Charting EDA: Characterizing Interactive Visualization Use in Computational Notebooks with a Mixed-Methods Formalism

by

Dylan Wootton

S.B., University of Utah (2019)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by:	Dylan Wootton Department of Electrical Engineering and Computer Science May 10, 2024
Certified by:	Arvind Satyanarayan Professor of Electrical Engineering and Computer Science Thesis Supervisor
Accepted by:	Leslie A. Kolodziejski Professor of Electrical Engineering and Computer Science Chair, Department Committee on Graduate Students

Charting EDA: Characterizing Interactive Visualization Use in Computational Notebooks with a Mixed-Methods Formalism

by

Dylan Wootton

Submitted to the Department of Electrical Engineering and Computer Science on May 10, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

ABSTRACT

Interactive visualizations are powerful tools for Exploratory Data Analysis (EDA), but how do they affect the observations analysts make about their data? We conducted a qualitative experiment with 13 professional data scientists analyzing two datasets with Jupyter notebooks, collecting a rich dataset of interaction traces and think-aloud utterances. By qualitatively coding participant utterances, we introduce a formalism that describes EDA as a sequence of analysis states, where each state is comprised of either a representation an analyst constructed (e.g., the output of a data frame, an interactive visualization, etc.) or an observation the analyst made with a representation (e.g., about missing data, the relationship between variables, etc.). By applying our formalism to our dataset, we are able to identify that interactive visualizations, on average, lead to earlier and more complex insights about relationships between dataset attributes compared to static visualizations. Moreover, by calculating metrics such as revisiting count and representational diversity, we are able to uncover that some representations serve more as as "planning aids" during EDA rather than tools strictly for hypothesis-answering. We show how these measures helped identify other patterns of analysis behavior, such as the "80-20 rule", where a small subset of representations drove the majority of observations. Based on these findings, we offer design guidelines for interactive exploratory analysis tooling and reflect on future directions for studying the role that visualizations play in EDA.

Thesis supervisor: Arvind Satyanarayan
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

I am deeply grateful to my advisor, Dr. Arvind Satyanarayan, whose brilliance and gentleness have shaped my journey as a researcher. Arvind, you inspire me not only through your exceptional intellectual contributions but also through your compassionate approach to research. When I've faced difficulties, your advice to "be gentle with myself" reminds me that research is a marathon, not a sprint, and that self-compassion is a necessity, not just a nice-to-have. I feel lucky to be your advisee, and I'm looking forward to the work we will do together.

I owe a debt of gratitude to my family, whose unrelenting support has been crucial in my pursuit of research. I am particularly blessed to have parents whose dedication to achieving their own goals has instilled in me a strong ambition. Mom and Dad, your efforts to provide a nurturing environment for me have left an indelible mark on my character. To my sister Rachel, words fall short of expressing how profoundly you have influenced my life. Your support and love have been my constants in this ever-changing journey.

I extend my thanks to my collaborators—the MIT Vis group, Miriah, Alex, Carolina, Amy, and Evan. You all have sparked my interest in visualization and supported my transition from biology to computer science. You have shown me that visualization isn't just the coolest work you can do, but also that it attracts some of the best people.

I would also like to express my thanks to my friends and the broader community here in Boston. While there are too many individuals to name, I want to highlight the community organizations that have made Boston feel like home. A special thank you to Stonewall Sports, where I found an incredible group of friends through dodgeball and kickball leagues. Additionally, I am grateful for the warm welcome and *invigorating* early morning workouts at the MIT Rowing Club. My appreciation also extends to my Unitarian Universalist congregation and the Cambridge Insight Sangha, both of which have been instrumental in helping me cultivate a spiritual practice to hold me steady.

This thesis was made possible through the support of the NSF Graduate Research Fellowship Award 2141064 and the NSF CAREER Grant 1942659.

Biographical Sketch

Dylan Wootton is a PhD student at the Massachusetts Institute of Technology (MIT) in the Computer Science Artificial Intelligence Laboratory. Under the supervision of Dr. Arvind Satyanarayan, Dylan's research aims to enhance the effectiveness of data communication through the design of interactive data visualizations. His work notably incorporates principles from cognitive science and anthropology to study effectiveness of interactive interfaces.

Previous to MIT, he was a software engineer for Microsoft where he built data analytics systems and conducted visualization research at the University of Utah. His work in data visualization has received numerous honors including academic conference awards and grants from National Geographic and the National Science Foundation Graduate Research Fellowship Program.

Contents

List of Figures

Chapter 1

Introduction

The research literature widely considers interaction to play a central role in effective visualization for exploratory data analysis (EDA) [1], [2] because it supports a "dialogue between the analyst and the data" [3]. Recent empirical results, however, suggest a less clear picture. For instance, Mosca et al. found that adding interactivity to static visualizations does not improve participants' accuracy in Bayesian reasoning tasks [4] while Theis et al. found no significant difference in participants' error rates when working with interactive or static visualizations of uncertainty [5]. And, through a contextual inquiry with professional data scientists, Batch & Elmqvist identify a gap where interactive visualization is primarily used to communicate results at the end of an analysis, rather than as a medium for conducting the analysis itself [6].

We hypothesize two diagnoses for these discordant bodies of results. First, much of the work demonstrating the value of interactive visualization in EDA is conducted within systems purpose-built to support this activity (e.g., Tableau [6], Voyager [7], VisTrails [8], among others [9]). As a result, participants are not able to "opt out" of the modality and conduct their analysis through other means (e.g., via code). Second, although existing approaches largely recognize that analysis is a *situated* activity — that is, it involves human analysts working in a particular context, making observations with various representations of data — thus far, these methods often focus on one aspect of this behavior rather than synthesizing across it. For instance, thematic analyses have usefully been used to identify patterns of analytic behaviors [9], but it can be difficult to describe how these patterns manifest with particular interactive representations. On the other hand, quantitative approaches (e.g., interaction telemetry and provenance [7], [10]) capture detailed information about how analysts use particular representations but, without bringing qualitative approaches to bear, can have trouble disambiguating observations — for instance, does hovering over a visualization indicate hesitation, gesticulation, or hypothesis testing? Recent "insight"-based approaches [11] have come perhaps the closest to capturing the richness of analytic activity, but are presently focused on a narrow band of activity: quantitative insights arrived at through data transformation.

To more deeply study how choices of data representations (including interactive visualizations) affects EDA, we conducted a qualitative experiment [12] with 13 data science professionals using Jupyter notebooks. Participants were asked to complete two analysis tasks: the first with a lightweight library for authoring *static* visualizations, followed by a second with an extended library including *interactive* visualizations. Given their widespread use, Jupyter notebooks afford a more real-world context to study analytic behavior and, critically, do not *presuppose* the value of interactive visualization. Thus, across both tasks, participants were free to forego visualization and interaction altogether, and simply author Python code using any third-party libraries they wished.

We collected a rich corpus of participants' think-aloud utterances alongside log-telemetry data of their notebook activity, and conducted a mixed-methods analysis that joined qualitative content analysis [12] with quantitative analysis of interaction traces. We express our mixed-methods analysis through a novel formalism that describes EDA sessions as a sequence of analysis states. Each analysis state is either the representation an analyst constructed (e.g., the output of a dataframe, or an interactive visualization), or an observation an analyst made (i.e., an utterance they issued about one or more representations). We identified 15 distinct types of utterances, grouped into four categories: *dataset* utterances about its size or orientation, or whether there was any missing data; *variable* utterances about its distribution or outliers; *relationship* utterances that expressed concepts including strength, directionality, and clustering; and *process* utterances that described intended analysis steps, or commented about meta characteristics about a representation.

Our mixed-methods analysis, and resultant formalism, reveals that observations follow distinct temporal patterns during EDA (§ 5.1). Analysts tend to address dataset-level metadata early on, while variable distributions and relationship insights occur throughout the analysis. Notably, interactive visualization accelerate relationship utterances, with these statements occurring 15% earlier than under the static condition. In analyzing representation telemetry (§ 6.0.2), we define a series of quantitative metrics including *revisit count*, or the total number of times a participant hovered over a representation; *output velocity*, or the number of representation instances created per unit time; and, *representational diversity*, or the number of unique representation types created during an analysis.

We use these metrics to investigate patterns of broad exploration, revealing why certain participants are able to achieve broad coverage during their EDA (section 6.0.4). Finally, by bridging think-aloud utterances and representation telemetry, our formalism helps uncover patterns in representation usage such as the 80-20 rule of representation use (section 6.0.2) and the propensity to use all-attribute representations as aids to plan their analyses (section 6.0.2).

Taken together, our work contributes to calls for "deepening [the] theoretical foundation" of exploratory data analysis [13].

Chapter 2

Related Work

Over the past decade, studies of Exploratory Data Analysis (EDA) have sought to understand visualization use in analysis contexts [7], [9], [14]–[20]. To study the open and iterative nature of EDA these studies leverage a wide diversity of methodological approaches.

Interviews and Surveys: Interviews have been a primary tool to study data scientist workflows. Kandel et al. conducted foundational work understanding the stages of data science work [21]. They interviewed data scientists across various enterprise organizations outlining five key job responsibilities: *discovery*, *profiling*, *data wrangling*, *modeling*, and *reporting*. These elements are central to data science activities. Further refining this understanding, Wongsuphasawat et al. conducted interviews that revealed a more detailed set of 16 analytic behaviors, such as *converting data formats* and *examining bivariate plots,* providing a set of commonly complete tasks in data scientist's workflows [22]. Interviews also enable researchers to investigates attitudes towards particular EDA tools, such as Batch's [6] work to understand the "Interactive Visualization Gap" in EDA.

Furthermore, when conducting empirical studies, surveys are used to follow up an exploratory analysis session [16]–[18]. Most commonly, surveys include questionnaires like the NASA-TLX[23] for understanding subjective workload during a task [18] or Likert scale questions to elicit preference using a particular tool [16], [17].

Interaction Traces Recording interaction logs of system use is a common method to characterize an analysts EDA. We refer to [24] for a full classification of the ways that interaction traces are analyzed to inform visualization systems and discuss major uses here.

Interaction traces used from demonstrating feature usage [7], [16], [17] to characterizing more complex patterns of action sequences [10]. Interaction logs have also been used to create metrics to assess exploratory behavior with particular visualization systems and reveal how user characteristics influence exploratory patterns [25]. Often, interaction traces are used in conjunction with other characterization strategies. For example, they are often used in Attribute methods to demonstrate when a particular set of attributes is "considered" ranging from hovering over particular visualizations [7] to visualization creation in tableau [15].

Attribute Methods Attribute based methods operationalize EDA in terms of the number of attribute sets that are explored [7], [15], [16], [26]. The most recent use of attribute characterization was by Battle and Heer to create *analysis search trees* showcasing exploratory behavior during Exploratory Visual Analysis, a subset of EDA [15]. Such an approach defines depth in terms of the length of the longest search branch (maximum number of attributes encoded in a visualization) and breadth in terms of the number of leaf nodes in an analysis search tree (the number of attribute set branches explored). Operationalizing EDA in this way revealed particular attribute sets that appeared to be key "analysis-states" during participant analyses. Such models also found that exploration in tableau is primarily depth oriented.

Thematic Analysis: Thematic analysis approaches seek to identifying occurrences of broad behavioral patterns[9], [19], [27] of participant use of a tool. We find these approaches typically involve participants thinking aloud in order interpret the meanings of behaviors given their context. For example, Kale et al. [9] investigated the effect of a exploratory tool that enables *model-checking* through a within subjects comparison with data analysts. Using thematic analysis, they characterize how the patterns of analysis shifted when the model-checking functionality was introduced. Such observations revealed that this functionality "structure participants' thinking around one or two long chains of operations" giving rich classifications of the analysis behavior. Thematic analysis does not seek to characterize the content of entire analysis session, choosing instead to focus more on larger themes that were observed during exploration.

Insight Methods The closest approach to our work is that of insight methods [28]. Insight methods seek to identify the insights made by a participant during an EDA, often by employing a think-aloud process [27], [28] or eliciting insights as a part of open-ended free responses [29]. This unstructured data is then extracting into individual *units of analysis*. The units are then coded based on their semantic content– such as *Generalization* or *Hypothesis* [30]. At this individual unit level, additional coding passes can be done to extract additional information from the insights, such as if they are broadening or deepening [16] or whether they are factually correct [31]. Using these coded utterances, existing methods then often aggregate these insights across entire analysis session, computing metrics such as time-to-first insight and total number of insights [20], [27], [28], [30], [31].

