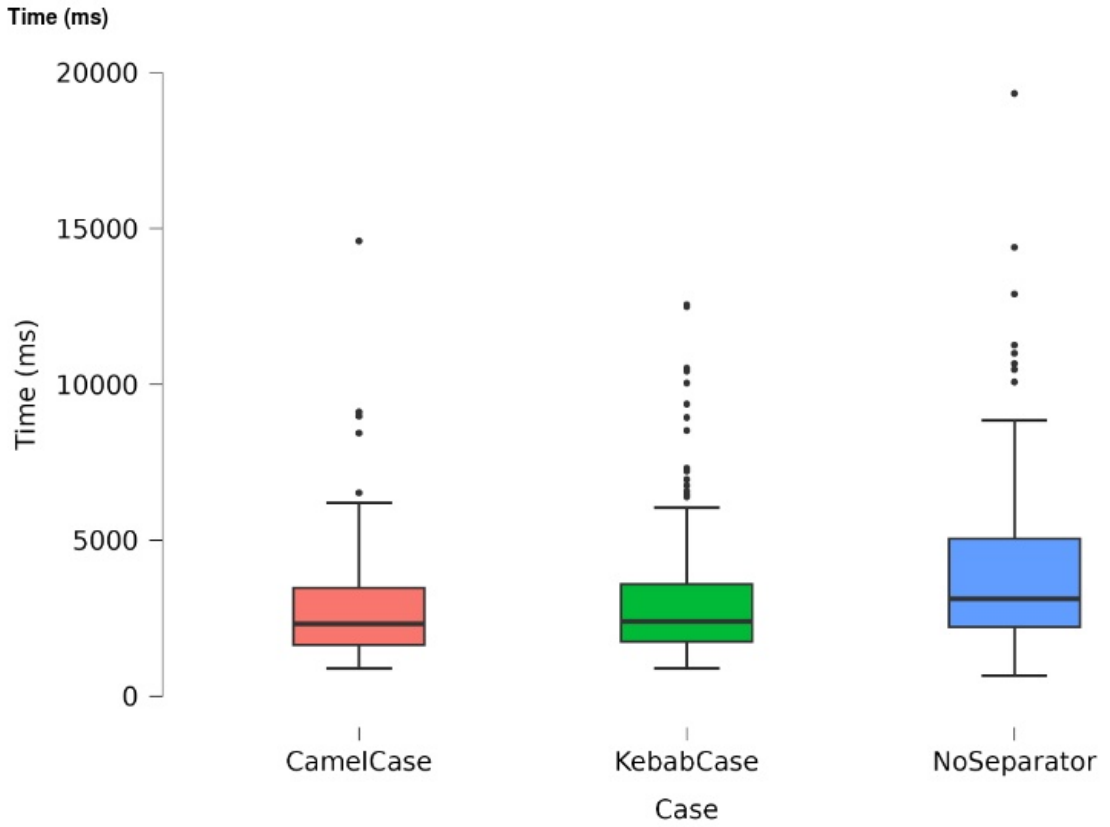
The survey was designed to be user-friendly, with clear instructions provided at the beginning to guide participants through the process. Ethical considerations, such as informed consent and the protection of participant privacy, were addressed in accordance with research guidelines.

Data collected during each session, including response times, were later analyzed to draw conclusions about the participants' ability to identify words in different formatting styles. The use of the online survey platform allowed for efficient data collection and streamlined the experimental process.

3. Results

3.1 Visual Overview

Graphs for the time of each question can be found in Appendix A.
In this section, we believe the most relevant data to show is a plot of the time distribution using the various cases.



3.2 Descriptive Statistics

To provide a comprehensive overview of the time distribution using each separator in various scenarios, we will present five key measurements: the minimum, first quartile, median, third quartile, mean, standard deviation, and maximum. The minimum and maximum values offer insights into the lower and upper bounds of processing times with a specific separator.

The first quartile signifies the value below which 25% of the data points reside, while the third quartile indicates the value below which 75% of the data points fall. The median provides an understanding of the central value within the distribution of processing times.
In addition to these traditional measures, we will include the mean, which represents the arithmetic average of the time distribution. The mean provides a measure of the central tendency, giving an overall sense of the average performance with the chosen separator. Furthermore, the standard deviation, a measure of the dispersion or spread of the data, complements the mean by indicating the extent to which individual processing times deviate from the average.

