SoK: A Systematic Review of Filtering in Information Visualization and Visual Analytics

Supplemental Analyses

Carlotta Guiso,
Marco Angelini

Contents

1 Introduction

In this supplementary material, we would like to expand further on our research presented in the main scientific article, providing further details on the process of cataloging the papers and sharing some interesting analyzes and aspects that emerged during this process, but which were not included in the final version of the main article. These additional details offer a more complete and in-depth view of our study and may be of interest to readers and researchers who want a more detailed understanding of our work.

During the paper cataloging phase, we examined a wide range of documents related to our area of study, systematically analyzing their characteristics, the methods used, the results obtained and other relevant factors. This process allowed us to identify key trends, patterns, and aspects that contribute to our understanding of the research field. However, due to space limitations in the main article and the need to focus on the main topics, we could not discuss in detail all the interesting observations and secondary analyzes we discovered during this phase.

This additional material was also created to report all the results of the analyzes carried out for each individual taxonomic group and is an essential component of the scientific article on filtering in information visualization and visual analytic systems. Its creation was motivated by the need to enrich and consolidate the research presented in the article itself. This material was specifically developed with the aim of offering a detailed and in-depth analysis of the significant differences that emerged during the study of the various taxonomic groups examined. In doing so, it contributes substantially to ensuring the transparency, completeness and reproducibility of the scientific research that has been conducted.

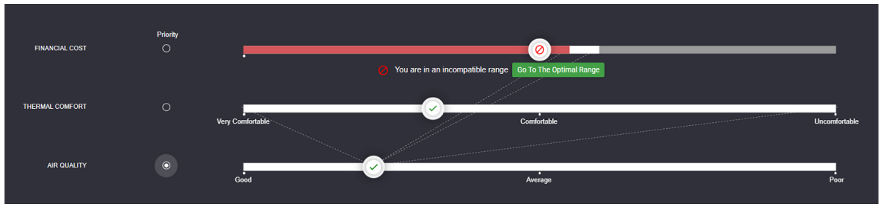
Transparency is fundamental because it allows other scholars to critically evaluate the methodology used and the data collected, allowing them to verify and validate the results presented in the article. This aspect is crucial for the credibility of scientific research and for promoting progress in the field of the study of visual systems.

Furthermore, the complete exposure of the analysis results for each taxonomic group allows other researchers to further deepen their work by conducting further research based on this detailed information. This can contribute to further discoveries and developments in filtering in visual systems, thus enhancing the scientific knowledge base available to the scientific community and interested public.

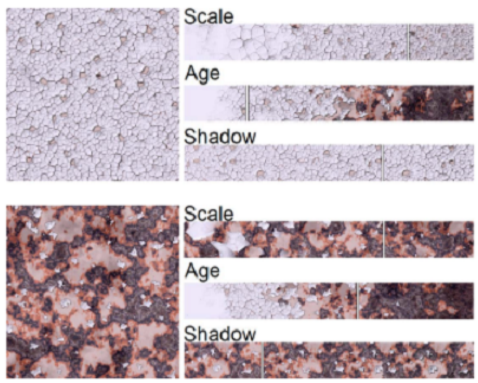
2 Enhancing Filtering Papers

Within the vast panorama of research in visual analytic systems, papers emerge that present themselves as active contributions to the very evolution of the filtering discipline. These articles represent a conscious effort to address the challenges inherent in visual data analysis by devising and applying new models and filtering techniques. These papers do not just present simple applications of existing filters; rather, they aim to challenge the pre-existing paradigm and suggest innovative ways to optimize filtering operations. The authors of these contributions try to explore new frontiers, exploring how filters can be adapted, combined or completely reinvented to obtain better and more accurate results. The presence of these papers within the research panorama demonstrates the vitality and commitment of the scientific community in making the visual analysis of data ever more accurate and effective.
During the cataloging of the papers, two areas were identified in which relevant contributions regarding filtering were identified. The first area concerns the implementation of new types of sliders. These interactive tools have captured the attention of numerous researchers who have proposed new types of sliders to refine and personalize the filtering experience.
Laurillau et al. [1] is a clear example that falls into this area. They develop a new type of slider, TOP-Slider (Figure 1a), which is composed of a set of parallel sliders enriched with visual features. One aspect to highlight is that there is a relationship of interdependence between the criteria that make up the filter. When a criterion is modified, the system automatically adapts to the new conditions set, reflecting the cross-influences between the same criteria.
Lasram et al. [2] instead propose visually guided sliders for procedural textures in order to simplify parameter selection and offer users greater control over the result (Figure 1b). Each filter is composed of a series of parameters such as the blur intensity or the direction of the relief and their previews allow users to quickly reveal the possible effects of the changes made.
Other authors have contributed to the development of new types of sliders in the context of visual data analysis [3, 4, 5]. By introducing these new slider variants they have broadened the landscape of options available to designers.

The second area is about viewing on sets. This visualization approach allows for a clear and



(a)



(b)

Figure 1: (a) TOP-Slider implementation [1], (b) Two possible texture configurations, proposed by Lasram et al. [2], by changing the slider parameters.

intuitive visualization of the intersections and relationships between the variables present within a dataset. These visualizations allow for a clear representation of how different variables or filters overlap or intersect, providing a broad overview of data complexity.

For example Yalçın et al. [6] introduce AggreSet, an interactive visualization method for exploring relationships within multivariate datasets (an example can be seen in Figure 2). This method enhances existing set visualization techniques through scalable data aggregation, providing a simple and intuitive interface, and presents the aggregated data as linked summaries that respond to user interactions. There are several filtering activities such as filterable visual distributions, filtering the elements of an aggregate or instead, for the set-lists, there are three filtering modes: union, intersection and exclusion.

Another paper that falls into this area is that of Lex et al. [7]. They introduce UpSet, a visualization technique for analyzing sets, their intersections, and aggregates. From the Figure 3 it can be seen that the matrix layout is used to display the intersections of the sets and introduces aggregates through grouping and queries, finally, the sets and their attributes are

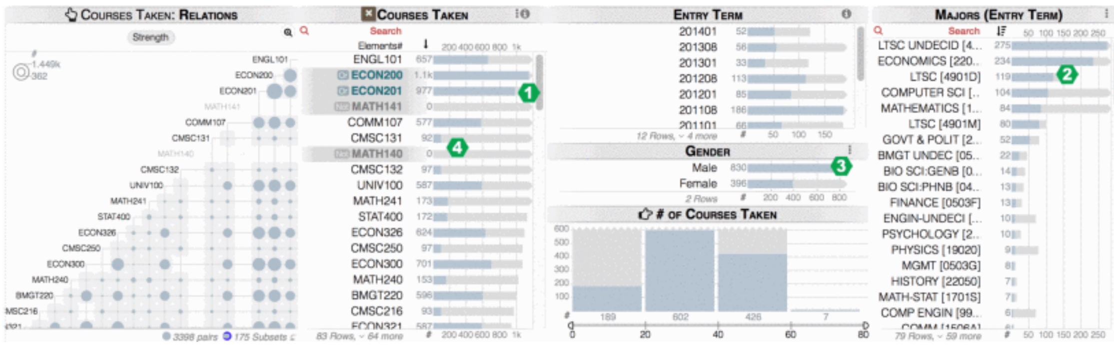


Figure 2: Exploring a dataset using AggreSet [6].

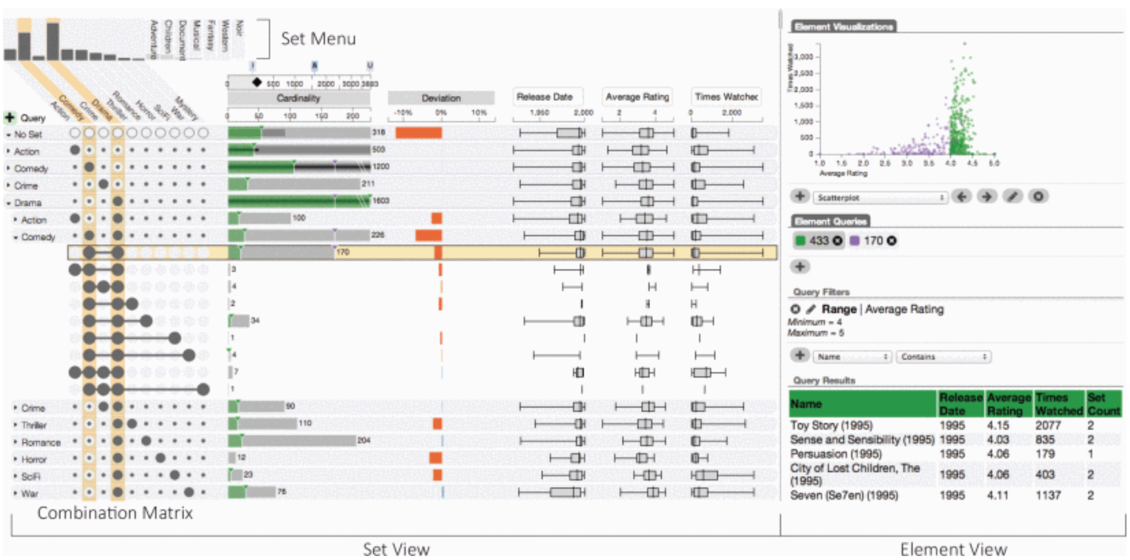


Figure 3: UpSet interface [7].

displayed separately and queries based on attribute intersections, aggregations or filters can be propagated between both views. To create a query, the user defines the filters that must be applied so that only the elements that satisfy those certain conditions remain. The filters that can be found are string attribute filters, maximum and minimum and range filters and if multiple filters are used for a query then the result of that query is the intersection of the results of the individual filters.

3 Full Analysis Description

Chapter 3 provides a complete overview of all analyses conducted. In this chapter, all aspects of the analyses are presented and explored in depth (even those not present in the paper) and allow readers to understand each analysis in detail, allowing them to delve deeper into specific areas of interest through reference to the analysis numbers. Figure 4 shows a hierarchical approach used to organize analyses. To ensure the reproducibility of the results, all analyses were performed using scripts that are made available in the supplemental material in the "code" folder.

