"It's a grid, it's a form, it's a spreadsheet!": Revisiting Sensemaking with Spreadsheets for Presentation and Data Analysis

Pingjing Yang
University of Illinois (UIUC)
py2@illinois.edu

Sajjadur Rahman
University of Illinois (UIUC)
srahman7@illinois.edu

Mangesh Bendre
University of Illinois (UIUC)
bendre1@illinois.edu

Joon Sung Park
University of Illinois (UIUC)
jp19@illinois.edu

Kevin Chen-Chuan Chang
University of Illinois (UIUC)
kcchang@illinois.edu

Karrie Karahalios
University of Illinois (UIUC)
kkarahal@illinois.edu

Aditya Parameswaran
University of Illinois (UIUC)
adityagp@illinois.edu

ABSTRACT

Spreadsheets are one of the most popular tools for ad-hoc management, exploration, and analysis of data. Even so, it is not clear whether the tabular layout of spreadsheets is a natural fit for sensemaking. To understand (i) the association between tabular representation of data within spreadsheets and participants' sensemaking process, and (ii) the challenges faced in the participants' spreadsheet task workflows, we conducted a study with 32 participants having three different levels of expertise—experienced, intermediate, and inexperienced. We studied the impact of the participants' expertise on the associations and the workflows through sketching and computational tasks, respectively. Our study revealed that about 30% of the participants opted for non-tabular representation, all of whom were either intermediate or inexperienced. However, after providing the contextual information about the data, approximately 60% of the participants altered their representations from non-tabular to tabular. In the data analysis tasks, the participants faced three major challenges: information overload, loss of context, and incomprehensible documentation. Based on our study, we propose to augment existing interfaces by introducing overviews, and context within spreadsheets while providing automated suggestions to help participants in finding appropriate workflows for computation tasks.

KEYWORDS

Spreadsheets; Sensemaking; Data Representation; Workflow.

1 INTRODUCTION

Spreadsheet tools are commonly used for ad-hoc management, exploration, and analysis of data by inexperienced and experienced users alike. Microsoft Excel, a popular spreadsheet tool, boasts a user base of over 750 million [2], about 10% of the world's population. With uses ranging from managing medical records to preparing financial reports, the popularity of spreadsheets is indisputable. Nardi and Miller [18] attributed the success of spreadsheets to two factors: presentation and computation. Spreadsheets *present* data in a tabular layout consisting of a collection of cells referenced by two dimensions: row and column. Additionally, formula-based *computation* allows users to perform complex tasks without low-level programming knowledge.

Despite the huge popularity of spreadsheets, performing management, exploration, and analysis tasks using this tabular layout is sometimes challenging [7]. These challenges include locating features within the spreadsheet interface and finding required functions to solve a task [7]. This earlier work exploring spreadsheets focused on everyday use and practice. Given the importance of presentation, we began by exploring the efficacy of the data representation with respect to the existing tabular layout from a sensemaking perspective."Sensemaking is the process of searching for a representation and encoding data in that representation to answer task-specific questions." [24] This often involves the exploration of interrelated activities that involve collecting, organizing and creating representations of complex information sets, and seeing how this leads to understanding, sometimes with ill-structured problems. In the context of spreasheets, it is not clear whether the tabular layout adopted by spreadsheets mirrors how people naturally represent or make sense of data or facilitates the data computation.

Furthermore, we explore workflows of spreadsheet users' typical "recipes" for performing computation, given their tabular model. Successful execution of computational tasks requires identifying the correct workflows, i.e., the sequence of steps for solving a task [14]. For example, Hendry and

Green [14] uncovered the *sum blocks problem*in spreadsheets. In this case, participant expectations for what would happen using the auto-fill dragging feature to add cells within specific blocks did not align with the actual outcome. They identified an inaccuracy–the use of the the auto-fill feature instead of explicit formulas for each block. The focus of Hendry and Green's work was on the debugging process. We extend this group to understand common workflow patterns for basic spreadsheet tasks and to understand how these workflows differ across participants' varying levels of expertise.

It is likely that people's levels of expertise impact how they represent data, how they think about data and how they plan and execute computational tasks. By understanding how users' expertise correlates with their mental model of data, designers will be able to build tools that are usable by a wide spectrum of users with different degree of experience with spreadsheets.

To further our goals, we conducted an in-lab interview study to understand (1) people's perception of data representations and how they adapt to tabular layouts, and (2) their choice of task workflows and the challenges they face while accomplishing tasks. More specifically, we focused on finding answers to the following research questions:

RQ1 (a) How do people represent previously unseen data (that they can relate to and understand)? (b) How do people's representations change after they learning why a dataset was collected?

RQ2 Can people's initial data representations be accommodated by the tabular representation of spreadsheets?

RQ3 In a tabular layout: (a) How do people plan and accomplish computational tasks? (b) What are the challenges people face when completing specific computational tasks?

RQ4 Do the answers to *RQ1*, *RQ2*, and *RQ3* vary with people's level of expertise in using spreadsheets?

The study consisted of two phases: (1) sketching tasks to understand data representations of a dataset with and without knowing the goal of data collection, and (2) performing computational tasks. While the sketching phase explored *RQ1*, and *RQ2*, the computational tasks phase explored *RQ3*. To address *RQ4*, we recruited participants that fell into three different levels of expertise: experienced, intermediate, and inexperienced (see Section 3). Our study revealed that people's perceived data representation varied with their level of expertise and, for some participants, deviated from the tabular layout of spreadsheets—approximately 30% of the participants, none of whom were experienced, opted for non-tabular representations. However, after understanding the purpose of the data collection, approximately 60% of the participants altered their representations to a tabular layout. Analysis of the second phase of the study revealed three significant challenges participants faced in performing tasks: information overload, lack of context, and incomprehensible documentation. Looking forward, to address these three challenges, we propose (1) the inclusion of a visual overview and contextual information of and about the data and (2) automated mechanisms for suggesting task workflows.

2 RELATED WORK

Existing spreadsheet research explores different aspects of users' experience with spreadsheets from their mental models for sensemaking to the challenges spreadsheet users encounter. In this section, we discuss the related work in more detail.

