

Appendix A: Case Study

WYTIWYR offers users a flexible approach to chart retrieval by using both query chart and user intent, thereby addressing various issues related to traditional retrieval and expediting the design process. In conventional retrieval, not all query chart attributes may be relevant to the user’s intent, and some preferred attributes may even be absent from the chart. To address these issues, our system disentangles and combines attributes to ensure that the retrieved results encompass all attributes that reflect the user’s intent. Additionally, during the process of creating a well-designed visualization chart, many inspirations may arise, but bringing those inspirations to fruition can be time-consuming. Retrieval is a fast way to validate those inspirations.

To demonstrate the effectiveness of our system, we have designed several scenarios that may arise during the retrieval process. These scenarios include original attribute changes, existing attribute deletions, new attribute additions, and attribute transfers, which correspond to figures (a), (b), (c), and (d), respectively. In each scenario, we present the intent-free retrieval results in the first line, while the second line displays the results with user intent. For example, in Fig. 1 (a), we replace the horizontal layout with a vertical layout as per the user’s intent. In Fig. 1 (b), our retrieval results do not include numbers, while the query and intent-free retrieval return a heatmap with numbers. Similarly, in Fig. 1 (c), our retrieval system strives to retrieve results that preserve the original structure in the query chart while also adding a new attribute, represented by the “with dot”. Finally, in Fig. 1 (d), we use a timeline as the query chart and retrieve charts of other types, with the aim of transferring the style of the query chart to other charts. It is worth noting that attribute transfer is particularly intriguing, as it enables the transfer of attributes across different chart types. This can greatly assist in maintaining perceptual consistency and constructing a dashboard with multiple charts. In fact, one of our users pointed out this feature’s usefulness in this regard.

In addition, our system encompasses fuzzy retrieval as part of its functionality. Based on user feedback, we have discovered that not all users have the explicit goal of retrieving specific attributes such as color, trend, etc. As demonstrated in Fig. 2, we showcase two cases with more implicit intent. These cases demonstrate that our system is capable of identifying the underlying data representation and visual perception within the chart.

Appendix B: Negative Examples

As an initial step towards exploring visualization chart retrieval using pre-trained vision-language models, there is significant room for improvement. We present several typical negative examples to highlight the limitations of the current approach.

User Intent or Query Chart The trade-off between the query chart and the user’s intent can significantly impact the retrieval results. As illustrated in Fig. 3, we modified the original attribute of a query chart from blue to “pink” and compared the retrieval results with different prompt weights. In the second and third lines, we used different μ values of 8 and 5, respectively, which control the weight trade-off between the user-selected explicit attributes and prompts. It is apparent that retrieval results with higher μ values are more

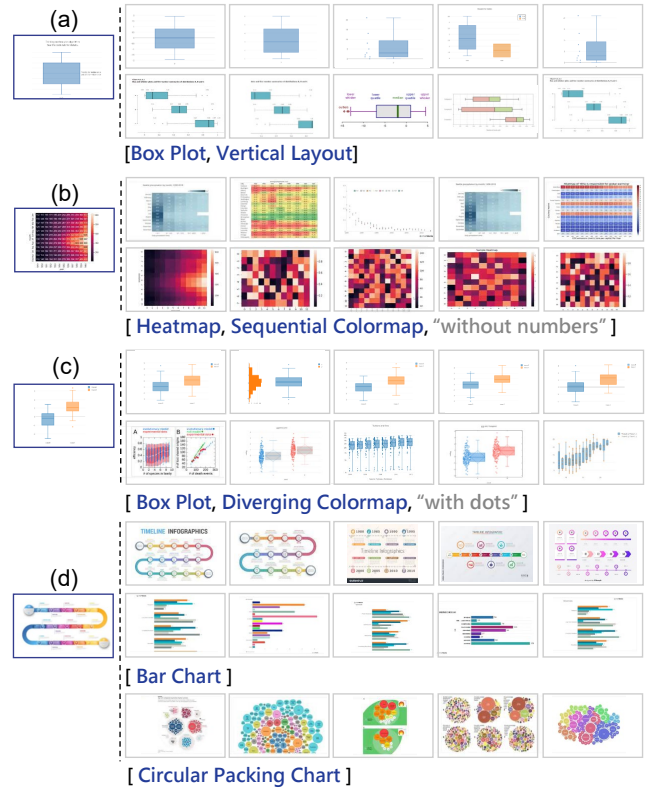


Figure 1: **Four cases for extending the design space by explicit attributes.** (a) original attribute changes, (b) existing attribute deletions, (c) new attribute additions, (d) attribute transfers.