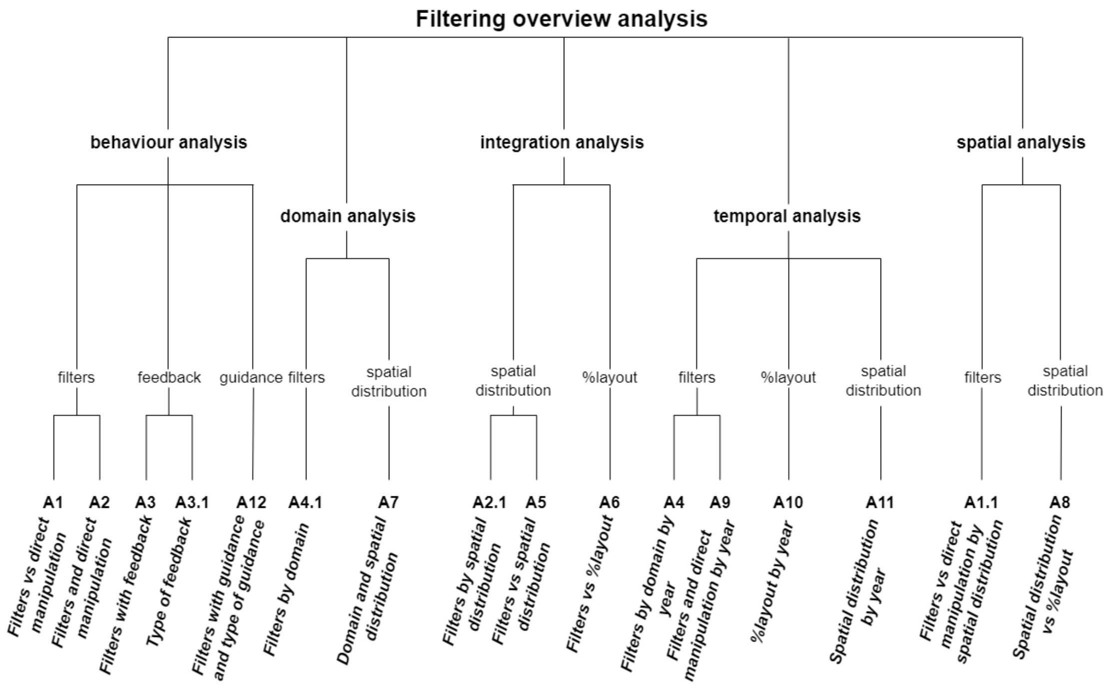


Figure 4: Overview of designed analyses

3.1 Behavior Analysis

The main objective of this analysis is to identify the behavioral patterns that emerge from the surveyed papers.

Analysis of the relation of filters with respect to direct manipulation (A1)

This analysis relates the number of filters and the number of direct manipulations. The main objective of this analysis is to verify if there is an inverse proportionality between the number of filters used and the number of direct manipulations.After running the analysis on the whole corpus, a slightly downward trend was noted between the two variables (see Figure 5).

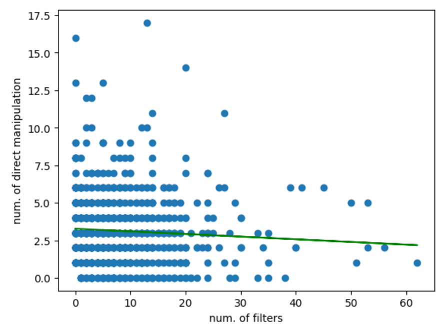


Figure 5: Relationship between filters and direct manipulation discussed in analysis A1

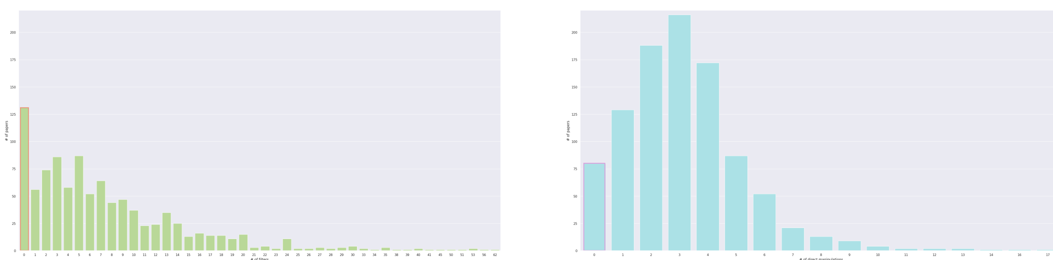
However this trend is not pronounced enough. To further investigate the relationship, two additional analyses were executed based on the statistical distributions for direct manipulation and filters: *(i)* Removing just the outliers from both distributions, a steeper decline was noted, suggesting that in some situations, a significant increase in filters might actually lead to a decrease in direct manipulations. *(ii)* Considering only the values inside the boxes (second and third quartiles), a flat trend was observed.

These three analyses combined led us to conclude that there seems to be no significant relationship between the number of filters and the number of direct manipulations and only slight inverse proportionality is observable when the number of filters or direct manipulations grows significantly. The analysis clearly demonstrated that the relationship between the number of filters and the number of direct manipulations is more complex than initially assumed. A clear and unique inverse proportional relationship between these two variables did not emerge. The absence of a clear inverse relationship emphasizes the need to consider additional variables that could affect how users use direct filters and manipulations. Finally, a further suggestion may be to adopt an approach that allows users to choose between filters and direct manipulations, based on their preferences and the characteristics of the analysis they are performing, this could improve the overall experience and encourage more efficient use visual systems.

The analysis was extended to each individual taxonomic group, to examine whether the trends identified in the global analysis were consistent across all groupings. Interestingly, for some taxonomic groups show results that differ from those found overall, for example, in that include the similar filters a directly proportional trend can be noted.

Analysis of filters and direct manipulation distribution (A2)

In this analysis, a detailed study on the frequency distribution of the filters and direct manipulation has been conducted, with the aim of determining interesting average trends in their use in visual systems.



(a) (b)

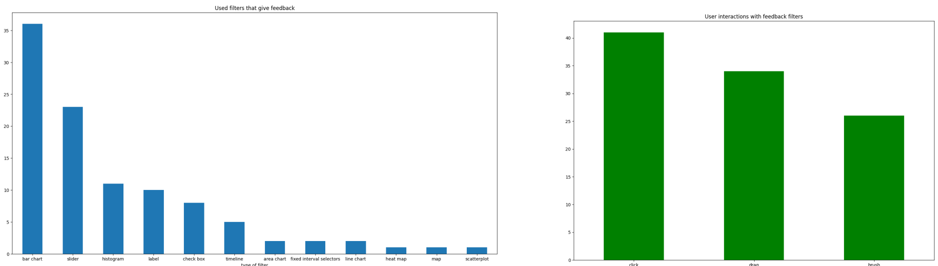
Figure 6: (a) papers for each number of filters, (b) papers for each number of direct manipulations

Concerning filters (Figure 6a), a strong concentration of papers use a number of filters that vary between 1 and 14 (median 7.5), suggesting that in system design, a manageable number of filters are often employed rather than a multitude of options. An exception is represented by papers that map data set features to single filters, for which this number increases significantly. Regarding direct manipulation(Figure 6b), a similar trend is observed. Most papers use a number of direct manipulations that vary between 1 and 6 (median 3.5), furthermore, it is evident that some papers do not implement direct manipulation at all, suggesting that this interaction may not be relevant or necessary for all visual systems.Interestingly, in the case where there are no direct manipulations, a greater variability is observed in the distribution of filters. This might suggest that without direct manipulations, designers of visual systems choose to implement full interactivity (filtering included) through filters. These filters could be used to customize the analysis or provide users with more options to explore the data.Adding too many filtering options or complexity in direct interactions can improve functionality but complicate the user experience so a balance must be found between advanced functionality and efficiency.

It is important to underline that the analysis conducted was also extended to each individual taxonomic group.The results obtained from these detailed analyzes within taxonomic groups were congruent with those of the overall analysis.

Analysis of filters with feedback and their interactivity (A3)

Through this analysis, we seek an in-depth understanding of the types of visual encodings used to convey user feedback through filters and what is the predominant interaction that allows users to interact with these filters.



(a) (b)

Figure 7: (a) Types of filters with feedback used, (b) interactions used for filtering.

Figure 7a shows that bar charts and visually enriched sliders are the most commonly used visual encodings to provide feedback to the user, followed by histograms, labels, and checkboxes which indicates that in any case there is a variety of approaches in implementing filters with feedback. Furthermore, it is interesting to note that new types of filters with feedback are also emerging, in particular, we can see the first adaptation of geographic maps [8] and scatterplots [9] as new types of filters used to provide feedback to users. This suggests that there may be growing interest in experimenting with different forms of visualization to improve user experience in visual systems. For what concerns interactivity Figure 7b shows three predominant types of interaction: click (41), brush (34), and drag (26). This result suggests exploring more advanced and customized interaction techniques for filters such as radial brush or lasso, which allow users to make more precise and detailed selections in the data while still acting on the filters.

This analysis was extended to each individual taxonomic group. However, in contrast to what was previously stated, some taxonomic groups, such as in the group direct manipulation dominated → aggregated filters or filter interface dominated → distributed filters, show results with different distributions from those obtained in the global analysis.

Analysis of the types of feedback (A3.1)

This analysis focuses on the semantic content of the filtering that the different types of feedback support.

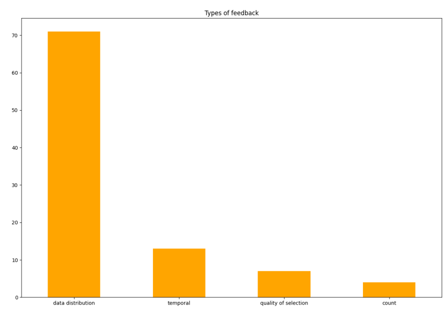


Figure 8: Types of feedback used in filters.

Figure 8 shows that "count" type feedback, acting simply on the cardinality of the data item selected, is little used as a type of feedback. This indicates that filters often provide more detailed information about the distribution of the data. The description of the data distribution is the most often used. The analysis highlights the presence of emerging trends regarding the feedback provided by the filters. It shows a growing interest in exploiting temporal data distribution and more interestingly supporting the feedback on the data selection quality given a score function [10]. The analysis conducted on each individual taxonomic group was congruent with that of the overall analysis.

Analysis of filters with guidance and associated forms of guidance (A12)

The analysis focuses on the filters that offer a form of guidance to users and on the different types of guidance implemented in these filters. Given the limited presence of just two contributions, we proceed to analyze them directly. Leite et al. [11] use (Figure 9) sliders, box-shaped glyphs, and networks that are enriched by some colored checkmarks which are deactivated when they do not match the filters that have been defined by guiding the user in the selection of filters. Angelini et al. [12] uses explicit arrows to guide the user in the expansion/reduction of the size of the data selection through a multi-feature filtering interface. While the former paper presents an example of prescriptive guidance (by removing options and leaving just the valid ones), the latter provides orienting guidance. It is too early to

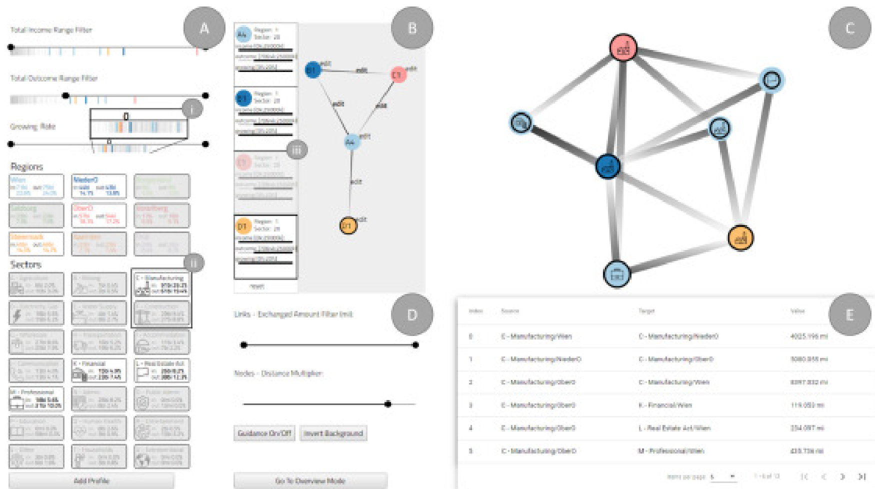


Figure 9: System interface implemented by Leite et al. [11]: section A and B show the implemented filters with guidance.

draw conclusions, but it is interesting to note that some contributions start to include guidance even for filtering. This analysis, although the number of papers considered is limited (only two), suggests that filter design can be tailored to specific application needs, therefore highlighted that there is no one-size-fits-all solution for filter guidance. It is important to consider the specific needs of the application and target audience to determine what type of guidance is most appropriate. Finally, given the limited amount of research considered in the analysis, it may be useful to conduct further experimental studies to evaluate the effectiveness of different forms of guidance in optimizing user-filter interaction and further improve the design of filters with functionalities. support.