We differentiate our approach from previous insight methods through the use of qualitative content analysis to record both what is said and what interface features were used to make such utterance. By explicitly linking the *Observation* to the *Representation* used to make it. As a result, we can compute aggregated information about insights during analysis conditions 5 but also investigate how insights are formed using particular representations (§6.0.1). This approach lets us investigate the role of specific visualizations on the EDA process, such as how analysts make 80% of their utterances from 20% of their representations (§6.0.2). They also cannot illuminate how insights differ between interactive and static visualizations when they are used within a single analysis session (§6.0.3). Additionally, our use of qualitative content analysis allows us to capture a wider, more nuanced range of insights, revealing how specific visualizations trigger particular types of observations (§6.1).

Chapter 3

Methods

Our study design was driven by our motivating research question: **How might interactive visualizations affect behavior during exploratory data analysis in computational notebooks?** This inquiry is both *descriptive* and *comparative*. We wish to describe how data scientists conduct analyses (the steps taken, representational choices, and inferences made) and to compare how these behaviors unfold with *static* vs. *interactive* visualizations. Thus we adopted a hybrid design, combining task observation and semi-structured interviews situated in the structure of a repeated-measures experiment: an approach described in the mixed methods literature as a *qualitative experiment* [12].

3.0.1 Study Design, Procedure, and Participants

Our independent variable is **representation interactivity** with two levels: *static* and *interactive*. We use a repeated-measures (i.e. within-subjects) structure where we measure participant behavior in two tasks (static, interactive), and with two datasets that are counterbalanced in their assignment across the two tasks. Note that we *did not* counterbalance static/interactive task order because the interactive features necessarily built upon knowledge of the static visualizations. Participant engaged in a 90-minute (recorded) video-conference divided into four parts. The lead author acted as interviewer and began with introductions and informed consent, before facilitating two *EDA Sessions* followed by an interview.

Each *EDA Session* began with an introduction of the (static/interactive) features of the visualization library (*Features Intro*), followed an opportunity for the participant to explore the new APIs via sample code (*Features Tutorial*). Next, participants were given a notebook with a dataset and scenario for an *Analysis Task*, and asked to complete an exploratory analysis in approximately 25 minutes while sharing their thoughts aloud as if they were explaining their work to a junior colleague. The structure of the *Static* tasks were identical. The dataset was counterbalanced across participants. Each session concluded with a semi-structured interview and debrief where participants offered feedback on the visualization library, and were asked about their experience with static and interactive visualizations, and how they conduct exploratory analyses as a part of their occupation. We recruited 16 participants through social media, personal networks, and crowdwork platforms. Two participants were involved in pilot studies to refine data collection procedures. Of the 16 participants who completed the study, three were excluded due to either incomprehensible think-aloud responses, or an insufficient level of Python proficiency. Our resultant pool comprised 13 participants: 4 women, 8 men, and one person who identified as non-binary; participant ages ranged between 27 and 41 years (average age 31). All participants regularly conducted exploratory data analysis using Jupyter notebooks as part of their occupation. Their most common job title was Data Scientist (5), followed by PhD Candidate (3), Software Developer (2), Data Analyst (1), Economist (1), and Statistician (1).

3.0.2 Controlling for Library Expertise with Altair Express

For our research question, while participants were free to use any third-party Python package as part of their data analysis, it was important that they all used the same visualization library in order to facilitate comparisons between participants' behaviors. However, this introduces a confound: participants' existing familiarity and expertise with visualization packages. To control for their prior expertise, we opted to develop a novel visualization package to establish a common baseline of relative novelty for all participants.

Our library, called *Altair Express (ALX)*[1] is a Python-based visualization package that offers a high-level declarative API for specifying interactive visualizations. In contrast to the composable approach of the existing Altair visualization package (and its underlying grammar Vega-Lite [32]), ALX instead provides a *typology* of visualizations and interaction techniques — an approach we chose to reduce specification friction analysts might face during EDA. We surveyed existing Python-based chart typologies (e.g., Plotly Express, seaborn, etc.) and implemented the set of statistical charts we hypothesized to be most relevant to EDA including: `barplot`, `countplot`, `hist`, `jointplot`, `lineplot`, `heatmap`, `pairplot`, `profile`, `scatterplot`, and `stripplot`.

ALX's interaction typology is defined in terms of *effect-action* pairs: an *effect* is the change to the data or encodings that occurs when a user performs an interaction (e.g., showing a tooltip, zooming into a region, etc.); an *action* is the event that triggers the interaction (e.g., clicking, brushing, etc.). ALX's interaction typology comprises: `highlight_brush`, `filter_brush`, `tooltip_hover`, `pan_zoom`, `filter_slider`, `filter_type`, `highlight_color`, and `highlight_point`.

Using the `+` operator, visualization and interaction types can be composed together. For instance, `alx.highlight_brush() + alx.scatterplot(data, x='Weight', y='Horsepower')` produces a scatterplot of the `Weight` and `Horsepower` of cars; users can brush the scatterplot highlighting selected points in blue and dimming the rest to gray. Using `+`, users can add multiple interaction techniques to a single visualization, or concatenate multiple static and/or interactive visualizations together to produce a custom dashboard. ALX implements these interactive visualizations via Vega-Lite [32].

Finally, besides its specification language, ALX implements a handful of features designed to address limitations researchers have identified of using interactive visualizations in computational notebooks [6], [17]. For example, with ALX, analysts can "copy-and-paste" with an interaction technique in order to extract the underlying selection: when a selection is made — for instance, by clicking on a point, dragging a slider, or brushing — the analyst

[1] The name was chosen to mirror the relationship between Plotly and Plotly Express. That is, *Altair : Altair Express :: Plotly : Plotly Express.*

can press `control + c` on their keyboard to copy the pandas query necessary to select the data. This query can then be pasted into the subsequent cells in the notebook to filter down to the selected data for further investigation or charting.

3.0.3 Data Analysis Procedure

We applied an inductive content analysis [33], [34] to the rich stream of video and think-aloud data our participants produced. We split transcripts of the video recordings into discretized units of meaning we call *utterances*. And, using participants' screenshare, mouse gestures, and linguistic prosody, we additionally coded what representations participants used in the process of making a particular utterance. We limited the scope of our coding to only include the *Analysis Tasks* — thus, we excluded utterances participants made when they were familiarizing themselves with ALX's features, debugging, or during the post-interview.

The first and second authors followed an inductive process consistent with the application of grounded theory in HCI [33], [35] to develop a codebook for categorizing participants' utterances. This processes involved eight iterations of independent coding centered on: (1) developing structure, (2) aligning criteria, and (3) reconciling discrepancies. In the final round of reconciliation, the first and second authors independently coded a random sample of 100 utterances, to calculate an Inter-Rater Reliability (IRR) measure of Krippendorf's $\alpha = 0.85$.[2]

[2] Krippendorf's alpha is the recommended IRR metric for multi-code structures where more than one can can be applied to one observation. Using a more generous alternative we calculate reliability of (*Observed Agreement*=0.87). In both cases our IRR passes normative thresholds of reliability [36].

Chapter 4

A Formal Description of EDA Sessions

We express the results of our mixed-methods analysis through the formal description shown in Figure 4.1. We find an EDA session progresses through a sequence of analysis States. Each State can either be a standalone Representation (e.g., a visualization, dataframe printout, etc.) or be a verbal Observation that an analyst makes about a particular Representation. For each representation, we collect a variety of Telemetry data but our analysis focuses only on HoverWindows (i.e., time spans of when a participant hovered over a given representation) — we leave other abstractions that can be derived from telemetry data to future work.

Observations more richly associate Representations with the verbal Utterances — we separately code for RepresentationalUsage to be able to distinguish observations made with static representations from those made with interactive representations, but where the interaction was not used to make the observation. We use the term Utterance rather than *insight* or *inference* to indicate that, even with the context of the participant's screenshare, mouse gestures, and linguistic prosody, we cannot precisely determine the participant's state of knowledge. Thus, we work to interpret as much of each utterance's semantic content as possible via our qualitative coding procedure. As Figure 4.2 shows, this procedure yielded 16 types of Utterances spread across four categories: utterances about the overall Dataset including its size, orientation, quality, provenance, and metadata; utterances about individual Variables including about the distribution of data values (e.g., mix, max, outliers) and the shape of this distribution; utterances about Relationships between variables including whether any relationship exists and, if so, what form, strength, and direction this relationship takes; and, finally, utterances about the overall analytic Process including statements about intended next steps or remarks about representations that are not about depicted data.

We find this formalism offers us unique affordances when applied to analyze EDA activity. Consider the following vignette which is inspired by behavior we witnessed our study participants engage in:

> Ada, a professional data analyst, is tasked with investigating a customer purchase behavior dataset that includes *customer age*, *purchase history*, *product categories*, and a *customer satisfaction rating* (scaled 0-10). Ada creates a data profiler, a multiview visualization with concatenated univariate histograms. While examining the distribu-

> tions, she spots missing values for *satisfaction ratings*. Using a crossfilter interaction, she brushes over this region of missing data and observes a slight shift towards a higher average *customer age*. Intrigued, Ada hypothesizes that older customers might be less likely to provide satisfaction ratings, and creates a scatterplot of *satisfaction ratings* vs. *customer age*. This scatterplot reveals that there is indeed a cluster of older customers with missing ratings data. Ada isolates this cluster with a brush, and examines the associated customer details in a table, noting that a significant portion of these customers purchased products in a specific *product category*.

Using attribute-based metrics [7], [15], we might view Ada's EDA as a three-step process: analyzing *all* attributes with the profiler; then analyzing *age* and *rating* specifically with the scatterplot; and finally, returning to *all* attributes with the data table. While helpful, this approach makes it difficult to identify that Ada did not ever actually analyze particular attributes (e.g., *purchase history*) despite their inclusion in certain representations (i.e., the profiler and data table). Moreover, by being representation-agnostic, attribute-centric metrics treat the profiler and data table as functionally equivalent and, as a result, miss nuance around the different ways Ada used these two views — for instance, that she brushed the profiler view to reveal a relationship between *age* and *satisfaction* versus examining the table in a more record-by-record fashion. These issues are compounded when applying attribute-centric metrics to analyze interactive visualization as the space of possible observations is greatly expanded [37].

