TX 7007 - Final Report
Data Analysis upon collaborative support

Arthur Buisson
Barnabé Sellier

21 June 2023

1 Introduction

Our TX project centered around a product from the Digital Hall of the Innovation Center at the University of Technology of Compiègne (UTC). The **TATIN** (Table Tactiles INterractive) product, with a link to the website indicated here: *Tatin.fr*.

This project aims to improve the preliminary design phase of mechanical engineering projects by designing and developing **a collaborative platform**. The project utilizes interactive tables (one vertical and one horizontal interacting through gateways) with various functionalities. The project stands out for its **multi-user**, multi-device, multi-template, and innovative approach in utilizing interactive surfaces for professional use. The project involves members from UTC laboratories specializing in engineering design, computer science/human-computer interaction, and cognitive sciences. The support is presented in the following form, visible in Figure 1:

Figure 1: Support **Tatin Table**

To gather data prior to our project, several sessions that took place in 2019 and 2018 around this support were observed and analyzed. We thus had access to different types of data (more or less complete):

- **Log data**, specifically the numerical data associated with the TATIN project tables.
- **Manual coding data**, which are manually entered data characterizing information every 25 minutes of a session about a student.
- Data related to interactions between **People & People**, **People & Objects**, **People & Poses**.
- Videos of students interactions during a sessions. It helped us to understand better our datas.

During our analysis, we were supervised by **Mr. Thierry GIDEL** and **Mr. Claude Moulin**, whom we warmly thank for their invaluable help throughout the project. We started the project in February 2023 as part of TX project number 7007 for a duration of 5 months.

2 Log Analysis

2.1 Data architecture

2.1.1 Sessions

To better understand the data comparison, we need to provide a brief description of the following:

- **Session 2**: Session 2 of the Kit for the A19 semester. For this semester, the students attending sessions 2 to 5 are the same.
- **Session 3**: Session 3 of the Kit for the A19 semester.
- **Session 4**: Session 4 of the Kit for the A19 semester.
- **Session 5**: Session 5 of the Kit for the A19 semester.
- **Session 6**: Session 1 of the Kit for the A18 semester.
- **Session 7**: Session 2 of the Kit for the A18 semester.
- **Session 8**: Only session of the Kit for the P19 semester.

2.1.2 Data attributes

Our dataset is organized as we mentioned in the introduction, based on multiple sessions divided by semesters at UTC. In each log table (where each row corresponds to an interaction to an object created on the table) of a session, we find the following columns, explained below:

- *datetime*: Timestamp in the format `M/d/yy HH:mm:ss` when the action occurred. We did not focus on this attribute as the column attribute was simpler to use. However, on a larger dataset (which may exist someday), it could be more relevant.
- *seconds*: A simpler time scale starting from when the machine was powered on.
- *sender_user*: Attribute containing the name of the machine operator. We did not use it in our analysis.
- *sender_agent*: Attribute characterizing the device on which an action was recorded. Some data cleaning was required, but in the end, we identified three types of devices: the vertical board, the horizontal table, and users' smartphones (used much less frequently than the other two).
- *sender_id*: Similar to sender_user, not very useful for our analysis.
- *id*: Identifier of the object recorded by the user. This attribute is very useful for tracing the lifecycle of an object and the different functions it may have performed.
- *documents*: Secondary attribute of little interest for the project.
- *portail*: Identifier of the portal recorded if it was logged during a session. This allowed us to identify the movement of an object from one screen to another.
- *container*: Identifier of a container, an object that can contain multiple sub-objects (post-its, links, etc.).
- *action*: Indicates the type of action associated with the log. There are three types of actions:
  - `Create`: Creating an object.
  - `Update`: Modifying an object, which accounts for the majority of logs (80%).
  - `Delete`: Deleting an object, unfortunately very rarely encountered in the logs (>1%).
- *model*: Various sub-actions that further characterize the *action*. For example, we have elements such as (`Transform`, `Color`, `Drawing`, etc.).

- *item_type*: Type of object associated with the log (mostly a post-it, but sometimes an item with a layer or an image).
- x: Object's coordinates on the table/board along the length.
- y: Object's coordinates on the table/board along the width.
- *rotation*: Object's rotation in degrees on the scale (-180°, 180°).
- *scale*: Object's size measured as the ratio of its size to a newly created object.
- *text*: Text contained in the object, if any.
- *value*: Object's color, stored differently (in `JSON` or `HTML` code depending on the sessions).
- *session* : number of the session associate to the row.

2.2 Data Analysis

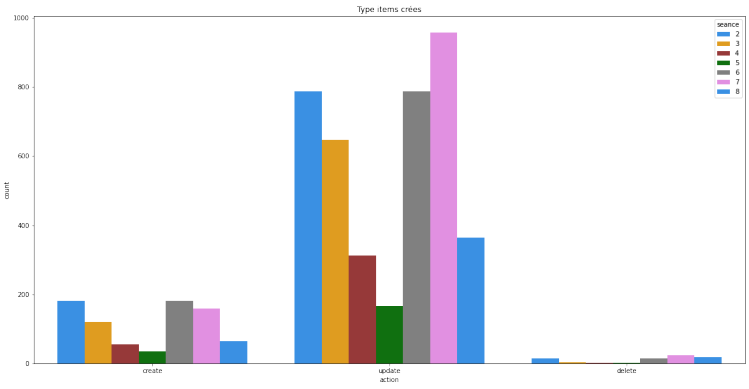
2.2.1 Cleaning Data

We first need to clean our data, as expected. To mention the few changes made (likely fewer than for the other two datasets), we corrected values in the following columns:

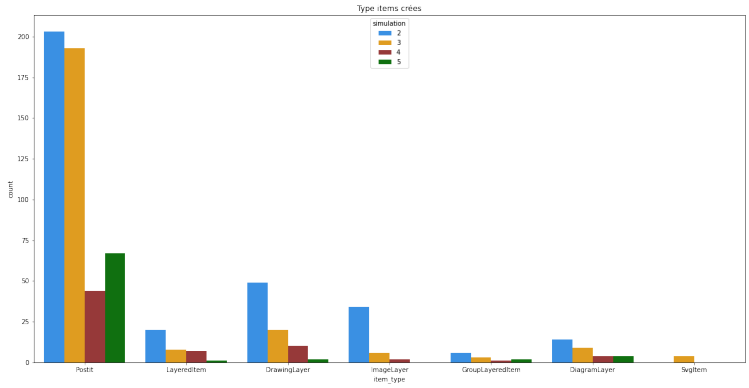
- *item_type*: Indeed, we had to make a choice when two types of items were provided in the column. For example: LayeredItem:DrawingLayer. Therefore, we decided to keep only one type, in this case, DrawingLayer.
- *sender_agent*: The room identifiers were recorded in the attributes (e.g., UbikeyOfficeHorizontal/1.4.4 (HALLNUMTBL05)). Since we assume that the room does not have an impact on the session, we cleaned the data to keep only the following attributes: *Horizontal* & *Vertical*.

2.2.2 Descriptive analysis

To better understand our dataset, we started by visualizing simple data, differentiating them by session. For example, we wanted to determine the types of actions (see Figure 2a) and the most prevalent types of objects (see Figure 2b) in the database, with a distribution by session as shown below:

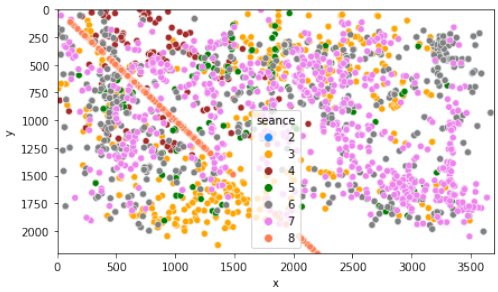


(a) Distribution of the number of actions per session

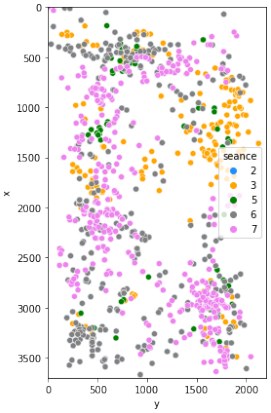


(b) Distribution of the number of items per session

We can already observe that the logs are unevenly distributed across sessions (for example, session 2 contains more logs than session 4). Furthermore, the majority of actions are of type *Update*, and the majority of items are *post-its* in the dataset. Finally, we were also eager to observe the distribution of logs on the horizontal table and the vertical board (which we divided based on...), so a preliminary representation is shown below (4b & 4a):



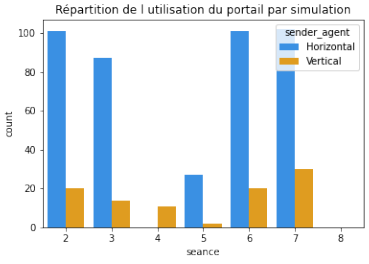
(a) Distribution of the actions for the vertical table



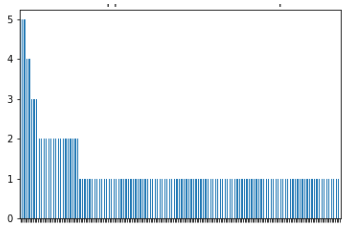
(b) Distribution of the actions for the horizontal table

Finally, we can deduce very little from these data, except that most areas of both tables are used (except perhaps the center of the horizontal table, which is understandable given its difficulty of access). We also notice that we sometimes lack information, such as the horizontal logs for sessions 4 and 8.

Furthermore, as we also want to study the interactions between the tables, we wanted to represent the data related to the portals as shown below:



(a) Distribution of the portal activitions by tables



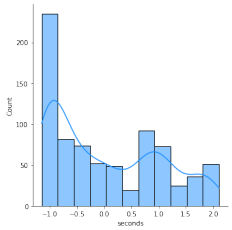
(b) Distribution of the number of travels by unique items

As expected, there are more journeys from the horizontal table to the vertical board than vice versa. As Mr. Moulins and Mr. Gidel presented to us, the horizontal table is associated with **Creation**, while the vertical board is a place for **Collaboration**. Thus, we have here evidence that individual ideas often written on the horizontal support are often presented to the entire group on the vertical board. Furthermore, as we can observe in Figure 4a, objects mostly make a single journey. This is consistent with Figure 4b, showing the movement from the horizontal support to the vertical support.