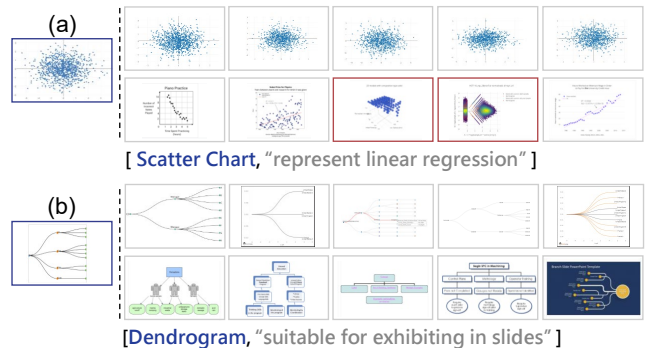
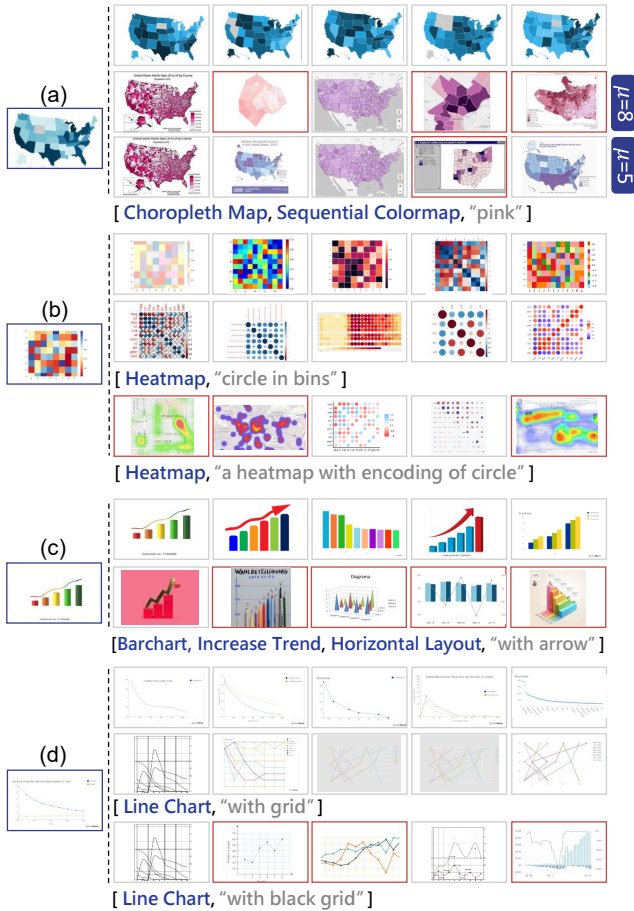


Figure 2: **Two cases for fuzzy retrieval by implicit intent.** The prompt in (a) and (b) provide abstract information about the content of the chart, without specifically pointing out any explicit attributes. The negative cases are in red borders.

likely to display a pink color, while retrieval results with lower μ values preserve more structural information, rather than color.

Prompt Sensibility Fig. 3 (b) illustrates that varying prompts can result in different outcomes. The model favors prompts that feature visual elements over professional term such as “encoding of circle”, as CLIP models are trained on natural rather than visual data. In the third row of Fig. 3 (b), two types of heatmaps are visible in



a sample of five images for each of the 18 types of charts in our collected data. Each set of five examples comprises three synthesized charts from the Beagle dataset (in green borders) and two from our collected real-world examples (in blue borders). These test examples are displayed in Fig. 4 and Fig. 5.

Figure 3: **Four negative examples of WYTIWYR system.** (a) illustrates the trade-off between query chart and user intent, (b) shows that using different prompts can result in varying outcomes, and (c) and (d) demonstrate that the alignment between the text and visual features of the CLIP model still needs improvement. The negative cases are in red borders.

the retrieval results due to the lack of a more detailed category in our dataset. This can lead to inaccurate or incomplete results, which we plan to address in the future.