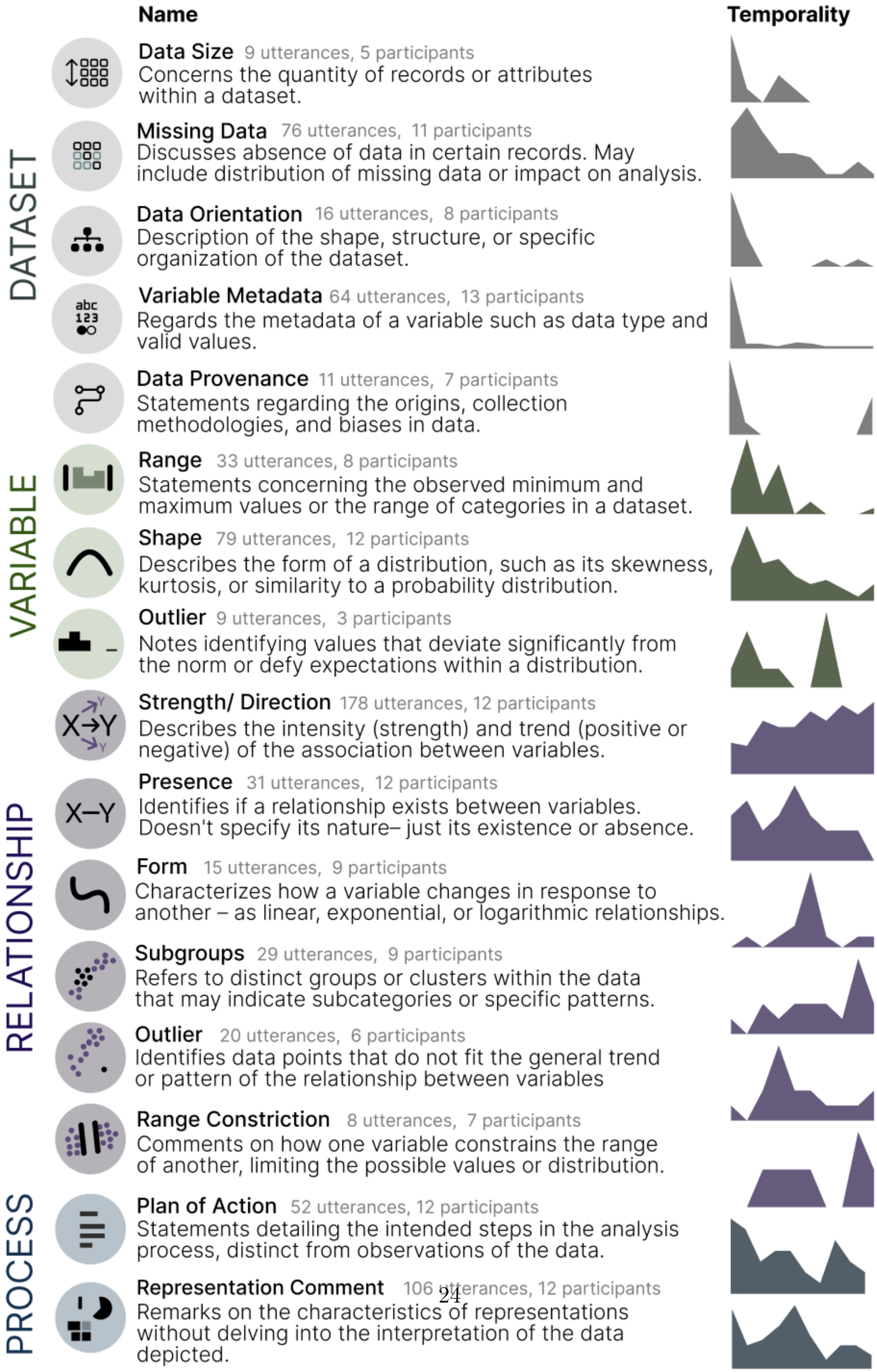
Task- and some insight-based methods often do not account for representation either. As a result, they ignore analytic expressions that are not verbalized and instead latently conveyed via the representation — that is, the act of making a chart is intrinsically an inquiry, even if it is not used to make an observation out loud. Moreover, depending on the granularity of task/insight codes, these methods may miss important nuance in Ada's activity. For instance, with the protocol followed by Zgraggen et al. [31], one might label Ada's analysis as a series of *Distribution Shape* insights followed by two *Correlation* insights — a strategy that collapses insights about "clusters" and "correlations" together. More recent insight-based approaches, such as the formalism developed by Battle & Ottley [38], begin to address many of these shortcomings — for instance, they formalize an `AnalyticKnowledgeNode` to encompass data relationships and transformations. While this method would be able to capture much of Ada's activity (e.g., interactive brushing as issuing a series of data queries), it is focused only on describing the *quantitative insights* a participant might make about a dataset.

In contrast, our formalism separately records the representations Ada constructed, the utterances she verbalized, and links the two sets together as a series of observations. Thus, according our formalism, Ada's analysis session would be represented as in figure 4.3.

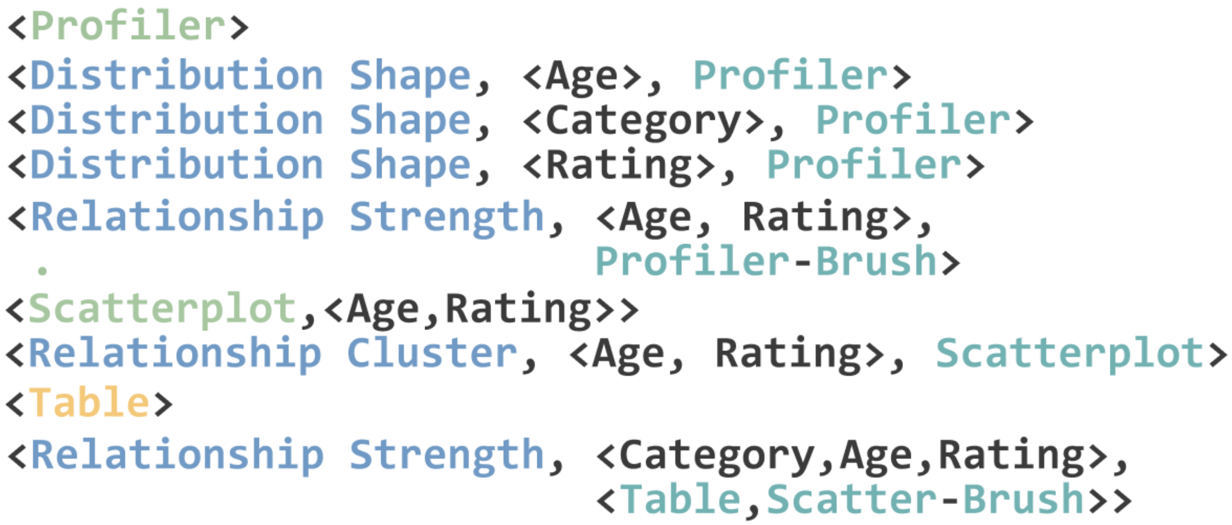
As we see, our formalism better reflects the situated nature of EDA — that observations (or tasks or insights) occur *with* representations, and that non-verbalized representations can play important roles in an analysis session. As a result, our formalism is able to surface a different set of patterns in analyst behavior than traditional methods including detecting usage rates of visualization types (§ **??**), identifying a visualization's role within analysis planning (§ 6.0.2), and revealing the 80-20 rule of EDA (§ 6.0.2).

```
State := <StateType, Timestamp>
StateType := Output | Observation

Output := <Representation, Telemetry>
Representation := Visualization | Dataframe | ValueCount |
                  CodeCell | Column | Info | Describe
Telemetry := HoverWindow[]
HoverWindow := <StartTime, StopTime>

Visualization := <Chart, Interaction{}>
Chart :=  <ChartType, Encodings{}>
ChartType :=  Scatterplot | Countplot | Profile | Pairplot |
              Barplot | Histogram | Lineplot | Stripplot |
              Heatmap | MultiView
Encodings :=  <Channel, Attribute>
Channel :=  x | y | color | ...
Interaction := <Effect, Action>
Effect := Highlight | Group | Filter | PanZoom | Tooltip
Action :=  Hover | Click | Drag | Type

Observation := <Utterance, RepresentationUsage[]?>
Utterance := <UtteranceType, Attribute[]?>
UtteranceType := Dataset | Variable | Relationship | Process
Dataset := Data Size | Missing Data | Data Orientation |
           Variable Metadata | Data Provenance
Variable := Range | Shape | Outlier
Relationship:= Strength and Direction | Presence | Form |
               Subgroups | Outlier | Range Constriction
Process := Plan of Action | Representation Comment
Attribute := DataAttribute | DerivedAttribute

RepresentationUsage := <Output, InteractionUsed>
InteractionUsed := TRUE | FALSE

revisitRate(HoverWindow) := COUNT(HoverWindow)
hoverTime(HoverWindow) := SUM((StopTime - StartTime)[])
representationDiversity(Session) := UNIQUE(Representation)
representationVelocity(Session) := COUNT(Output)/Duration
```

Figure 4.1: A formal definition of EDA sessions in terms of analysis states that comprise either a representation alone (e.g., a visualization, dataframe output, etc.) or an observation made with one or more representations. Italics indicates terminal symbols.



	Name	Temporality
DATASET	**Data Size** 9 utterances, 5 participants Concerns the quantity of records or attributes within a dataset.	
	Missing Data 76 utterances, 11 participants Discusses absence of data in certain records. May include distribution of missing data or impact on analysis.	
	Data Orientation 16 utterances, 8 participants Description of the shape, structure, or specific organization of the dataset.	
	Variable Metadata 64 utterances, 13 participants Regards the metadata of a variable such as data type and valid values.	
	Data Provenance 11 utterances, 7 participants Statements regarding the origins, collection methodologies, and biases in data.	
VARIABLE	**Range** 33 utterances, 8 participants Statements concerning the observed minimum and maximum values or the range of categories in a dataset.	
	Shape 79 utterances, 12 participants Describes the form of a distribution, such as its skewness, kurtosis, or similarity to a probability distribution.	
	Outlier 9 utterances, 3 participants Notes identifying values that deviate significantly from the norm or defy expectations within a distribution.	
RELATIONSHIP	**Strength/ Direction** 178 utterances, 12 participants Describes the intensity (strength) and trend (positive or negative) of the association between variables.	
	Presence 31 utterances, 12 participants Identifies if a relationship exists between variables. Doesn't specify its nature– just its existence or absence.	
	Form 15 utterances, 9 participants Characterizes how a variable changes in response to another – as linear, exponential, or logarithmic relationships.	
	Subgroups 29 utterances, 9 participants Refers to distinct groups or clusters within the data that may indicate subcategories or specific patterns.	
	Outlier 20 utterances, 6 participants Identifies data points that do not fit the general trend or pattern of the relationship between variables	
	Range Constriction 8 utterances, 7 participants Comments on how one variable constrains the range of another, limiting the possible values or distribution.	
PROCESS	**Plan of Action** 52 utterances, 12 participants Statements detailing the intended steps in the analysis process, distinct from observations of the data.	
	Representation Comment 106 utterances, 12 participants Remarks on the characteristics of representations without delving into the interpretation of the data depicted.	



```
<Profiler>
<Distribution Shape, <Age>, Profiler>
<Distribution Shape, <Category>, Profiler>
<Distribution Shape, <Rating>, Profiler>
<Relationship Strength, <Age, Rating>,
 .                        Profiler-Brush>
<Scatterplot,<Age,Rating>>
<Relationship Cluster, <Age, Rating>, Scatterplot>
<Table>
<Relationship Strength, <Category,Age,Rating>,
                        <Table,Scatter-Brush>>
```

Figure 4.3: Example of Ada's analysis session encoded in our formalism. For clarity, we have omitted some levels of nesting for the formal description of this example

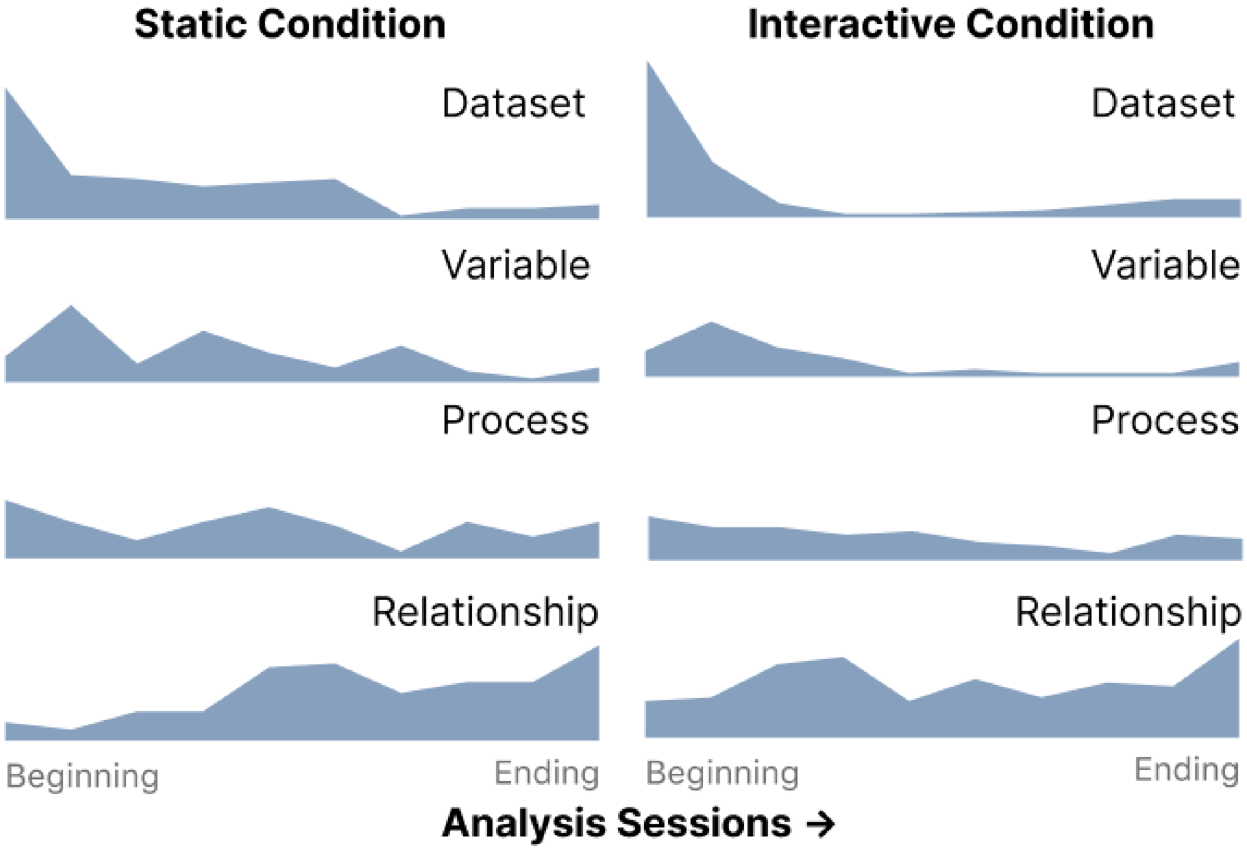


Figure 5.1: Frequency of high-level utterances categories over time across static and interactive charts.