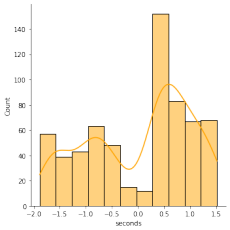
2.3 In-depth Analysis

2.3.1 Time analysis

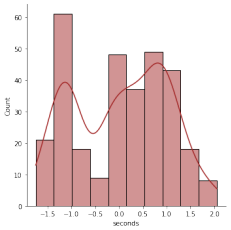
Now that we have identified the key elements from our data table for the next stage of our study, we attempted to compare them in relation to the time variable. To do this, we treated the *seconds* column for each session as a normal distribution, which allowed us to compare the interactions across sessions. In the end, based on feedback from our project advisors, we realized that this method may not have been the most effective. However, it did enable us to highlight certain elements, as shown in the figures below:



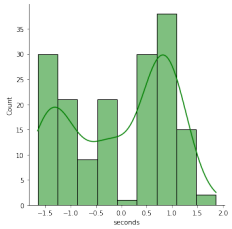
(a) Distribution of the interactions with the device for the session 2



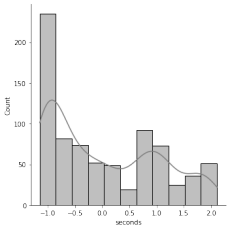
(b) Distribution of the interactions with the device for the session 3



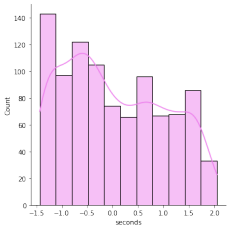
(c) Distribution of the interactions with the device for the session 4



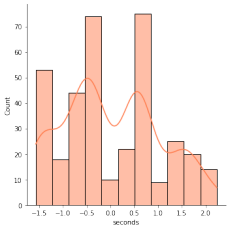
(d) Distribution of the interactions with the device for the session 5



(e) Distribution of the interactions with the device for the session 6



(f) Distribution of the interactions with the device for the session 7



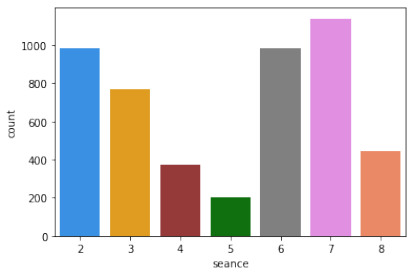
(g) Distribution of the interactions with the device for the session 8

From the preceding graph, we can deduce some information that will be crucial for the next stage of our analysis:

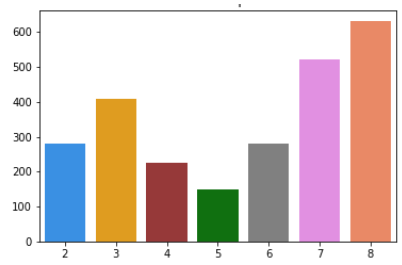
- We observe **2 peaks of interactions** with the machine in each session. In parallel, there is **a period of lower activity** between them.
- These two periods are generally present in the normal distribution on the interval [-1.5, -0.5] and [0.5, 1]. Scaling it back to a percentage scale, this corresponds to the intervals **[5%, 35%]** and **[70%, 90%]**.

2.3.2 Word Analysis

In parallel, we observe certain peaks in the figures (e.g., at the beginning of session 2, see Figure 5a), which lead us to believe that some data may be distorted. Indeed, it is possible that some post-its were created but did not play a significant role in the session (as the horizontal table can serve as a testing ground for the user to better familiarize themselves with the medium). We set out to find a way to **filter out** these erroneous values. For this purpose, we turned to the *text* column of the dataframe. If a post-it contained text at least once, it had at least some utility compared to an object that was used once and then destroyed. This filtering method proved fruitful and allowed us to obtain a more accurate representation of our data. Its effectiveness can be observed in the following figures 6a and 6b:



(a) Distribution of the number of items per session



(b) Distribution of the number of filtered items per session

Now, let's go deeper. Since we have successfully filtered our data based on the text and have a better understanding of this column, we can try to extract information from it. We decided to study the maximum length (i.e., the number of words) of an object, often a post-it. After extracting this data, we categorize our words into 3 categories:

- **Short**: objects containing only one word.
- **Medium**: objects containing between 1 and 5 words.
- **Long**: objects containing more than 5 words.

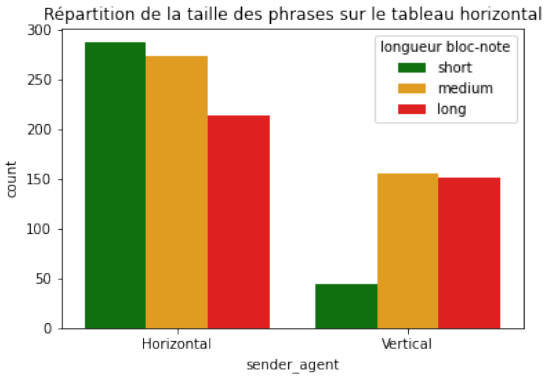


Figure 7: Distribution of the different types of words per tables

Subsequently, we studied the distribution of these object types across the two tables and obtained the image 7.

From the previous figure, we can deduce that the vertical table allows for the exposure of generally more consequential ideas (the proportion of *medium* and *long* words is much higher compared to the horizontal table). Conversely, on the horizontal table, words are generally shorter. We can even go further and speculate that the cause of this distribution may arise from the fact that multiple short objects have given rise to a longer object (thus highlighting the role of **collaboration** on the vertical table). We also reduced our dataset to the session 2 to 5 in order to aligny with the other data present in the other folders.