Mental Models for Computational Tasks

One approach to sensemaking involves exploring the the generation of people's mental models as a representation of the world being studied [6]. A mental model is constructed when an individual forms a preliminary understanding of how something works based on their past experiences. This approach is commonly used in the fields of cognitvie science and developmental psychology [27]. These mental models, however, can be difficult for the users to express verbally due to the unfamiliarity of terminology or background knowledge. Using sketching approaches to elicit mental models, Poole et al. [23] were able to identify levels of user expertise and their understanding of household networks. Friedman et al. found that people across diverse communities incorrectly understood internet connection security [12]. Similarly, Eslami et al. found that people were not aware of algorithmic curation of news feeds in social media through a mental model elicitation task [10].

Sketching reveals the internal model which affects an individual's impression of and interaction with the external world. Tohidi et al. [26] further believes that this method is an efficient and effective complement to other usability tests such as think-aloud techniques.

In the context of spreadsheet mental models, Kankuzi et al. [15] found that participants mainly used real-world and problem domain concepts (e.g., a car) when explaining their spreadsheets. Hendry et al. [14] interviewed ten ordinary users and found even simple spreadsheet formulas were hard to create and understand without excessive knowledge of the data itself. Barry et al. [16] found significant mental model differences between experienced and inexperienced spreadsheet using a survey methodology but did not provide reasoning from the participants. How spreadsheets actually *support* the sensemaking of data has not been addressed in the existing literature. The present study begins to address this goal.

Cognitive Analysis for Computational Tasks

Cognitive analysis aims at understanding complex tasks that require significant cognitive activity from users. Russell et al. [25] introduced a notional model of an analyst's sensemaking which includes two main loops: a foraging loop and a sensemaking loop. In the foraging loop, people work on collecting and evaluating necessary data for analysis. The person's subsequent sensemaking process attempts to find the right representation to complete their task. Our sensemaking mental model study design follows this cognitive model.

Spreadsheets have long been used for computational tasks [18]. Recent work showed that spreadsheets are used by researchers to perform meta-analyses [19]. For specific domains such as medicine, spreadsheets are widely used for computation. Peakall et al. [22] introduced a package for Microsoft (MS) Excel for genetic analysis. Zhang et al. [28] developed PK-Solver for pharmacokinetic and pharmacodynamic analysis in MS Excel. While executing computational tasks on spreadsheets has been common practice in many domains, applying cognitive analysis for computational tasks in the context of spreadsheets is less explored.

People's Spreadsheet Challenges

Norman [20] investigated the first electronic spreadsheet VisiCalc and attributed its popularity to its ability to flexibly organize terms and figures. Nardi et al. [18] stated that a spreadsheet's visual formalism is the reason for its success. Despite this success, researchers have also reported challenges people face with spreadsheets. Mack et al. [17] characterized scalability issues spreadsheet users faced through online forum mining. Panko et al. [21] pointed out that people overestimated their spreadsheet quality. The present study elicits challenges in using spreadsheets to represent and analyze data.

3 STUDY DESIGN

To answer our research questions, we designed a two-phase laboratory interview study. While the first phase addressed *RQ1* and *RQ2* with a two-session sketching exercise, the second phase addressed *RQ3* by asking participants to solve eight computation tasks. We recruited participants with different levels of expertise to address *RQ4*.

Phase 1: In the first session of the sketching exercise, participants were provided with a description of a dataset, and were asked to sketch a representation. We described the dataset in more detail in a later section. In this case, people imagined their own personal goals for the data. In the second session, we described a specific purpose for the creation of the dataset to the participants and asked them to sketch a representation given the additional information. The purpose of asking participants to sketch with and without specific background information was to understand how the availability of additional specific information affected data representations.

Phase 2: For the computational tasks, we asked participants to plan their step-by-step approaches for each of eight tasks while thinking aloud. We describe these task in more detail later in the paper. We then asked them to execute their proposed approaches and accomplish the tasks using a spreadsheet. The purpose of this two-stage propose and execute approach was to see how participants' models guided expected workflows and if those workflows could be operationalized in practice.

Participants

We recruited participants from a Midwestern city using email lists and flyers posted in public libraries, coffee shops, and local restaurants. Our goal was to recruit diverse genders, ages, and professional backgrounds. 32 people participated in the study–17 female, 15 male. Participants ranged from 19 to 46 years in age (μ = 27, σ = 7), had a variety of educational backgrounds from high school to Ph.D., and a variety of professions, including accounting, graphic design, economics, IT, and social work.

The study lasted 58 minutes on average, with a standard deviation of *nine* minutes. Each participant received $10/hour as compensation at the end of the study.

The Dataset

We used a third party dataset from Airbnb [4]. Their stated purpose [3] for aggregating the data was to investigate a concern that the majority of Airbnb listings were entire homes rented all year round, which violated the notion of "shared economy" and disrupted communities. This publicly available dataset was encoded as a .xlsx spreadsheet, consisting of 142,042 rows of rental listings with 16 information categories (columns) that described each listing. The information categories included price, host name, last review, and minimum nights. Data types also varied across categories and included text and numeric types.

We chose the Airbnb dataset for multiple reasons. First, the data is publicly available on the Internet and it is intended for general consumption [4]. Second, the structure of the data was complex enough to allow us to reasonably expect a variety of sensemaking processes in our experiments, yet not so complex as to confuse participants. Finally, the data was relatable, requiring no specialized knowledge to understand.

Procedure

Pre-Study Survey. We reached out to the participants 24 hours before the study and asked them to fill out a questionnaire to assess expertise level. This survey was derived from Lawson et al [16]. We used the following multiple-choice questions to classify participants into three experience levels:

(1) Please classify your experience with spreadsheets. ($Q1$)
(2) How large are the models you normally create? ($Q2$)
(3) How important are spreadsheets in your daily work? ($Q3$)

Table 1: The selection criteria of experienced and inexperienced participants for $Q1$,$Q2$, and $Q3$.

