

Visual Layering and Visual Clustering – A Model of Graph Complexity for Business Data

Fang Chen
Fang.chen@umanitoba.ca
University of Manitoba
Canada

Limin Zhang
Limin.zhang@ndsu.edu
North Dakota State University
USA

Abstract

Businesses, large or small, often create graphs/charts (e.g., pie charts, column charts, line charts) to gain insights from their business data for decision making. Before a user decides to use a chart on a presentation, report or website, he/she needs to make sure that the chart is not too complex for fast and accurate processing. In our teaching of data visualization, we have seen charts that do not violate any established design principles but are too complex for quick processing. We have also observed that the number of data points do not necessarily make a chart more complex or difficult to understand. What determines the graph complexity is how data elements are displayed or plotted as visual clusters and visual layers. In this paper, we present a new model of graph complexity. The model specifies that there is a visual order for visual elements. From higher to lower order, there are visual layers, visual clusters, and visual elements. A visual layer consists of one or more visual clusters, which, in turn, consists of one or more visual elements. We present a pilot study that tests this model.

Keywords: data visualization, graph complexity, visual comprehension, visual cluster, visual layer, Model of Visual Clustering and Visual Layering.

1. INTRODUCTION

Data visualization has been a fast-evolving field in the past decade. Many business schools now include data visualization as a course module or a stand-alone course in the curriculum. Teaching data visualization consists of software skills and visual design. Students need to learn how to use different software packages to create visuals, whereas the visual design is about how to choose the appropriate chart types to reveal particular data patterns from a data set regardless of the software used.

Both authors of this paper have had at least 4 years of teaching data visualization. Six types of charts are commonly taught in the business

context: comparison, time trend, part-whole relationship, correlation, distribution, and location/map. The classic charts for comparison are column and bar charts. Line charts are considered the best to illustrate time trends. Pie charts, donut charts, tree map charts, and waterfall charts are used to illustrate part-whole relationships. Scatter plots are for correlations. Histogram and box plots are used to show distributions. Map charts plot data on a geographic map.

The following lists some of the visual design principles that we taught:

- Each chart must be independent, which means there should be a concise, descriptive

chart title, clear labels for X axis and Y axis, and a clear legend. A user should be able to understand the chart accurately without having to read the source data.

- The chart type must be appropriate for a particular data pattern. For example, use a pie chart to illustrate a part-whole relationship.
- Use color, text, alignment, and white space appropriately. For example, don't use red color for positive profit. Data should be sorted before creating the charts.

These principles may sound like common sense, but it is not uncommon to see graphs used by organizations that do not follow these principles. For example, the following chart shows the Computer Human Interaction (CHI) conference expenses over years, copied from CHI website (<https://sigchi.org/conferences/conference-history/chi/>).

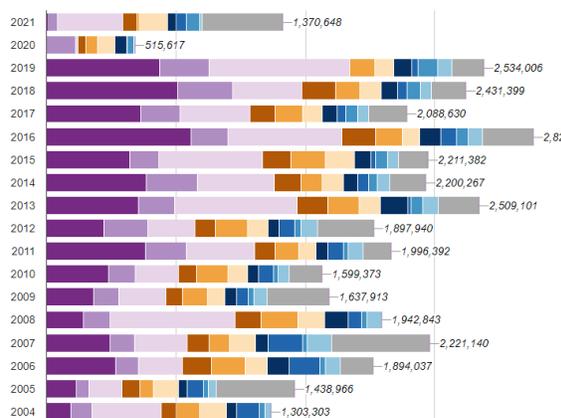


Figure 1: CHI Expenses

This chart does not have a legend on the chart or at the website, and the numbers do not have the "\$" sign in front of them to indicate dollar values. Since this is not a standalone chart, readers will not understand the chart without notes or references to the original data. All we can guess is the total amount spent each year. If this is what the graph designer wanted the reader to know, then a line chart illustrating the total expense over the years would be a better choice.

CHI is well known for its research in computer human interaction, which includes the interface/visual design. This example indicates that the afore mentioned design principles, even though sound as common sense, were not universally applied as a design routine. The issue of chart design, in fact, was not uncommon in organizations. The author of Functional Art

pointed out that many visualization designers treated users/readers as editors, expecting the readers draw the conclusion from charts with a variety of visual elements without clear visual organizations. He emphasized that visual designers should be clear about using visualization to support communication (Cairo, 2012, p. 217).