3.2 Integration Analysis

This analysis examines the integration of different categories of data and information within systems, revealing combination preferences among the most common categories or data sources and providing a deeper understanding of the data itself.

Analysis of filters distribution based on spatial distribution (A2.1)

In the context of analysis A2.1, the investigation carried out in analysis A2 was further explored, focusing on the frequency distribution of the filters and on the trends relating to different types of spatial distribution.Figure 10 shows several significant trends;it can be observed that most of the papers tend to use a relatively low number of filters which varies from 1 to 20 filters, while it is evident that there are few papers that use a high number of filters. Also, when a few filters are used, there is a tendency to prefer a unified layout over the others. As the number of filters increases, the distributed and distributed compact layouts tend to rise in usage. For a very high number of filters, a marked preference for a specific type of spatial distribution becomes less evident.

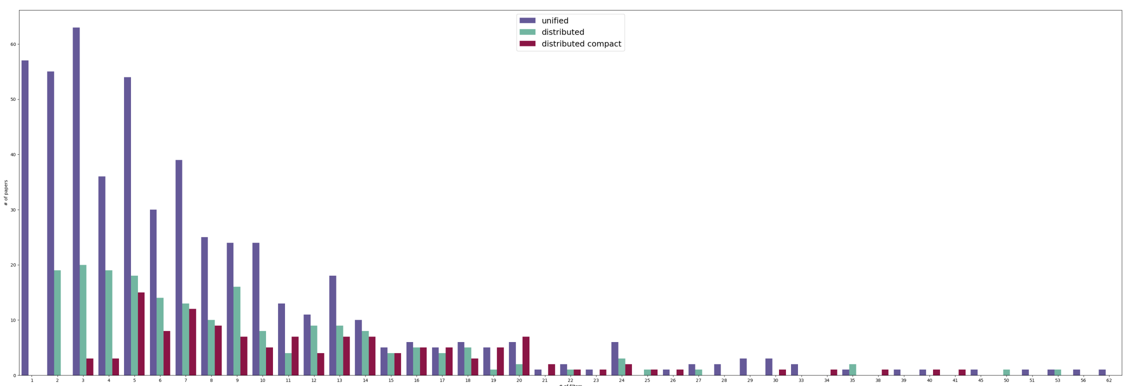


Figure 10: Distribution of filters divided by spatial distribution

This suggests that, in the design of visual systems, the choice of spatial distribution could be significantly influenced by the complexity of the interface due to the number of filters involved.

Analysis of the presence of filters for each spatial distribution (A5)

The A5 analysis represents a deepening of the previous analysis. The main objective is to examine whether there is a relationship between the number of filters present in a system and the type of distribution chosen for these filters.

From the table in Figure11a, some interesting trends emerge: it confirms that as the number of filters increases, systems tend to distribute these filters more widely within the interface. To avoid bias due to the presence of outliers or extreme cases, we conducted the analysis on the box of the boxplot representing the frequency distribution of filters, split into equal parts (seeFigure 11b). The trend confirms that as the number of filters increases, the difference between the unified filters and the distributed ones (which include distributed and distributed compact) seems to be reduced, almost vanishing. However, analyzing the individual columns shows some interesting variations. For example, the number of systems

Filters	unified	distributed	distributed compact
few (>=0 & <2)	57	0	0
average (>=2 & <=10)	350	137	62
many (>10 & <=29)	88	52	57

(a)

divided equally	unified	distributed	distributed compact
I (>=2 & <=4)	154	58	6
II (>4 & <=7)	123	45	35
III (>7 & <=10)	73	34	21

(b)

Figure 11: (a) Filter distribution table for each spatial distribution, (b) filter distribution table with boxplot divided equally.

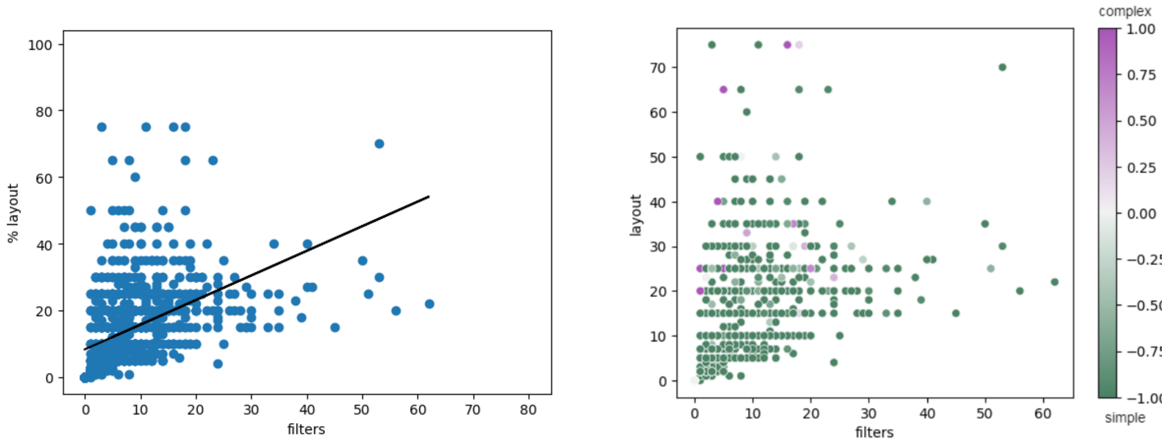
with unified filters appears to decrease, while those with compact distributed filters begin to increase before stabilizing, while distributed ones experience only small changes. Ultimately, this analysis offers important insights for the design of visual analytics systems, highlighting the importance of considering the number of filters and their spatial distribution to improve the usability and effectiveness of the user interface. The implementation of filters within a system can vary greatly depending on the number of filters involved, suggesting that designers may take different approaches based on the specific needs of their system.

Analysis of the relation between filters and occupied space (A6)

The analysis focuses on exploring the relationship between two variables: the number of filters present in a system and the percentage of the layout that these filters occupy. The main goal is to check if there is a direct proportional relationship between these two variables. Figure 12a shows that as the number of filters increases, the percentage of layout occupied by them tends to increase in a directly proportional way. This means that when more filters are added to a system, they require more space in the user interface.

This relationship was subsequently analyzed by separating systems with a predominance of simple filters from systems with a predominance of complex filters. Figure 12b) shows results plotted in five intervals defined as follows by the ratio of these two types of filter:

- Range -1, -0.75:This range consists of systems where simple filters clearly dominate over complex ones. The percentage of occupied layout varies on average from 5% to 20%. In other words, in systems with a predominance of simple filters, there is little occupancy from filters.



(a) (b)

Figure 12: (a) Relationship between filters and % of layout, (b) Color scale scatterplot.

- Range -0.75, -0.25:In this range, simple filters continue to dominate, but with a slightly less pronounced trend than in the previous range. The percentage of occupied layout increases slightly, varying on average from 17% to 25%.

- Range -0.25, 0.25:This range represents systems where the number of simple and complex filters is balanced or comparable. In this situation, an increase in the occupied layout is observed, with a percentage that varies on average between 15% and 30%. This indicates that a balanced combination of simple and complex filters can lead to a higher occupation of the interface space.

- Range 0.25, 0.75:In this range, systems begin to show a predominance of complex filters over simple ones. The percentage of occupied layout varies on average from 15% to 25%.

- Range 0.75, 1:This range consists of systems where complex filters significantly dominate. The percentage of occupied layout varies on average from 15% to 32%. Here, a steep slope of the curve is observed, which suggests that as the filters increase in complexity and number the percentage of the layout occupied by the filters increases significantly.

This analysis provided a clear understanding of the relationship between the number of filters and the percentage of layout occupied within the visual analytics systems interface that can inform visual system designers.

It is essential to highlight that the detailed analysis was extended to each individual taxonomic group, in order to examine whether the trends identified in the global analysis were consistent across all groupings. However, in contrast to what was previously stated, some taxonomic groups show results that differ from those found overall.

3.3 Spatial Analysis

This analysis explores the spatial interactions between different variables, providing an essential perspective to understand the mutual influences between them.

Analysis of the relation of filters with direct manipulation based on spatial distribution (A1.1)

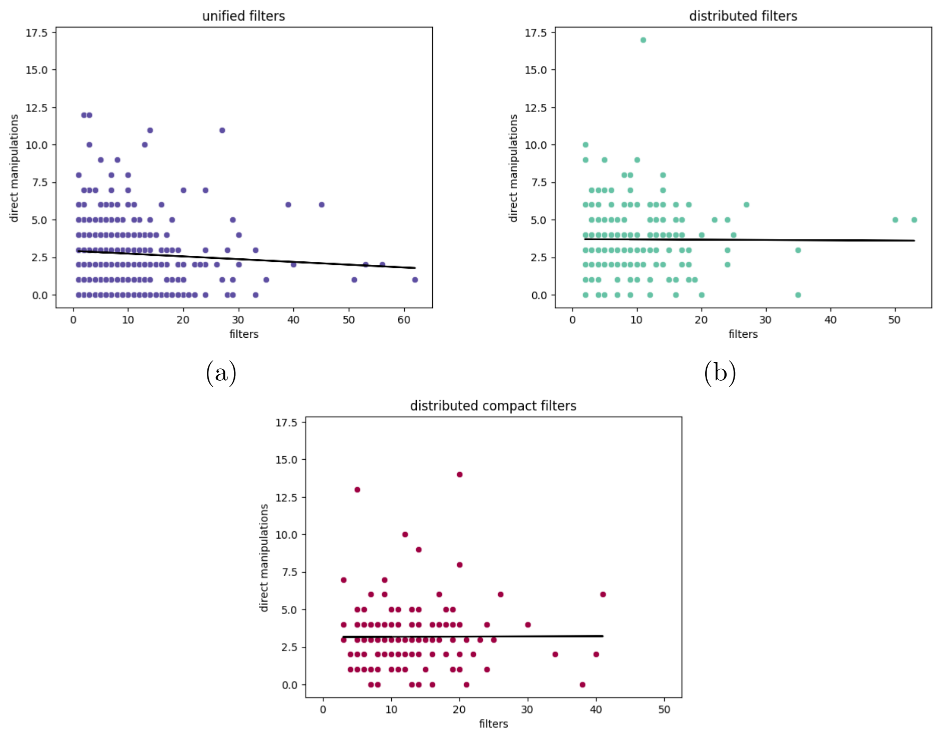
This analysis represents an insight into the work done in the A1 analysis. The main objective of this analysis is to examine whether there is an inversely proportional relationship between the number of filters used and the number of direct manipulations considering the different spatial distributions of the filters. The relationship between the number of filters and the number of direct manipulations seems to show a slight decline, indicating that there may be some type of relationship between these variables, but it is not strong enough to state with certainty that there is an inverse proportionality.