Experience Level	Q1	Q2	Q3
Experienced	high expertise	exceed 10,000 cells	critical
Inexperienced	little or no expertise, some expertise;still a beginner, or extensive experience	under 1,000 cells	unimportant or moderately important

The options for $Q1$ included: having (a) little or no expertise, (b) some expertise; still a beginner, (c) extensive experience, and (d) high expertise. The option for $Q2$ included: creating a spreadsheet with (a) under 100 cells, (b) 101 to 1000 cells, (c) 1001 to 10,000 cells, (d) 10,001 to 100,000 cells, and (e) over 100,000 cells. The option for $Q3$ included: are (a) unimportant ,(b) moderately important,(c) very important, and (d) critical. Table 1 shows how we classified our participants into the experienced and inexperienced groups. Participants who did not satisfy the criteria for these two groups were classified as having intermediate experience. Six participants were classified into experienced group, eight were classified into inexperienced and the other 18 were classified into intermediate group.

After participants arrived in our lab, we began with an approximately 10-minute ice-breaking discussion where we asked each participant to introduce how they used spreadsheets, what type of operations they executed on spreadsheets, and how they collaborated with others on spreadsheets.

Phase 1: Think-aloud Sketches and Explanations. The first part of the study was a sketching task to investigate participants' mental models of data representations.

We provided each participant with a document listing the information categories of interest. Alongside this document, participants were presented with a screenshot from the Airbnb website so that participants could view these information categories in the context of the site to better understand their role and meaning. Example information categories included city, name, host name, neighbourhood, and price.

The participants were asked, "Based on the names and description of the data categories, how would you organize this data? You can organize the data based on your own scenario and are free to organize it any way you like." After the first sketching session, we conducted a short interview to understand the participants' data representation intentions. More specifically, we asked the participants whether they had any specific task in mind as they were sketching the representations. We wanted to see if they had any personal goal as they reasoned about the data representation.

We then presented each participant with a document explaining that the dataset used in the study was created from the Airbnb website so that the public could investigate violations of city shared economy conduct–specifically whether the majority of Airbnb listings were entire homes rented as hotels throughout the year instead of as a shared room in a home. We then asked the participants to sketch a data representation again, now with investigating this specific goal in mind. In this manner, we can compare if having a different goal alters one's representation of the data.

After the second sketching session, we asked participants again to describe their representations and compare the second sketch with the first sketch. In the cases where participants changed their representations, we further asked them to explain the reasoning behind their change.

We then presented each participant with the actual dataset as an Excel spreadsheet. They had the opportunity to explore the dataset (e.g., . apply functions, navigate, highlight) for approximately five minutes. After, we asked our participants if they could convert their sketches into a spreadsheet format. For the participants who answered "yes", we further asked them to explain the conversion process. We call this conversion of the sketched data representation into a spreadsheet representation *transformability*.

Phase 2: Spreadsheet Computational Tasks. The second phase of the study focused on computational tasks and consisted of two sessions, (a) planning and (b) execution of workflows. Participants were presented with eight tasks. These tasks were selected to cover a range of commonly used spreadsheet tasks as reported in Lawson et al. [16] and to vary in complexity.

The eight tasks were presented with increasing level of difficulty. Simple tasks included asking participants to calculate the mean value for a collection of cells (e.g., "What is the average(Mean) of the listing prices?", "What is the price of the listing with the ID equaling 14491416?", " How many listings in the Harlem neighborhood have a price under 100 dollars?"). Complex tasks combined some of the simple operations (e.g., "What ratio of the listings in New York City are frequently rented? A frequently rented listing has fewer than 60 days of availability in 365 days.", "Find the city with the second-highest average listing price of listings.", " Highlight all Shared room listings with a yellow background. How did you approach this?"). We included an open question in our study to see what operations people would use to explore a dataset (e.g., "If I want to go to travel in San Francisco next month, which listing would you recommend for me? I want to spend 3 nights there. Give me 2 candidates and reasons."). Our final question was to ask people to explicitly use the VLOOKUP Up Function, "How can you use the VLOOKUP function to return the listing's price by inputting the listing id. (You are free to decide where to put the function). Please write down your approach."

In the planning session, participants were first asked to write down the steps they would take to perform these tasks

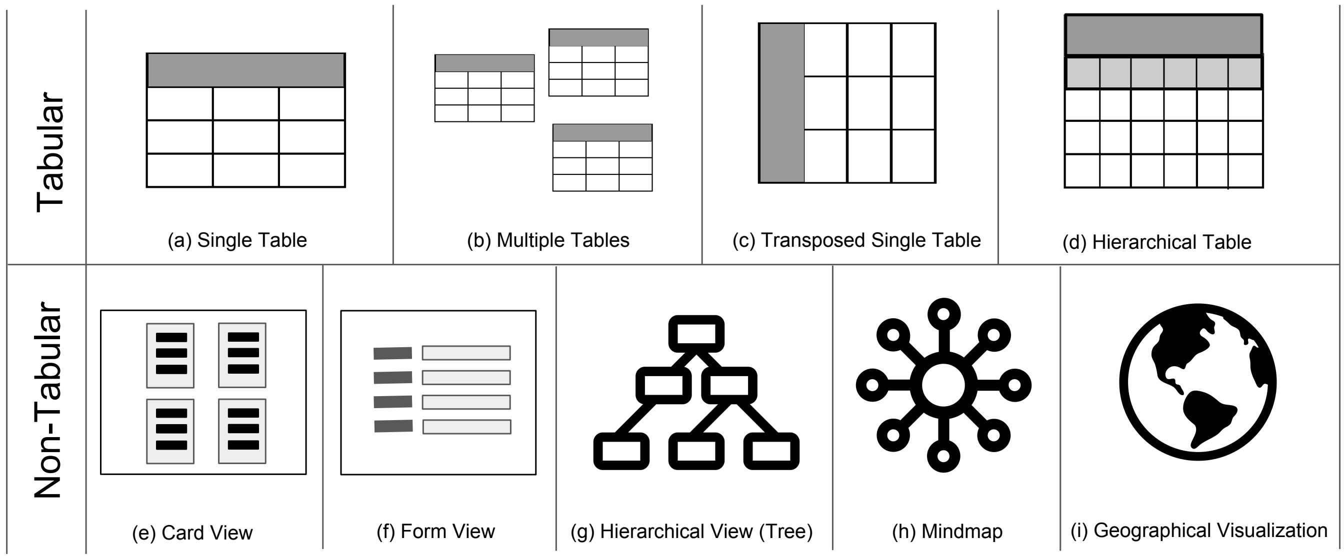

Figure 1: Illustrations of sketches created by participants.

in a think-aloud fashion. The planning session was followed by the execution session where participants were asked to execute their workflows on the actual dataset in a spreadsheet while they explained the rationale of each step.