Model effectiveness Fig. 3 (c) and (d) indicate that our system still struggles with accurately aligning text and visual features. In Fig. 3 (c), the visual feature of an arrow is projected onto a pencil, cone, polyline with dots, and a small arrow icon, respectively. In Fig. 3 (d), our system successfully retrieves a line chart "with grid", but is not robust enough to retrieve more fine-grained results, such as "with black grid". These limitations suggest that there is significant room for improvement in vision-language models.

Appendix C: Testing Examples in Preliminary Study

To improve our understanding of user intents, we conducted a preliminary study aimed at identifying which chart attributes users tend to focus on. In order to minimize any possible subjective bias during the chart selection process for testing, we randomly selected

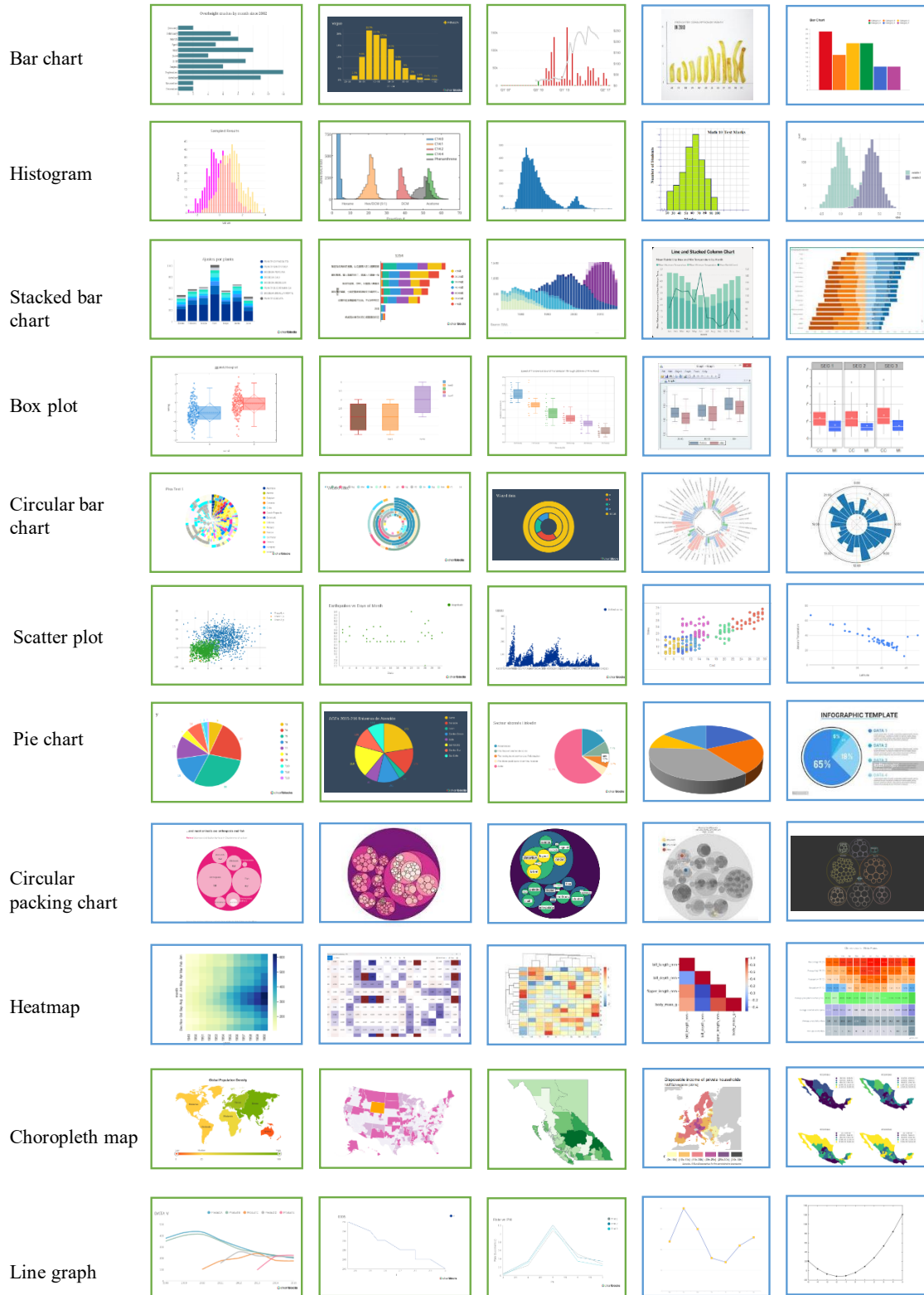


Figure 4: Five examples for each chart type in the preliminary study.

