# Multiple Workspaces in Visual Analytics

Maha El Meseery, Yuyao Wu, Wolfgang Stuerzlinger

*School of Interactive Arts + Technology (SIAT), Simon Fraser University,* Vancouver, Canada
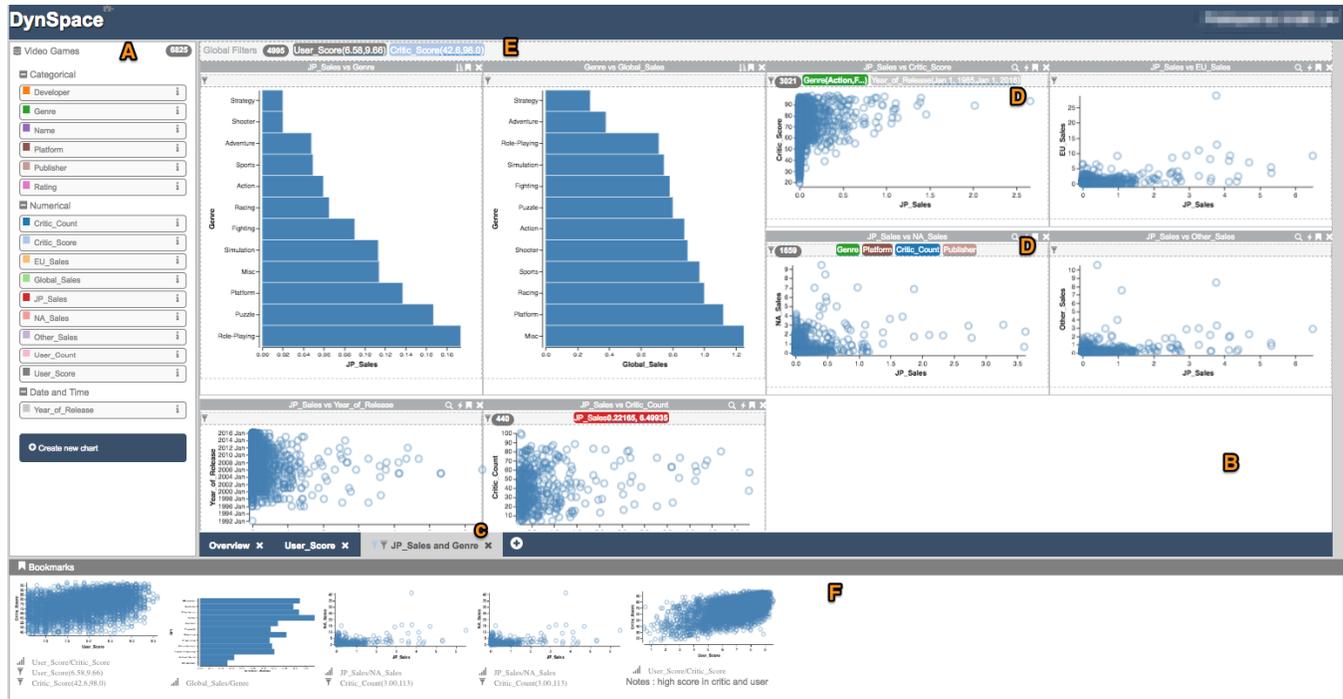
w.s@sfu.ca

Figure 1. DynSpace User Interface, with two main panels, the Data Dimension and the Workspace Panel. The Data Dimension Panel (A) displays the dataset name, total number of data records, and the dimensions. Clicking on a dimension will highlight all charts that show that dimension. The info symbol beside a dimension generates chart recommendations based on that dimension in a new workspace. The Workspace Panel (B) consists of multiple analytical workspaces, represented as tabs (C). Each workspace can contain an arbitrary number of charts and all the charts inside a workspace are linked. The Local Filter Bar (D) at the top of each chart shows all applied filters color-coded by dimension. These filters can be created, copied, or removed through simple drag-and-drop operations. Filters in the Global Filter Bar (E) at the top apply to all charts inside that workspace. At the bottom, a Bookmark Panel (F) displays user bookmarks, which contain a thumbnail, a description of the chart specification, filters, and bookmark notes, and a button to navigate to the context of its source workspace.

*Abstract*—**Exploratory visual analysis is an iterative process, where analysts often start from an overview of the data. Then, they often pursue different hypotheses through multiple rounds of interaction and analysis. Commercial visual analytics packages support mostly a model with a single analysis path, where the system view represents only the final state of the users' current analysis. In this paper, we investigate the benefit of using multiple workspaces to support alternative analyses, enabling users to create different workspaces to pursue multiple analysis paths at the same time. We implemented a visual analytics prototype for multiple workspaces using a multi-tab design. The results of our user studies show that multiple workspaces: enable analysts to work on concurrent tasks, help organizing an analysis, and make it easy to revisit previous parts of their work.**

*Keywords—Visual analytics, multiple analysis alternatives*

## I. INTRODUCTION

In current visual analytics (VA) systems, users explore complex datasets through dynamic visualizations and a series of interactions by selecting, filtering, manipulating, and drilling down into the details of the data. Through such interactions users explore and analyze the data in an iterative process through multiple rounds of questions and answers [1], [2]. Popular VA tools support a common interaction pattern where the user focuses first on single visualizations and then builds an overview, i.e., dashboard, out of multiple visualizations [3].

Even though current systems allow users to explore and analyze data in various ways, there is little support for all aspects of the knowledge generation process [1]. Current tools do not support concurrent states during the analysis process, as any modification to the visual encoding or filtering to the data is globally applied throughout the application. Other systems like Voyager 2 provide an overview of the data first, then permit analysts to explore the data [4], but again provide only a single-state interface in which any changes to the filtering or the selected dimension will change the whole system view. In this context, consider a scenario where an analyst is working on a hypothesis on a dataset they have been working on for several days, only to discover that there are (still) potential

outliers/errors in the data. Current VA tools do not make it easy to non-destructively evaluate *how* different options such as removing, transforming, or retaining such data points affects previous analysis paths and their results. To pursue multiple alternative analysis tasks, users often have to redo part of their analysis or backtrack previous interactions to follow other hypotheses. In contrast, concurrent workspaces have the potential to enable users to create multiple analysis paths that can be explored in parallel, without affecting previous work.

Here we investigate the affordance and benefits of multiple concurrent workspaces in VA systems. We define a single workspace as a fully editable dashboard or container that holds multiple coordinated views at the same time with the associated state (i.e., data filtering and visual encoding), and one or more analytical goals (i.e., data wrangling, exploration, clustering). As such, multiple workspaces allow users to see multiple states of their work and thus support different analytical tasks. Additionally, maintaining separate workspaces allows users to access multiple aspects of their analysis simultaneously. E.g., users can go back and experiment with filters created during data wrangling to remove outliers, without affecting other analyses. With multiple concurrent analysis paths, another main characteristic of workspaces is the ability to revisit and reuse part of users' analysis from the different workspaces. As noted by Shneiderman [5], supporting users in their search for answers, locating findings, and organizing them in more efficient ways is extremely valuable feature in the visual exploration process.

To investigate the value of looking at different analysis aspects in parallel, we created DynSpace, a multi-workspace VA tool with a simple multi-tab interface. We choose the tab interface as a simple representation of different analysis processes, while still being integrated in a single application to facilitate sharing of data and charts. Also, many other tools use a tab interface and users are familiar with it. The multi-tab interface allows users to create multiple workspaces that are independent from one another, but user can still interact with each of them in the same way. We conducted two user studies to explore how users interact with such multiple workspaces.

Based on DynSpace, our VA prototype implementing concurrent workspaces using a multi-tab interface, the main contributions of our work are: 1) design considerations for multiple, concurrent VA workspaces; 2) a first user study investigating the benefit of multiple workspaces; and 3) a comparative study to compare a single state interface with multiple concurrent workspaces.

## II. BACKGROUND

Our works is informed by prior work in VA on dashboards, support for the VA process, and multiple tab interfaces.