2.4 Development of the main idea

2.4.1 Exposition

We now have multiple elements that allow us to develop our main idea regarding the dataset associated with the TATIN table logs.

Let's recall our **Observation**: There are two privileged moments of interaction between users and the tablet (5% and 30% of the session, then 70% and 90% with a low activity period between 40% and 60%). We formulate a **preliminary hypothesis**: Each of these time periods is related to one of the tables (horizontal and then vertical). And we establish a **method** to validate our hypothesis:

- Generate a common time scale based on the percentage relative to the total duration of the session (which will be more relevant than our previous normalized scale).
- Measure various indicators:

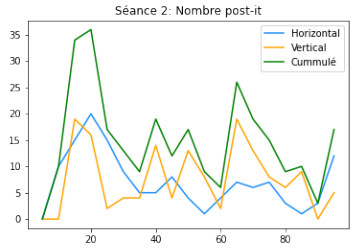
  - Item Index (number of objects created or modified during the interval).
  - Screen Index (distribution of these new interactions between the Vertical and Horizontal screens).
  - Word Index (average number of words present on an object).

- Observe the active regions on the tablet: position and average direction of movement.
- Consider interaction with the portals.

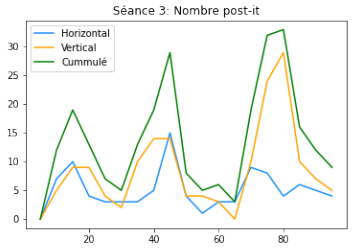
2.4.2 Indidcator n°1 - Item Index

The number of created or modified items varies throughout the session, and this interaction count is indicative of the usage of the platform. We used this data to confirm our previously stated hypothesis.

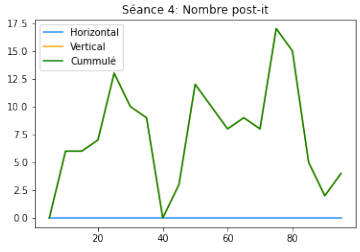
We plotted the number of interactions (creation or modifications) with each of the two tables, as well as the cumulative count, based on the previously established percentage scale. The following curves were obtained:



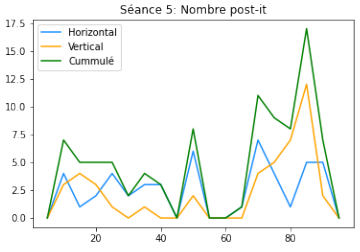
(a) Distribution of the item index in session 2



(b) Distribution of the item index in session 3



(c) Distribution of the item index in session 4



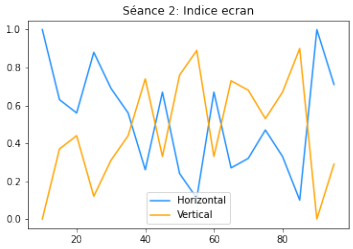
(d) Distribution of the item index in session 5

We can easily identify these two curves in sessions 2 and 3, which align with the previously mentioned periods. In the other two sessions, a third peak seems to appear, but it appears to be more of an exception than the norm, as

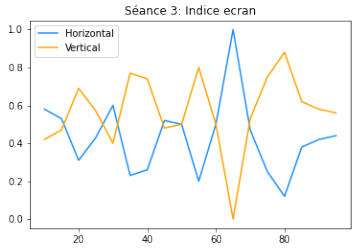
indicated by the curves 5e, 5f, 5g. This once again validates our hypothesis. If we examine these two moments more closely, they are visible in both the horizontal and vertical support curves. Lastly, what we previously referred to as the "crest" moment is an exposure period for the horizontal table, reaching peaks in the interval [40%, 60%] (visible in sessions 3 and 5).

2.4.3 Indicator n°2 - Screen index

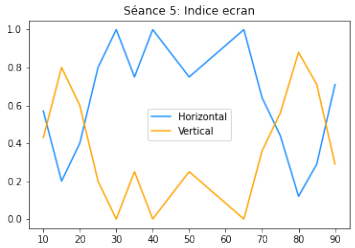
We also studied the distribution of portal usage (taking into account the direction of travel between the table and the board) based on the percentage scale for each session. We present our results below:



(a) Distribution of the screen index in session 2



(b) Distribution of the screen index in session 3

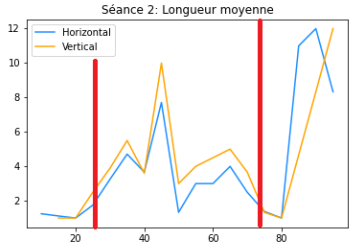


(c) Distribution of the screen index in session 5

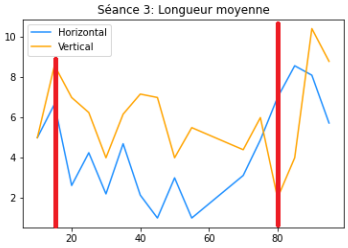
If we consider the two periods that we already mentionned, on the previous curves, we can easily observe that during the later period, the vertical board is more utilized. Thus, we conclude that at the end of the session, the vertical board is more frequently used, as expected, as students synthesize their work and likely present their results to their supervisors. However, we cannot draw a conclusion about the initial moment as the curves overlap and do not allow us to reach a conclusion.