To gain a more detailed understanding we split the analysis into three sub-analyses based on the spatial distribution of the filters:

- In the subanalysis dedicated to unified filters (Figure 13a) there is a slight inversely proportional trend between the number of filters and the number of direct manipulations, however the results of this subanalysis did not show a significant change in the trend compared to the analysis overall.

- In the subanalysis dedicated to distributed filters (Figure 13b), a flat trend can be seen, therefore as the number of filters increases there is no decrease in the number of direct manipulations. This trend suggests that the existence of an inversely proportional relationship between the two variables cannot be affirmed.

- Even in the sub-analysis dedicated to distributed compact filters (Figure 13c) we notice a flat trend, therefore also in this case there is no significant variation between the number of compact distributed filters and the number of direct manipulations and it cannot be affirm the existence of an inversely proportional relationship.



(a) (b) (c)

Figure 13: Relationship between numbers of filters and direct manipulation: (a) unified filters, (b) distributed filters and (c) distributed compact filters

Therefore, both the overall analysis and the sub-analyses conducted on the spatial distribution of the data did not allow us to state with certainty that there is an inversely proportional relationship between these two variables. This means that as the number of filters increases, there is not necessarily a consistent decrease in the number of direct manipulations, and vice versa.

Therefore, the importance of considering additional factors that can influence the relationship between direct manipulation and filters emerges. This more nuanced approach can contribute to a more complete understanding of the dynamics between filters and direct manipulations.

Analysis of the effects that spatial distribution has on the percentage of layout (A8)

In this analysis, the three types of filter distribution are investigated with respect to the percentage of layout they occupy in the viewport. It allows us to understand if there is a correlation between these two variables and if the choice of a type of distribution affects the occupied space.

- The analysis found that distributed filters take up less space, ranging from 5% to 15%. (Figure 14a).

- Distributed compact filters appear to take up much more space, ranging from 15% to 27%. This could be because the combination of a grouped and a distributed part or two grouped parts distributed in the viewport take up more space than distributed filters (Figure 14b).

- The unified filters were found to be more variable in terms of occupied space (Figure 14c), ranging from 5% to 40%, with the presence of many outliers (even up to 80% of occupation). Removing them the distribution seems to stay with a maximum number of 10 filters and occupation up to 25% of the layout. This suggests that unified filters require a significant amount of space from the outset, but also feature a compact swathe with a concentration of systems using them more heavily, before a progressive relaxation of distribution (Figure 14d).

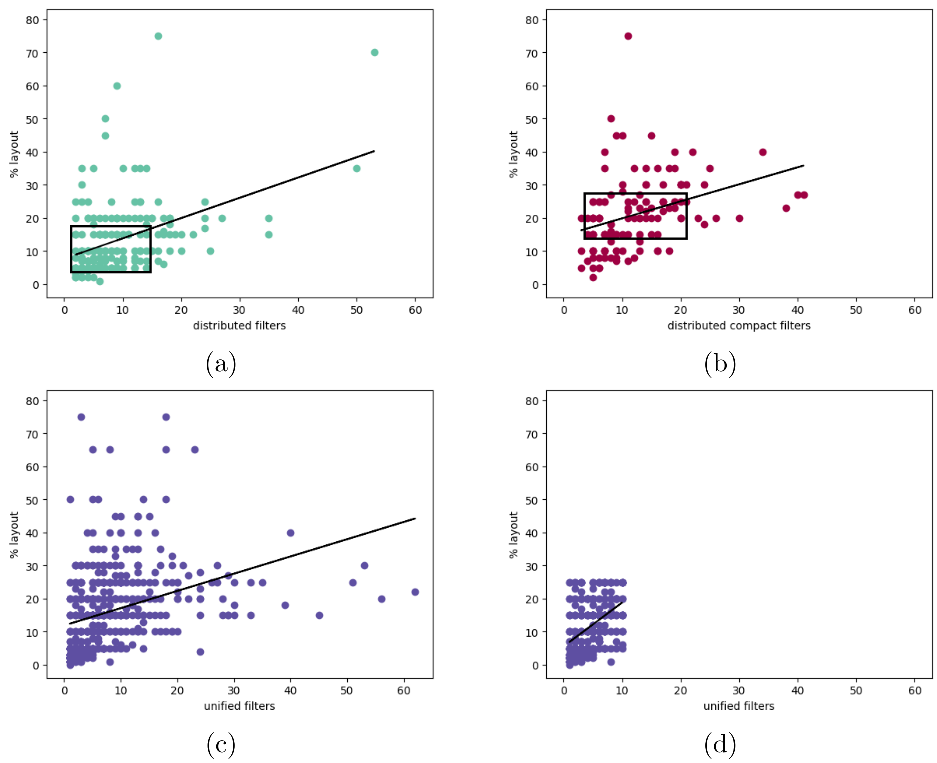


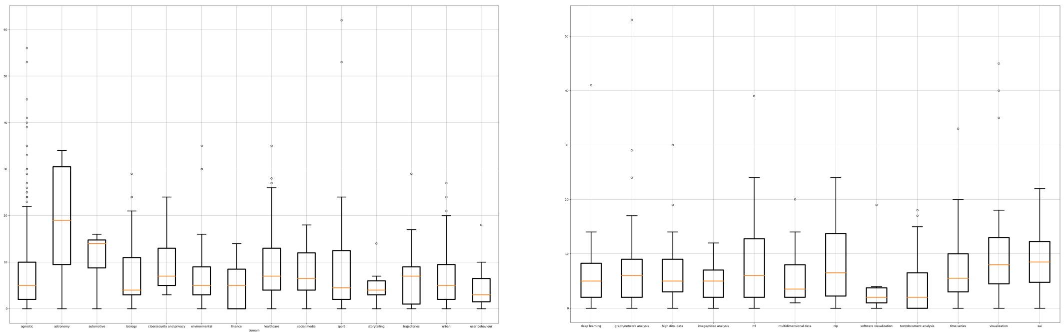
Figure 14: Relationship between % of layout and: (a) distributed type of filters, (b) distributed compact type of filters and (c) unified type of filters, (d) centered unified filters

3.4 Application Domain Analysis

This analysis is dedicated to the relation between filter presence and application domains.

Analysis of the presence of filters in the application domains (A4.1)

The main objective of this analysis is to determine how many filters are implemented within different areas of interest or subject fields. It is clear that, regarding the domains (Figure 15a), astronomy and automotive emerge as those that use a significantly higher number of filters than the others. The other domains, however, use an average number of filters ranging from 0 to 15, suggesting that their application is more limited or specific. We even analyzed the trend of the "domain agnostic" solutions by their discipline (Figure 15b). In this case, it turns out that disciplines such as machine learning, Explainable Artificial Intelligence (XAI), visualization, and Natural Language Processing (NLP) tend to use a higher number of filters, often exceeding 10 filters on average per publication. This suggests that



(a) (b)

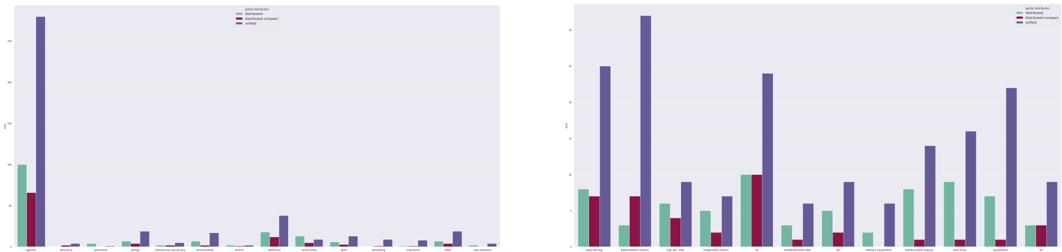
Figure 15: Boxplots showing the distribution of filters (a) for each application domain, (b) for each disciplines

these disciplines require a greater number of filter elements for visual analysis purposes than other disciplines.

Analysis of spatial distributions by domain (A7)

This analysis focuses on analyzing the spatial distribution of systems within each application domain. The main objective is to examine how many systems fall into a given spatial distribution within each domain, in order to understand the specific distribution trends for each application domain.In the context of application domains (Figure 16a), clear trends emerge in the spatial distribution of filters, the majority of scientific works in each domain seem to prefer a unified layout for filters. The other two layouts are less present, with distributed compact as clearly the least present. The results for disciplines (Figure 16b) show similar trends to those of application domains, here too there is a predominance of works that prefer a unified layout for filters. However, we also see a greater use of distributed and distributed compact layouts with respect to application domains.

Overall, the analysis results reveal how the spatial distribution of filters is influenced by both the application domain context and the discipline.



(a) (b)

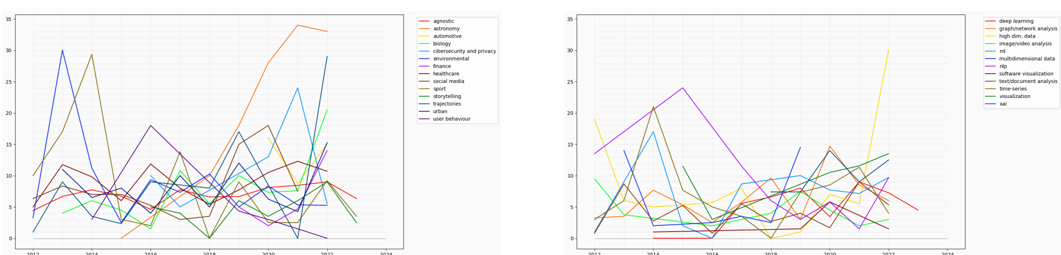
Figure 16: Spatial distribution of filters (a) for each application domain, (b) for each disciplines

3.5 Temporal Analysis

This analysis aims to investigate the temporal trends of the previous analyses.

Temporal analysis of filters divided by application domain (A4)

The main objective of this analysis is to detect and understand how the use of filters varies over time within different disciplines or areas of interest.Figure 17a shows that the majority of application domains tend to use a number of filters that fluctuate between 5 and 15 over the years. In the "environmental" and "sport" domains, we initially notice the use of a high number of filters, but over time, domains such as "astronomy," "biology," "cybersecurity," and "trajectories" tend to stabilize instead. instead they seem to increase the use of filters over the years. Looking at disciplines shows a relatively constant trend over time, with an average of filters between 5 and 10. Some disciplines show interesting behaviors; disciplines



(a) (b)

Figure 17: Temporal distribution of filters (a) for each application domain, (b) for each disciplines

such as "natural language processing (NLP)," "machine learning (ML)," and "time-series" initially appear to use a higher number of filters than average, but then tend to level off. The

”high dimensional data” discipline, however, shows an initial increase in the use of filters, after which it tends to stabilize in the medium range with a further increase in recent years.