### A. Dashboards

Dashboards are not a new concept, and the most common definition of a dashboard in several fields is a visual display of the most important information, consolidated and arranged on a single screen so the information can be monitored at a glance [6]–[8]. In VA a dashboard is defined as a "faceted analytical display", which holds multiple charts, together with the

interaction techniques to support efficient analysis [6]. This paradigm has recently become more popular, e.g., in Tableau [3] and SAP Lumira [9], which both support construction of dashboards from individual charts and other interactions such as filter, drill-up/down, sort, and drag-and-drop. Via rich interactions on a single chart, users can get very detailed information about the shown data. Even though one can create multiple dashboards, these systems all maintain a consistent filter state across all dashboards. Our work enables users to manage different, independent workspaces for a dataset. VizDeck presented drag-and-drop construction of dashboards from recommendations [35]. Other systems, such Data Voyager [10], supports a breadth-first strategy by showing multiple recommended charts in the initial view to give users an overall sense of what their data looks like. Related work focuses on recommending charts, e.g., [36], [37]. Data Voyager 2 [4] adds facilities for drill-down, filtering and related actions but does not support multiple dashboards.

### B. VA process support

While current VA systems, such as Tableau, facilitate exploring and analyzing data in an iterative process, there is little support for the whole knowledge generation process the users employ during their analysis of data [1], such as backtracking and investigating multiple parallel scenarios. In a purely iterative process, users can easily lose track of what they were doing and/or forget previous insights. In recent years, various VA systems have been developed to aid users in exploring, analyzing, and understanding data [1], [3], [11]–[13]. Some work discussed the various design trade-offs for developing a VA history tracking mechanism [13]. Few systems have developed a mechanism to keep track of users' interactions and store it in a history that enables users to review, recall, and retrieve their actions [13], [14]. Others focused on discussing and investigating how to support and capture insight provenance [15]. HARVEST tried to automatically capture a higher level of user intents and provided insight provenance based on user actions [16], [17]. Yet, few have explored the idea of multiple concurrent analyses or alternative hypotheses, where users can "branch out" to different alternative scenarios, backtrack, copy parts of one analysis to another scenario, or go through different analysis phases, with the potential to revisit each phase while still keeping other results accessible.

### C. Multiple tab interface

While multiple researchers investigated multi-tab browsing in the web browsing environment, few studies focused on web applications that include multi-tab interfaces [18]. While using the web browser to open multiple version of an application could permit multiple states and versions of the data, a single web application with multiple tabs does not provide the same affordance. A study showed that, in comparison with traditional multi-tab browsing, using a tabbed interface with explicit support for a user goal reduced the degree of disorientation and helped in gathering information [19].

## III. DESIGN CONSIDERATION AND GOALS

Our goal is to enable users to use multiple workspaces as a medium to reflect about their analytical tasks and test their

hypotheses. Current VA solutions afford users only a single workspace, which leads to (mostly) linear workflow. Yet, analysts backtrack, reuse parts of previous workflows, and investigate different alternative, concurrent scenarios. Such non-linear processes are not well supported in current VA systems. Based on our analysis of previous work and several iterations and pilots on DynSpace, we identified several design considerations for such multiple concurrent workspaces.

***Separate Analysis Paths.*** Workspaces should allow analysts to explore and pursue different analysis paths. As such, each workspace should maintain separate state with respect to the data, including filters, the view and encoding, and the analysis state. Different distinct states in the workspace allows analyst to pursue different analytical investigations. Dynamic brushing and data queries represent elements of the current state of user analysis [2]. Additionally, workspaces can also support different phases of analysis such as data wrangling, data overview, and clustering.

***Shared Context.*** While the state of each workspace needs to be distinct to support concurrent alternative paths, there has to be common context information, which should be shared between the different workspaces. Additionally, shared bookmarks with associated state information helps the analyst to keep track of the different concurrent analyses.

***Easy Navigation between Workspaces.*** An analyst should be able to quickly navigate between the different workspaces to switch between concurrent and/or alternative analyses. For example, the user may investigate an interesting analysis path, seek additional data perspectives in a different workspace, and backtrack to see previous results or even start the whole process over when they run into a dead end, but still re-use parts of the previous workflow in their current attempt. Support for all such behaviours, necessitates efficient and quick methods to navigate between the different concurrent analyses.

***Revisiting and Retrieving.*** To support users in investigating multiple concurrent hypotheses they need to be able to go through subsequent rounds of questions and answers, revisit a previous view, recall a previous state, and compare different aspects of the data [13]. Supporting these interactions enables the user to keep track of their process by allowing them to review, recall, reuse, and revisit their steps. Workspaces should allow easy mechanism to recall and restore part of the analyst previous work.

Finally, we see recommendations as an integral supporting part of multiple workspaces, especially for the exploration phase, but do not focus on this topic in our work.

## IV. PROTOTYPE DESIGN

To explore the multiple workspaces design space, we created a new VA prototype, *DynSpace*. Like current state-of-the-art VA tools, DynSpace supports both open-ended exploration and focused data analysis through multiple coordinated views for each workspace, view recommendations, and manual view specification. DynSpace's user interface consist of three main panels: Data Dimensions, Workspaces (see Figure 1), and Bookmarks. The Data dimension panel shows information about the dataset and the dimensions. The

workspaces panel holds the multiple workspaces through a multi-tab interface (Figure 1). At the bottom of the DynSpace window, the Bookmark panel shows previously saved charts.

We implemented our VA prototype as a web application on a Node.js server with MongoDB running in the backend to support data collection. We chose a web application because it of its ubiquitous accessibility, flexibility in terms of both client and server-side technologies, and ease of adding more features in the future. All the functionality is implemented in JavaScript and jQuery. We used gridstack.js for the layout and jQuery UI for all drag-/droppable widgets. For charting, we used Vega.js [20], a higher-level visualization specification language on top of D3.js [21]. Vega constructs a specification and then passes it off to the runtime to visualize data, which makes it fast, customizable, and reusable. Vega also provides an abstraction layer for both rendering and event processing, which improves performance and platform flexibility.

### A. Data Dimensions and Charts

Dimensions are data attributes in the dataset that can be used to create charts. DynSpace groups dimensions into Categorical, Numerical, and Temporal. To uniquely identify each dimension throughout the system, we color-code them through a small square besides the dimension name. To enable quick visual search for data attributes, clicking on a dimension will highlight all charts that show that dimension. To support quick exploration of a specific dimension, users can click on an "Explore" icon beside each dimension, which will automatically create a new workspace with multiple charts focused on this dimension. Dimensions are drag- and reusable, and can be used to create charts, change axes, and create filters.

As our main target audience is novice users, who are typically unfamiliar with advanced chart types, DynSpace defaults to using only histograms, bar charts, and scatterplots. Histograms are used for single dimensions, scatterplots for two numerical dimensions; and bar charts are created when one dimension is categorical and the second numerical [3]. By default, the bars are vertical, except if there are more than 20 values for a categorical dimension, in which case the bars will be displayed horizontally. To facilitate pattern exposure, bars can be sorted in ascending order through the sorting icon at the top-right of a chart. On hovering, tooltips are used to show all details of the individual data point, including statistical information, such as the mean and median (i.e., for bar charts).

### B. Multiple workspaces

The Workspaces panel consist of multiple workspaces in a multi-tab interface, which are designed to support *separate analysis paths*. Each workspace can contain an arbitrary number of views or charts that display a different aspect of the common dataset. To allow users to explore relations between the different views, all views inside a single workspace are coordinated through brushing/linking and filtering.

For the initial workspace, an Overview workspace is automatically generated once a dataset is loaded. Currently, this overview shows charts for all potential relations between different data dimensions, through straightforward permutation of dimensions. We chose this simple mechanism to provide a

quick starting point for analysis, without focusing on a better recommendation engine.

To allow users to organize their work, to *facilitate navigation between workspaces* and to enable *reuse and retrieving*, DynSpace permits users to create an unlimited number of workspaces and to switch between, rename, and delete workspaces. Workspaces can contain an unlimited number of charts, which can be copied to another workspace through drag-and-drop onto the tab-title, which maintains all filters for a chart. This functionality enables users to organize their exploration by creating workspaces as they, e.g., focus on a single aspect (without affecting other analysis results), move into new questions, or into different exploration and analysis paths and phases. Beyond automatic chart generation, users can also create new charts manually in the currently active workspace. We use drag-and-drop of dimensions for creating or modifying charts dimensions as an intuitive and familiar interaction for most users. DynSpace automatically (re-)sets the chart type according to the "dropped" dimensions. To permit users to see more detail, chart can be resized.