Sketch and Task Interview Analysis. The interviews from the first sketching exercise, the second sketching exercise, the tabular transformation question, and the eight task exercises were transcribed. To conceptualize the main themes that emerged from each of the two sketching tasks, two researchers used open coding to label categories and sub-categories that emerged from the transcripts and associated sketches using NVivo [9]. The two researchers then iterated on these themes to agreement and then used axial coding to identify relationships among the open codes. The two researchers used the same process to extract themes in the transformation questions and in Phase Two of the study for each of the eight tasks, respectively. A codebook was developed of sketching categories, common intentions in data organization, transformability of data organizations, and common solving patterns and challenges [5].

4 RESULTS AND DISCUSSION

In this section, we report the findings from the sketching and computational tasks and identify the key contributions.

Phase 1: Data Representations

*Types of Representations (*RQ1a*). RQ1a* addresses how individuals represent previously unseen data and how their representations of the data changed after being presented with the purpose for which the data was collected. After a manual coding of participants' sketches and corresponding interview transcripts from the think-aloud exercise, nine styles of data representations emerged. These types fell into two broad groups-tabular and non-tabular representations (see Figure 1).

Tabular representations cover four types of representations. Single Table (ST) representations contain a single header of labeled columns, with each data point occupying one row (1 (a)). Multiple Table (MT) representations use two or more single-table representations (1 (b)). Transposed Single Table (TST) representations transpose (or flip) the columns and rows from Single Table representation(1 (c)). Finally, Hierarchical Table (HT) representations nest headers and attributes (1 (d)).

Additionally, five types of non-tabular representations emerged (see Figure 1). Card View (CV) representations include information attributes grouped into drawings of blocks as containers for information attribute (1 (e)). Form View (FV) representations include HTML form-like data (1 (f)). Hierarchical View (HV) representations include all types of data representations with multiple levels of data, such as trees or other nested structures (1 (g)). Mindmap (MM) representations contain a network-like structure, where concepts are connected to each other by edges (1 (h)). Geographic Visualization (GV) representations include a geographical map along with any other types of representation (1 (i)). In fact, one participant's sketch depicted North and South America(P7).

The Single Table and Multiple Table representations (see Table 2) were the preferred types of data representations across the two sketching sessions. (51.5% overall before participants knew the purpose of the dataset and 50% overall

Table 2: The percentage of each type of representation sketched by the participants before and after knowing the purpose for a dataset.

Representation	First-session sketches	Second-session sketches
Multiple Tables	27.27%	18.75%
Single Table	24.24%	31.25%
Transposed Single Table	9.09%	18.75%
Hierarchical Table	9.09%	12.50%
Card View	9.09%	3.13%
Hierarchical View	6.06%	9.38%
Mindmap	6.06%	6.25%
Form View	6.06%	0.00%
Geographic Visualization	3.03%	0.00%

after). The findings show that people employed a variety of data representation types, with a majority choosing tabular representations and not conflicing with the majority of users' mental models.

*The Impact of Purpose on Representations (*RQ1b*).* After the first sketching task, participants were asked to explain the tasks that they had in mind as they sketched their representation. We coded these tasks into six categories: data exploration (35.5%), search (19.4%), data organization (19.4%), database creation (3.2%), data cleaning (3.2%), and web design (3.2%). Data exploration includes exploring the data from a data analysis perspective. Search refers to task oriented data exploration, such as finding an appropriate listing. Data organization refers to arranging data in a well-structured representation. Creating a database refers to creating a database schema. Data cleaning involves removing missing data and making the data consistent. Web design refers designing a web page with the required data.

Users with different levels of experience exhibited different preferences. The most common task in mind for inexperienced participants was a search task (four out of eight). P23: "I was kind of thinking in terms of finding a particular place to stay." Experienced participants displayed a more analytical view. Three out of six experienced participants chose data organization as the task they had in mind. P15 explained his goal for representing the data was to "...organize the back end of having this show up on like the Airbnb site..." For intermediate participants(ten out of eighteen), data exploration was the primary purpose. P10 reported,"I guess you can see all your different hosts and how many listings each one has..."

Figure 2(a) shows, for inexperienced users, two out of three non-tabular representations (HV and FV) were used for search tasks. Inexperienced participants felt non-tabular representations were appropriate for Search."(Mindmaps) make it easy to search for information based off these key concepts." (P28) Participants utilized the spatial characteristics of representations to locate the information they need. Figure 2(b) indicates that the non-tabular representations were preferred by intermediate participants for data exploration tasks(five out of nine). Figure 2(c) shows experienced participants preferred tabular representations in both sketching exercises(four out of six).

*The Impact of Expertise on Representation (*RQ4*).* Participants of different skill levels displayed different preferences when sketching the representations (Table 2). Before informing the participants about the purpose of the data creation, 69.7% of the data organizations were tabular. The inexperienced and intermediate participants generated all of the non-tabular representations (30.3%).

Among the participants, 37.5% (three out of eight) of the inexperienced and 38.9% (seven out of eighteen) of the intermediate participants used a non-tabular layout for representing data when they did not know the purpose of creating the dataset. None of the experienced participants used non-tabular representations for their first sketch.

Experienced and inexperienced participants showed different behavior in the second sketching session, i.e., after the data creation purpose was explained. Figure 3 illustrates how participants modified their sketches before and after the purpose was explained. The coordinates of each data point represent the transformation between the first sketch (x-axis) and the second sketch (y-axis). For example, Figure 3(a) shows that participants who used Form View (FV) for their first sketch used either Multiple Tables (MT) or Transposed Single Tables (TST) for their second sketch. More broadly, Figure 3(a) shows that all inexperienced participants who used non-tabular representations (three) in their first sketch used tabular representations in their second sketch. In contrast, Figure 3(c) shows that the experienced participants made the fewest modifications, using tabular layouts across both sketches.