Temporal analysis of filters and direct manipulation (A9)

The main objective of this analysis is to examine how the use of filters and direct manipulation vary over time. To graphically represent this analysis, bar charts were created in which there are five distinctive bands relating to filters or direct manipulations. This banding was based on the boxplot statistics corresponding to each variable (Figure 18b, 18d).

Observing the temporal trend of the filters (Figure 18a), it can be noted that, as regards the

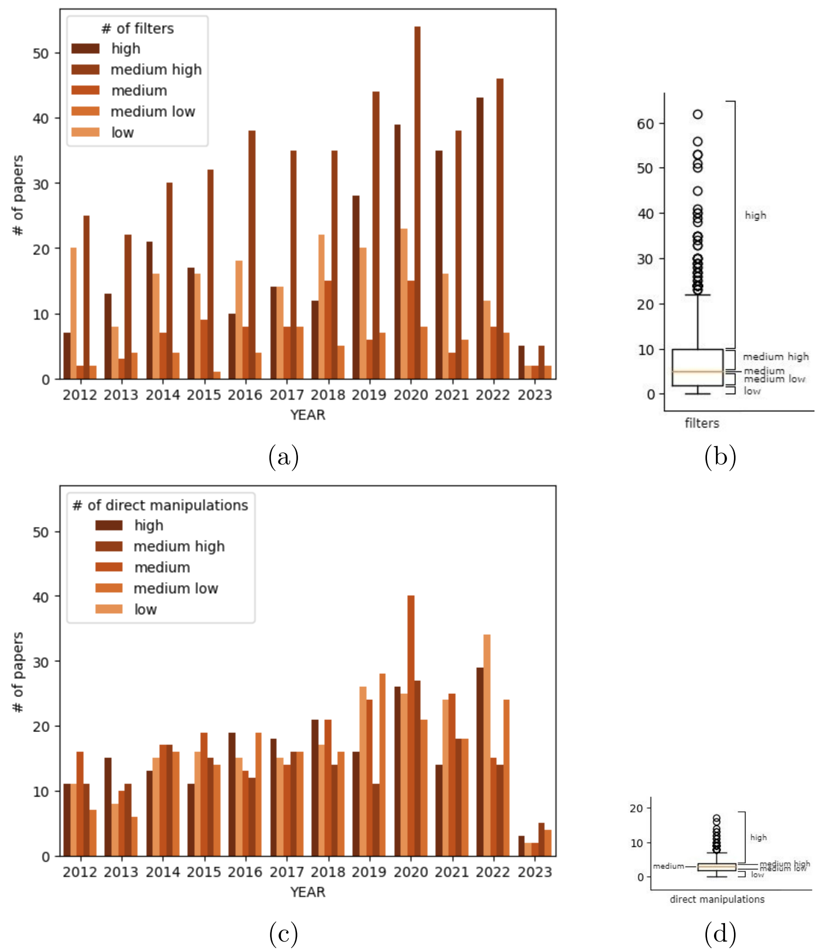
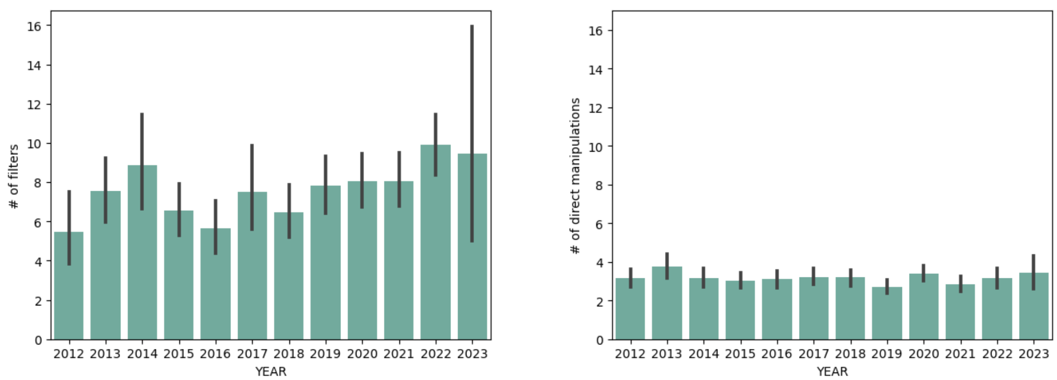


Figure 18: (a) Bar chart showing the temporal trend of filters, (c) bar chart showing the temporal trend of direct manipulations (2023 is still in progress so its data is partial), (b, d) boxplots showing how the five intervals were defined

"low", "medium" and "low medium" filter bands, there is stability over the years, there are variations but they are quite slight. However, we can observe a notable increase in papers that implement systems with a number of filters falling into the "medium high" and "high" categories. As regards the temporal trend of direct manipulations (Figure 18c), until 2018 all bands show a relatively constant trend. Furthermore, for each year it can be noted that all bands are made up of a similar number of papers. From 2019 a significant change begins to emerge, we can notice an increase in papers that use a number of direct manipulations falling within the "medium" and "medium high" range, but these values subsequently tend to decrease.



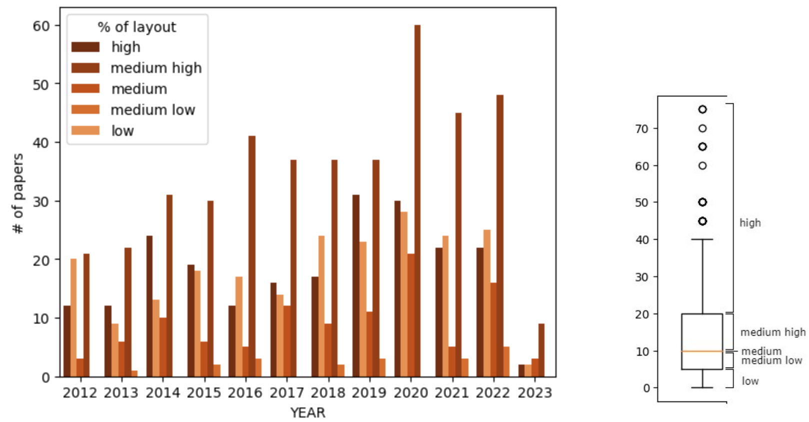
(a) (b)

Figure 19: Temporal distribution of the average of: (a) filters, (b) direct manipulations

The analysis actually seems to suggest that there is a growing interest in adopting more filtering elements within visual systems.
The analysis conducted was also extended to each individual taxonomic group. The results obtained were congruent with those of the overall analysis.

Temporal analysis on the layout occupied by the filters (A10)

This analysis focuses on the temporal trend of the percentage of layout occupied by filters within systems. To represent this analysis graphically, a bar chart was created (Figure 20a), in which, to facilitate understanding of the data, five layout percentage intervals were created for each year based on the data distribution shown in the boxplot (Figure 20b). It can be noted that there has been a significant increase in works that tend to dedicate space to filters, represented by the "medium high" range. This suggests that more and more papers



(a) (b)

Figure 20: (a) Bar chart showing the temporal trend of the % of layout occupied by the filters (2023 is still in progress so its data is partial), (b) boxplot showing how the five intervals were defined

analyzed have adopted an approach that assigns a substantial part of their interface to filters. Additionally, there has been an increase, although less noticeable, in the frequency of papers falling into the extreme "low" and "high" ranges. The "medium" and "medium low" bands appear to maintain more or less the same trend over time, with some fluctuations.
To complete the analysis, a further bar chart was created which represents the overall average of this percentage for each year (Figure 21).

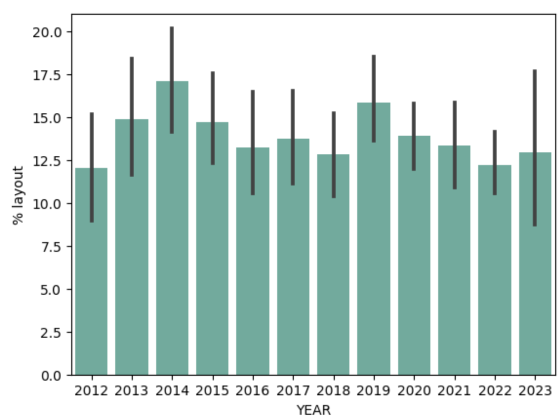
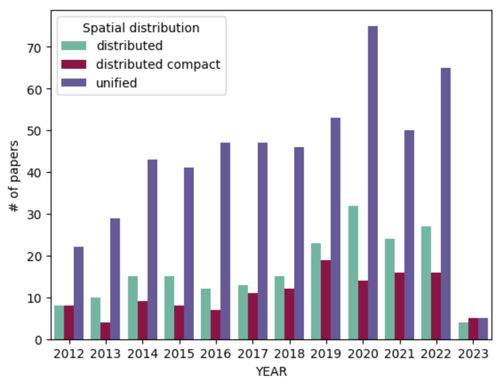


Figure 21: Temporal distribution of the average of the % of layout occupied by the filters

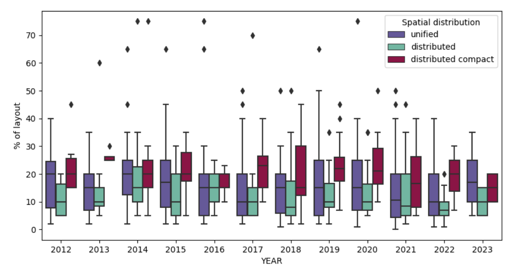
The analysis conducted was also extended to each individual taxonomic group. The results obtained were congruent with those of the overall analysis.

Temporal analysis of the spatial distribution (A11)

The analysis focuses on the temporal trend of the spatial distributions of filters within systems. Figure22 shows a marked rise in the bar relating to unified distribution, indicating that there is a notable increase in the number of systems dedicating a single unified space for filters over the years. On the other hand, regarding distributed and compact distributed filters, an increasing trend is still observed, but less pronounced. This suggests that although the preference for unified distribution is increasing, there is still a general trend towards the use of distributed or distributed compact filters.



(a)



(b)

Figure 22: (a) Temporal distribution of spatial distribution of the filters within the systems, (b) temporal trend of the percentage of layout occupied for each spatial distribution.

A further analysis was carried out which relates the spatial distributions and the percentage of layout occupied by the filters, as represented in Figure22b, and provides further details on the evolution of these variables over time. Interestingly, compact distributed filters tend to take up more space over the years than the other two layouts. This may be because, while they are distributed, they contain both a unified part and a distributed part, which may require more overall space in the user interface. On the other hand, distributed filters seem to occupy less space and show a more constant trend, while unified filters show greater variability over time.

4 Analysis by Taxonomy Groups

4.1 Analysis of the inverse proportionality of filters with respect to direct manipulation (A1)

The study links the amount of filters and direct manipulations and is critical for understanding how these two variables are connected and how they might impact each other inside visual systems.