To minimize chart rearrangement/resizing effort, each workspace uses a non-overlapping grid layout, initially with 4x4=16 charts. This provides an adequate number of simultaneously visible charts in the window, while still showing some amount of detail. To maintain the grid layout and minimize chart layout changes, charts snap to the grid. During a rearrangement of a chart or a chart resize, we minimize the number of (other) charts that change position.

### C. Drag-and-drop Filters

DynSpace supports two types of filters – global and local ones. Global filters apply to all charts inside a workspace while local filters apply to a single chart. Both types of filters are shown as a small rounded rectangle color-coded by dimension, which displays the name of that dimension and the extent of the filter. They are displayed in the respective filter bar at the top of each chart or workspace, shown in Figure 1(D) respectively (E). Multiple local filters can be applied on a chart at the same time. If there is not enough space to show all of them, the local filter representations compact automatically. Then details of each filter are available through hovering [22]. To support awareness of global filter existence, a small color-coded indicator is shown next to the tab name, see Figure 1(C).
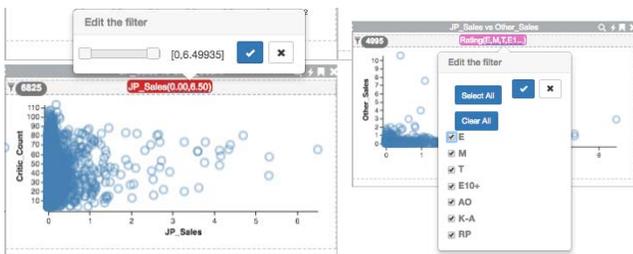


Figure 2. Filter editing dialogs for numerical and categorical dimensions.

DynSpace supports two ways of creating a local filter. First, a filter can be created via (rectangle) dragging and selecting an extent directly in the chart, i.e., brushing, and then clicking the filter icon on top-left corner of that chart. This will create a new chart with the corresponding brush applied as a filter.

Compared to the approaches supported in commercial VA software, this enables very quick filter creation. Second, DynSpace also permits the user to drag-and-drop a dimension onto a filter bar, which will then create a filter with the full extent of the dimension values in the dataset. Filters can be edited by clicking on them, which will show a pop-up dialog box with a value slider for numerical, calendar ranges for temporal, and a checkbox list for categorical dimensions (see Figure 2). Filter can be removed by drag-and-drop to empty space or by hitting the backspace key.

All filters are draggable. Any dropped filter is copied to the target (including to other workspaces, where they act as global filters). Adding multiple filters for the same dimension will merge the filters by intersecting the range of these filters (i.e., using the and operator to merge the different ranges). When the user drops a filter into the global filter bar, this filter is automatically removed from all the individual charts in that workspace, which avoids problems with duplicate filters.
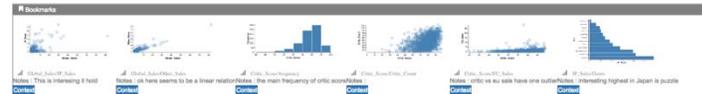


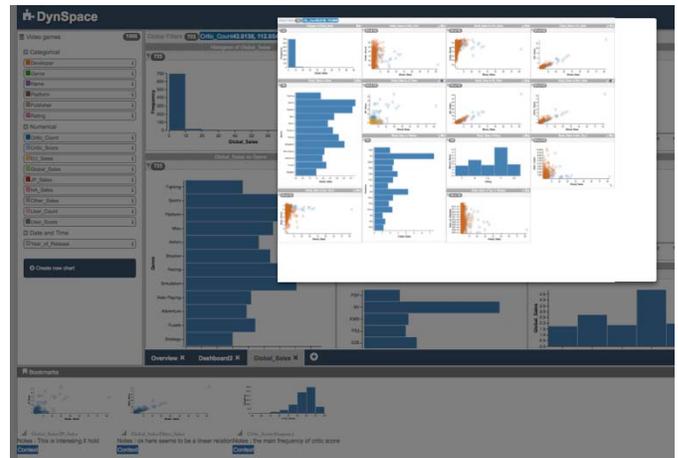Figure 3. Bookmark Panel with multiple chart thumbnails.



Figure 4. Display of context for a bookmark.

### D. Context Bookmarks

To support users in saving and keeping track of their insights and to help them maintain a *shared context*, DynSpace facilitates bookmarking of charts, through the "bookmark" icon of any chart, with an option to annotate the new bookmark. All saved bookmarks appear in the, initially minimized Bookmark panel (see Figure 3), which holds bookmarks in linear display in order of creation. Each bookmark shows a chart thumbnail, the chart dimensions and applied filters, any user annotation, and a context button. Clicking on the context button shows the state of the workspace at time of bookmark creation, providing additional context (see Figure 4). This avoids the issue that traditional bookmarks do not carry information about the analysis context, such as other related views that were co-present in the workspace with the saved chart, which can help the user to recall insights better.

## E. Usage Scenario

Here we describe how an analyst can benefit from using multiple workspaces in DynSpace. Our use case is based on the Superstore Sales dataset [23], with 18 dimensions and 4104 data records in total and modeled on an analysis scenario one of the authors encountered.

Jane is a new data analyst and needs to report to her boss at next week's meeting about sales in their stores. After the dataset loads, she wants to know what factors might affect the profit of stores. Thus, she clicks on the explore icon beside the Profit dimension. A new workspace with the charts related to profit appears in the Workspace Panel. After browsing the different charts, she is curious as to which factors created a profit or a loss. Thus, she creates a second, *separate analysis path* by clicking on explore profit again to create another workspace, then re-names the first workspace Profit and the second Loss respectively. She then drags the Profit dimension onto the global filter bar in each of these two workspaces, clicks on the created filter, and edits the filter extent to be positive and negative, respectively. This filters all charts inside these two workspaces accordingly. The dimension name in the global filter bar and the small filter indicator next to the tab name show the successful filtering.
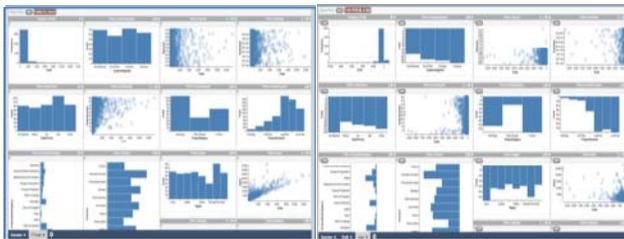


Figure 5. The profit and loss workspaces. The left is the profit workspace and the right the loss one.

Then, she focusses her analysis on several charts related to Profit, such as Profit vs. Shipping_Cost, Profit vs. Region, and Profit vs. Discount. She wants to compare Shipping cost for loss and profit, so she opens a new workspace and calls it "Compare". Easily *navigating between workspaces*, she duplicates charts from the Profit and Loss workspace into the new one and creates a side-by-side comparison (note that global filters are instantiated as local filters in this operation to maintain view consistency). Figure 6 shows this Profit vs Loss comparison in a workspace.
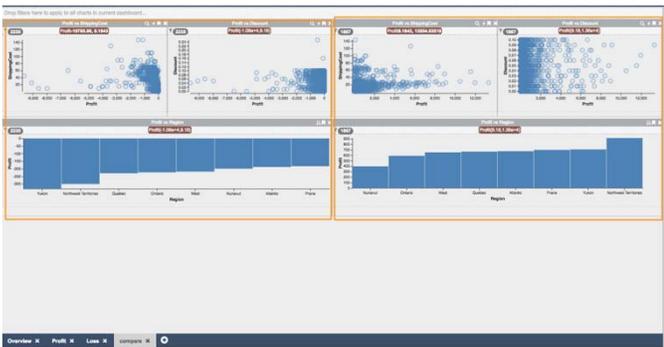


Figure 6. Comparison of shipping cost, regions, and discount for loss (left) and profit (right) using different local filter for the same charts.