Some software applications, such as Power BI and Tableau, present suitable charts automatically for users to choose from based on the datasets. However, among the multiple alternatives, there could be subtle yet importance emphasis each chart type may illustrate. For example, in the CHI conference expenses chart, a line chart may be more appropriate than a stacked bar chart. However, this level of decision is based on the users' purpose of communication. More importantly, if users choose to use other visualization tools such as R or Python, they are completely left on their own to make decisions about the chart types. As a result, it is necessary to teach how to choose appropriate chart types based on not only the underlying relationship among data points, but also the purpose of communication.

In addition, we also taught Gestalt principles of visual perception. "Gestalt" is German for "unified whole". The first Gestalt Principles were proposed in the 1920s by German psychologists Max Wertheimer, Kurt Koffka and Wolfgang Kohler. (Gordon, 1989). The Gestalt principles specify how human brains find order in disorder among visual objects. In other words, "the mind 'informs' what the eye sees by perceiving a series of individual elements as a whole" (Gestalt Principles).

Some of these principles are similarity, proximity, closure, continuity, and connection (Knaflc, 2015, pp. 74 – 80). For example, Gestalt principle of similarity states that when things appear to be similar to each other (in color, shape, or size), they should be grouped together. For example, a user will see columns on the left image, but rows on the right image (Figure 1).

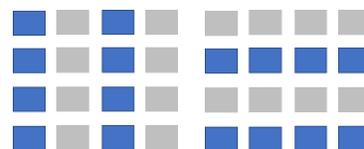


Figure 2: Gestalt law of similarity

We chose to teach Gestalt principles of visual perception because it is one of the most important

foundations of understanding how people visually perceive visual objects (e.g., Arnheim, 1969; Gordon, 1989; Ware, 2008, Elder, et al. 2012). In addition to general visual design, it has been applied to web interface design (e.g., Cao, 2015; Hensley, 2016a, 2016b), and data visualization design (e.g., Cairo, 2012; Knaflic, 2015).

During our teaching, we have observed that some graphs are complex to process, even if all design principles are followed. We have also observed that the sheer number of data points alone does not determine whether a chart is difficult to process or not. For example, Figure 3B has 5 times more data points than Figure 3A, however, Figure 3B is easier to understand than Figure 3A, and it is also more pleasant for readers to look at 3B than 3A.

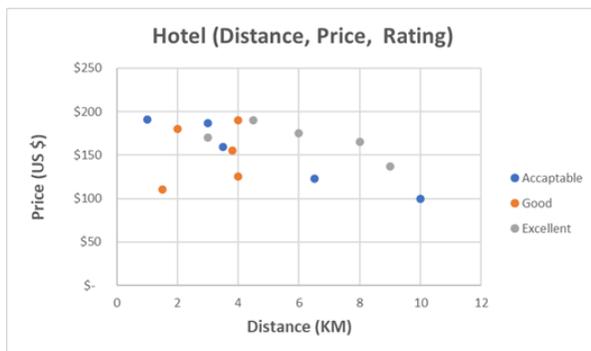


Figure 3A: A scatter plot with 15 data points

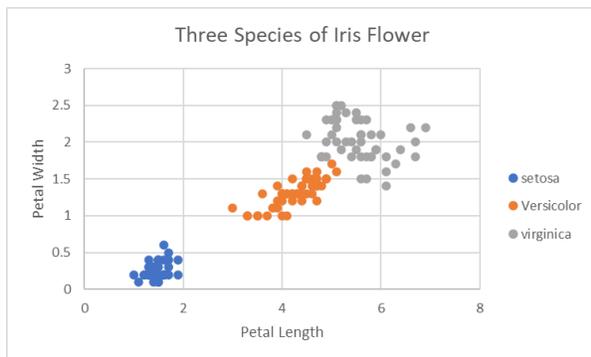


Figure 3B: A scatter plot with about 90 data points

The graph complexity, we believe, lies in the complexity of underlying datasets. We speculate that the patterns of data points and how the patterns visually appear on the chart may play a more important role than the number of the data points. The question is: which arrangements of data points make a chart easier or harder to understand? To answer this question, we develop

a model of graph complexity named the Model of Visual Clustering and Visual Layering.