Chapter 5

Characterizing Analyst Utterances

In this section, we report the temporal and sequential patterns we identified with participant Observations.

5.0.1 Temporal Patterns

As the area charts in Figure 5.1 and Figure 4.2 show, we find that while analysts' processes align *in aggregate* with traditional, linear EDA models (from dataset exploration to individual variable analysis and then relationship exploration [21]), the analysis process is both more fluid and sensitive to interactivity than rigid interpretations of those models would suggest. To examine analyst processes, we calculated the median moment through the analysis session (expressed as a percentage) in which analysts made Observations across our four UtteranceTypes: Dataset (13.43%), Variable (25.60%), Relationship (56.86%), and Process (40.18%).

In particular, interactive EDA sessions prompted earlier observations about Relationships in the data (IQR 28%-75% through a session) compared to static EDAs (IQR 43%-85%). We hypothesize that the use of interactive profilers, featuring cross-filterable univariate histograms, encouraged analysts to explore relationships sooner. Our subsequent findings of analysts switching from static to interactive profilers (§ 6.0.3) support this: many participants shifted from Variable to Relationship utterances almost immediately upon encountering the interactive profiler. This finding opens questions about whether the affordances (or presence!) of interactive profilers enables bypassing distribution analysis, and whether we can articulate the tradeoffs of such process changes. More broadly, the presence of relationship utterances across both static and interactive EDA sessions suggests that analysts are willing, perhaps even eager, to explore Relationships before fully developing a mental model of individual Variables.

5.0.2 Sequential Transitions

During their analyses participants made seven different types of utterances on average. Looking at the sequential transitions between utterances reveals a number of common analysis motifs [9].

Tour-Driven Exploration (Fig 5.2-1): Frequent self-transitions between similar utterance types (e.g., multiple consecutive utterances focused on Relationship strength) suggests that analysts often adopt a systematic "touring" approach during EDA. This finding aligns with concepts of univariate and bivariate tours **Lux**, [9], where analysts methodically explore specific aspects of individual Variables and their Relationships. However, we observed self-transitions extending beyond Relationship analysis to include utterances about `Missing Data` and `Variable Metadata`. This suggests that "touring" behaviors are broader than previously described [9].

Column- vs. Row-Centric Missingness (Fig 5.2-2): The most common transition between utterances types was moving from `Missing Data` to `Distribution Shape`. The design of profilers presented missing data alongside the columns data distribution subtly promotes a column-centric view of missingness. However, as a counter-example, P10 investigated missingness as a characteristic of individual data records (rows), skipping the profiler entirely. Visualizing the missingness per record on a scatterplot, he commented *"... most of the rows have no missing columns, and then they progressively have more and more. So I guess, depending on what the analysis we're gonna do is, we may or may not exclude data points."* This approach affords thinking about a different set of causes for missingness in the

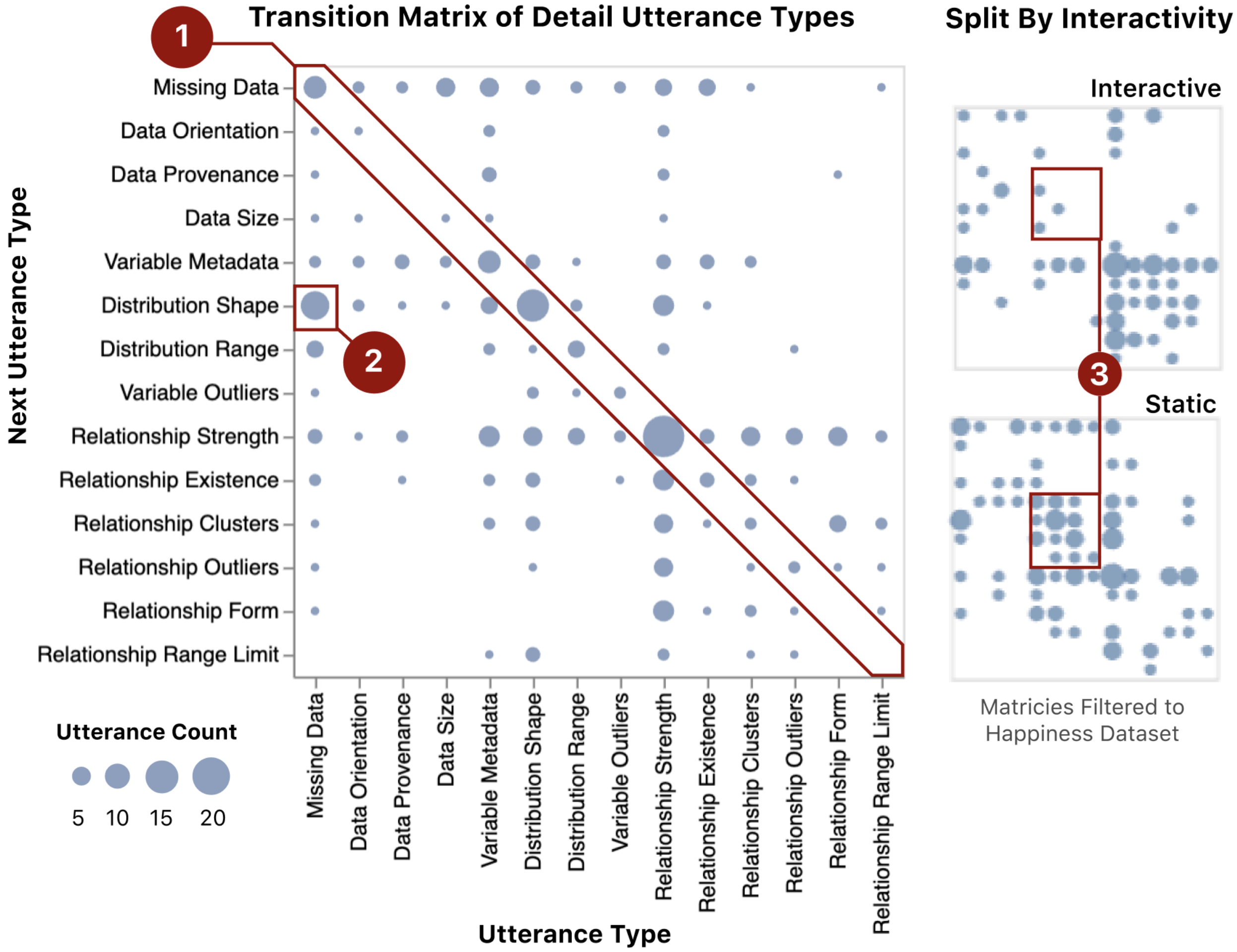


Figure 5.2: Left, transition matrix of sequential utterances; Right, happiness data transition matrix by Condition [Interactive, Static]. (1) The diagonal transitions represent repeated utterances of the same semantic type. (2) The large amount of `Distribution Shape` utterances after a `Missing Data` utterance, show how univariate profilers prompted consideration of missing data at a column instead of a row level view. (3) *The Variable Gap*, whereby participants inside of the interactive condition with the happiness dataset tended to skip over distribution characterization.

data generation process and raises a design question: how might we design profile representations that encourage both column-level and row-level consideration of missing data?

The "Variable Gap" and Interactive Profilers (Fig 5.2-3): In the happiness dataset, many participants skipped characterizing Variables, instead immediately focusing on Relationships. This caused a *Variable Gap* between conditions, visible in the transition matrices. This shift in focus often coincided with the use of interactive profilers. For example, participant P5 initially followed a variable-first pattern in her static analysis, narrating out 6 distributional utterances about her variables using the profiler. Upon beginning her interactive analysis, she immediately began making relationship utterances by cross filtering on the profiler view (see § 6.0.3 for more information).

Chapter 6

Characterizing Representations and Usage

We now turn to describing which Representations were constructed, and how RepresentationUsage impacted Observations.

6.0.1 Temporality, Diversity, and Velocity

Across all Sessions our participants constructed a total of 1169 Outputs and, on average, an individual analyst constructed 44 different outputs. The most common output was the execution of an arbitrary piece of Python code, followed by Visualizations. Looking at when participants made these outputs (Fig. 6.1-left) reveals interesting temporal patterns. Code cells were used most frequently near the beginning and end of analysis sessions and often involved functions that checked the measure of central tendancy of attributes. Starting approximately 15% of the way through their sessions, participants began to switch to favoring visualizations — these representations were then used to form the bulk of their subsequent observations. Based on this data, we compute two metrics of representation construction: representationDiversity, which is the number of *unique* representations constructed in an analysis session; and, representationVelocity, or the rate at which representations were constructed over course of an analysis session. As Figure 6.1 shows, these metrics are moderately correlated (Pearson's $r = 0.47$); we discuss their role within analysis sessions in a subsequent section (§ 6.0.4).

Examining the intersection of ChartTypes and Observations (Fig. 6.1(right)) reveals expected and surprising usage patterns. For instance, unsurprisingly, scatterplots were most frequently used to make utterances about Relationships while profiler views helped participants make Variable utterances. However, as this heatmap shows, participants would frequently use charts beyond their intended purposes or in ways that break with best practice. For instance, Variable utterances constituted only 42% of observations made with profiler views — even though, ostensibly, this is the core purpose of a columnar distribution of data values. Similarly, in contrast to visualization theory and visualization recommender systems, which emphasize perceptual effectiveness, participant P9, a data science instructor, specifically created a representation she called a *"spaghetti plot"* — a line chart with 180 different series overplotted. Ahead of creating the chart she commented *"It's going to be a bad idea"*,

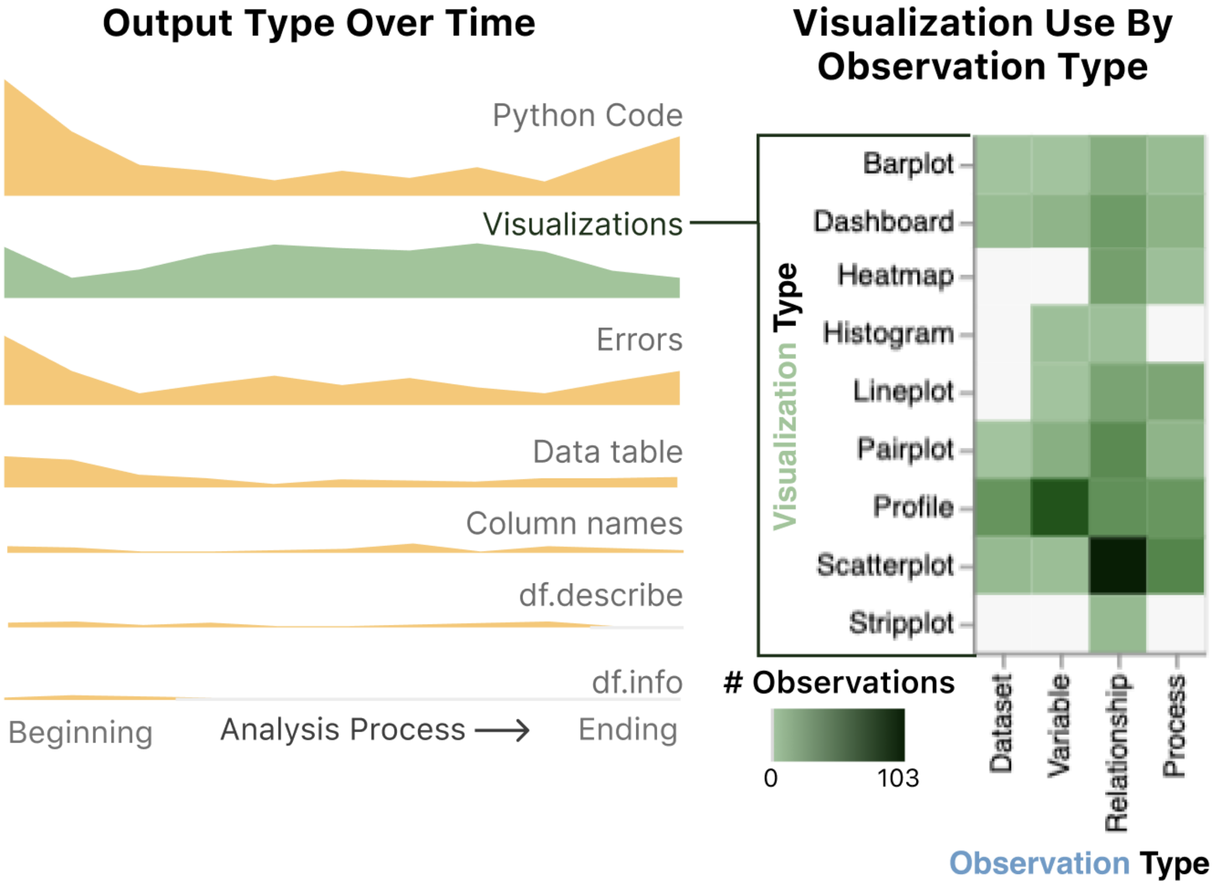


Figure 6.1: Left: Representation creation over time. Right: Heatmap of the number of times different Visualizations were used to make an Observation, according to UtteranceType.

but persisted precisely because she wanted to ensure that the plot itself *was ineffective*, as a gut check.