A noticeable exception was $P15$, an experienced participant, who sketched a Multiple Table (MT) in the first sketching session, an then opted to use a Hierarchical View (HV) sketch for the second sketching session. $P15$ reported: "...I was thinking more of just maintaining a list of accurate information on this", when explaining the first sketch. However, after knowing the purpose of creating the dataset, $P15$ reported: "...I was thinking more of how it would be visualized, or how someone might want to work with it, how they might filter it or sort it to see ... to answer specific questions that they had." This

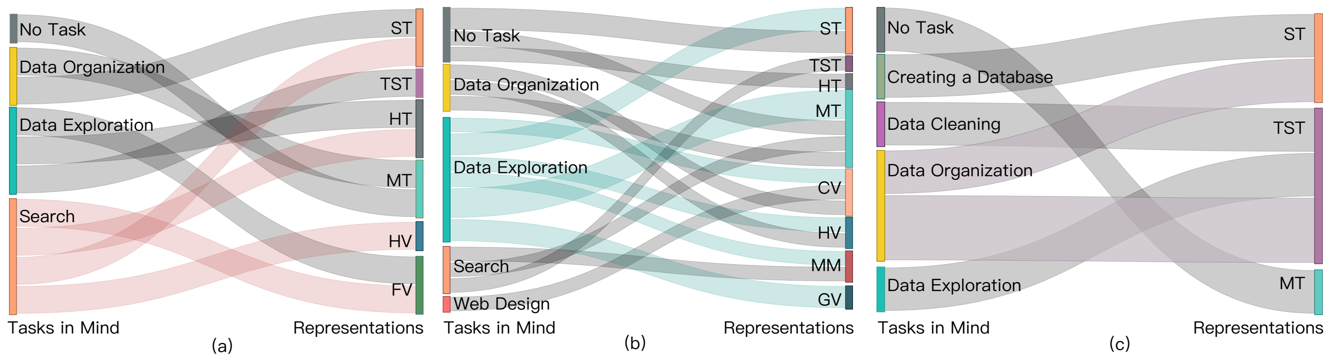

Figure 2: The conversion from people's task in mind to the types of organization they made. (a). Inexperienced group. (b). Intermediate group. (c). Experienced group.

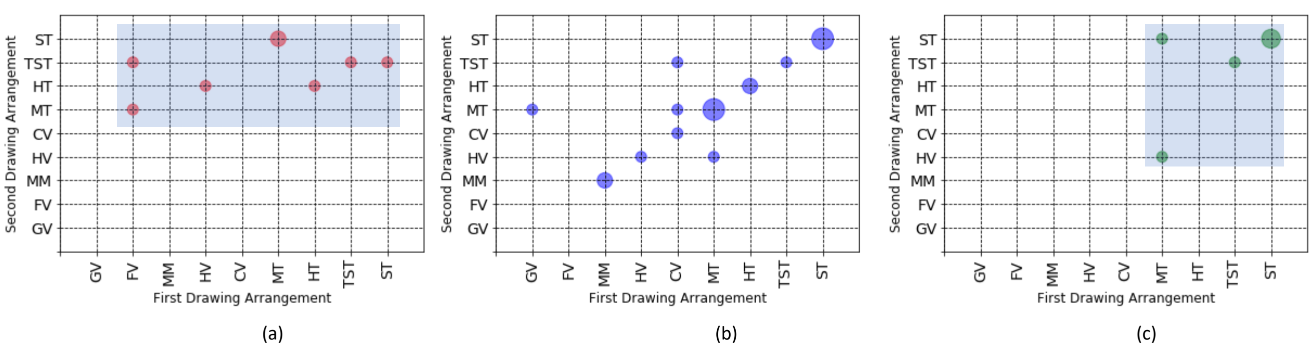

Figure 3: The transformation of sketching before and after knowing the purpose of creating the dataset across three experience level groups: (a) Inexperienced group, (b) Intermediate group, (c) Experienced group.

observation suggests experienced users, with flexible data representations, are skilled in accomplishing tasks in different ways.

The findings suggest that most experienced participants opt for tabular representations regardless of the tasks in mind. It may be that the experienced participants know that tabular representations, with the right queries, can be used for many computational tasks. Describing previous experiences, *P*9 mentioned writing database SQL queries within an Excel cell. "I used it to basically create a dashboard so it would refresh, it would pull up query from some site internal to ..." Inexperienced participants, on the other hand, often changed their data representations as the purpose changed.

We further investigate the reasons why people changed their data representation (see Figure 3). It appears participants' task in mind altered after we provided the purpose of creating the dataset. *P*23 reported: "...For the individual traveler wanting to avoid a mega-host, I would say it's gonna be almost identical to what I had in the first version up there. For the regulator trying to identify who are the mega-hosts, I think it's kind of a different arrangement (representation) of the data." *P*7 reported: "This is more in line with a different interpretation. This, I was thinking we were going to show this to a consumer who wants to buy a listing. That's why it's arranged this way. This will be in more in line with the objective of looking at each host and making sure no one host owns too many listings, and is concentrated on a neighborhood. This is more in line with the goal. It's also more elaborate."

Table 3: The transformability (whether the sketches can be converted to a tabular layout) of participants' sketches across three experience groups.

Task	Data Representation	Experienced	Intermediate	Inexperienced
Sketching1	Non-tabular	NA	0.71	0
	Tabular	1	0.82	0.8
Sketching2	Non-tabular	1	0.8	NA
	Tabular	1	0.85	1

*Transformability of Representations to Spreadsheets (*RQ2*).* Table 3 indicates that none of the inexperienced participants who opted for non-tabular representations were able to convert their sketches to spreadsheets. They reported the main reason was the lack of an interface supporting non-tabular data representation in spreadsheet software. P26 reported "...it's not a grid, it's just a form, so in order to put your data in rows and be able to sort on it, or it just wouldn't work... " P2 stated"I just, it's hard to imagine anything that would drop down in a spreadsheet..." P29 described the challenges as "Too many files. Too many things to reference. That would just be terrible to work with." The comments from the participants indicate that spreadsheets have poor support for inexperienced participants to incorporate non-tabular representations.