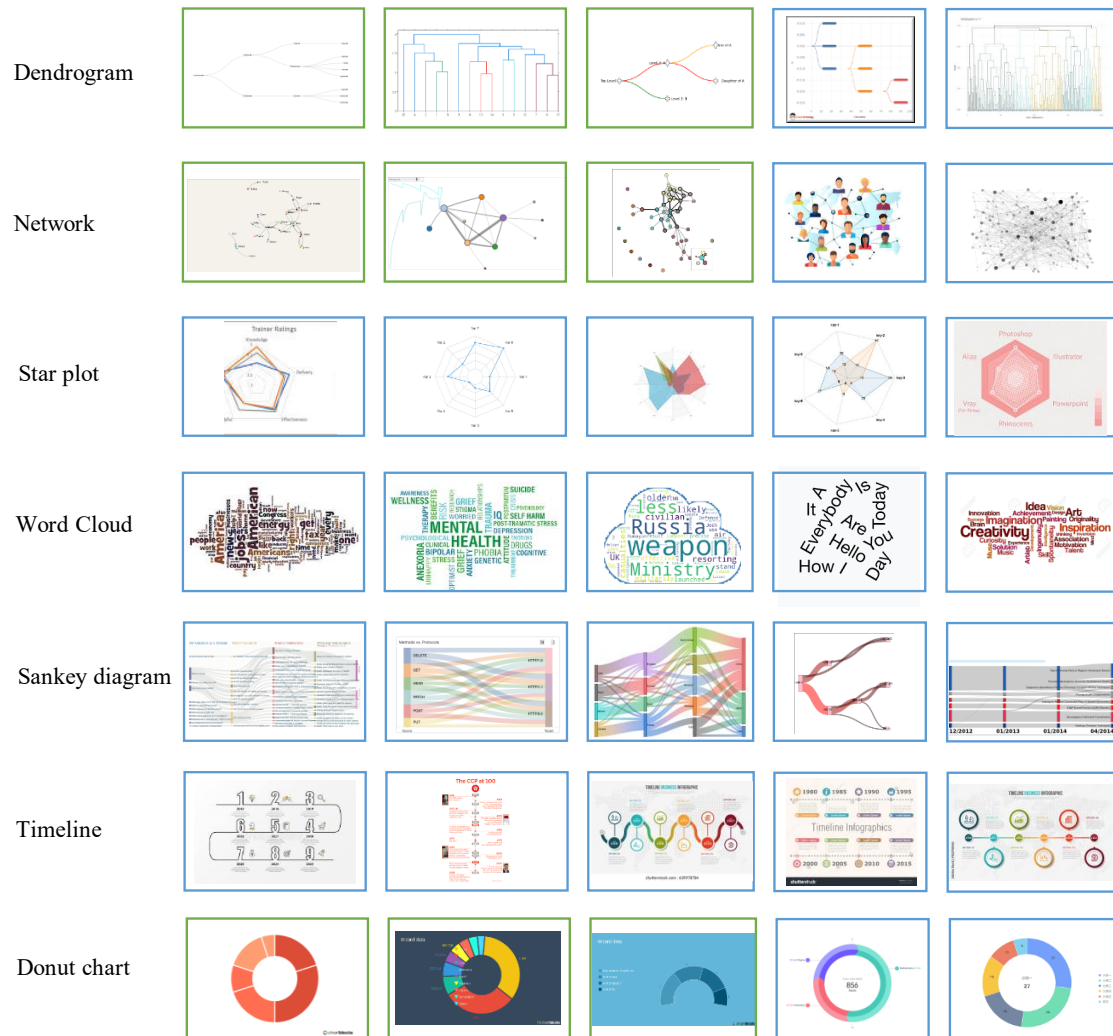


Figure 5: Five examples for each chart type in the preliminary study.