6.0.2 Hover Patterns and Observations

We can also connect Representations and Observations through Telemetry data, by calculating two metrics: hoverTime, the total time an analyst spent hovering over a representation; and, revisitCount, or the number of times an analyst hovered over a particular representation.

The '80-20 Rule': Why Some Visualizations Matter More

Our analysis reveals a 80-20 threshold in how participants use representations during EDA. The top 20% of most frequently hovered representations (*top-20*) accounted for 79% of total hoverTime and 75% of observations. Representations in the *top-20* had hover durations of at least 30 seconds and an average of 2.8 Observations each, indicating deep engagement. In contrast, the bottom 80% of representations (*bottom-80*) saw significantly less use, with an average of just 0.2 observations per representation. We identify two key differences

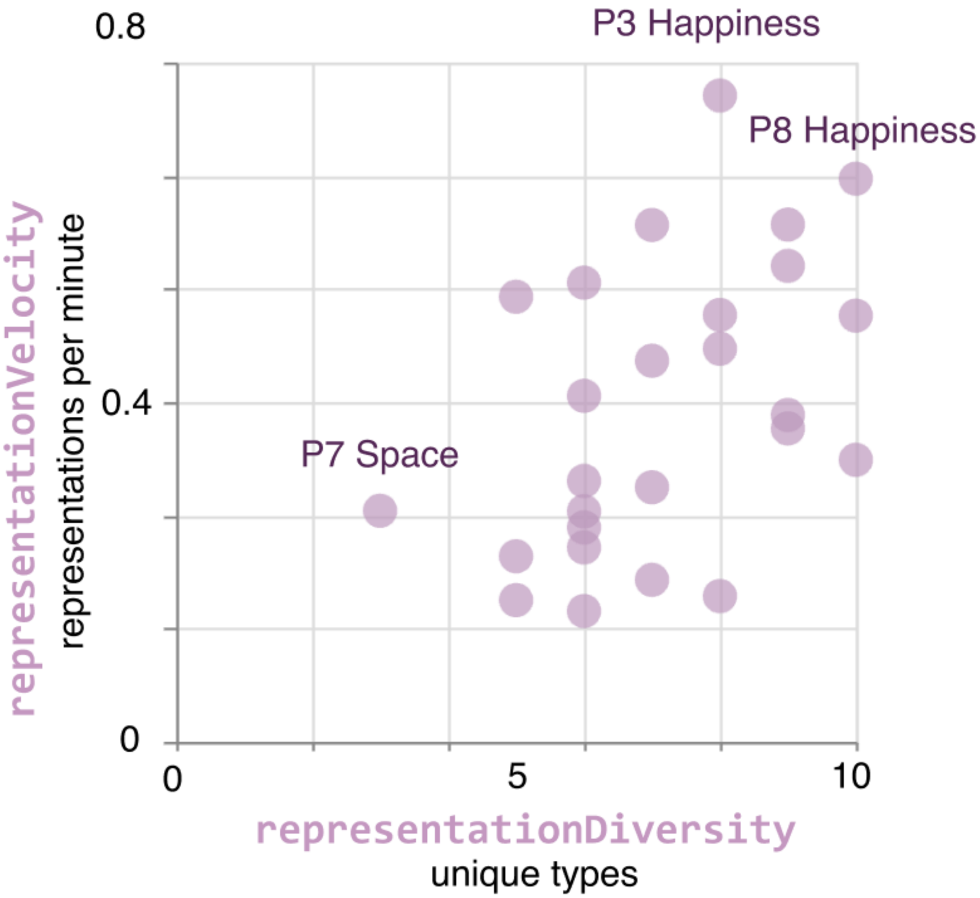


Figure 6.2: Scatterplot of computed representationDiversity and representationalVelocity for each analysis session.

between these two sets that sheds light on analyst preferences: the ability to encode multiple attributes simultaneously, and the role of interactivity.

Representations displaying information about multiple variables simultaneously (e.g., profilers, correlation heatmaps, pairplots) were more common within the *top-20*. These *all-attribute representations* made up only 2% of the *bottom-80* but constituted 22% of the *top-20*, an 11X increase. Analysts engaged with these visualizations through the "touring" behaviors we previously described in § 5.0.2, systematically exploring them and commenting on different variable combinations approximately every 5-15 seconds. This is underlined by the longest average hover times observed for all-attribute visualizations (67 seconds for profiles, 75 seconds for heatmaps, and 169 seconds for pairplots). Similarly, we see a marked decrease in hoverTime with `Code Cells` used for quick statistical checks (from 48% of the *bottom-80* to 9% of *top-20*, averaging 4.9 seconds of hovering per representation).

Interactive visualizations were more prevalent within the *top-20* (24% of the *top-20* vs. 16% of the *bottom-80*). Analysts particularly favored the `highlight_brush` as it enabled cross-linking data subsets across multiple charts. This technique was used in over 56% of interactive representations in the *top-20*, compared to 37% in the *bottom-80* Similarly, the `filter_brush` technique, which filters out all non-selected data marks from view, was used in 30% of the interactive scatterplots found within the *bottom-80*. However, `filter_brush` went to 2% in the *top-20*, a likely side effect of filtering obscuring important context in

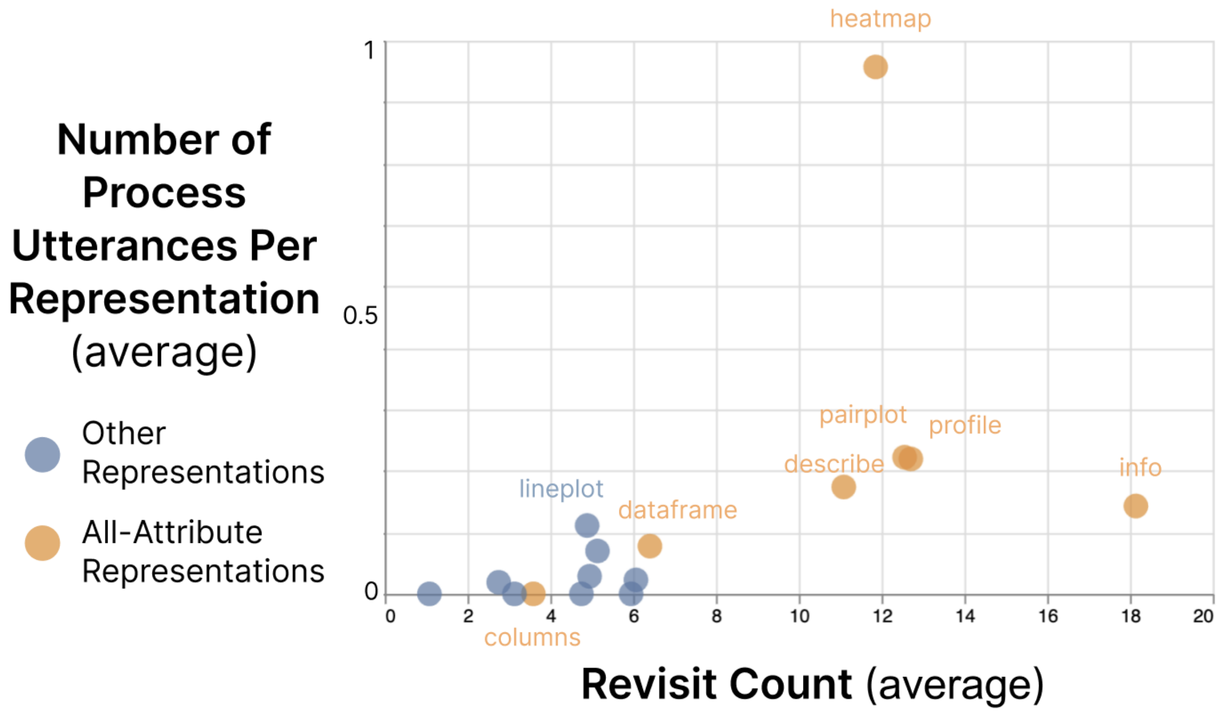


Figure 6.3: Average revisitCount and count of `Plan of Action` utterances by Representation. Representations are colored by whether or not it is an *all-attribute representation*. Representations to the left are typically one-off question-answering tools whereas representations to the right are frequently revisited when deciding analysis paths.

standalone charts.

Finally, `pan_zoom` interactions, was the second most popular interaction technique in *bottom-80* – present in over 31% of interactive representations. However, declined to just 18% in the *top-20*. Analysts consistently struggled to find effective use for pan-zoom interactions, suggesting a lack of intuition for how this technique could enhance their analysis workflows. Out of the 16 instances in which pan-zoom was used, we observed only one instance where it successfully uncovered an insight that would have been difficult to obtain otherwise. In this case, participant P10 zoomed into a dense, overplotted region of a scatterplot to gain more resolution, and was able to reveal a pattern in the depicted data. However, even this success story was marred by discomfort — P10 added pan-zoom to a set of horizontally arranged scatterplots that shared a common y-axis; thus, the coordinated scrolling of all scatterplots made him feel disoriented, prompting him to request *"can we turn that off?"*

All-Attribute Visualizations Aid Planning

As Figure 6.3 shows, high revisitCounts (over 10 times) indicate that a representation serves as an process planning tool, aiding analysts in orienting themselves and preparing their next actions. For example, participant P5 created a correlation heatmap to identify the most strongly correlated attributes within her dataset. She frequently returned to this visualization, using it as a guide for selecting specific attributes to investigate further: *"let's look at the one that is most positively correlated, which seems to be log GDP per capita. So*

I'll start with that variable". This led her to generate scatterplots and custom dashboards for deeper exploration. This approach proved effective, leading to 23 observations. Such action-planning is not restricted to only visual all-attribute representations — participants frequently revisited data frame outputs (including `df.describe`, `df.info`, and simply the tabular preview output) to formulate their plans. For instance, P11 read through the individual values of a dataframe printout, commenting: *"Of course, we cannot say for the whole thing [based on just the shown rows]. So my strategy will be like going through each of the variable here, and do the summary statistic."* Looking across all Observations tuples in our dataset, all-attribute representations are associated with `Plan of Action` utterances at a rate of 5x more than other representations.

6.0.3 An Interactive Draw Towards Complexity

When using interactive visualizations, we observed shifts in the types and number of attributes that analysts considered. An *attribute addition* behavior appears prominently in analysts who are progressively exploring data relationships of escalating complexity, moving utterances from univariate distributions to bivariate relationships, and further morphing into multivariate analyses. For example, in the static condition participant P5 used the profiler visualization to analyze the univariate distributions of her columns, making 6 utterances about their distributions. At the beginning of the interactive session, she created an interactive version of the profiler visualization, and immediately began using it to analyze relationships — brushing on the chart to examine a target population and generating 6 new utterances about that population's relationship to variables. This pattern of behavior persisted across datasets for other participants (Fig. 6.4 (left)).