In this comparison she notices the three regions with the highest profits and is becomes interested in exploring the sales in these regions. She creates a new workspace to explore these sales. Sorting the bar chart – Profit vs Region, she creates another local Profit filter through selecting the top three bars with a mouse drag and creates a filter through the filter icon. Moreover, she considers that Unit_Price might affect Profit and manually creates a new chart showing Profit vs Unit_Price. After that, she brushes on various charts to check the corresponding data points highlighted in related charts (Figure 7). She creates the filter and drags it to the global filter bar to apply it to the whole workspace.
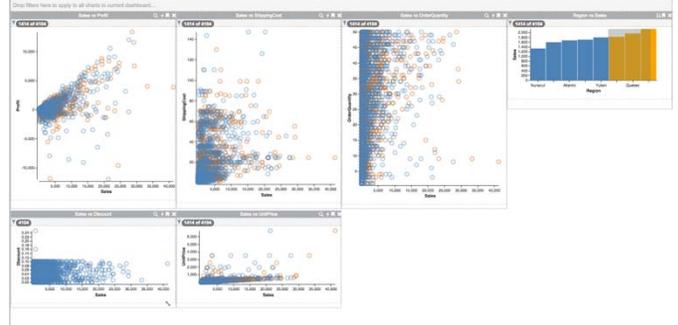


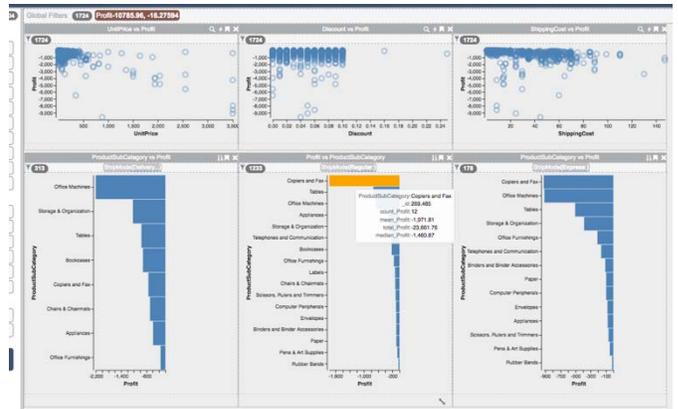Figure 7. Highlighting the 3 highest profit regions while exploring sales.



Figure 8. Comparison of different shipping modes and drilling down.

She notices Profit did not have a clear linear relationship with Unit_Price, Shipping_Cost or Discount, and hypothesizes that there might be a combination effect. However, the stores who have a relative low profit also have a relative high Shipping_Cost. She also notices Office Supplies and copiers and fax machine made the lowest profit. Thus, she speculates that the profit of Office Supplies was low because it had a high shipping cost. With that thought, Jane creates a new workspace, names it "Profit & Shipping", and applied a global filter to remove all loss records. *Revisiting and retrieving work*, she copies three charts of Profit vs Shipping_Cost from previous workspaces and applies a local filter of Ship_Mode. She also creates a Profit vs Product Sub Category chart and duplicates the Ship_Mode filter via drag-and-drop from one chart to another. Subsequently, she edits the three charts to filter for Ship_Mode = Express Air, Regular Air and Delivery Truck, respectively, to enable her to perform a visual comparison. She notices that most of the products were shipped with Regular

Air, and that Delivery Truck was the second-most used shipping method, but that it had the highest shipping cost, as per another new chart she created for Shipping_Cost vs Ship_Mode. Drilling down to the details form each of the Product Sub Category she finds out that for Regular Air shipping there is a very high loss for Copiers and Faxes machine with a small number of orders. Considering that this part of the business also made little profit, she comes to the temporary conclusion that the high shipping cost and the limitations of the shipping method might be a bottleneck that should be addressed.

## V. STUDY METHODOLOGY

The goal of the first study was to explore and understand how users can benefit from multiple concurrent workspaces in DynSpace when exploring and analyzing data analysis tasks. The study attempted to answer two main questions:

- **RQ1**: What usage patterns are observable while users complete a data analysis tasks in multiple workspaces?

- **RQ2**: Does an analyst perceive interacting with DynSpace to be easy, flexible, and not distractive?

For this study, we decided to employ a qualitative user study approach to enable us to arrive at a deeper understanding of the different possible analysis pattern that analysts use in DynSpace. Due to the exploratory nature of this study, we adopted a case study approach to obtain in-depth information about users' analysis strategies.

### A. Participants

In total, we recruited 26 participants in total, 6 males and 20 females. All participants were undergrad students with ages ranging between 18 and 22 years. About 73% of them had no experience with VA, and 19% had less than 6 months.

### B. Procedure and Apparatus

The study ran in a lab environment on a dual monitor desktop computer. The participant used the main 27" monitor to interact with DynSpace in a Google Chrome browser. The second 21.5" monitor was used to display the online survey system, which was used to present the tasks and gather answers in the pre- and post-task questionnaires. Screen capture software was used to record participants' interactions with DynSpace. One observer/investigator was always present to collect observations, to provide instructions at the beginning, and technical support. The study lasted about an hour.

The study started with a pre–study questionnaire to collect information about participants and their background experience with simple charts (i.e., bar and scatterplot charts), VA tools, as well as related software. Next, each participant was given a quick training session with a five-minute tutorial video and a set of practice exercises on a sample dataset. After that, the main task scenario and study dataset was presented to the participants and they were asked to complete five analysis tasks. The system logged all interactions during their use of DynSpace. For each task, participants were asked to create charts, to explore and analyse the data, and to enter their insights into the online survey. Participants were informed that

they would need to explain the steps of their analysis at the end of the study, to simulate a scenario where one has to explain later what one did, say to a superior. After finishing all tasks, participants completed a Likert type questionnaire that recorded participants' opinions on the overall DynSpace prototype as well as specific features. At the end of the study, an interview was conducted to clarify experimenters' observations, collect additional feedback about the prototype, identity problems, and elicit ideas for future improvements.

TABLE I.       TABLE 1 THE TASKS PARTICIPANTS PERFORMED DURING EACH SESSION, AND THE PURPOSE OF EACH TASK.

| Task Scripts | Purpose |
|---|---|
| Your friend Susan wants to buy a new car, but she is not that familiar with cars. So, she asked you to offer some suggestions. We have already loaded a car dataset for you to explore. Please show Susan some charts to ensure she can understand how and why you came up with your suggestions. In contrast to browser usage, we suggest not to close tabs with important results. | Scenario introduction. |
| Q1 Susan seems to be on a budget these days, so she cares that the car is reasonably priced. Show her a chart with only the 3 cheapest Vehicle Types and write down the Vehicle Types' names in the blank below. | Checks how users create a chart, and if they can easily filter for useful information. |
| Q2 Susan wants to know the relations between Horsepower and other car properties. | Checks how users will analyze a given data perspective through *navigation between workspaces*. |
| Q3 Susan also wants to explore what impacts the Retail Price. Therefore, she wants to know the relations between Retail Price and other car properties. | Checks how users will analyze other data perspectives with different workspaces for *separate analysis paths*. |
| Q4 She came back to you again and asked about the relations between Highway MPG and other car properties to explore the trade-offs. As mentioned before, she again wants you to focus only on 3 cheapest Vehicle Types. | Checks how users create & edit a filter and copy it. Checks how users *revisit and retrieve* previous work. Checks how users perform comparisons. |
| Q5 Susan usually uses the car to take her family on short trips on weekends, so she may consider only SUVs or Minivans, with Highway MPG between 20 and 30, and Horsepower between 300 and 500 to save money. Does she have any options? | Checks how users drill down to details and interact with global and local filters. |

### C. Dataset and Tasks

The Cereal sample dataset [24] used during the training session contains information about 100 cereals with 16 dimensions. The formal analysis session used the Car dataset [25], which contains information about 428 cars with 14 dimensions. Both datasets were new to the participants and complex enough to allow for detailed exploration and analysis within the allotted time. Yet, they were small enough to enable participants to arrive at meaningful conclusions within the time constraints of our study. We presented the task scenario to each participant and gave them a brief explanation of the dimensions in the dataset. We asked participants to use DynSpace to create visualizations to complete a list of specified tasks. Table 1 provides details and the purpose of each task relative to our

research questions. The first three tasks encourage the participant to explore different aspects of the data and to use *navigation between workspaces*, with the third looking at *separate analysis paths*. The last two tasks ask the participant to filter and drill down into the details of the data, which also *revisiting and retrieving* previous work. Participants were free to create a *shared context* through bookmarks.

### D. Data Collection & Analysis

We used different analysis methods for the interaction logs, screen-capture videos, observations, questionnaires, and the open-ended interview. First, we analysed the interaction logs collected from DynSpace. Then, we used a thematic analysis to identify the different usage patterns from the screen capture and an open-ended coding approach [26] to categorize different patterns. Our analysis focused on how each participant used the multiple workspaces to complete their tasks. The time to completion for each analysis task was recorded by the survey system automatically, while the accuracy of task results was checked manually by the researchers. The questionnaire and interview data were analyzed to understand participants' opinions about DynSpace (see RQ2).