2.4.4 Indicator n°3 - Words index

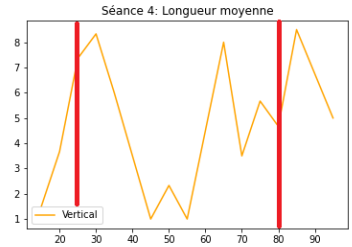
Finally, in order to analyze our sessions, we studied the word index, which represents the average word length on an object during a given period. We depict this index on the same scale as before, adding in red the moments where interaction peaks are reached:



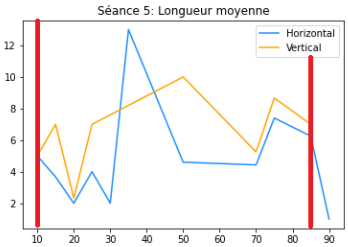
(a) Distribution of the word index in session 2



(b) Distribution of the word index in session 3



(c) Distribution of the word index in session 4



(d) Distribution of the word index in session 5

From these figures, we deduce two points:

- Firstly, the peaks in the word index are **brief**. This means that the word index rarely remains high, indicating that users of the TATIN table consistently tend to synthesize their thoughts and ideas.
- Additionally, we observe that **strong interaction periods are often followed by an increase in the word index**. This is quite consistent and reassuring: If users interact strongly with the platform, there is a high chance that they are more creative and present ideas, perhaps in more detail, which increases the average word count.

Conclusion In conclusion, our analysis focused on understanding the data related to sessions and data attributes in order to gain insights into the interactions and patterns within the dataset. We provided a description of each session, highlighting the semesters and sessions associated with the Kit. We also discussed the various data attributes present in the dataset, such as timestamps, sender information, object identifiers, actions, models, and more.

Moving on to the data analysis, we started with cleaning the data by addressing specific issues in columns like item type and sender agent. We then performed descriptive analysis to gain a better understanding of the dataset. We visualized the distribution of actions and item types across sessions. We also analyzed the distribution of actions on the horizontal table and vertical board.

For a more in-depth analysis, we explored the data in relation to time and conducted a time analysis by treating the seconds column as a normal distribution. This allowed us to compare interactions across sessions and identify two peaks of interactions with a period of lower activity between them. We also performed word analysis by filtering out objects without any text, resulting in a more accurate representation of the data. This filtering method helped us extract insights from the text column, such as the maximum length of an object.

Finally,we developped 3 index to assure our first thoughts. First, the Item Index confirmed our hypothesis with distinct peaks in interactions with each table during these periods. Then the Screen Index indicated a higher utilization of the vertical board in the later session period. To conclude, the Word Index showed brief peaks in word count, indicating users' tendency to synthesize ideas. Strong interaction periods were associated with increased word counts. These findings provide valuable insights for understanding user behavior and can inform design improvements for the TATIN table, enhancing the overall user experience.

Overall, our analysis provided valuable insights into the dataset, revealing patterns and trends in the interactions and attributes of the data. These findings can serve as a foundation for further exploration and understanding of the dataset, potentially leading to more meaningful conclusions and applications in the future.

3 Atomic Actions and Competency Analysis

3.1 What are Atomic Actions and Competency Analysis ?

In order to be able to find patterns in a meeting, atomic actions were saved into a file using a VOTT model. They represent little gesture, that are related to an interaction with another person or objects such as the TATIN table. They are divided in three categories : Person to person interaction, Person to object interaction and pose. Those categories have subdivided in *tag name* that represent the atomic action itself. Person to person: *talk to person*, *look at person*, *laugh*, *listen to*, *talk to group*, *point at person*

Person to object: *look at notebook or smartphone*, *look at table*, *write on the table*, *use portal*, *point at object*, *look at board*, *write on the board*, *read*, *look at slide*, *take notes*

Pose: *enter*, *sit down*, *hand gesture*, *stand up*, *walk*, *stand board*, *exit* and *move post-it*.

Competency Analysis are competencies manually evaluated during a certain time span. There are two kind of competency analysis. The first one is evaluated every 25 minutes and is divided in five skills each subdivided in three skill types, which are: Regulation: *engagement*, *responsibility* and *coordination and work evaluation* Communication: *communication*, *listening* and *reactions/feedback* Team Work: *teamwork*, *group objective* and *social* Social Intelligence: *interpersonal conflict*, *emotional needs of colleagues* and *speaking time share* Constructive Conflict: *keep shared notions*, *argumentation*, *opening*

The other kind of competency analysis is divided in five categories: Collaboration, Cooperation, Coordination, Communication, and Individual Work and are measured every 30 seconds.

3.2 Data Architecture

Initially, there was a different file for each atomic action type, in an Excel format. To make the data more easily readable we combined those files into one in csv format, including every session. Each line of this csv is represented by 11 variables which are:

- *timestamp*: The moment in second where the atomic action was detected
- *seance*: The session where it was detected
- *atomic_action_type*: Type of the atomic action ie personperson, personobject or pose.
- *root_asset_id*: Variable associated to the VOTT model, it represents the id of the video of the atomic action

- *asset_id*: Is unique for every second of the videos, if two rows share a same asset_id it means they happened at the same time
- *region_id*: Id of the region where the atomic action was detected
- *tag_name*: Name of the atomic action
- *bounding_box_top*: The top y coordinate where the atomic action took place in the video
- *bounding_box_left*: The left x coordinate where the atomic action took place in the video
- *bounding_box_width*: The width of the box where the atomic action took place in the video
- *bounding_box_height*: The height of the box where the atomic action took place in the video