For this analysis it was initially hypothesized that as the number of filters increased there would be a decrease in the number of direct manipulations, suggesting that the introduction of more filters could partially replace the need for direct manipulations by users. However, the results obtained from the initial analysis contradicted this hypothesis, showing a flat trend between the number of filters and the number of direct manipulations. This means that, in general, the increase in filters has not significantly affected the amount of direct manipulations used in visual analytics systems.

However, subsequently rerunning the analysis for each taxonomic group (Figure 23, 24), a more detailed picture of the relationship between filters and direct manipulations emerged. It became evident that although the overall analysis may show a flat trend, when analyzing the data more granularly, different trends can be detected. Many of these sub-analyses have shown that, in reality, there is a directly proportional relationship between the number of filters implemented and the number of direct manipulations. In other words, as the number of filters increases, the number of direct manipulation interactions by users also increases.

These findings highlight the importance of examining data in detail and not relying only on aggregate results, and suggest that introducing more filters may not necessarily replace direct manipulations, but rather may encourage more complex interactions between the user and the visual analysis system.

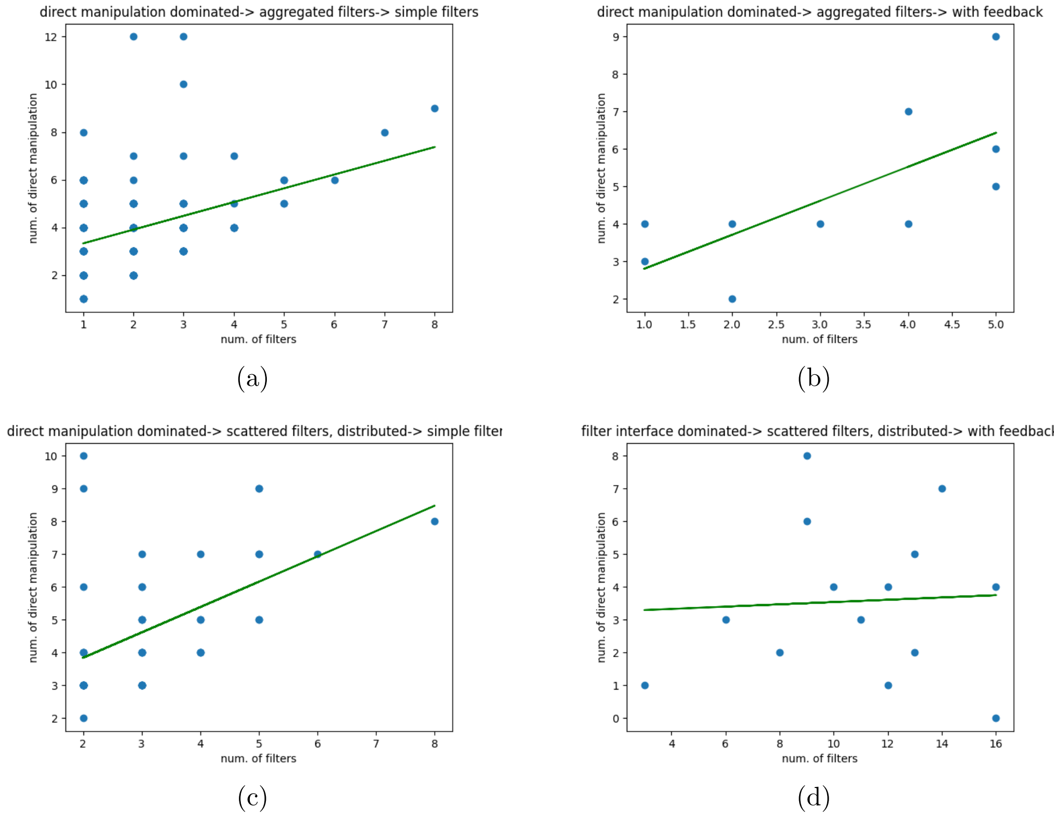


Figure 23: A1 analysis by taxonomic groups.

4.2 Analysis of filters and direct manipulation distribution (A2)

In analysis 2, the frequency distribution of the filters and direct manipulation was examined in detail, but the results obtained by rerunning the analysis for each taxonomic group are reported below. This allowed us to obtain a comprehensive overview of the trends in these categories.

The results obtained from the analysis by taxonomic groups (Figure 25, 26) were extremely consistent with those found in the global analysis. In general, it is observed that the number of filters used usually varies from 1 to 14, while the number of direct manipulations varies between 1 and 6. However, the analysis of papers that do not use any filters (with 0 filters) remained unchanged as these papers are commonly grouped into a single category within the taxonomy. As regards the analysis of papers that do not present any direct manipulation, a

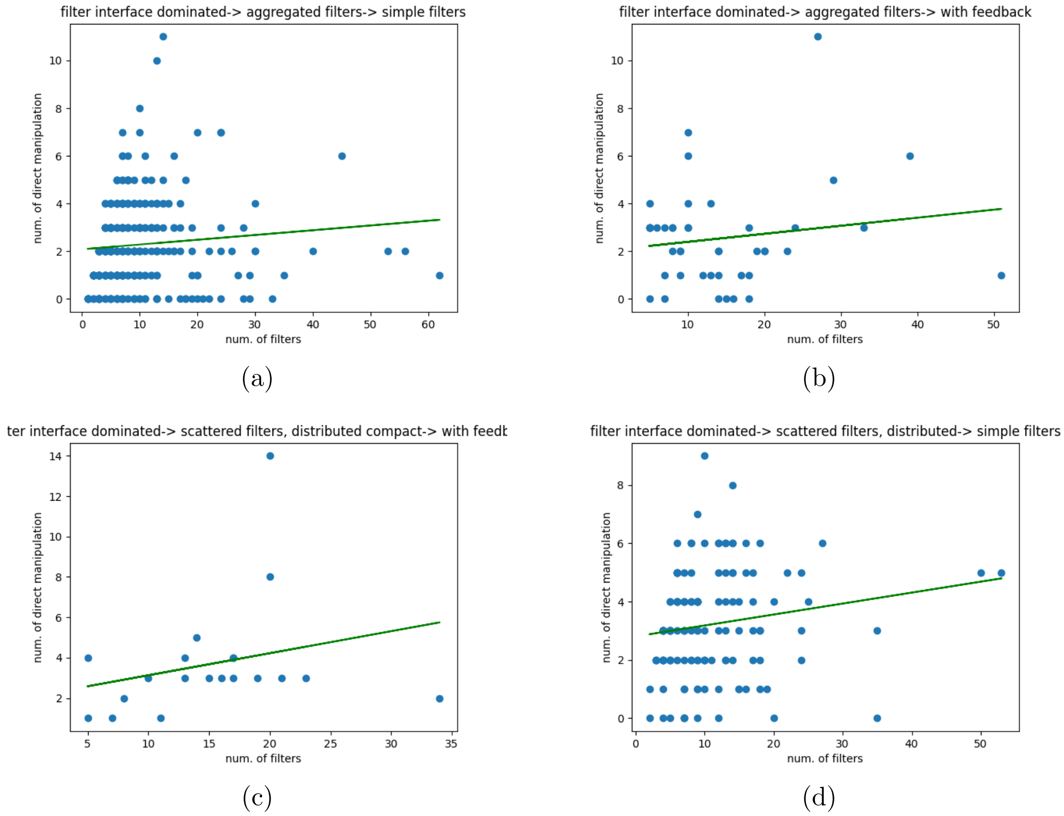
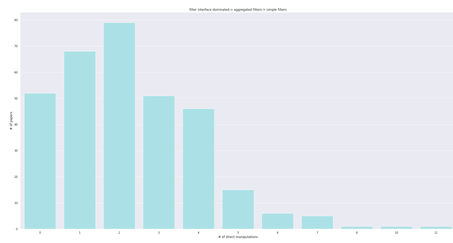
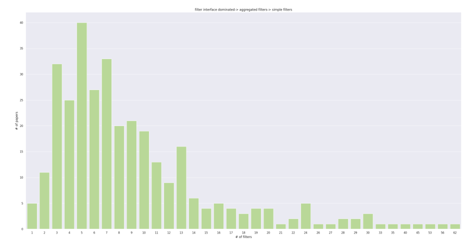


Figure 24: A1 analysis by taxonomic groups.

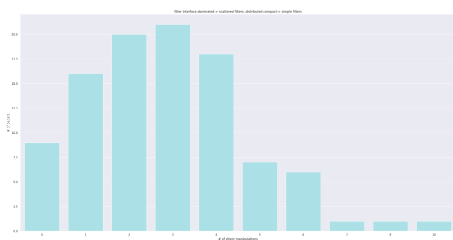
greater variability is noted in the data compared to what was observed in the global analysis. This suggests that, within the scientific community there are different approaches and trends, while some works may opt for a limited use of filters, others may adopt a broader and more varied approach in interacting with filters.



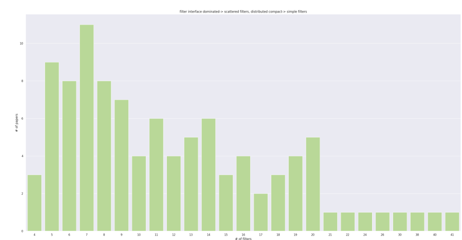
(a)



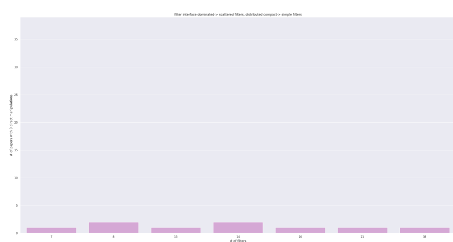
(b)



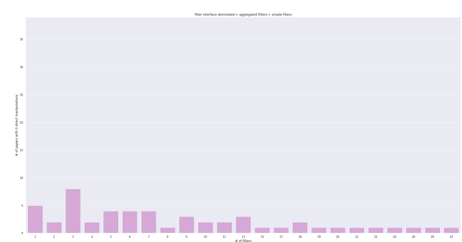
(c)



(d)



(e)

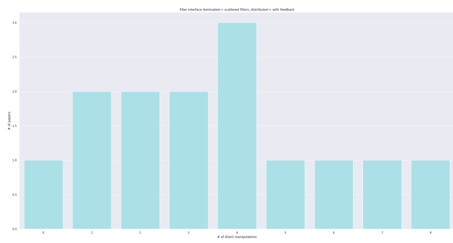


(f)

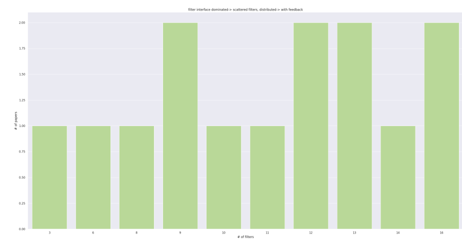
Figure 25: A2 analysis by taxonomic groups.