### E. Findings

We focused our analysis on the main interactions related to workspace management and the overall analysis process. Analysing two core workspace actions identifies that participants created between 2 to 10 workspaces (Avg: 5, SD: 2.18, Median: 5) and re-used previously created charts between 1 and 30 times (Avg: 6.6, SD: 7.9, Median: 3), with most people reusing at least one or two charts.
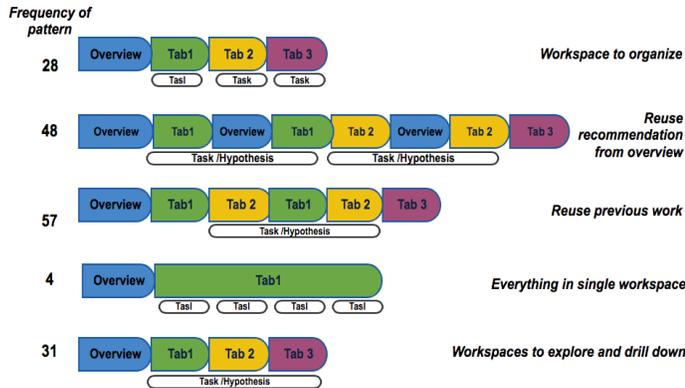


Figure 9. Illustration of the different usage patterns and their frequency. The bar underneath each line represent the task(s) or hypotheses that were investigated. The coloured blocks are abstract representations of the used workspaces. One can see how the workspace usage varies across different investigation patterns.

The main actions we coded during the process were: creating a new chart, creating a new workspace, switching the workspace, reusing a chart, adding a filter, coping a filter, sorting, and bookmarking. Tagging these interactions in the video and system logs, we created a list of common patterns and then clustered and merged them to finally identify common patterns. Using an open-ended coding approach [26], we identified six main patterns that participants used with multiple workspaces. While we observed multiple patterns, some

patterns were used more frequently than others. Figure 9 illustrates the different patterns and the frequency of each pattern observed in the study. The most frequent pattern observed is reuse from a previous workspace. More than half of the participants employed the multiple workspace patterns. The least used patterns are ones that use a single workspace (i.e., everything in single workspace, and comparing charts in single workspaces), and workspace for recovery, see Figure 10.
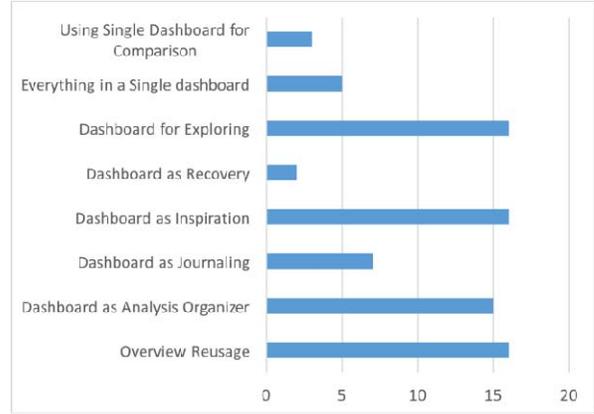


Figure 10. Number of participants that used each workspace pattern.

#### 1) Overview Workspace as starting point

In this pattern, participants start each task by searching the overview workspace to locate charts that could benefit their current task. They highlight the dimension they want explore or just visually scan the overview workspace. If they locate a chart, they copy it to a new workspace to investigate it in more details. This pattern was observed in multiple participants. For example, at the start of the third task, participant P11 opened the overview tab, highlighted the "Retail_price" dimension, and dragged two charts, Hwy_MPG vs. Retail_price and Vehicle_Type vs. Retail_price, to a new workspace to start the analysis. In the interview, Participant P11 reflected on that stating *"[The] Overview background is needed to create [a] new chart, since we are going to create something, some background information is necessary"*. The participant repeated this pattern for each task, creating a new chart, returning to the overview, then focusing on a specific dimension, then moving charts to the new workspaces. This demonstrates that participants considered the overview to provide a form of "context" for the data. Participant P18 mentioned that the overview helped in focusing their analysis: *"First, I go back to the Overview. I can focus on the chart that has been highlighted and it's easy for me to drag it into my analytic [tab]"*. This pattern was observed in 61.5% of the participants, which indicates the importance of the initial overview of the data.

#### 2) Workspaces as Analysis Organizers

Participants used multiple workspaces to organize their analysis by creating a new workspace for each analytical task. While a lot of participants created workspaces to organize their work, some (P1, P8, P13, and P23) even (re-)named them to reflect their current tasks. Participants used each workspace to focus on a specific analytical task, which – in their stated opinions – made their analysis easier and more efficient. For example, P11 indicated that *"I think I can say those tabs will*

*make my data more organized. Because let's assume I combined all the answers to [all] the questions in one tab will make it more complicated. So just isolating, creating new tab will be helpful for efficient analysis"*. Similarly, participant P8 created five different tabs and renamed each of them manually as Cheap, Power, Budget, Hwy_MPG, and SUV or Minivan, respectively. Also, participants found that they could benefit from grouping charts and globally filter and manipulate them simultaneously, as Participant 8 mentioned: *"I felt it's easier for me to have everything similar in the same category, and then I like the tab part because once you have your tabs, you can put the global filter on it."*

### 3) Reuse from Previous Workspaces

Some participants found that the multiple workspaces useful for saving and reusing their work process. While participants created multiple charts in each workspace, the ability to copy one chart from one workspace to the other was frequently used to support their analysis in multiple ways. Some participants preferred the drag-and-drop interaction of coping charts rather than creating a new chart, for example participant P14, who said *"I prefer to create a tab, and drag charts into it"*. Others found different analytical reasons to reuse charts from pervious charts. We observed three specific types of reuse between the multiple workspaces:

**Workspace for Inspiration**. When a participant seemed to be stuck when reflecting on the task questions, they went back to previous workspaces and searched for chart and dimensions that would help them. Participant P23 created a new workspace for her last task, and clicked the "Create new chart" button, but she did not know which data attributes to explore. She clicked on the overview workspace, scrolling up and down to get some ideas, and went back to her working workspace completing the creation of that chart. Another participant, P11 had already explored several perspectives of the data by drilling-down with filters. When he did not know what else could be explored, he went back to previous workspaces to collect more ideas and went on to explore new options.

**Workspace for Journaling**. Some participants kept the previous workspaces as a journal to potentially go back and retrieve previously created and analysed charts. The difference from the previous pattern is that here the analyst intent is to reuse a specific chart they previously created. Participant could filter on a chart and then copy it to another tab. For example, P11 created a new workspace for the last task and went back to their Vehicle_Type workspace, which was created previously, and dragged a chart (Vehicle_Type vs Dealer_Cost with filter on Vehicle_Type=Minivan, Wagon, Pickup) from there to the current working workspace. The participant then copied the filter from that chart and used it as a global filter.

**Workspace as Recovery**. Some participants made mistakes through wrong selections in filtering or by creating/copying unhelpful charts. P15 created a new tab and started from scratch for an analysis task by going back to "correctly working" workspaces and grabbing useful results from there.

### 4) Everything in a Single workspace

Another pattern we observed is using only a single workspace to complete the whole analysis. Here, a participant will create multiple charts and explore different aspect, drill down to details without ever leaving that workspace. They started each task by creating new charts, then used multiple local filters to drill down to details. As the number of charts increased, more and more scrolling became necessary when adding new charts or to change the layout to focus on a specific chart. When asked about their process during the interview, participant P7 answered that *"I am just lazy to create a new tab"*, and P3 *"I like just being able to see everything all at once, [I do not] like having to make a new tab for every chart, like you can see a lot of charts at the same time really easily"*. This pattern was observed in 19% of the participants, but each time for different reasons. While DynSpace is designed so that user can analyze with multiple workspaces, the system still supports users that complete their analysis in single one.