Competency analysis also required some adjustments and was converted in csv. At the end we obtain this architecture:

- *skill_type*: the subcategory of a skill
- *skill*: one of the five skills described above
- *student*: the student the grade is attributed to
- *note*: grade between -2 and 2
- *seance*: session in which the skill was observed
- *timespan*: multiple of 25, period in which the skill was observed

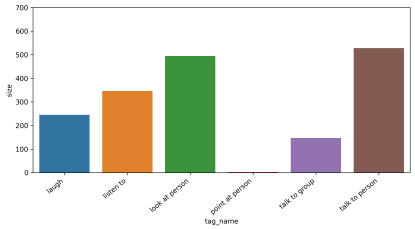
For the second kind of competency analysis, we have a row each time an observation was made (every 30 seconds), so the columns are *time*, *Collaboration*, *Coordination*, *Communication*, *Individual Work* and *session*. There also other columns such as *Teacher Intervention* or *Technical Issue* that I didn't study here.

3.3 Visualization of the Data

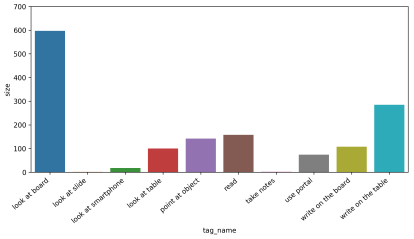
3.3.1 Atomic Actions

In order to get familiar with the data and to have a better understanding of them, we did an exploratory analysis that consist of making some simple visualization. On the figures above, you can see the distribution of atomic actions, and which ones are the most present among sessions.

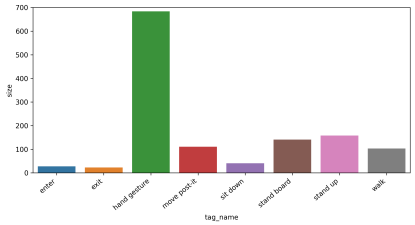
We can see that the person to person atomic actions are the most present, with a total of 1763, and almost all of the tag name are well represented. In the other categories, only one is very present: *look at board* for person to object and *hand gesture* for posture. What we can guess from theses graph is that during sessions the most present kind of interaction is person to person interaction which makes sense as the goal of a meeting is to communicate.



(a) Distribution of person to person atomic actions



(b) Distribution of person to object atomic actions



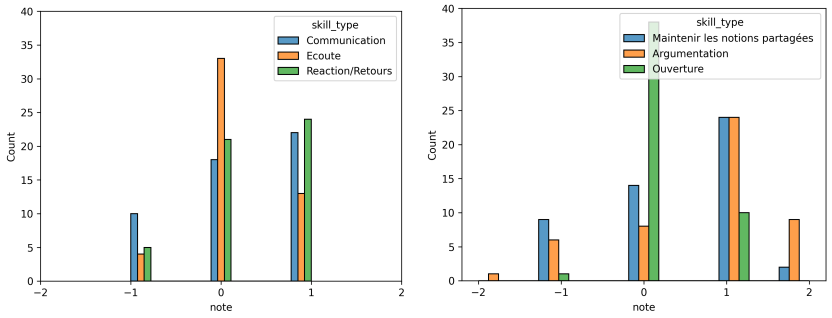
(c) Distribution of pose atomic actions

3.3.2 Competency Analysis

After having analyzed atomic actions I decided to make a first visualization of competency analysis too as the two seemed linked together. Competency analysis represent the score of competency in five categories: Communication, Regulation, Team Work, Social Intelligence and Constructive Conflict. If an action or a situation shows a participant having one of the categories, the scores is incremented by one. On the figure below you can see the overall score in each category.

The competency that the best score in total is Constructive Conflict with a total score of 62, then Communication with a score of 40.

As I explained before, in the manual coding folder we also have the competency analysis of Collaboration, Cooperation, Coordination, Communication and Individual Work. I did a visualization for these competencies too, with the idea to encounter the results found by in a thesis made on the TATIN table and the meeting. This result states that Individual Work leads to Communication, then leads Coordination which leads Cooperation and then to Collaboration. To do so I regrouped the score of those competencies by duration of 10 minutes, and the goal is to see if the score of each competency follow the scheme indicated by the thesis. I split the results by session so the results would not be mixed. Figure 13 shows the different scores from session 4 and as we can see, we do not get the expected result. The first problem is that Communication and Individual Work are always the highest. We can argue that Communication is present in the others ones and that Individual Work could also be in Collaboration. And in fact we can vaguely see the pattern if we do not include Communication and Individual Work, Coordination has the first highest score,



(a) Distribution of Communication score

(b) Distribution of Constructive Conflict

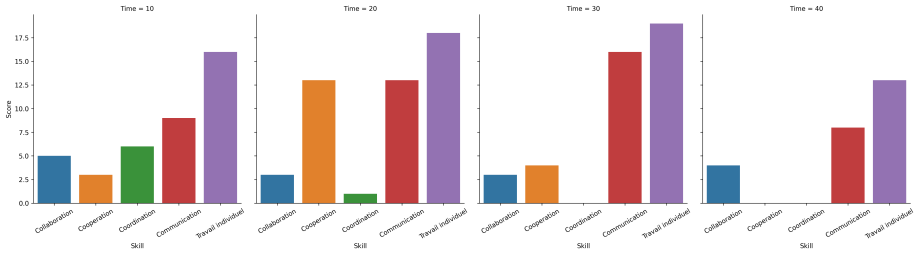


Figure 13: Distribution of Collaboration, Cooperation, Coordination, Communication and Individual Work during session 4

then it's Cooperation and at the end Collaboration. But the session 4 is the only one with this kind of result, and it's not enough to confirm the hypothesis.