On the other hand, all of the experienced participants could convert their sketches into a spreadsheet. The dominant strategy they used was to rearrange columns of data or to use multiple sheets to support their representation, such as Multiple-Table (MT) representations. P1 responded "Oh, I would just have to rearrange the columns by cutting and pasting." One explanation for this is that experienced participants had a better understanding of using spreadsheets to represent data since they naturally connected their tasks in mind to a tabular format.

Phase 2: Computational Tasks

In this section, we analyze how people solve computational tasks and the challenges they encountered.

*Expertise vs. Task Completion (*RQ3a*).* Unsurprisingly, experienced participants successfully completed more tasks than inexperienced participants (3.3 to 2.63 on average). For the purpose of simplicity, we denote the i-th task as Ti from here on. Table 4 summarizes the average number of steps required by the three different expertise groups of participants. However, for each question, there were instances where participants were unsuccessful in completing the task. Therefore, we show the average number of steps for successful and unsuccessful attempts separately. Moreover, $T7$ was an open-ended question and $T8$ explicitly asked participants to use the *VLOOKUP* function to see if they either knew the function's usage or were able to use online resources to successfully apply the function. Therefore, we exclude these two tasks while preparing Table 4. For each task, we highlight the expertise group who took the least number of steps on average to successfully complete the task, in boldface. Moreover, if there were no instances of successful or unsuccessful attempts for a task for an expertise group, we denote those instances as '-'.

Table 4 shows that, except for $T1$, the experienced users took the least number of steps on average to complete a task.

Table 4: Number of steps taken to solve computational tasks across three expertise level. E = Experienced, I = Intermediate, IE = Inexperienced, S = Successful, and U = unsuccessful.

Task ID	E		I		IE		Overall	
	S	U	S	U	S	U	S	U
T1	2.33	-	**2.17**	2.67	3.00	2.00	2.30	2.40
T2	**1.25**	2.00	1.50	2.33	2.25	1.00	1.58	2.00
T3	**3.00**	2.00	3.75	3.29	6.00	3.33	3.75	3.13
T4	**2.00**	3.60	6.00	3.50	6.00	3.00	5.33	3.48
T5	**2.00**	1.75	4.33	2.80	-	2.50	3.40	2.57
T6	**2.00**	1.50	2.17	1.57	3.00	2.33	2.22	1.75

One experienced participant ($P15$) used a combination of mathematical functions to solve the first task which contributed to the increase in the average number of steps for the first task for experienced participants. As the complexity of the tasks increased from the $T1$ to $T5$, the average number of steps increased for the intermediate and inexperienced participants. However, experienced participants were able to complete the tasks with a relatively small number of steps irrespective of the task complexity.

*Expertise vs. Task Workflow (*RQ3a*).* We applied a frequent sequential pattern mining algorithm (CM-Spade [11]) on a collection of workflows to discover the most frequently emerging sequences of each participant group for each task. We further coded the frequent patterns obtained from the mining algorithm into six high level categories: Navigation and Selection; Sorting and Filtering; Search, Find and Look Up Function; Highlight and Conditional Formatting; Statistical, Math and Pivot Table; Spreadsheet Editing Operations. Navigation refers to scrolling and moving from one spreadsheet region to another, and Selection refers to selecting a cell or a collection of cells within the spreadsheets. Spreadsheet editing operations refer to sheet creation, deletion, and cut or copy, and paste to move subsets of data to a new sheet. The rest of the categories refer to default spreadsheet operations [1].

Table 5 summarizes the findings we obtain by mining the workflows. Our mining revealed that people from different levels of expertise used similar operations to accomplish tasks. The percentage of use for each feature—Selection and Navigation, Sorting and Filtering, Search, Find and Look Up Function, Highlight and Conditional Formatting—is similar across the three participant groups (see Table 5).

The Table 5 indicates that most participants knew which operations to conduct in the spreadsheet. It is interesting to note that experienced users finished more tasks than inexperienced users, but, as Table 5 indicates, they were able to do this using similar features with lesser number of step than

Table 5: Percentage of features used across three groups

Feature Category	Experienced	Intermediate	Inexperienced
Navigation and Selection	21%	31%	24%
Sorting and Filtering	16%	17%	15%
Search, Find and Look Up Function	18%	17%	21%
Highlight and Conditional Formatting	12%	10%	12%
Statistical, Math and Pivot Table	31%	21%	13%
Spreadsheet Editing Operations	0%	1%	7%

Table 6: Percentage of challenges and high level themes we identified

Theme	Examples	Percentage
Incomprehensible documentation	Applying functions	29%
Loss of context	Sorting ordering Maintain state	10%
Omission of critical steps	Incomplete operations	21%
Information overload	Scalability-Navigation/Selection	25%
I don't know	"I have no idea how to approach it!"	16%

other users. This suggests that experienced users may use features more effectively than other user groups.

*Task Strategies Across Expertise Groups (*RQ3a*).* Participants followed two different types of execution strategies when performing the computational tasks: in-situ and ex-situ. Ex-situ strategies refer to the cases where participants created an extra sheet to complete the task. In the ex-situ execution, participants used a combination of different spreadsheet editing operations. For in-situ tasks, participants chose to work within one sheet. Experienced participants preferred solving tasks in-situ. Conversely, inexperienced participants frequently used spreadsheet editing functions to solve tasks ex-situ (four out of eight). Only one participant from the intermediate group chose to perform tasks ex-situ.

We can attribute inexperienced user's tendency to perform tasks ex-situ to being unable to process the volume of data. *P*2: "Yeah, I didn't know just how huge this file is... " They also reported that they did not want to work on the same spreadsheet for fear of contaminating the data. P32: "That's where I would control, copy, paste. Because within this dataset I don't know how to leave this sorted and then just sort the second column I want. That's why I control, copy, paste it.". Therefore, they chose to create a new sheet, cut or copy subsets of data to the new sheet, and then perform their operations.