### 5) Workspaces for Comparisons

Participants could use multiple charts in a workspace to compare different aspects or different subsets of the data. Here, the user filtered two (or more) charts showing the same data with different (local) filters on the same dimension and then moved them side by side to visually compare them. For example, participant P12 placed two charts related to Retail_Price next to each other and discovered interesting insights from comparing the two charts. Similarly, participant P3 mentioned in the interview that DynSpace made it easy to compare saying *"if you had two specific things to compare, it's easy to see the two parts matched [in] the filters"*. Participant P20 resized two charts to compare details. Participants also indicated the interactions made the comparison they wanted easier. For example, P8 said *"I think it's pretty interesting because sometimes you want to compare something, it's pretty easy to drag the stuff and like then compare it"*.

### 6) Workspace for Exploration

Sometimes, a workspace was also used to explore specific aspect of the data and to drill down to details. We observed that several participants created multiple charts to explore specific dimension in a single workspace (e.g., Retail price vs Engine size, vs Cylinder, and price vs Horse power). While this is similar to using workspaces to organize the analysis, this pattern focuses on using a single workspace on drilling down to a single dimension or aspect of the data. This strategy allowed them to use the workspace to explore the different aspect of this specific dimension. Participant P6 explained *"it's clear to use different tabs when you are [not] focusing on the same subject"*. Participants P8, P13, and P26 renamed the workspace to match the dimension they were exploring.

### F. Participant feedback

The responses to the interview question provided additional feedback on the participant opinions' and added some suggestions. Overall, DynSpace received mostly positive comments. Participant P1 said that *"it's efficient to explore large amount of data, if the data information is very large, it's convenience to use it."* Participant 19 mentioned *"if someone needs to analyze a lot of Giga really quickly, […] it's a really good idea, because it's very visual [and] the visual part made it very easy to interact with."* Participant P5 replied about multiple workspaces, *"it's [easier] for me to analyze different purposes, different goals, answering different questions."* Participant P6 mentioned similarly *"it's clear to use different*

*tabs when you are focusing on [a different] subject. If I am doing task 2, I won't be like influenced by task 1."*

Participants also suggested modifications for DynSpace. While most participants had no problem interacting with DynSpace, some faced problems interacting with some aspect of the system. Participant P22 suggested adding scrollbars to bar charts when faced with the problem of visualizing vehicle names (i.e., more than 30 distinct values). Participant P3 suggested another filter icon saying, *"the filter symbol, it did not [look] it will make a filter"*. Participant P4 suggested to add multiple selection features mentioning that *"I cannot select two vehicles at the same time"*. Even though they could edit the filter to select specific categories, the ability the select categories from the chart itself was also seen as important.

### G. Other observations

While participants started their analysis using their own individual style, as the task progressed they were all able to utilize DynSpace to explore the data, filter and zoom, and drill down to details, which allowed them to explore the data and complete the tasks. While some users liked to see multiple recommendations, our observations show that others preferred a clear, empty workspace, where they could manually specify charts. Given the choice of having multiple recommendations or creating a new empty workspace for an uncluttered view in DynSpace, most participants choose to create new workspaces to do their own exploration. This confirms that (given the opportunity) many people use *separate analysis paths*, and that *easy navigation between workspaces* is critical.

Also, participants revisited and retrieved previous results, especially from the recommendations. Yet, easy, quick manual specification of charts enables users to explore data quickly. As the effort of creating new chart decreases, users seem to find the cost of creating and searching for bookmarked states, i.e., managing a *shared context*, or revisiting previous states to be higher. In other words, we observed that the quick chart generation feature reduced the need for reusing previous chart and states. For example, Participant P8 created more than 8 charts in each of 4 workspaces rather than searching previous workspaces or using bookmarks. This "lesson learned" implies that the need for bookmarking decreases when chart creation becomes very easy. Finally, while it is interesting to see that many participants actually used multiple workspaces in our hour-long user study – where they had only little incentive to use this feature (and some indeed choose not to use multiple workspaces). We believe that this indicates that participants see inherent value in multiple workspaces.

## VI. COMPARATIVE STUDY

In this second user study, we wanted to assess the effects of supporting multiple concurrent workspaces in comparison to a baseline single-workspace version of our prototype. For this, we compared the full version of DynSpace, which provides both bookmarking and multiple workspaces, with a baseline version that supports only bookmarking and provides only a single, main workspace. To enable simultaneous exploration of multiple dimensions, the baseline system allows users to specify multiple dimensions using a check box interface. The

goal of the study was to understand the impact of the availability of multiple workspaces on the analysis process and to compare it in a fair manner to a single state system.

We conducted a controlled within-subjects lab experiment. We used insight-based evaluation, similar to the work on Keshif [27]. However, because our goal was to investigate how participants track their analysis process, we recruited participants with prior VA experience. The comparative study had two factors (Tool, data), each with two levels (tool: baseline, full DynSpace, data: birdstrikes, video games), which resulted in four experimental conditions. We counterbalanced the presentation order of tools and data sets across participants.

We recruited 8 participants using public message boards and group lists (7 males and 1 female). Participants were graduate students in an HCI-related program (2 aged 29-40, 6 aged 23-28). Participants had a wide range of VA experience: 1 participant had more than three years, 3 more than a year, and 5 less then 6 months of experience. Most participants (7) had previous experience with Tableau or similar software.

We used two datasets: the birdstrike database and the video games sales and rating database. The bird strike dataset is a redacted version of FAA wildlife data which have been used in similar evaluations [4]. It consists of 3041 records and 13 dimensions (8 categorical, 1 temporal, and 4 numerical). The video game dataset [28] contains games sales, rating, platform, and genre data scrapped from the VGChartz website [29] and consists of 3340 records with 16 dimensions (6 categorical, 1 temporal, and 9 numerical). The study apparatus and lab conditions were similar to the first study

### A. Study procedure

Before the beginning of the study, participants answered a pre-study questionnaire to collect background information, including experience in data analysis and VA. Participants were randomly assigned to one of two groups, where one group started with the full DynSpace system and the other with the baseline. Each participant conducted two 30-minute analysis sessions (excluding training and practice), one with each system, each time with a different dataset. After the questionnaire, each participant was introduced to the datasets and asked to complete two consecutive sessions. Each session was divided into two phases, an open-ended exploration phase and another with focused questions. During the open-ended exploration, we asked participants to explore the dataset and note down interesting patterns, trends, and insights in an online system. They were asked to use the think-aloud protocol. After 15 minutes, the online survey system moved to the question answering task and the participants were asked to investigate five specific hypotheses for each dataset. For instance, hypotheses such as "Do critics and normal users usually have similar opinions towards a video game?" for the video games dataset, and "Bird strikes happening at night result in higher total damage costs in California than Pennsylvania" for the bird strikes dataset. Participants were also (again) asked to keep track of their work as they were going to be asked to summarize their work at the end of the session. After completing the first analysis session, the participants were asked to complete a second session with the other system.

In the exploration phase, participants performed unguided, self-driven exploration without any specific task for 15 minutes. The researcher attending the session acted as help and did not enforce any recommendation or direction to the analysis. However, the researcher encouraged verbal reflection as per the think-aloud protocol and reminded participants when communication stopped. The researcher also recorded the time when they wrote down each insight (through creation of a bookmark), without limiting the number of insights they could record. The screen capture software also recorded all verbal insights that participants mentioned during their exploration.

After the participants completed both sessions, they were asked to complete a survey that captured their preference for exploration and also their opinions about the usability of each system. At the end, each participant was asked to elaborate on their analysis process and their observations in a short semi-structured interview. We also asked about their impressions on the use of multiple workspaces versus bookmarks.

### B. Data collection and analysis

As in the first study, we used a variety of data collection analysis methods. We collected interaction logs, screen-capture videos, observations, questionnaires, and performed an open-ended interview. To create a concise interaction sequence for each participant, we integrated the interaction logs with the screen capture coding. High level interactions like creating charts, adding filter, and bookmarking charts were extracted from this log. From the screen capture, we added context information about the current participant hypothesis or insight related to these interactions. The result is a single high-level interaction log with the system, which we used in our analysis.

The insights were extracted from the transcribed screen capture video. To code the insights, we used a coding system similar to the one reported in Keshif [30]. We coded each insight based on its characteristic and the state of the system at that time. Additionally, most participants voiced questions and the hypothesis they were exploring. We coded these hypotheses with similar characteristics. The final list of characteristics for insights was as follow: type of statement (i.e., insight or questions); text of the insight; time of insight; correctness of the insight based on the current selected data; feature of the insight (i.e., fact, min or max, distribution, and correlation); relevant dimensions; current selected dataset subset; which dataset used (birdstrike or video games); and finally, which system was used (Baseline and DynSpace). As the sessions were divided into exploration and focused question answering phases, the insight coding was only applied to the 15-minute exploration at the beginning of each system session. However, we collected the participants' answers to the specific questions and computed the error rate in answering these questions, too.