We also observed shifts in behavior prompted by filtering interactions in scatterplot (Fig. 6.4 (right)). Prior to the interactive session, we observed participants discussing bivariate relationships using scatterplots; however, when interaction was added, their utterances tended to focus on the multivariate relationships. Multiple participants used brushes to extract subsets from data clusters and pursued analysis paths to differentiate that cluster from the rest of the data. Another case of this was the use of the `filter_slider`, an interaction technique which filters the chart to only the data value present in a particular value on a slider query widget. The shift we observe between these interactive and static charts presents the allure of interactive representations, seemingly pulling analysts towards investigating more complicated relationships even when those interactions are not actively being used.

However, attribute addition behavior was not observed equally across data types. On average, our participants used interactive visualizations for multivariate (often all continuous variables) and continuous x continuous bivariate relationships (Fig. 6.4 (right)). However we note the overall patterns are most salient at the aggregate level and the participant level contains sparsity in the utterances made for each participant for a given data type. Thus while we chose to report the results to fully describe the behavior that we saw, such descriptions warrant additional investigations to understand the role that interaction may play in drawing analyst hypotheses towards more multivariate and complex relationships and if such patterns exist over longer periods of time.

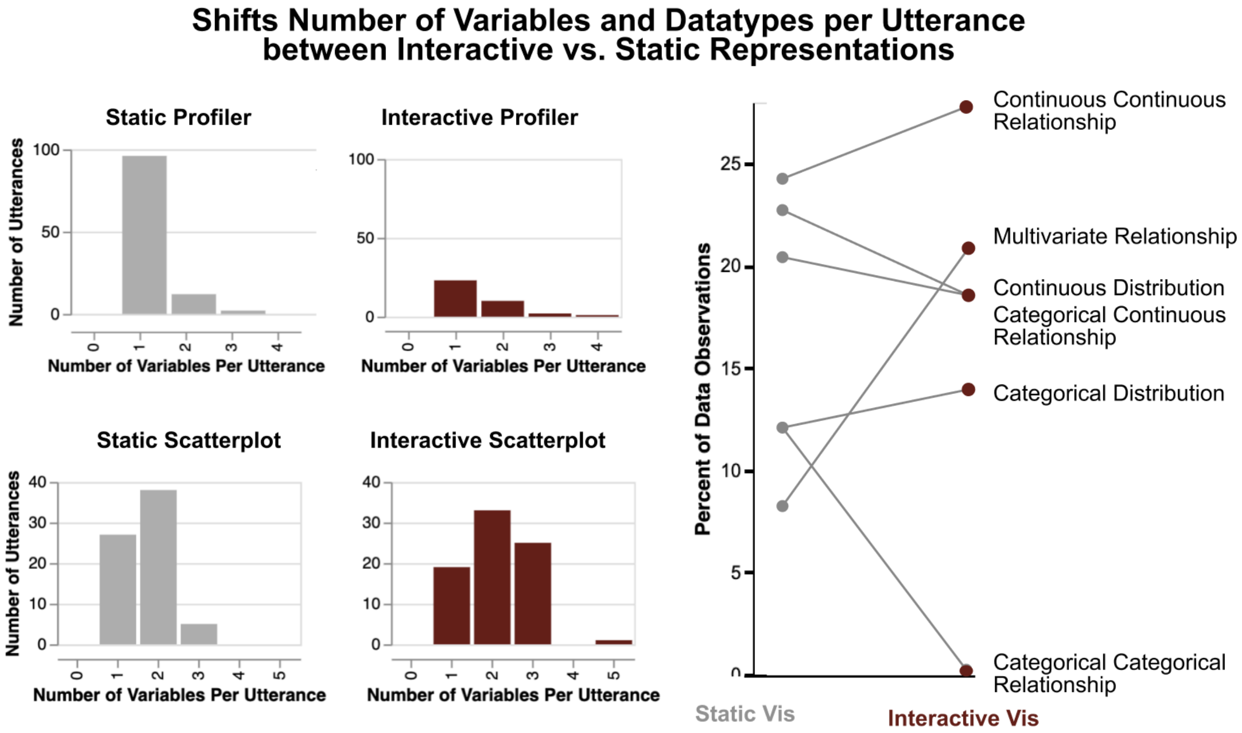


Figure 6.4: Left: Numbers of Utterances between static and interactive versions of the profiler and scatterplot visualizations. The most common interaction techniques used for these representations were cross-filter brushes, tooltips, and filtering sliders. Right: Comparison of which data types were discussed in utterances in static vs. interactive visualizations. The plot indicates that there is a steep rise in multivariate relationship utterances in interactive visualizations

6.0.4 Patterns of Broad Analysis Space Exploration

Previous studies of EDA offered characterization of analysis sessions based on the number of attributes analysts reasoned about [7], [15]. We extend this type of analysis, applying it to our definition of State which encompasses the Representations analysts constructed, and how RepresentationUsage affects Observations analysts made. Follow Battle et al.'s method [15], we created binary histograms that represent whether or not our participants made a particular observation (e.g. observed the relationship between *happiness* and *GDP*). Using the total count of observations made by our participants, we can compute for each participant what percent of total states visited, allowing us to rank our participants by the breadth of exploration. For example, participant P9, a data science instructor, made the most extensive Dataset observations across both static and interactive conditions (Fig. 6.5 (1)). These observations occurred as P9 began each of her analysis sessions with a variable metadata tour: systematically going through each attribute in the data dictionary, spending time discussing what the variable meant and her opinions on its usefulness. Similarly, we observe the 5 participants who made the most Variable utterances (Fig. 6.5 (2)) did so in the static condition using profile visualizations.

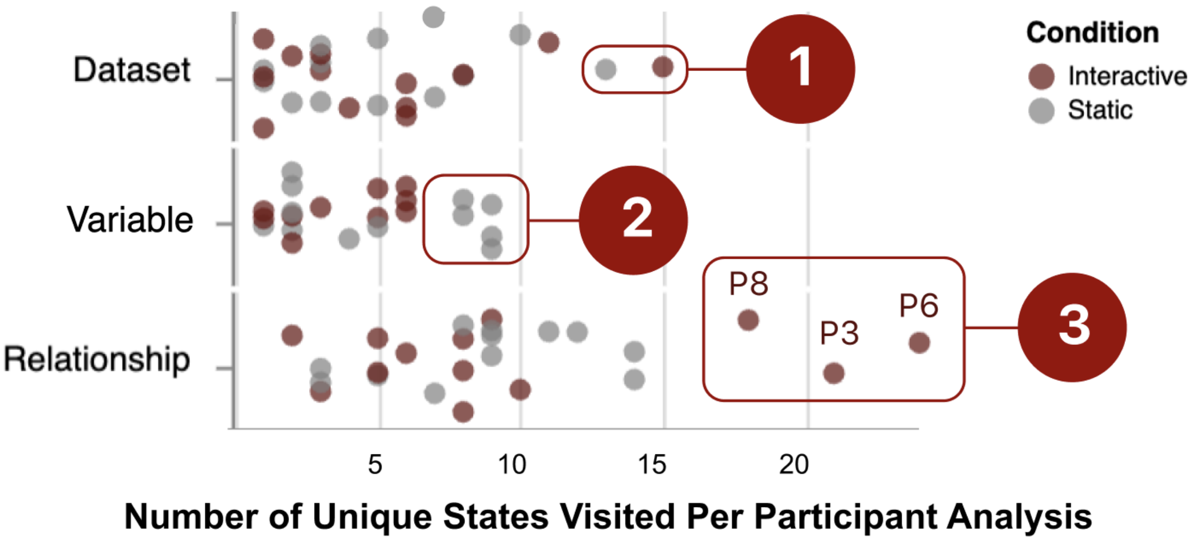


Figure 6.5: Stripplot of % of total unique Observations visited per analysis session, broken down by high level type and colored by Analysis Condition (Interactive or Static).

In contrast, approaches for exploring a broad set of Relationship observations (Fig. 6.5 (3)) reveals a diverse set of strategies. To investigate these patterns of exploration, we created attribute co-occurance heatmaps (Fig. **??**) to "fingerprint" and explain these strategies:

P8: Parameterized Search. Driven by a clear goal and an aversion to *"mindless"* exploration, P8 adopted a systematic, iterative approach reminiscent of a parameterized search through Representations and Encodings. She explored the analysis space by cycling through which attributes were mapped to encodings (e.g., `scatterplot(y=happiness, x=column[index])`), methodically investigating potential relationships between each attribute and the outcome variable. When she encountered specific patterns of interest, she then modified her scatterplot, adding interactions such as brushes and tooltips to investigate outliers and subsets. The resultant fingerprint visualization depicts a focused analysis centered around the outcome variable, with some targeted off-diagonal probes into the country, investigated using tooltips and brushes.

P3: Iterative Deepening. P3's approach was guided by emergent patterns in the data, resembling an iterative deepening search that is commonly used in graph traversal and game-playing algorithms. He generated scatterplots based on his intuition about interesting variable relationships, largely ignoring the outcome variable at the outset. This is reflected in his focus on predictor variables, evident in his thumbprint visualization. Upon noticing clusters within a plot, he investigated their characteristics, iterating through both interaction and encodings (adding tooltips, brushes and color encodings) to identify potential explanatory variables. This behavior is captured in his high representationalVelocity and representationDiversity (each the highest out of all 26 analyses as shown in 6.2), suggesting he wasn't wedded to a single visualization type but rather explored various options to gain deeper insights. This iterative deepening process ultimately led to a scattered thumbprint reflecting his serendipitous journey through attribute space, driven by unexpected findings.

P6: Heuristic-Guided Best First Search. P6 combined a methodical foundation with a responsive, opportunistic approach characteristic of best-first search [cite BestFS].

This approach prioritizes exploring the most promising nodes (observation-analysis states) within a search space based on a pre-defined heuristic. P6's analysis mirrored this approach. Her initial bivariate exploration using a cross-filtered profile set the stage for a more targeted investigation guided by a correlation heatmap. She guided her search through the correlation of attributes with her outcome variable, revisiting the correlation matrix 35 times during her analysis showcasing a high revisitCount. After locating a variable to investigate she would conduct a multivariate analysis using a custom interactive dashboard (made of a scatterplot, profile, and countplot). This approach is reflected in her high revisit count, which captured the number of times she revisited the correlation matrix. This strategy produced a cohesive analysis that investigated both direct predictors and potential confounders of the outcome variable, as evidenced by her targeted analysis along the bottom row and off-diagonal considerations in her thumbprint visualization.

6.0.5 Thinking in the Language of Interaction

In interaction design, perceived affordances [39] signal the operations a user believes are possible within an interface. Well-designed affordances establish interaction dynamics — the rules governing how users interact with the interface. Our study revealed that data scientists reasoned about these dynamics to generate new analytical hypotheses. In other words, they translated "the language of interaction" into novel analytical questions. As participant P8 described: *"My thought of intersecting High GDP and High Life-Expectancy [countries] happened precisely because there was interaction... I was thinking, 'Oh I wonder if multi-select works'... That is actually what led me to think, 'Oh this would also be interesting on an analytical level."'*. Later she commented that such an insight *"would not have occurred to me if not for the fact I was working with an interactive visualization."*

Participant P6's insights emerged from a similar process of experimentation. Having successfully used ALX's copy-and-paste technique to paste filters between charts, he began to consider the broader possibilities this interaction technique offered for filtering his current visualization. While browsing other charts, he stumbled upon a bar plot showing the count of records over time. Intrigued, he initially tested if the copy-and-paste mechanic would function in this context. However, a spark ignited: rather than a simple test of function, he realized it would be more insightful to filter on the most recent years of data. This act of guided experimentation, prompted by the affordances of the interaction design (rather than performing the interaction itself and observing any updates), led him to discover an unexpected trend in life expectancy over time.

These examples suggest that interactive features play a more generative role in analysis than typically acknowledged. While the literature often emphasizes interactions as a means to complete specific tasks, our observations reveal that rules instantiated during interaction design may be reasoned about to inform hypotheses that emerge. First, it makes potential suggestions for interaction design– how do we reify the constraints and rules of interaction dynamics, and in doing so how do different designs impact the reasoning about these rules? Secondly, studies of interaction should also investigate these topics. Rather than simply investigating the impact that visual cues do to influence interaction usage, [40], studies should also investigate how different cues shape how analysts talk about how interaction might be used. By recognizing the interplay between interaction mechanics and analytical cognition, we can pave the way for tools that more effectively partner with the analyst during the discovery process.