*Task Difficulties (*RQ3b*).* We identified four major challenges faced by our participants (The four challenges exclude the 16% of participants who reported that they had no idea how to approach the tasks).

The four challenges are incomprehensible documentation, loss of context, omission of critical steps, and information overload. Incomprehensible documentation refers to the situation where the participants were unable to understand the documentation related to function and therefore, failed to apply the function correctly. Loss of context refers to the scenario where people lost the context where they were as they were processing the data on spreadsheets. Omission of critical steps refers to the cases where people failed to execute their planned workflows. This issue occurred because participants missed critical steps when they planned the workflows in the planning session of the computational tasks. Information overload refers to the challenges appearing in selection and navigation operations (defined earlier). Examples include failure to select a collection of cells or failure to navigate to the desired location.

Loss of context Maintaining a context is a challenge for participants. For complex operations which require multiple steps, maintaining a small set of working data is helpful. P32 described that *One difficulty is keeping the city column sorted while also sorting the price column. Hence that is why I would Ctrl + C the city data into another sheet so that I can sort the price without other cities "interfering" in the sorting.* P8 reported that when he was creating a subset of data in another sheet, "Header was lost, (I need to) find an alternate method."

Incomprehensible documentation

When trying to understand the appropriate usage of spreadsheet functions, people struggled to understand the documentation of the functions, leading to the incomprehensible documentation problem. Inexperienced participants faced more difficulties in understanding the documentation of functions they did not know how to use.

Using Excel Support and Google searches did not adequately help participants to understand how to use functions. In all of the five attempts to use Google to solve tasks (searching for the usage of functions), only one pointed to a correct answer. This result indicates that effective training and documentation are still in short for participants. P27 watched online video from Microsoft official website for 10 minutes to learn how to use *VLOOKUP*. He still failed to use it in solving tasks.

Omission of critical steps

For some participants, planning approaches to solving tasks is manageable. However, when they executed their methods on the spreadsheet file, it turned out to be impractical if they missed a step. For example, P17 reported that when he planned to execute a conditional formatting, his planned method does not work because "First time, I didn't highlight the column before setting parameters and it didn't work".

Information overload Information overload denotes the scenario where participants were overwhelmed by the volume of data. *P*10 ascribed the failure to "I am not sure how to easily find the ratio without manually counting all the listings." *P*17 reported that "I am not sure how to highlight large amounts of data, except manually." As a pervasive challenge, it prevented participants to complete tasks they expected to succeed in a smaller size spreadsheet.

5 IMPLICATIONS

Data Represetation Sensemaking

People's representation of data may change in different stages of the sensemaking process. As our study shows, without knowing the background of the data, inexperienced participants preferred 'search' tasks to understand and make sense of the data. Understandably, a significant portion of those participants preferred visual representations of data to enable better comprehension of the data and easier search. However, upon discovering the background of the data, they tend to focus on the organization aspect of the data, and therefore, prefer tabular organization. This indicates that inexperienced users may comprehend previously unseen data better using visual representations compared to tabular representations. Existing works in designing overview+detail interfaces [8] also advocate providing overviews to reduce user's cognitive load in comprehending the data [13]. The tabular representation of spreadsheets presents data in detail whereas visual representations provide an overview. Therefore, providing a visual overview may accelerate the sensemaking process. However, current spreadsheet interfaces do not support additional visual representations to facilitate management and exploration of data.

Providing Overview and Context

As participants perform different tasks on the data, they face different challenges in interacting with and exploring the data. The interaction challenges primarily stem from the scale of the data. Even with the Airbnb data which consumes only 15% of Excel's supported scale (one million rows), people tend to struggle in performing simple tasks like selecting a range of data. The typical scrolling and windowing approaches of spreadsheets make it difficult for the participants get the overall context of the data and access large quantities of data. P32 indicated that "One difficulty is keeping the city column sorted while also sorting the price column, hence that is why I would Ctrl + C the city data into another sheet so that I can sort the price without other cities "interfering" in the sorting. Again, another difficulty is scrolling to highlight all the values needed" Therefore, providing features to assist participants in having an overview of the whole dataset will be helpful for participants. People can use this type of tools to dive into a sub-set of interesting data.

The exploration challenges stem from the lack of history and context. As people explore the spreadsheets and perform different tasks, they lose the context of where they were, where they currently are, and where they can explore next. P5 had to "freeze top row" when looking for the price column. Providing tools for people to maintain context and history can be helpful for data exploration and analysis.

A Workflow Assistant

Our sequential pattern mining reveals that inexperienced participants and more experienced participants do not use fundamentally different set of functions and task patterns to solve problems on spreadsheet applications. Rather, what these groups differ in is the efficiency of finding the right workflow. Using the same set of functions, experienced participants accomplish tasks with smaller number of steps than novice participants. To bridge the gap between experienced and inexperienced users, we need to ensure that finding the right workflow is intuitive for all users.

Existing spreadsheet documentation users explore to find the correct functions and workflow are somewhat incomprehensible. For example, *P*17 tried to use search engine to learn how to use the *VLOOKUP* function. After reading an online instruction, *P*17 said:"Yeah, I was looking at the instructions, and I was like okay, okay. I think I got it. Then I was ... let me move back. " However, after trying the instructions on the spreadsheet, *P*17 told us: "... I think it was just without doing it myself a few times I probably wouldn't be able to get it just reading it and doing it. I was actually thinking I wonder if Linda.com has a thing on how to use it..." Therefore, improving the accessibility of the available functions and explaining their use cases is an important area of future research. An example solution for this challenge can be to use a chatbot system that can dynamically understand user's task at hand and suggest suitable functions and workflows accordingly.

6 LIMITATIONS AND FUTURE WORK

This study aimed to understand how people use spreadsheets to meet data representation and computation tasks. It involved a limited number of participants who came from a small number of backgrounds. Future work can expand our scope to a more diverse population and take other factors, such as socioeconomic level, into the consideration.