4.3 Analysis of filters with feedback and their user interactions (A3)

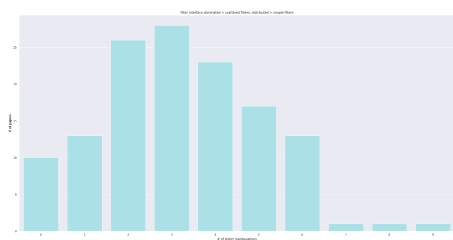
The analysis aimed to examine filters with feedback and user interactions when using these filters within data visualization systems. This analysis was extended to also consider taxonomic groups in order to better understand the differences and trends within each group. However, it is important to note that not all taxonomic groups (Figure 27) returned the same results in the analysis, there was significant variation between groups in terms of the types of filters used and interactions with users. For example, some groups may have a greater



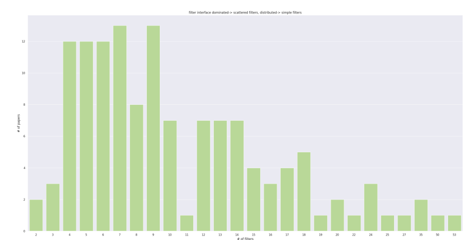
(a)



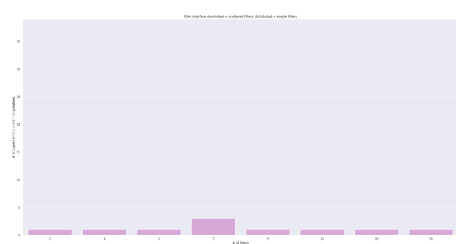
(b)



(c)



(d)



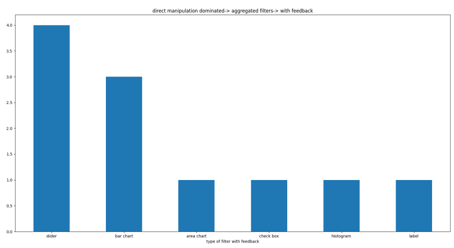
(e)

Figure 26: A2 analysis by taxonomic groups.

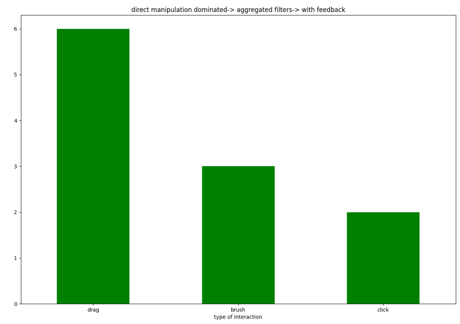
prevalence of sliders or labels as types of feedback, while others may show preferences for other forms of interaction.

The same goes for interactions with users, in fact, in some groups, there may be a predominance of brush or drag as the preferred interaction modes.

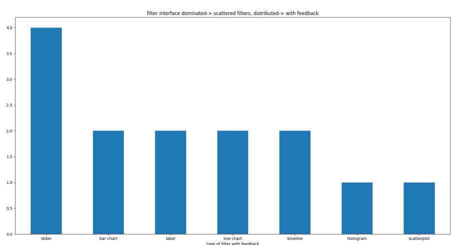
These differences between taxonomic groups suggest that preferences and practices in using feedback filters and user interactions may vary based on other parameters specific to each group, therefore, when designing data visualization systems or conducting research in certain sectors, it is important to take into account the specific preferences and tendencies of the taxonomic group of reference.



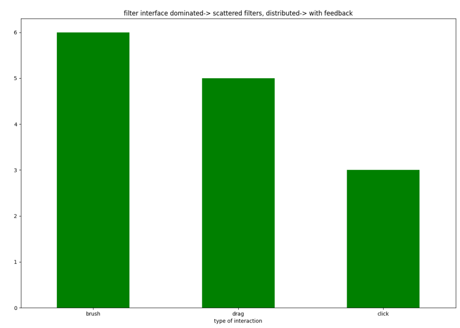
(a)



(b)



(c)

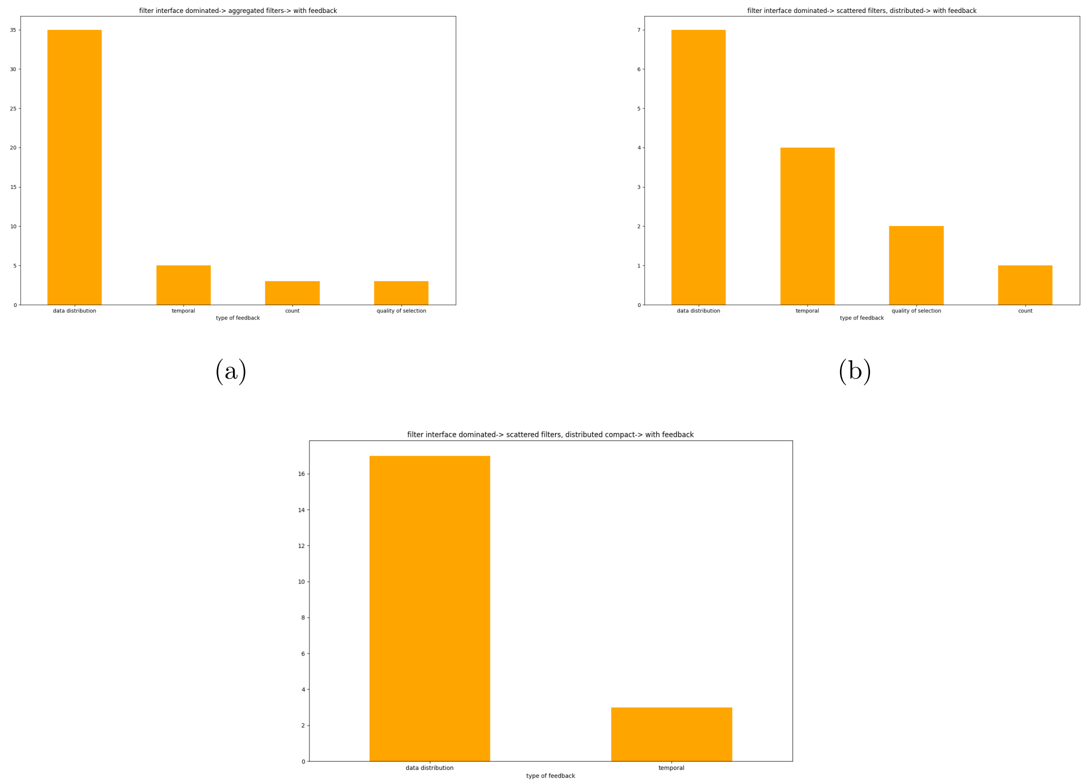


(d)

Figure 27: A3 analysis by taxonomic groups.

4.4 Analysis of the types of feedback (A3.1)

The analysis focused on examining the types of feedback implemented in filters within data visualization systems. This analysis was also extended for the taxonomic groups, in order to evaluate whether there were significant differences in the types of feedback implemented.
The results obtained from this analysis for each taxonomic group (Figure 28) demonstrated consistency with the results of the global analysis; for each taxonomic group, a prevalence of "data distribution" was found as a type of feedback implemented in the filters.
This result suggests that, despite variations between groups in other aspects of feedback filters, the "data distribution" feedback typology appears to be widely adopted and used consistently across all contexts analyzed.



(a) (b) (c)

Figure 28: A3.1 analysis by taxonomic groups.

4.5 Analysis of relation between filters and percentage of layout they occupy (A6)

The analysis aimed to establish the relationship between the number of filters within a system and the percentage of layout they occupy. This analysis was extended to also consider taxonomic groups, in order to evaluate whether there were significant variations in the results.

The results obtained from the global analysis showed a directly proportional trend between the number of filters and the percentage of occupied layout. In some taxonomic groups (Figure 29, 30), the results confirm the directly proportional relationship between the number of filters and the occupied layout, in line with the global analysis, however, in other taxonomic

groups, an inversely proportional trend was noted. This means that in these specific contexts, as the number of filters increases, the percentage of layout dedicated to them tends to decrease rather than increase.
These differences highlight that each taxonomic group may have different needs and preferences, which suggests that interface design strategies should be adapted appropriately to meet the specific needs of each group.

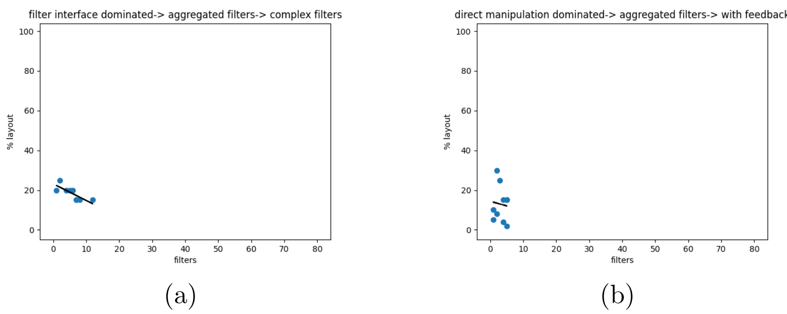


Figure 29: A6 analysis by taxonomic groups.

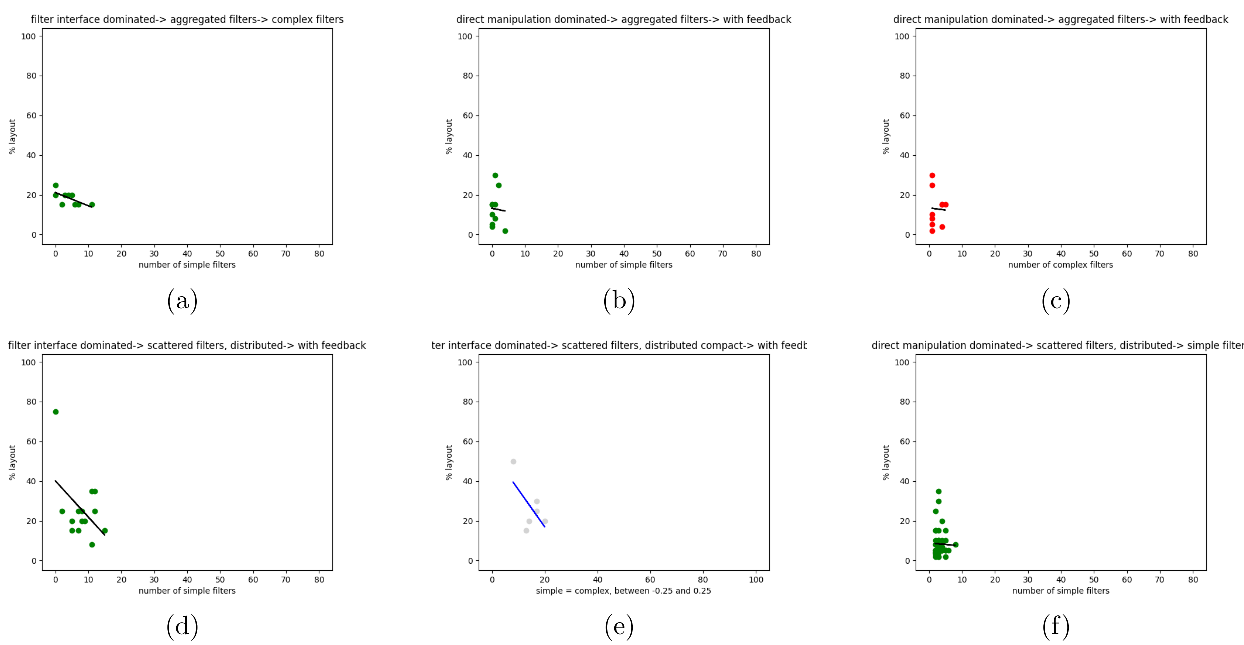


Figure 30: A6 analysis by taxonomic groups.