Chapter 7

Discussion and Future Work

In this paper, we conducted a qualitative experiment to richly characterize the *situated* nature of EDA. Through a mixed-methods analysis of participant utterances and telemetry data, we developed a formal description of EDA sessions as comprising a sequence of representation constructions or observations. We also record which observations are derived from which representations. We use our formalism to express a dataset of 26 analysis sessions conducted by 13 data science professionals. Our results reveal usage patterns of interactive visualizations like *attribute-addition* or *reasoning in the language of interaction*. Beyond interaction, our results reveal that analysts tend to have a small subset of representations that they use most frequently – a sort of *80-20 rule* for EDA. Finally, using our *formalism* we showcase how metrics like revisitCount, representationalDiversity, and representationalVelocity can be used to understand why certain participants have *broad coverage* during EDA.

7.0.1 Limitations

Although our approach has yielded useful insights about how data science professionals analyze data, we note that studying EDA in a laboratory context poses some inherent limitations. For example, think-aloud protocols may artificially structure thought processes that are more fluid in unobserved settings (e.g., participants may prioritize tasks that are easier to articulate and overlook more complex tasks) [41]. However, in comparison to post hoc reflections, thinking aloud provided *in situ* insights that captured important nuance, and aligns with approaches used in other studies [42].

Moreover, the 25-minute time limit per analysis may have limited the range of analyses participants chose to engage in. This time limit follows the design of prior visualization studies [7], [15], [17], and reflects a delicate balance in study design: longer sessions risk reducing engagement and may limit participation to only those who can allocate a prolonged period of time outside of their daily work. Research has shown that data scientists often encounter time-sensitive tasks in their work[22] and, in practice, we did not abruptly cut participants off. Thus, on average, participants took 29-minutes to complete an analysis.

Finally, being forced to use a new visualization library inevitably presented challenges to our users (and ALX's interactive capabilities may have also contributed a novelty effect if a participant was more used to constructing only static visualizations as part of their regular work). We sought to mitigate these effects in two ways. First, we allocated 20-minutes of the overall study to demonstrations and tutorials of the library. Second, ALX was intentionally designed as a visualization and interaction *typology* (as opposed to a more composable grammar) to minimize specification difficulty — with the terms of the two typologies designed to mirror common visualization and interaction design patterns. Our participants indicated that they use several of these patterns frequently. Perhaps more importantly, we argue that our approach of introducing a new visualization library brings the advantage of controlling for participant expertise: they did not bring any prior tool-specific habits that would have confounded our ability to compare their analysis sessions. Nevertheless, these analysis sessions reflect a "first-use study" commonly found in studies of EDA activity [7], [9], [17], [31].

7.0.2 Implications for EDA Tool Design

Our results suggest several opportunities for interactive visualization tooling to better support EDA. For instance, several of our participants engaged *touring* to systematically explore the data (§ 5.0.2). Yet, existing tools provide poor support for such activity, largely leaving analysts to drive interactions based on their priors and hypotheses they may wish to answer. Akin to visualization recommender systems [43], novel EDA tooling might instead leverage nascent grammars [44] to systematically enumerate the space of hypotheses that can be interactively reached with a given visualization, and proactively suggest particular analysis paths. By leveraging information scent [45], such tools could help analysts think more deeply in the *language of interaction* (§ 6.0.5) — that is, even if an analyst did not adopt a suggestion for an interactive path, the suggestion itself may prompt them to think in different ways.

Relatedly, we found our participants' use of visualizations as *action planning aids (§ 6.0.2)* striking. In computational notebooks, where visualizations are linearly presented, several participants were willing to pay a "scrolling tax" to reach these representations. While some research systems have explored mechanisms for making such representations more readily available (e.g., B2 stitches a visual analytics dashbaord alongside a traditional linear notebook view [17]), our results suggest a wider opportunity. In particular, while the research literature has identified the merit of overview+detail or focus+context techniques, few visualization libraries support them out-of-the-box. When they do, these techniques are supported in relatively limited ways (e.g., when panning/zooming a scatterplot or map). Our results suggest the need for more generalized support for wayfinding — especially to coordinate multiple separate visualizations. Here, we find the *interaction snapshots* developed by Yifan Wu et al. particularly promising [46].

Finally, the prevalance of Process utterances throughout the analysis sessions illustrates that participants engage in a level of metacognition — that is, thinking about their own thinking. How might interactive visualization and visual analysis tooling better support process reflections that seamlessly span visualization creation, interaction design, written code, and statistical output? Drawing on research in distributed cognition [47], we imagine that analysts might engage in valuable self-reflection when presented with displays of their analysis histories. Recent systems such as Lumos [48] have begun to explore this prospect, and we believe there is a rich research space to explore here [49]. For example, what constitutes a significant point in the analytical journey for retrospective purposes? Our formalism suggests that Observations and Representation creation are such key moments, but analysts may believe otherwise when reflecting on their own activity.

7.0.3 Towards Richer Methods for Studying Interactive Analysis as Situated Activity

Our work was motivated by a desire to study interaction as *situated activity* — that is, involving human analysts working in a particular context, externalizing their cognition through visual representations, and interactively making observations with them. While valuable, we believe this paper takes only an initial step towards this approach. To complement recent work that looks to scale-up our ability to study interaction (e.g., through benchmarks [50] and novel systems [10], [51]), we advocate for methods that allow us to study it *more closely*.

We find methods from sociolinguistics and linguistic anthropology used to analyze interpersonal interaction particularly compelling. For instance, discourse and conversational analysis [52] involves a meticulous examination of conversation transcripts, and has been used by researchers to make fundamental linguistic discoveries such as turn-taking [52]. While visualization researchers are beginning to draw on such linguistic theories to inform interaction design guidelines [53], we believe there is a ripe opportunity to adapt them for analyzing interactive behavior as well. For instance, the development of a specialized notation system was particularly crucial to the success of conversational analysis — allowing researchers to annotate linguistic features such as prosody, tone, pitch, pauses, and gaze. What would an equivalent notation for analyzing interaction look like? Similarly, systems for conversational analysis enable flexible definitions of analytic units and abstractions. In contrast, existing interaction provenance systems [54] largely follow a dichotomy of either low-level event logs (e.g., mouse movements, clicks, etc.) or high-level semantically meaningful events (e.g., filter, explore, etc.) — future systems must grapple with how to support more fluid analysis between these levels. Finally, as our study demonstrates, to "closely read" interactive behavior requires capturing a rich multimodal data streams. Simply concatenating and visually linking these streams together risks introducing ambiguities in understand the precise sequences and potential causal relationships between measures. Rather, akin to systems like ChronoViz [55], we envision future systems offering richer juxtapositions of this multimodal data.

References

[1] J. Heer and B. Shneiderman, "Interactive dynamics for visual analysis," en, *Communications of the ACM*, vol. 55, no. 4, pp. 45–54, Apr. 2012, ISSN: 0001-0782, 1557-7317. DOI: 10.1145/2133806.2133821. URL: https://dl.acm.org/doi/10.1145/2133806.2133821 (visited on 07/05/2023).

[2] J. van Wijk, "The value of visualization," in *VIS 05. IEEE Visualization, 2005.*, Oct. 2005, pp. 79–86. DOI: 10.1109/VISUAL.2005.1532781.

[3] J. Thomas and K. Cook, "Illuminating the Path: Research and Development Agenda for Visual Analytics," National Visualization and Analytics Center, Tech. Rep., 2005.

[4] A. Mosca, A. Ottley, and R. Chang, "Does Interaction Improve Bayesian Reasoning with Visualization?" en, in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan: ACM, May 2021, pp. 1–14, ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445176. URL: https://dl.acm.org/doi/10.1145/3411764.3445176 (visited on 09/14/2023).

[5] S. Theis, C. Bröhl, M. Wille, P. Rasche, A. Mertens, E. Beauxis-Aussalet, L. Hardman, and C. M. Schlick, "Ergonomic Considerations for the Design and the Evaluation of Uncertain Data Visualizations," en, in *Human Interface and the Management of Information: Information, Design and Interaction*, S. Yamamoto, Ed., vol. 9734, Series Title: Lecture Notes in Computer Science, Cham: Springer International Publishing, 2016, pp. 191–202, ISBN: 978-3-319-40348-9 978-3-319-40349-6. DOI: 10.1007/978-3-319-40349-6_19. URL: http://link.springer.com/10.1007/978-3-319-40349-6_19 (visited on 09/15/2023).

[6] A. Batch and N. Elmqvist, "The Interactive Visualization Gap in Initial Exploratory Data Analysis," en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 278–287, Jan. 2018, ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2017.2743990. URL: https://ieeexplore.ieee.org/document/8017577/ (visited on 12/01/2022).

[7] K. Wongsuphasawat, Z. Qu, D. Moritz, R. Chang, F. Ouk, A. Anand, J. Mackinlay, B. Howe, and J. Heer, "Voyager 2: Augmenting Visual Analysis with Partial View Specifications," en, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver Colorado USA: ACM, May 2017, pp. 2648–2659, ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025768. URL: https://dl.acm.org/doi/10.1145/3025453.3025768 (visited on 02/28/2023).

[8] L. Bavoil, S. Callahan, P. Crossno, J. Freire, C. Scheidegger, C. Silva, and H. Vo, “VisTrails: Enabling Interactive Multiple-View Visualizations,” en, in *VIS 05. IEEE Visualization, 2005.*, Minneapolis, MN, USA: IEEE, 2005, pp. 135–142, ISBN: 978-0-7803-9462-9. DOI: 10.1109/VISUAL.2005.1532788. URL: http://ieeexplore.ieee.org/document/1532788/ (visited on 03/23/2024).

[9] A. Kale, Z. Guo, X. L. Qiao, J. Heer, and J. Hullman, *EVM: Incorporating Model Checking into Exploratory Visual Analysis*, en, arXiv:2308.13024 [cs], Aug. 2023. URL: http://arxiv.org/abs/2308.13024 (visited on 12/02/2023).

[10] C. Nobre, D. Wootton, Z. Cutler, L. Harrison, H. Pfister, and A. Lex, “reVISit: Looking Under the Hood of Interactive Visualization Studies,” en, in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan: ACM, May 2021, pp. 1–13, ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445382. URL: https://dl.acm.org/doi/10.1145/3411764.3445382 (visited on 09/17/2021).

[11] L. Battle and A. Ottley, *What Exactly is an Insight? A Literature Review*, arXiv:2307.06551 [cs], Jul. 2023. URL: http://arxiv.org/abs/2307.06551 (visited on 09/15/2023).

[12] S. Robinson and A. L. Mendelson, “A Qualitative Experiment: Research on Mediated Meaning Construction Using a Hybrid Approach,” en, *Journal of Mixed Methods Research*, vol. 6, no. 4, pp. 332–347, Oct. 2012, ISSN: 1558-6898. DOI: 10.1177/1558689812444789. URL: https://doi.org/10.1177/1558689812444789 (visited on 06/26/2023).

[13] J. Hullman and A. Gelman, “Designing for Interactive Exploratory Data Analysis Requires Theories of Graphical Inference,” en, *Harvard Data Science Review*, Jul. 2021. DOI: 10.1162/99608f92.3ab8a587. URL: https://hdsr.mitpress.mit.edu/pub/w075glo6 (visited on 10/19/2021).

[14] W. Epperson, V. Gorantla, D. Moritz, and A. Perer, “Dead or Alive: Continuous Data Profiling for Interactive Data Science,” en, *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2023, ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2023.3327367. URL: https://ieeexplore.ieee.org/document/10301695/ (visited on 12/08/2023).

[15] L. Battle and J. Heer, “Characterizing Exploratory Visual Analysis: A Literature Review and Evaluation of Analytic Provenance in Tableau,” en, *Computer Graphics Forum*, vol. 38, no. 3, pp. 145–159, Jun. 2019, ISSN: 0167-7055, 1467-8659. DOI: 10.1111/cgf.13678. URL: https://onlinelibrary.wiley.com/doi/10.1111/cgf.13678 (visited on 09/13/2023).

[16] A. Sarvghad, M. Tory, and N. Mahyar, “Visualizing Dimension Coverage to Support Exploratory Analysis,” en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 21–30, Jan. 2017, ISSN: 1077-2626. DOI: 10.1109/TVCG.2016.2598466. URL: http://ieeexplore.ieee.org/document/7534787/ (visited on 03/23/2024).