In the ice-breaking section of the study, participants were asked to introduce their daily practice with spreadsheets. This question could imply that tabular representation should be selected for sketching, which may have primed users to use spreadsheets to represent the data. We told each participant that "You are allowed to organize the data in any form you like" to compensate for the priming effect. Additionally, we asked participants to bring a recently used spreadsheet file for ice-breaking. However, not everyone brought one file to the study, so excluded this task from our study.

7 CONCLUSION

We conducted an interview study with participants of various spreadsheet expertise levels to better understand how well spreadsheets' data representation align with people's perception of the data. Additionally, we investigated how intuitive it is to perform computational tasks on spreadsheets. We performed a series of sketching exercises and asked users to perform eight computational tasks to address both questions. Our results from the sketching exercise reveal that participants expected data representation at times deviate from the tabular representation of spreadsheets and their expertise contribute to such deviation. To accommodate different styles of thinking, spreadsheets can benefit by providing a visual overview and context of the data. Our results from the computational task exercise indicate that current documentation of spreadsheet capabilities is not very comprehensible, particularly for inexperienced users. Further research is needed to better understand how to intuitively inform users of how to utilize spreadsheet capabilities.

REFERENCES

[1] Excel help center. https://support.office.com/en-us/excel. Accessed: 2018-09-20.

[2] How finance leaders can drive performance. https://enterprise.microsoft.com/en-gb/articles/roles/finance-leader/how-finance-leaders-can-drive-performance/.

[3] The inside airbnb about page. http://insideairbnb.com/about.html.

[4] The inside airbnb dataset. http://insideairbnb.com/get-the-data.html.

[5] Boyatzis, R. E. *Transforming qualitative information: Thematic analysis and code development*. sage, 1998.

[6] Brock, M. E., Vert, A., Kligyte, V., Waples, E. P., Sevier, S. T., and Mumford, M. D. Mental models: An alternative evaluation of a sensemaking approach to ethics instruction. *Science and engineering ethics 14*, 3 (2008), 449–472.

[7] Chambers, C., and Scaffidi, C. Struggling to excel: A field study of challenges faced by spreadsheet users. In *Visual Languages and Human-Centric Computing (VL/HCC), 2010 IEEE Symposium on*, IEEE (2010), 187–194.

[8] Cockburn, A., Karlson, A., and Bederson, B. B. A review of overview+ detail, zooming, and focus+ context interfaces. *ACM Computing Surveys (CSUR) 41*, 1 (2009), 2.

[9] Edhlund, B., and Mcdougall, A. *Nvivo 11 Essentials*. Lulu.com, 2016.

[10] Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., and Kirlik, A. First i like it, then i hide it: Folk theories of social feeds. In *Proceedings of the 2016 cHI conference on human factors in computing systems*, ACM (2016), 2371–2382.

[11] Fournier-Viger, P., Gomariz, A., Campos, M., and Thomas, R. Fast vertical mining of sequential patterns using co-occurrence information. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer (2014), 40–52.

[12] Friedman, B., Hurley, D., Howe, D. C., Felten, E., and Nissenbaum, H. Users' conceptions of web security: a comparative study. In *CHI'02 extended abstracts on Human factors in computing systems*, ACM (2002), 746–747.

[13] Grudin, J. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2001), 458–465.

[14] Hendry, D., and Green, T. Creating, comprehending and explaining spreadsheets. *Int. J. Hum.-Comput. Stud. 40*, 6 (June 1994), 1033–1065.

[15] Kankuzi, B., and Sajaniemi, J. An empirical study of spreadsheet authors' mental models in explaining and debugging tasks. In *Visual Languages and Human-Centric Computing (VL/HCC), 2013 IEEE Symposium on*, IEEE (2013), 15–18.

[16] Lawson, B. R., Baker, K. R., Powell, S. G., and Foster-Johnson, L. A comparison of spreadsheet users with different levels of experience. *Omega 37*, 3 (2009), 579–590.

[17] Mack, K., Lee, J., Chang, K., Karahalios, K., and Parameswaran, A. Characterizing scalability issues in spreadsheet software using online forums. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM (2018), CS04.

[18] Nardi, B. A., and Miller, J. R. *The spreadsheet interface: A basis for end user programming*. Hewlett-Packard Laboratories, 1990.

[19] Neyeloff, J. L., Fuchs, S. C., and Moreira, L. B. Meta-analyses and forest plots using a microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. *BMC research notes 5*, 1 (2012), 52.

[20] Norman, D. *The design of everyday things: Revised and expanded edition*. Constellation, 2013.

[21] Panko, R. R., and Halverson Jr, R. P. Are two heads better than one?(at reducing errors in spreadsheet modeling). *Office Systems Research Journal 15*, 1 (1997), 21–32.

[22] Peakall, R., and Smouse, P. E. Genalex 6: genetic analysis in excel. population genetic software for teaching and research. *Molecular ecology notes 6*, 1 (2006), 288–295.

[23] Poole, E. S., Chetty, M., Grinter, R. E., and Edwards, W. K. More than meets the eye: transforming the user experience of home network management. In *Proceedings of the 7th ACM conference on Designing interactive systems*, ACM (2008), 455–464.

[24] Russell, D. M., Stefii, M. J., Pirolli, P., and Card, S. K. The Cost Structure of Sensemaking (1993). 24–29.

[25] Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K. The cost structure of sensemaking. In *Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems*, INTERCHI '93, IOS Press (Amsterdam, The Netherlands, The Netherlands, 1993), 269–276.

[26] Tohidi, M., Buxton, W., Baecker, R., and Sellen, A. User sketches: a quick, inexpensive, and effective way to elicit more reflective user feedback. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, ACM (2006), 105–114.

[27] Vosniadou, S., and Brewer, W. F. Mental models of the earth: A study of conceptual change in childhood. *Cognitive psychology 24*, 4 (1992), 535–585.

[28] Zhang, Y., Huo, M., Zhou, J., and Xie, S. Pksolver: An add-in program for pharmacokinetic and pharmacodynamic data analysis in microsoft excel. *Computer methods and programs in biomedicine 99*, 3 (2010), 306–314.