### C. Findings

Participant most frequently reported min-max features (54%), describing how these data points have the highest or lowest value. For instance, Participant D4 noted that *"Shooter games have the most critics."* Comparisons between two data entities was not as frequent, with only 14.4% of the insights. *"Fighting is better scored than Shooters, yet it achieves less in*

*sales"* is an example of a comparison insight. Insights about correlations between two dimensions were next with 12.9%. Examples include D4 stating *"Users and critics have a nice understanding"*, referring to a linear relation between User-score and Critic-score. Participants also commented on the distribution of data in 8.63% of the insights. Most insights (90%) were created in the default overview of the datasets during the exploration. Few insights were generated from a subset or facet of the data.

While there is individual variation in how participants interacted with the systems, we observed fairly consistent overall patterns. Users started by exploring a specific aspect of the dataset by focusing on a single dimension. In the baseline system, more than half of the participants (5/8) started by selecting dimensions to explore. After that, they either viewed the recommendations or created new charts based on the same dimensions they were exploring. They used the recommendations to start asking questions about the data, which they then tried to answer by creating new charts or exploring other dimensions. In the full DynSpace system, more than half of the participants (5/8) started by creating an empty workspace and populated it with new charts of the dimensions they were exploring. The remaining participants (3/8) choose to use the explore dimensions' feature and viewed the recommendations in the created workspaces.

Participants created on average 5 workspaces during their session in the DynSpace condition, which supports our design guidelines for *easy navigation between workspaces* and *separate analysis paths*. Participants fluidly navigated between these workspaces. While many participants remembered the workspace where a given result was stored, we observed that some participants adopted a "search through workspaces" approach when trying to *revisit and retrieve* previous work. During the exploration phase, the average number of bookmarks created in the baseline system was 3.5, which again points to *shared context* usage. In DynSpace only 2.3 bookmarks were created on average, as users depended more on the multiple workspaces. Similar to our first study, and while users used multiple workspaces to organize and explore different aspect of the data, they often opted to create new charts rather than *revisited and retrieved*/reused previously created charts in their free exploration of the data. In the baseline, we noticed that participants repeatedly (M=8.1, SD=6.3) used the "explore dimension" feature more than once − rather than bookmark a specific chart. On the other hand, participant favoured creating new workspaces and creating new charts there in DynSpace, without the aid of recommendations.

During the interview, participants clarified their opinions about the multiple workspaces, bookmark features, and the filtering. The reasons for using multiple workspaces were consistent with our observations in the first study. Participants used multiple workspaces to reduce clutter, organize analysis, journaling or tracking of analysis, and to investigate multiple concurrent analyses, i.e., *separate analysis paths*. For instance, participant D4 explained using workspaces to explore two analysis paths: *"Tabs are also effective [if] you have [a] set of analysis [here] and [a] set of analysis there."* As for bookmarks, some participants explained that they used the bookmarks only to save their insights.

## VII. Discussion

Throughout both studies participants used the overview workspace frequently to support their exploration process, but also created new workspaces, mainly to drill down into details. After all, our motivation for multiple workspaces was enabling a better approach to organize the *analysis of separate paths*, as well as support for *re-visitation and retrieval* of previously created charts/results. Participants frequently recalled seeing a chart previously, looked for it, and then reused that chart by copying it to new workspaces. While most participants created multiple workspaces to complete different aspect or paths of their analysis, a few used a single workspace. It seems possible that the overhead in the context switch between different workspaces is the reason for working with a single workspace, which emphasizes that *easy navigation between workspaces* is important. Our findings for the support of *shared context* are mixed, but the value of enhanced bookmarking mechanisms may become more evident with charts that are more complex to create. Overall, our results still support our main hypothesis that users can benefit from a multiple-workspaces design and that they make it easier to pursue different analytical tasks independently. While it seems conceptually simple to add support for multiple workspaces to VA tools, our results show that their addition improves the workflow and supports the overall VA process better than current solutions. We acknowledge that more advanced chart types, such as parallel coordinate plots, can show more information, which could facilitate some (but not all) of the tasks we investigated. Yet, fewer people are able to quickly understand such charts, and the effort to create them is generally also larger. Also, reuse of partial results out of parallel coordinate plots is not trivial, especially if global and local filters are used in parallel.

An alternative approach for complex VA workflows is to use a large display, where analysts can organize their process [38]. Yet, such displays are currently still expensive and require a lot of desktop real estate. Moreover, the effort to organize charts on a large canvas (regardless if it is zoomable or not) can become a challenge. Our workspaces with non-overlapping grid layout offer a compromise between the overly restrictive single-workspace model and the freedom that a large canvas affords. After all, as the user has to manually manage overlap and resizing, a large canvas pushes more of the layout effort to position the content on the canvas onto the user. It is instructive to contrast our approach with the naïve approach of conducting an analysis multiple times and just comparing the results. As we observed that participants frequently recalled seeing previous, relevant results and retrieved and reused them, we believe that there is inherent value in support for multiple workspaces. Especially for direct comparisons, it is efficient to copy a workspace and to just change a global filter, which then in effect "re-runs" a whole analysis path to generate a different result – while still making each step of the path followed evident. Also, multiple workspaces enable users to organize their analysis work in a natural way, similar to how one starts a new page when trying a new idea.

## VIII. Impact And Potential Open Questions

***Consistent vs. Separate data state:*** Participants were able to explore different aspect of the data by creating workspaces with different global filters. Other systems maintain a consistent state across tabs to aid users in drilling down to various aspect of the data and avoid confusion or conflict within the data. Yet our participants were easily able to handle multiple tabs with different global filters. Further investigation on the impact of separate or consistent data state on the analysis process is thus warranted.

***Branching and Merging:*** Our results show the potential of using multiple workspaces in VA systems, where users employ the different workspaces to work on different analytical tasks, compare different aspect of the data, and revisit previous work. Users were able to revisit a previous, separate analysis path or aspect of their analysis easily through the multiple-workspace design. It would be interesting to investigate the use of multiple workspaces to support more complete tracking of interaction history and insight provenance. For example, a VA tool could enable users to branch from the current state of the analysis to pursue another analytical path or hypothesis. Using multiple workspaces in a branch-explore-merge analogy could also be used to support collaborative VA, where users can branch workspaces, explore on their own, then merge workspaces to combine and compare results. Moreover, such as system could check for duplication/redundancy of chart sequences between workspaces, to identify work sequences that could be merged. We intentionally did not explore this in our current prototype, as a downside of this approach is the increased overhead needed to manage the branching and merging process.

***Provenance:*** Insight provenance supports a wide range of use-cases, such as understanding a user's analytical reasoning process, e.g., [16], supporting collaboration, e.g., [31], improving exploration, e.g., [32], and enabling re-visitation and reuse, e.g., [33]. While multiple workspaces can support insight provenance, full support for provenance would also make the user interface more complex.

***Collaboration:*** While the focus of our work is on a single user, the concept of using multiple workspaces might assist collaborative work. The ability to separate workspaces could be expanded to support multiple users, e.g., by supporting awareness of others' interaction. Mechanisms similar to branch-explore-branch [34], could be employed in a multiple-workspaces interface. However, additional mechanisms are needed to fully support collaborative interaction [2].

## IX. Conclusion

We investigated the affordance and the benefit of using multiple concurrent workspaces in Visual Analytics tools. Based on design considerations for concurrent workspaces, in VA we presented an implementation of multiple workspaces with a multi-tab interface in DynSpace, which supports both exploration and focused question answering. We conducted two user studies to examine how users use multiple workspaces while exploring and analyzing data. The first study identified five main usage patterns for multiple workspaces, while the second, comparative study provided further evidence of the value of multiple workspaces for data exploration.

Our work opens many avenues for further investigation, including other design, user interface, and implementation choices for multiple workspaces in VA tools. While the focus

of our work was on how users interact with multiple workspaces in short analysis tasks, the same mechanisms might also support the knowledge generation process in other scenarios. As visualization chart scalability is orthogonal to our work, we deliberately limited the number of records in the datasets in our evaluation and focused on datasets with higher dimensionality. We plan to revisit this decision in future work.