	Camel Case	Kebab Case	No Separator
Minimum (ms)	892.00	895.00	662.00
1st quartile (ms)	1640.00	1745.00	2225.75
Median (ms)	2318.50	2394.00	3125.00
3rd quartile (ms)	3471.00	3591.25	5049.50
Maximum (ms)	14599.00	12555.00	19331.00
Mean (ms)	2834.11	3193.47	4069.68
Std. Deviation (ms)	1846.27	2270.82	2786.81

3.3 Inferential Statistics

In our test, we utilized three inferential statistics to delve deeper into the results.
The t-value is a measure that measures the difference between the means of two groups, considering the variance within each group. A higher t-value signifies a more substantial difference between the means.
The p-value, another crucial measure, represents the probability of obtaining results as extreme as the ones observed, assuming the null hypothesis is true. A lower p-value suggests stronger evidence against the null hypothesis, indicating greater statistical significance.
Finally, we also used the Cohen's d. Cohen's d acts like a ruler for measuring how big the difference is between two groups. It helps us understand how important or practical that difference is. It gives us a standard way to figure out the significance of what we observe. A higher Cohen's d indicates a larger effect size and vice versa. Together, these statistical measures contribute to a comprehensive evaluation of the study's findings.

Of course, we found these measurements for each couple: Camel case – Kebab case, Camel case – no separator, Kebab case – no separator.

	t-value	p-value	Cohen’s d
Camel Case – Kebab Case	- 1.77	0.08	- 0.14
Camel Case – No Separator	- 5.76	< 0.001	- 0. 47
Kebab Case – No Separator	- 4.04	< 0.001	- 0.33

4. Discussion

4.1 Compare Hypothesis to Results

Using the data and statistics from section 3, we can draw several conclusions regarding the experiment.

Since every one of the three t-values turned out to be negative, it suggests that our sample means fall below what we would expect based on the null hypothesis, or the average we were anticipating. Since we notice higher absolute values for these t-values (especially for Camel Case – no separator and kebab case – no separator), it indicates a more pronounced and robust signal against the null hypothesis.
Interpreting the Cohen's d values, the negative signs indicate that, on average, the values in the first group are lower than those in the second group. Examining the actual values, we observe a small effect size between Camel case and Kebab case. Additionally, there are medium effect sizes between Camel case and No separator, as well as between Kebab case and No separator. This means that while the negative sign tells us about the direction of the difference, the magnitude of the effect sizes gives us a sense of how substantial these differences are. Specifically, the differences between Camel case and Kebab case are relatively small, whereas the differences between each of these cases and No separator are more substantial.

Finally, as we can see from section 3.3, the p-values can help us to conclude that there is a significant difference in response time between using any kind of separator respect of not using a separator between the words. For this reason, we reject the null hypothesis.
Observing the data collected for every single separator (both in section 3.2 and from Appendix A), we can notice a sligthly higher performance in the use Camel case respect Kebab case. We can see this from the values of the minimum and the quartiles. Despite this, the difference is too little to say which one of the two is actual better than the other.
Observing the p-value between these two groups, we can state that there is not a significant difference in response time between using the Camel case separator or the Kebab case.

4.2 Limitations and Threats to Validity

This study is not without its limitations, and certain factors pose potential threats to its validity. Firstly, the use of only 15 test subjects, while offering valuable insights, raises concerns about the adequacy of the sample size for obtaining accurate and reliable results.

Another potential challenge to validity is the variability in the environments where participants conducted the tests. The fact that each participant undertook the test at their convenience introduces a confounding variable. The changes in environment could have had an impact on the test results, adding a layer of complexity to the interpretation of findings.

4.3 Conclusions

From this study, we can draw the conclusion that the human mind demonstrates improved performance in reading tasks when words are separated by a separator, both in general reading contexts and within programming code. This finding holds true irrespective of the specific type of separator used.