As I could not find the expected result with Competency Analysis, I chose to focus on Atomic Actions and see if I could get something interesting. To do this, I chose to regroup some atomic actions together and form the categories. As we don't have contextual information about an atomic action, it was not possible to put them in the categories Collaboration, Cooperation or Coordination. So I split them into Individual Work and Communication, plus Neutral for the ones that did not go in neither of them.

- Communication: *listen to, look at person, use portal, write on the board, point at person, hand gesture, point at object, stand board, talk to group, talk to person*
- Individual Work: *take notes, read, write on the table, look at notebook or smartphone, move post-it, look at table*
- Neutral: *enter, exit, laugh, look at board, sit down, stand up, walk, look at slide*

The idea is that Communication would include Collaboration, Cooperation and Coordination, so hat after a peak of Individual Work we should always see

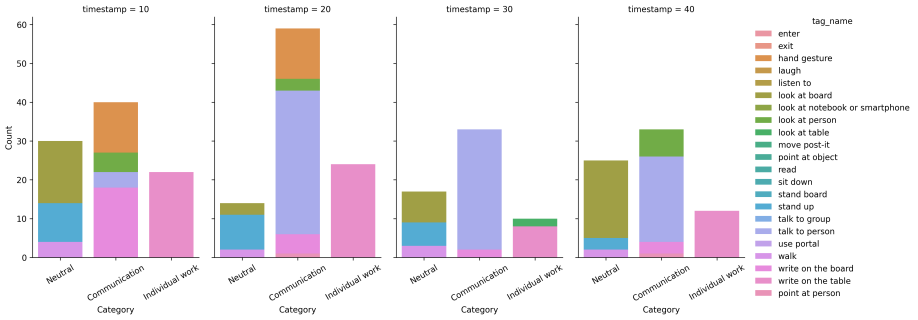


Figure 14: Neutral, Communication and Individual Work atomic actions in time for session 4

a peak of Communication. The result for session 4 is shown in Figure 14. We do not get the result expected, Communication is always the most present and we can not identify a period which was dedicated to Individual Work.

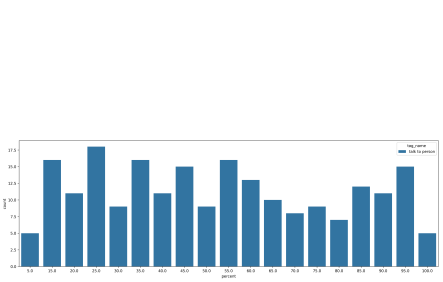
The conclusion to this section is that with the only data we have, it's difficult to encounter the result of the thesis. The problem can also come from my categorization that was not necessary perfect and arbitrary for some of the atomic actions such as *laugh* or *look at board*. Maybe a combination of atomic actions and the competencies - other than Collaboration, Cooperation, Coordination, Communication and Individual Work - could have been more interesting as we would have had more data to study.

3.4 Analysis of Atomic Actions and Log

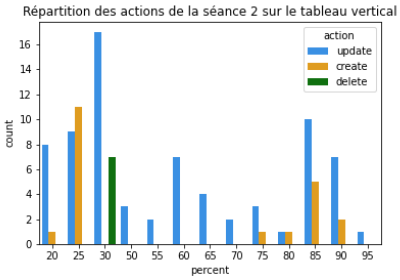
After trying to encounter already proved pattern, I focused on trying to find new ones. To do so I first realized videos that showed the apparition of atomic actions during time to see if recurrent patterns where appearing. There is a videos for each session and they are accessible from Nextcloud. The idea was to see if a recurrent succession of atomic actions was showing up. But with the high number of atomic actions it was hard to see a pattern stand out. Therefore, I changed my approach with the idea of identifying pattern that would have an impact on the use of the TATIN table.

I compared each atomic actions and the log of creation or modification of an element in time for each. I noticed that most of the time the atomic action *talk to person* seemed to precede a creation or an update in the logs. We can see that applied on session 2 on Figure 15a and Figure 15b where a peak of *talk to person* action seems to precede the peak of creation and update.

To have a better visualization of this phenomenon, I crossed the curves of *talk to person* and *creation or modification* for each session. I should also mention that the logs were split between the TATIN table (horizontal) and the TATIN board (vertical). I made this visualization for each session and for both the horizontal and vertical table. They seemed to confirm my hypothesis especially