## REFERENCES

[1] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim, "Knowledge Generation Model for Visual Analytics," IEEE Trans. Vis. Comput. Graph., vol. 20, no. 12, pp. 1604–1613, Dec. 2014.

[2] J. Heer and M. Agrawala, "Design considerations for collaborative visual analytics," Inf. Vis., vol. 7, no. 1, pp. 49–62, Feb. 2008.

[3] J. D. Mackinlay, P. Hanrahan, and C. Stolte, "Show me: Automatic presentation for visual analysis," IEEE Trans. Vis. Comput. Graph., vol. 13, no. 6, pp. 1137–1144, 2007.

[4] 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," in Proc. CHI Conference on Human Factors in Computing Systems, 2017, pp. 2648–2659.

[5] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations," in Proceedings of IEEE Symposium on Visual Languages, 1996, UMCP-CSD CS-TR-3665, pp. 336–343.

[6] S. Few, Information dashboard design. O'Reilly, 2006.

[7] S. Mittelstadt, M. Behrisch, S. Weber, T. Schreck, A. Stoffel, R. Pompl, D. Keim, H. Last, and L. Zhang, "Visual Analytics for the Big Data Era - A Comparative Review of State-of-the-art Commercial Systems," in Proc. IEEE Conf. on Visual Analytics Science and Technology, 2012, pp. 173–182.

[8] M. Elias and A. Bezerianos, "Exploration views: Understanding dashboard creation and customization for visualization novices," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 6949 LNCS, no. 4, pp. 274–291, 2011.

[9] "SAP BusinessObjects Lumira." [Online]. Available: http://saplumira.com/. [Accessed: 20-Mar-2017].

[10] K. Wongsuphasawat, D. Moritz, A. Anand, J. Mackinlay, B. Howe, and J. Heer, "Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations," IEEE Trans. Vis. Comput. Graph., vol. 22, no. 1, pp. 649–658, 2016.

[11] D. Gotz, Z. Wen, J. Lu, P. Kissa, M. X. Zhou, and S. X. Liu, "HARVEST: Situational Visual Analytics for the Masses," in Proceedings of the First InternationalWorkshop on Intelligent Visual Interfaces for Text Analysis, IVITA, 2010, pp. 1–8.

[12] M. Dörk, N. Henry Riche, G. Ramos, and S. Dumais, "PivotPaths: Strolling through Faceted Information Spaces," IEEE Trans. Vis. Comput. Graph., vol. 18, no. 12, pp. 2709–2718, 2012.

[13] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala, "Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation," IEEE Trans. Vis. Comput. Graph., vol. 14, no. 6, pp. 1189–1196, Nov. 2008.

[14] C. Dunne, N. H. Riche, B. Lee, R. A. Metoyer, and G. G. Robertson, "GraphTrail: Analyzing Large Multivariate, Heterogeneous Networks while Supporting Exploration History," in Proc. CHI Conference on Human Factors in Computing Systems, 2012, pp. 1663–1672.

[15] C. North, R. Chang, A. Endert, W. Dou, R. May, B. Pike, and G. Fink, "Analytic Provenance: Process + Interaction + Insight," in Proc. CHI Conference on Human Factors in Computing Systems, 2011, pp. 33–36.

[16] D. Gotz and M. X. Zhou, "Characterizing Users' Visual Analytic Activity for Insight Provenance," Inf. Vis., vol. 8, no. 1, pp. 42–55, Jan. 2009.

[17] D. Gotz, Z. Wen, J. Lu, P. Kissa, M. X. Zhou, N. Cao, W. H. Qian, and S. X. Liu, "Interactive Poster: HARVEST-Visualization and Analysis for the Masses," pp. 1–2, 2008.

[18] J. Huang and R. W. White, "Parallel browsing behavior on the web," in Proceedings of the 21st ACM conference on Hypertext and hypermedia, 2010, pp. 13–18.

[19] L. Li, G. Chen, L. Li, and G. Chen, "A Web Browser Interface to Manage the Searching and Organizing of Information on the Web by Learners," Educ. Technol. Soc., vol. 13, no. 4, pp. 86–97, 2010.

[20] A. Satyanarayan, R. Russell, J. Hoffswell, and J. Heer, "Reactive Vega: A Streaming Dataflow Architecture for Declarative Interactive Visualization," IEEE Trans. Vis. Compter Graph., vol. 22, no. 1, pp. 659–668, 2016.

[21] M. Bostock, V. Ogievetsky, and J. Heer, "D3; Data-Driven Documents," IEEE Trans. Vis. Comput. Graph., vol. 17, no. 12, pp. 2301–2309, 2011.

[22] P. H. Nguyen, K. Xu, A. Wheat, B. L. W. Wong, S. Attfield, and B. Fields, "SensePath: Understanding the Sensemaking Process Through Analytic Provenance," IEEE Trans. Vis. Comput. Graph., vol. 22, no. 1, pp. 41–50, 2016.

[23] "Superstore Sales Dataset." [Online]. Available: https://community.tableau.com/docs/DOC-1236. [Accessed: 01-Mar-2017].

[24] "Cereal Dataset." [Online]. Available: https://perso.telecom-paristech.fr/eagan/class/igr204/datasets. [Accessed: 01-Jan-2017].

[25] "Car Dataset 2004." [Online]. Available: https://perso.telecom-paristech.fr/eagan/class/igr204/datasets. [Accessed: 01-Jan-2017].

[26] J. W. Creswell, Research design: Qualitative, quantitative, and mixed methods approaches. Sage Publications, 2013.

[27] M. A. Yalcin, N. Elmqvist, and B. B. Bederson, "Keshif: Rapid and Expressive Tabular Data Exploration for Novices," IEEE Trans. Vis. Comput. Graph., vol. PP, no. 99, p. 1, 2017.

[28] "Video Game Sales with Ratings," 2017. [Online]. Available: https://www.kaggle.com/rush4ratio/video-game-sales-with-ratings. [Accessed: 01-Aug-2017].

[29] "Video Game Charts," 2017. [Online]. Available: http://www.vgchartz.com/. [Accessed: 01-Aug-2017].

[30] M. A. Yalcin, N. Elmqvist, and B. B. Bederson, "Keshif: Rapid and Expressive Tabular Data Exploration for Novices," IEEE Trans. Vis. Comput. Graph., vol. 2626, no. c, pp. 1–1, 2017.

[31] A. Sarvghad and M. Tory, "Exploiting Analysis History to Support Collaborative Data Analysis," in Proceedings of the 41st Graphics Interface, 2015, pp. 3–5.

[32] A. Sarvghad, M. Tory, and N. Mahyar, "Visualizing dimension coverage to support exploratory analysis," IEEE Trans. Vis. Comput. Graph., vol. 23, no. 1, pp. 21–30, 2017.

[33] N. Kadivar, V. Chen, D. Dunsmuir, E. Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury, "Capturing and Supporting the Analysis Process," in IEEE Symposium on Visual Analytics Science and Technology, 2009, pp. 131–138.

[34] W. McGrath, B. Bowman, D. McCallum, J. D. Hincapié-Ramos, N. Elmqvist, and P. Irani, "Branch-explore-merge: Facilitating Real-time Revision Control in Collaborative Visual Exploration," Proc. 2012 ACM Int. Conf. Interact. Tabletops Surfaces, pp. 235–244, 2012.

[35] A. Key, B. Howe, D. Perry, and C. Aragon, "VizDeck: self-organizing dashboards for visual analytics", Proc. 2012 ACM SIGMOD, pp. 681-684, 2012.

[36] F. Bouali, A. Guettala, and G. Venturini, VizAssist: an interactive user assistant for visual data mining. Visual Computer, vol. 32, no. 11, pp. 1447-1463, 2016.

[37] M. Vartakm S, Rahman, S. Madden, A. Parameswaran, N. Polyzotis, SeeDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics, Proc. VLDB Endowment, vol. 8, no. 13, pp. 2182-2193, 2015.

[38] C. Andrews, A. Endert, and C. North, Space to think: large high-resolution displays for sensemaking. In Proc CHI Conference on Human Factors in Computing Systems. pp/ 55-64, 2010.