[17] Y. Wu, J. M. Hellerstein, and A. Satyanarayan, “B2: Bridging Code and Interactive Visualization in Computational Notebooks,” en, in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, Virtual Event USA: ACM, Oct. 2020, pp. 152–165, ISBN: 978-1-4503-7514-6. DOI: 10.1145/3379337.3415851. URL: https://dl.acm.org/doi/10.1145/3379337.3415851 (visited on 09/14/2023).

[18] K. Gadhave, Z. Cutler, and A. Lex, *Persist: Persistent and Reusable Interactions in Computational Notebooks*, en, Dec. 2023. DOI: 10.31219/osf.io/9x8eq. URL: https://osf.io/9x8eq (visited on 03/27/2024).

[19] X. Pu and M. Kay, “How Data Analysts Use a Visualization Grammar in Practice,” en, in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–22, ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3580837. URL: https://dl.acm.org/doi/10.1145/3544548.3580837 (visited on 09/14/2023).

[20] H. Guo, S. R. Gomez, C. Ziemkiewicz, and D. H. Laidlaw, “A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights,” en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 51–60, Jan. 2016, ISSN: 1077-2626. DOI: 10.1109/TVCG.2015.2467613. URL: http://ieeexplore.ieee.org/document/7192662/ (visited on 03/24/2024).

[21] S. Kandel, A. Paepcke, J. M. Hellerstein, and J. Heer, “Enterprise Data Analysis and Visualization: An Interview Study,” en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2917–2926, Dec. 2012, ISSN: 1077-2626. DOI: 10.1109/TVCG.2012.219. URL: http://ieeexplore.ieee.org/document/6327298/ (visited on 12/08/2023).

[22] K. Wongsuphasawat, Y. Liu, and J. Heer, *Goals, Process, and Challenges of Exploratory Data Analysis: An Interview Study*, arXiv:1911.00568 [cs], Nov. 2019. URL: http://arxiv.org/abs/1911.00568 (visited on 11/28/2023).

[23] S. G. Hart and L. E. Staveland, “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” *Human mental workload*, vol. 1, no. 3, pp. 139–183, 1988.

[24] K. Xu, A. Ottley, C. Walchshofer, M. Streit, R. Chang, and J. Wenskovitch, “Survey on the Analysis of User Interactions and Visualization Provenance,” en, Open Science Framework, preprint, Mar. 2020. DOI: 10.31219/osf.io/jux76. URL: https://osf.io/jux76 (visited on 05/23/2020).

[25] M. Feng, E. Peck, and L. Harrison, “Patterns and Pace: Quantifying Diverse Exploration Behavior with Visualizations on the Web,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 501–511, Jan. 2019, Conference Name: IEEE Transactions on Visualization and Computer Graphics, ISSN: 1941-0506. DOI: 10.1109/TVCG.2018.2865117.

[26] S. S. Alam and R. Jianu, “Analyzing Eye-Tracking Information in Visualization and Data Space: From Where on the Screen to What on the Screen,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 5, pp. 1492–1505, May 2017, Conference Name: IEEE Transactions on Visualization and Computer Graphics, ISSN: 1941-0506. DOI: 10.1109/TVCG.2016.2535340.

[27] A. Boggust, B. Carter, and A. Satyanarayan, “Embedding Comparator: Visualizing Differences in Global Structure and Local Neighborhoods via Small Multiples,” en, in *27th International Conference on Intelligent User Interfaces*, Helsinki Finland: ACM, Mar. 2022, pp. 746–766, ISBN: 978-1-4503-9144-3. DOI: 10.1145/3490099.3511122. URL: https://dl.acm.org/doi/10.1145/3490099.3511122 (visited on 03/31/2024).

[28] C. North, “Toward measuring visualization insight,” *IEEE Computer Graphics and Applications*, vol. 26, no. 3, pp. 6–9, May 2006, Conference Name: IEEE Computer Graphics and Applications, ISSN: 1558-1756. DOI: 10.1109/MCG.2006.70.

[29] C. Nobre, D. Wootton, L. Harrison, and A. Lex, “Evaluating Multivariate Network Visualization Techniques Using a Validated Design and Crowdsourcing Approach,” en, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, Apr. 2020, pp. 1–12, ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376381. URL: https://dl.acm.org/doi/10.1145/3313831.3376381 (visited on 09/14/2023).

[30] Z. Liu and J. Heer, “The Effects of Interactive Latency on Exploratory Visual Analysis,” en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2122–2131, Dec. 2014, ISSN: 1077-2626. DOI: 10.1109/TVCG.2014.2346452. URL: http://ieeexplore.ieee.org/document/6876022/ (visited on 03/31/2024).

[31] E. Zgraggen, Z. Zhao, R. Zeleznik, and T. Kraska, “Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis,” en, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal QC Canada: ACM, Apr. 2018, pp. 1–12, ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3174053. URL: https://dl.acm.org/doi/10.1145/3173574.3174053 (visited on 09/14/2023).

[32] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, “Vega-Lite: A Grammar of Interactive Graphics,” en, p. 10,

[33] H.-F. Hsieh and S. E. Shannon, “Three approaches to qualitative content analysis.,” *Qualitative health research*, vol. 15, no. 9, pp. 1277–88, Nov. 2005, ISSN: 1049-7323. DOI: 10.1177/1049732305276687. URL: http://www.ncbi.nlm.nih.gov/pubmed/16204405 (visited on 05/24/2014).

[34] E. Marsh and M. White, “Content analysis: A flexible methodology,” *Library trends*, vol. 55, no. 1, pp. 22–45, 2006, ISSN: 1559-0682. DOI: 10.1353/lib.2006.0053. URL: http://muse.jhu.edu/content/crossref/journals/library_trends/v055/55.1white.html (visited on 06/04/2014).

[35] M. Muller, “Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method,” en, in *Ways of Knowing in HCI*, J. S. Olson and W. A. Kellogg, Eds., New York, NY: Springer, 2014, pp. 25–48, ISBN: 978-1-4939-0378-8. URL: https://doi.org/10.1007/978-1-4939-0378-8_2 (visited on 09/11/2023).

[36] M. Lombard, J. Snyder-Duch, and C. Bracken, "Practical resources for assessing and reporting intercoder reliability in content analysis research projects," no. 2002, pp. 1–18, 2004. (visited on 06/06/2014).

[37] E. Jun, M. Birchfield, N. De Moura, J. Heer, and R. Just, "Hypothesis Formalization: Empirical Findings, Software Limitations, and Design Implications," en, *ACM Transactions on Computer-Human Interaction*, vol. 29, no. 1, pp. 1–28, Feb. 2022, ISSN: 1073-0516, 1557-7325. DOI: 10.1145/3476980. URL: https://dl.acm.org/doi/10.1145/3476980 (visited on 12/08/2023).

[38] L. Battle and A. Ottley, *A Programmatic Definition of Visualization Insights, Objectives, and Tasks*, arXiv:2206.04767 [cs], Oct. 2022. URL: http://arxiv.org/abs/2206.04767 (visited on 09/14/2023).

[39] D. Norman, *The Design Of Everyday Things*, English, Revised edition. New York, New York: Basic Books, Nov. 2013, ISBN: 978-0-465-05065-9.

[40] J. Boy, L. Eveillard, F. Detienne, and J.-D. Fekete, "Suggested Interactivity: Seeking Perceived Affordances for Information Visualization," eng, *IEEE transactions on visualization and computer graphics*, vol. 22, no. 1, pp. 639–648, Jan. 2016, ISSN: 1941-0506. DOI: 10.1109/TVCG.2015.2467201.

[41] S. Davies, "The Cognitive Psychology of Planning," en, in *Planning and problem solving in well-defin ed domains*, The Psychology Press, 2005, p. 43. URL: https://www.routledge.com/The-Cognitive-Psychology-of-Planning/Morris-Ward/p/book/9780415646772 (visited on 12/08/2023).

[42] R. Arias-Hernandez, L. T. Kaastra, and B. Fisher, "Joint Action Theory and Pair Analytics: In-vivo Studies of Cognition and Social Interaction in Collaborative Visual Analytics," en, *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 2011.

[43] U. I. D. Lab, *Draco: Representing, Applying & Learning Visualization Design Guidelines*, en, Library Catalog: medium.com, Oct. 2018. URL: https://medium.com/@uwdata/draco-representing-applying-learning-visualization-design-guidelines-64ce20287e9d (visited on 04/18/2020).

[44] A. Suh, Y. Jiang, A. Mosca, E. Wu, and R. Chang, *A Grammar for Hypothesis-Driven Visual Analysis*, arXiv:2204.14267 [cs], Apr. 2022. URL: http://arxiv.org/abs/2204.14267 (visited on 03/22/2023).

[45] W. Willett, J. Heer, and M. Agrawala, "Scented Widgets: Improving Navigation Cues with Embedded Visualizations," en, *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1129–1136, Nov. 2007, ISSN: 1077-2626. DOI: 10.1109/TVCG.2007.70589. URL: http://ieeexplore.ieee.org/document/4376132/ (visited on 03/26/2024).

[46] Y. Wu, R. Chang, J. M. Hellerstein, and E. Wu, *Facilitating Exploration with Interaction Snapshots under High Latency*, arXiv:1806.01499 [cs], Sep. 2020. URL: http://arxiv.org/abs/1806.01499 (visited on 03/31/2024).

[47] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. McCandless, “Edit wear and read wear,” en, in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*, Monterey, California, United States: ACM Press, 1992, pp. 3–9, ISBN: 978-0-89791-513-7. DOI: 10.1145/142750.142751. URL: http://portal.acm.org/citation.cfm?doid=142750.142751 (visited on 01/24/2024).

[48] A. Narechania, A. Coscia, E. Wall, and A. Endert, “Lumos: Increasing Awareness of Analytic Behavior during Visual Data Analysis,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 1009–1018, Jan. 2022, arXiv:2108.02909 [cs], ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2021.3114827. URL: http://arxiv.org/abs/2108.02909 (visited on 03/31/2024).

[49] R. Peng and 2018, *Simply Statistics: Divergent and Convergent Phases of Data Analysis*. URL: https://simplystatistics.org/posts/2018-09-14-divergent-and-convergent-phases-of-data-analysis/ (visited on 03/26/2024).

[50] S. Gathani, S. Monadjemi, A. Ottley, and L. Battle, *A Grammar-Based Approach for Applying Visualization Taxonomies to Interaction Logs*, arXiv:2201.03740 [cs], Apr. 2022. URL: http://arxiv.org/abs/2201.03740 (visited on 03/24/2024).

[51] D. Dotan, P. Pinheiro-Chagas, F. A. Roumi, and S. Dehaene, “Track It to Crack It: Dissecting Processing Stages with Finger Tracking,” English, *Trends in Cognitive Sciences*, vol. 23, no. 12, pp. 1058–1070, Dec. 2019, Publisher: Elsevier, ISSN: 1364-6613, 1879-307X. DOI: 10.1016/j.tics.2019.10.002. URL: https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(19)30237-2 (visited on 09/06/2021).

[52] H. Sacks, E. A. Schegloff, and G. Jefferson, “A Simplest Systematics for the Organization of Turn-Taking for Conversation,” *Language*, vol. 50, no. 4, pp. 696–735, 1974, Publisher: Linguistic Society of America, ISSN: 0097-8507. DOI: 10.2307/412243. URL: https://www.jstor.org/stable/412243 (visited on 03/29/2024).

[53] V. Setlur, M. Correll, A. Satyanarayan, and M. Tory, *Heuristics for Supporting Cooperative Dashboard Design*, arXiv:2308.04514 [cs], Aug. 2023. URL: http://arxiv.org/abs/2308.04514 (visited on 09/05/2023).

[54] A. Lex, “Opportunities for Understanding Semantics of User Interactions,” in *Workshop – Machine Learning from User Interactions*, Oct. 2021.

[55] A. Fouse, N. Weibel, E. Hutchins, and J. D. Hollan, “ChronoViz: A system for supporting navigation of time-coded data,” en, in *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, Vancouver BC Canada: ACM, May 2011, pp. 299–304, ISBN: 978-1-4503-0268-5. DOI: 10.1145/1979742.1979706. URL: https://dl.acm.org/doi/10.1145/1979742.1979706 (visited on 03/30/2024).

MIT-thesis-template/mitthesis-sample