(a) Distribution of *talk to person* atomic action in time during session 2



(b) Distribution of creation and modification on TATIN board during session 2

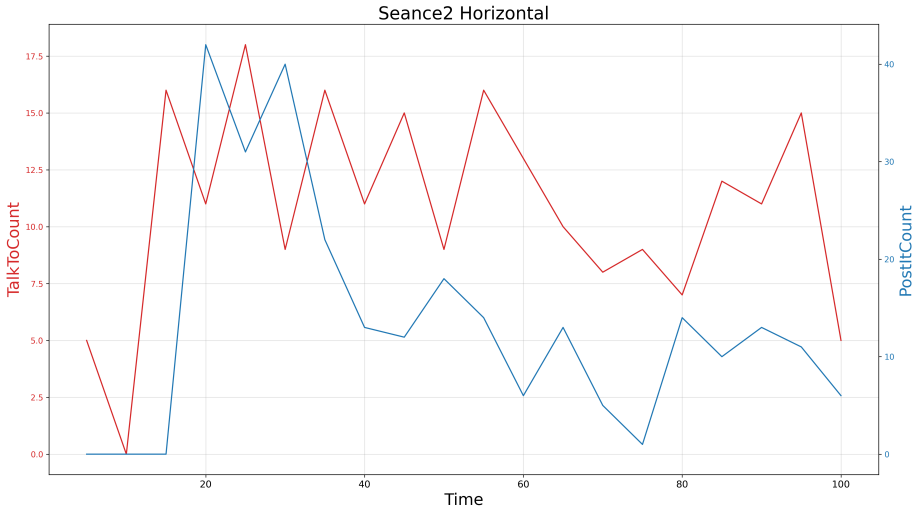


Figure 16: *talk to person* and *creation/modification* in time

for the horizontal table. You can see on Figure 16, that represents session 2, that the peak nearly always cross each other and *talk to person* peaks precede *creation or modification* peaks.

All those visualizations made so far confirm my hypothesis; but are not enough to prove it. To go further in the proof of my hypothesis I used the Granger Causality Test, which is a statistical method used to determine whether one time series can be considered a cause of another time series. This test compute a p-value that reject the null hypothesis X do not cause Y if it is lesser than the significance level, which is 0.05 here. The test compute an array 4 p-values (X causes X, X causes Y, Y causes X and Y causes Y), but we are only interested in one which is *talk to person* causes *creation or modification*. As for the previous visualization, I applied this test on each session (except session 4 because of missing logs) and each kind of table (horizontal and vertical). I said

	TalkToCount_x	PostItCount_Horizontal_x
TalkToCount_y	1.0000	0.0
PostItCount_Horizontal_y	0.0006	1.0

(a) Granger Test on the Horizontal Table for session 2

	TalkToCount_x	PostItCount_Horizontal_x
TalkToCount_y	1.0000	0.2203
PostItCount_Horizontal_y	0.0272	1.0000

(b) Granger Test on the Horizontal Table for session 3

	TalkToCount_x	PostItCount_Horizontal_x
TalkToCount_y	1.000	0.0005
PostItCount_Horizontal_y	0.035	1.0000

(c) Granger Test on the Horizontal Table for session 5

before that my hypothesis seemed more suitable to the horizontal table and the tests confirm it. Indeed the p-value is always lesser than 0.05 for the null hypothesis *talk to person* causes *creation or modification on the horizontal table*. I also wanted to see if creation or modification on the TATIN board (vertical) causes discussion, which would have been an interesting and an intuitive result but one of the p-value is higher than 0.05 which prevents me from confirm it.

In conclusion, thanks to several visualizations and to the Granger causality test I was able to find the pattern that discussion among people causes creation or modification of elements on the horizontal TATIN table. This result is very interesting and could be maybe extended to others atomic actions in order to find others patterns of causality between them and the use of the TATIN table.

4 Conclusion

In conclusion, our data analysis of the TATIN table logs has provided valuable insights into the usage patterns and interactions within the collaborative platform.

Through descriptive analysis, we gained an overview of the distribution of actions and objects across sessions, highlighting the prevalence of update actions and post-it objects. We also observed the utilization of both the horizontal and vertical surfaces, with the vertical board being a hub for collaboration and idea presentation. Furthermore, the analysis of portal activities revealed a predominant movement from the horizontal table to the vertical board, supporting the notion of individual ideas evolving into group discussions. Moving into in-depth analysis, we delved deeper into time-related aspects, investigating interaction peaks and periods of lower activity. Our examination of time distributions showcased distinct interaction patterns, with two prominent peaks observed in each session and interludes of reduced activity. This finding reinforced our hypothesis that these time periods corresponded to specific tables, namely the horizontal and vertical surfaces. Additionally, our word analysis unveiled the disparity in object lengths, indicating that the vertical board often facilitated more elaborate and consequential ideas compared to the horizontal table.

Moreover, by exploring atomic actions, such as item creation and modification, we confirmed the existence of the identified interaction periods. The item index curves consistently aligned with the previously observed peaks, providing further validation to our hypothesis. These insights contribute to a better understanding of how users engage with the collaborative platform and the dynamics of their interactions. Lastly, competency analysis empowered us to evaluate users' proficiency in utilizing the TATIN table. By examining the distribution of actions across sessions and analyzing the average number of words per object, we gained valuable insights into the users' competency levels and their ability to effectively utilize the platform's features. This analysis highlighted the importance of collaboration and knowledge sharing within the context of mechanical engineering projects.

In summary, our comprehensive analysis of the TATIN table logs has shed light on the usage patterns, interaction dynamics, and competency levels within the collaborative platform. The findings contribute to enhancing the preliminary design phase of mechanical engineering projects by providing insights into how teams can effectively leverage the TATIN table's capabilities. This study lays the foundation for further research and development in the field of interactive surfaces for professional use, opening up new possibilities for collaborative engineering design.