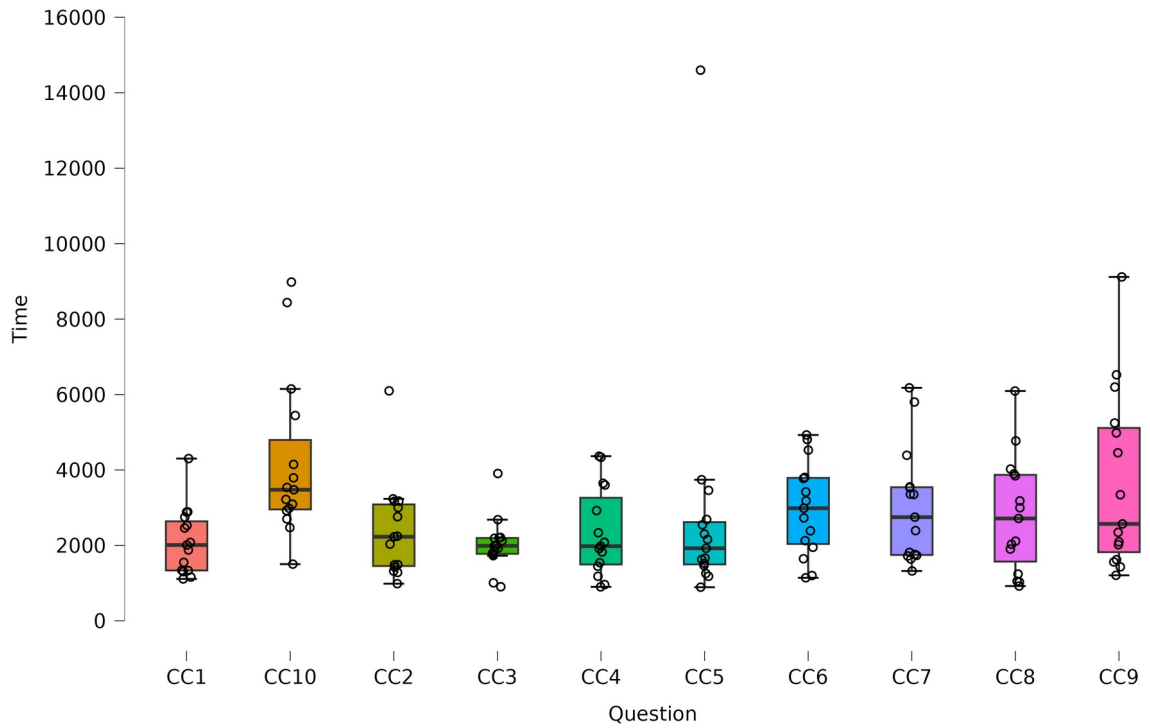
Appendix

A. Materials

NOTE: ALL THE VALUES IN GRAPHS AND TABLES ARE IN ms

	Mean	Std. Deviation	Minimum	Maximum	25th percentile	50th percentile	75th percentile
CC1	2105.20	879.90	1111.00	4302.00	1336.00	2010.00	2638.50
CC10	4190.47	2156.25	1503.00	8981.00	2954.00	3475.00	4795.00
CC2	2395.27	1289.39	985.00	6097.00	1453.50	2229.00	3086.00
CC3	2018.40	687.51	903.00	3906.00	1777.00	1989.00	2196.50
CC4	2335.93	1168.35	901.00	4365.00	1494.50	1979.00	3262.00
CC5	2868.47	3345.49	892.00	14599.00	1496.50	1923.00	2617.50
CC6	2973.60	1242.01	1141.00	4928.00	2038.00	2985.00	3789.00
CC7	3019.20	1511.62	1322.00	6177.00	1747.00	2749.00	3542.00
CC8	2787.00	1529.12	922.00	6093.00	1571.00	2714.00	3869.50
CC9	3647.60	2345.34	1207.00	9117.00	1820.50	2570.00	5113.50