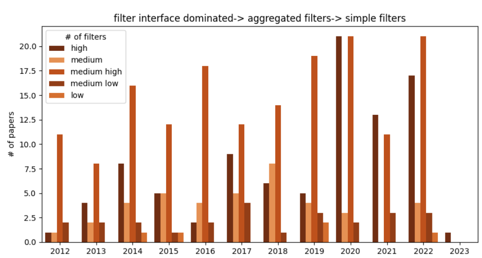
4.6 Temporal analysis of filters and direct manipulation (A9)

Analysis 9 analyzed the time course of filters and direct manipulation, and was extended to analyze each individual taxonomic group (Figure 31, 32). In general, the results obtained in these analyzes confirm what was observed in the global analysis.

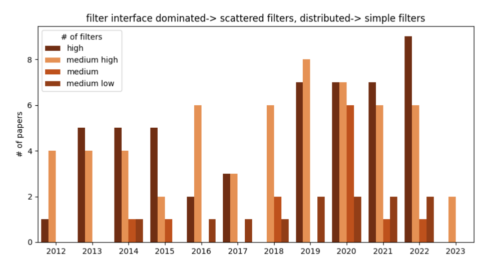
Regarding filters, there is clearly an increase in the number of works falling into the "high" and "medium high" categories over time, indicating a growing adoption of more complex filters in user interfaces. On the other hand, regarding direct manipulation, an increase in the number of jobs in the "low", "medium" and "medium high" categories is observed over the years.

It is interesting to note that even for the analysis of the average of the number of filters and the direct manipulation for each taxonomic group, results consistent with what emerged from the global analysis are achieved.

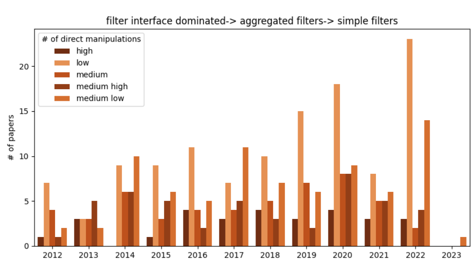
This consistency in results suggests that the general trends observed in the use of filters and direct manipulation are consistent across different categories of systems and applications, providing a more complete view of the evolution of user interfaces in this space.



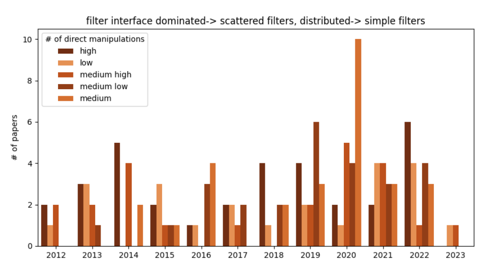
(a)



(b)

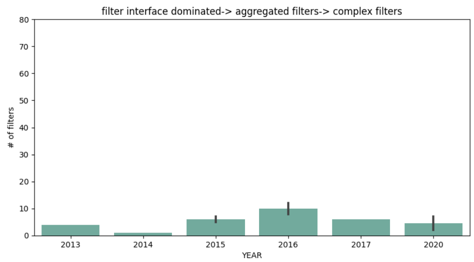


(c)

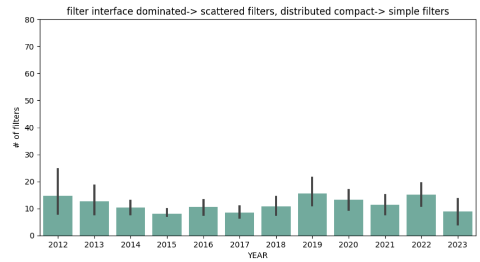


(d)

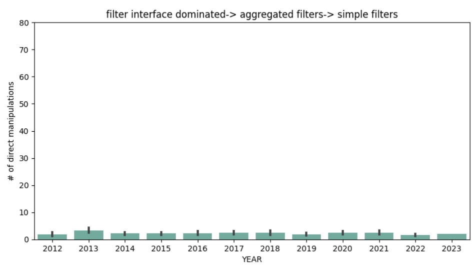
Figure 31: A9 analysis by taxonomic groups.



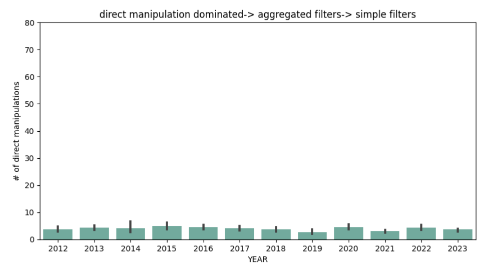
(a)



(b)



(c)



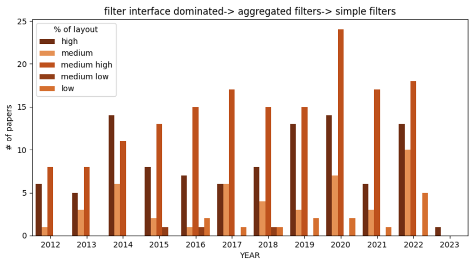
(d)

Figure 32: A9 analysis by taxonomic groups.

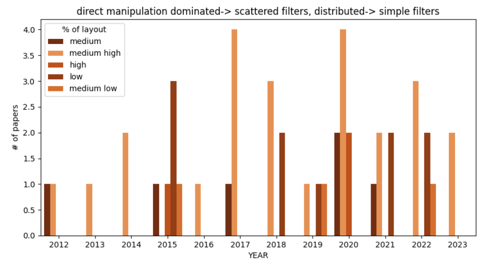
4.7 Temporal analysis on the layout occupied by the filters (A10)

Analysis 10, which focuses on the temporal trend of the percentage of layout occupied by filters, was also performed for each taxonomic group.

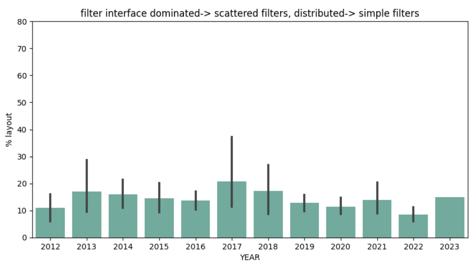
Interestingly, the results obtained in these analyzes for the different taxonomic groups (Figure 33) are consistent with what was observed in the global analysis. In particular, over time, there has been a notable increase in works that dedicate a percentage of layout to filters that fall within the "medium-high" range. Furthermore, an increase, although less evident, is observed in the "low" and "high" ranges, while, as regards the annual analysis of the average trend of the layout percentage, it is interesting to note that there is a fairly constant trend, with small fluctuations over time.



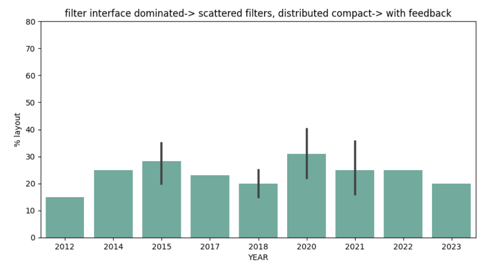
(a)



(b)



(c)



(d)

Figure 33: A10 analysis by taxonomic groups.

References

[1] Yann Laurillau et al. "The TOP-Slider for Multi-Criteria Decision Making by Non-Specialists". In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. NordiCHI '18. Oslo, Norway: Association for Computing Machinery, 2018, pp. 642–653. ISBN: 9781450364379. DOI: 10.1145/3240167.3240185. URL: https://doi.org/10.1145/3240167.3240185.

[2] Anass Lasram, Sylvain Lefebvre, and Cyrille Damez. "Scented Sliders for Procedural Textures". In: *EUROGRAPHICS short papers*. Proceedings of the Eurographics conference (short papers). Cagliari, Italy, May 2012. URL: https://inria.hal.science/hal-00748188.

[3] Benoit Otjacques et al. "MCS Filters to Express Partial Satisfaction of Criteria". In: *EuroVis - Short Papers*. Ed. by Miriah Meyer and Tino Weinkaufs. The Eurographics Association, 2012. ISBN: 978-3-905673-91-3. DOI: 10.2312/PE/EuroVisShort/EuroVisShort2012/049-053.

[4] Miriam Greis et al. "Input Controls for Entering Uncertain Data: Probability Distribution Sliders". In: *Proc. ACM Hum.-Comput. Interact.* 1.EICS (June 2017). DOI: 10.1145/3095805. URL: https://doi.org/10.1145/3095805.

[5] Géry Casiez, Sylvain Malacria, and Eva Mackamul. "Signifidgets : What you see is what widget !" In: *IHM 2023 - 34e Conférence Internationale Francophone sur l'Interaction Humain-Machine*. AFIHM and Université de Technologie de Troyes. Troyes, France, Apr. 2023. URL: https://hal.science/hal-04043346.

[6] M. Adil Yalçin, Niklas Elmqvist, and Benjamin B. Bederson. "AggreSet: Rich and Scalable Set Exploration using Visualizations of Element Aggregations". In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 688–697. DOI: 10.1109/TVCG.2015.2467051.

[7] Alexander Lex et al. "UpSet: Visualization of Intersecting Sets". In: *IEEE Transactions on Visualization and Computer Graphics* 20.12 (2014), pp. 1983–1992. DOI: 10.1109/TVCG.2014.2346248.

[8] Matthew Brehmer et al. "Matches, Mismatches, and Methods: Multiple-View Workflows for Energy Portfolio Analysis". In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 449–458. DOI: 10.1109/TVCG.2015.2466971.

[9] Marcos Lage et al. "StatCast Dashboard: Exploration of Spatiotemporal Baseball Data". In: *IEEE Computer Graphics and Applications* 36.5 (2016), pp. 28–37. DOI: 10.1109/MCG.2016.101.

[10] Changjian Chen et al. "Towards Better Caption Supervision for Object Detection". In: *IEEE Transactions on Visualization and Computer Graphics* 28.4 (2022), pp. 1941–1954. DOI: 10.1109/TVCG.2021.3138933.

[11] Roger A. Leite et al. "Hermes: Guidance-enriched Visual Analytics for economic network exploration". In: *Visual Informatics* 4.4 (2020), pp. 11–22. ISSN: 2468-502X. DOI: https://doi.org/10.1016/j.visinf.2020.09.006. URL: https://www.sciencedirect.com/science/article/pii/S2468502X20300371.

[12] Marco Angelini et al. "CrossWidgets: Enhancing Complex Data Selections through Modular Multi Attribute Selectors". In: *Proceedings of the International Conference on Advanced Visual Interfaces*. AVI '20. Salerno, Italy: Association for Computing

Machinery, 2020. ISBN: 9781450375351. DOI: 10.1145/3399715.3399918. URL: https://doi.org/10.1145/3399715.3399918.