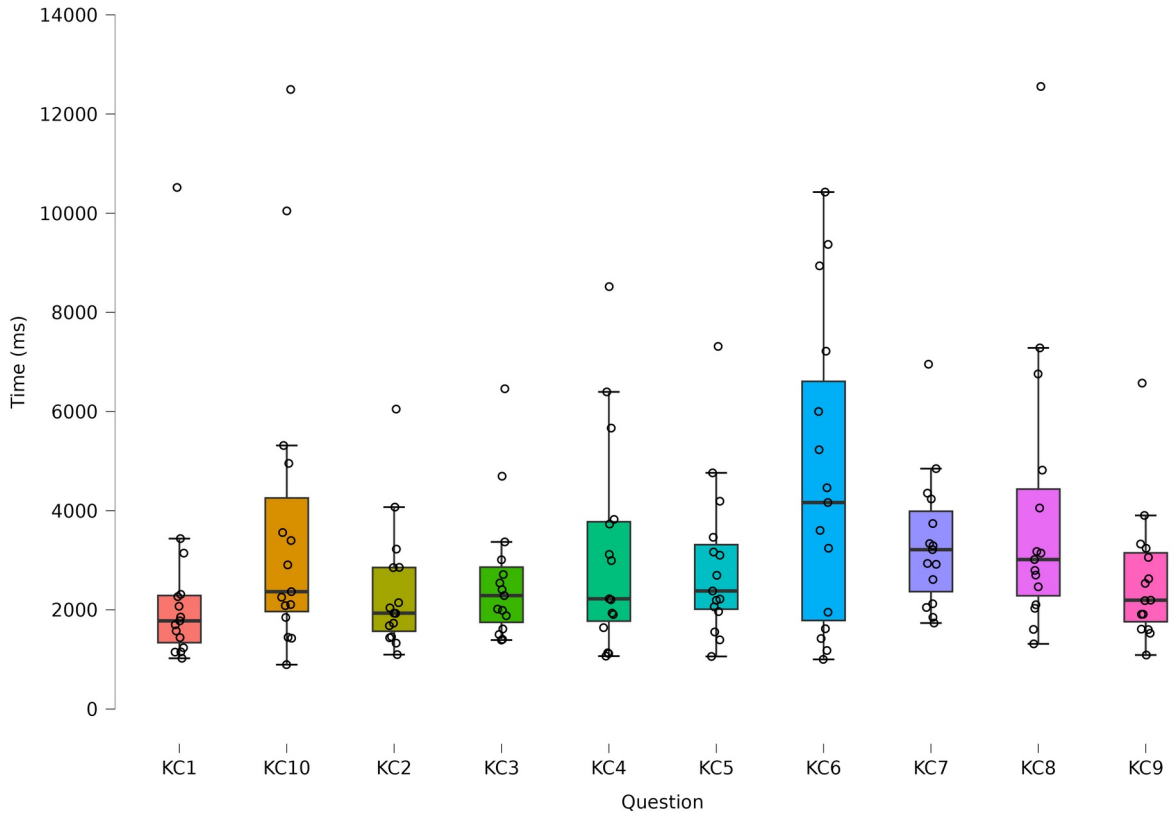
Statistics table per questions in Camel Case (all in ms)



Distribution time-question for Camel Case (ms)

	Mean	Std. Deviation	Minimum	Maximum	25th percentile	50th percentile	75th percentile
KC1	2443.93	2342.15	1023.00	10519.00	1338.50	1778.00	2289.00
KC10	3806.93	3305.06	895.00	12495.00	1966.00	2367.00	4256.50
KC2	2388.93	1294.60	1097.00	6049.00	1568.50	1934.00	2854.50
KC3	2619.07	1375.27	1391.00	6459.00	1748.50	2288.00	2861.50
KC4	3164.87	2172.81	1067.00	8519.00	1773.50	2223.00	3777.00
KC5	2901.47	1589.16	1060.00	7313.00	2013.50	2381.00	3314.50
KC6	4655.07	3143.76	1001.00	10426.00	1785.50	4165.00	6608.50
KC7	3346.53	1367.15	1734.00	6954.00	2367.50	3214.00	3988.50
KC8	3988.07	2934.23	1314.00	12555.00	2284.00	3015.00	4436.00
KC9	2619.87	1345.00	1089.00	6573.00	1760.50	2195.00	3149.50

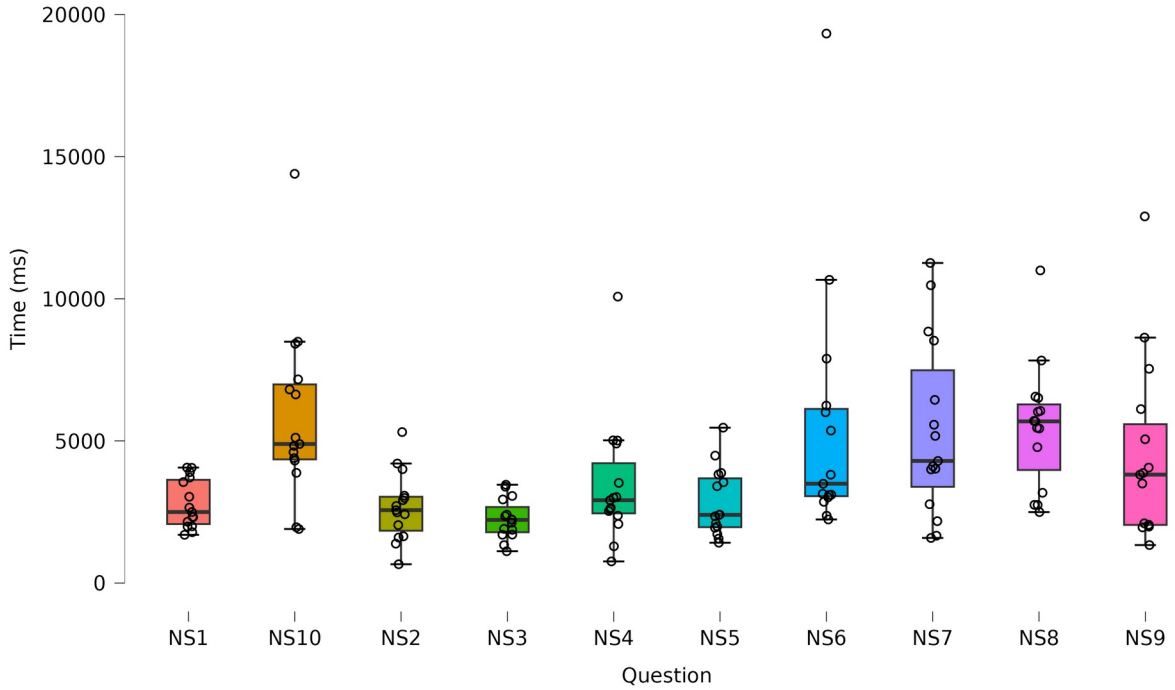
Statistics table per question in Kebab Case (ms)



Distribution time-question for Kebab Case (ms)

	Mean	Std. Deviation	Minimum	Maximum	25th percentile	50th percentile	75th percentile
NS1	2783.13	859.83	1697.00	4062.00	2075.50	2499.00	3631.50
NS10	5853.33	3082.60	1902.00	14397.00	4351.50	4892.00	6989.00
NS2	2668.53	1190.27	662.00	5312.00	1842.00	2565.00	3035.00
NS3	2255.93	703.20	1120.00	3459.00	1788.50	2222.00	2674.00
NS4	3451.20	2216.28	762.00	10077.00	2452.50	2914.00	4216.00
NS5	2832.73	1196.63	1419.00	5467.00	1966.50	2398.00	3683.50
NS6	5509.60	4488.68	2237.00	19331.00	3058.00	3495.00	6125.50
NS7	5395.67	3110.94	1587.00	11261.00	3386.00	4296.00	7486.50
NS8	5481.67	2212.53	2497.00	10999.00	3978.00	5688.00	6284.50
NS9	4465.00	3189.50	1337.00	12899.00	2046.00	3815.00	5589.00

Statistics table per question with no separator (ms)



Distribution time-question for no separator (ms)

B. Reproduction Package (or: Raw Data)

All the data and the, the graphics and the programs used are available in the following online Git-Hub repository: https://github.com/Ap0calypse2017/EEUSI inside the directory Assignment2.

The code used to create the experiment for is stored inside the Git-Hub repository: https://github.com/azanzi/case-survey (A connection between the two repository can be found inside the directory case-survey inside Assignment2)