We choose not to use the term “visual complexity” because visuals include graphs, graphics, photos, and paintings. There is a difference between graphs and graphics. Graphs refer to visuals that display data, whereas graphics refer to designs of architecture or landscaping (Graph vs Graphics, 2022). In this paper, we use the terms “graph” and “chart” interchangeably. However, we use the term “graph complexity” instead of “chart complexity”, because the term “chart” seems to suggest that the image has a row or column pattern. For example, we use the terms “column chart”, “line chart”, but “scatter plot” instead of “scatter chart”. Therefore, graph complexity is a more general term, and more appropriate.

On the other hand, we use the term “visual comprehension” in addition to the term “graph comprehension” because many cognitive processes are general for all types of visuals, not just graphs.

Before diving into the model, we may ask the question why it is important to understand and estimate the graph complexity? The answer lies in the designer’s objectives of using graphs. There are two common objectives of using graphs: 1) illustrating a data pattern so that it is easier to process and understand than using text and 2) attracting and retaining readers’ attention with beautiful visuals.

With the increasing digitization, people’s attention span is shortened. If a chart is too complex, readers may lose interest to process it or misunderstand it. In either case, we would fail the two objectives listed above. With the proposed model, we hope to provide some concrete measures that can be used to estimate graph complexity by graph designers.

We test our model with business datasets. The reason is that business data is usually in the format of tabular data, while scientific data and social media data may be in hierarchical format or free form format. The types of visuals for scientific data and social media data may be different. For example, social media data can be represented by network map, which is seldom used for business transaction data.

The outline of the paper is as the following. We summarize literature review related to visual comprehension in section 2 and present the research model and hypotheses in section 3,

followed by the experiment design and pilot study in section 4. We report and discuss the results and study limits in section 5, and present conclusion and future studies in Section 6.

2. LITERATURE REVIEW OF VISUAL COMPREHENSION

Data visualization has been extensively studied in multiple disciplines. These studies can be grouped into three major categories: technical-oriented studies, development of practical applications, and cognitive studies. Research in the technical category investigates how to render complex datasets into innovative visuals, focusing on encoding datasets into visuals. Computer graphics are in this research stream (e.g., Zhu & Watts, 2010; Chen et. al. 2021). The stream of research is about visual engineering, i.e., how to make charts/graphs.

The research in practical application assesses graph readers' performance in terms of speed and accuracy with different designs, or different perceptions with the same graphs (e.g., Talbot, et. al. 2014; Keogh et. al. 2020).

The research in cognitive category examines how graph readers understand visuals. In other words, how readers decode the visuals (e.g., Ratwani, et. al., 2008; Walden et. al., 2018; Bera, et. al, 2019). This stream of research studies the process that a reader's brain would go through in order to decode and understand the visuals.

Our research spans all three streams. We built the research model primarily on cognitive and technical components, and we test it as visual studies in the applications. The cognitive component for the model is Framework of Visual Cluster (Ratwani et. al. 2008). The technical component is Feature Layering, an ArcGIS technique of plotting data on maps. We borrow the definition of visual cluster and feature layer from geographic information systems (GIS) and extend the definitions to graphing the tabular data.

Framework of Visual Cluster (Ratwani et. al. 2008) suggests that visual clusters are a core component of graph comprehension. Graph readers form visual clusters by making eye fixations on the boundaries of visual clusters, followed by eye transitions between cluster to cluster and cluster to legend. This process is cyclical, and the number of iterations is dependent on the complexity of the graph. As a result, the design goal is to reduce the number of cycles of processing. Three guidelines are further

provided to achieve this goal: 1) visuals should be designed for graph readers to form visual clusters easily, such as highlighting the boundaries of visual clusters; 2) the association between the cluster and the legend should be intuitive to reduce the number of cluster-to-legend transitions (e.g., using the same color for both visual cluster and the legend); and 3) reduce the number of variables to be displayed on a visual (e.g., display fewer clusters on one visual).

From their findings, we can deduce that there are three major functions (or costs) for visual comprehension: cluster identification, cluster and legend association, and cluster to cluster comparison. We can also see how the authors' suggestions for visual design correspond with these three functions/costs. First, highlighting the cluster boundaries can reduce the cluster identification cost. Second, coding the cluster and the legend in the same color can reduce the cluster legend association cost. Third, reducing the number variables on a visual can reduce the cluster-to-cluster comparison cost.

Ratwani and his colleagues (2008) designed three experiments to test their framework. All experiment tasks used map charts, consisting of some U.S. counties. The counties in the same color formed one distinct visual cluster. This definition of visual cluster is natural and intuitive, but it is not applicable to datasets that do not have the geographical dimension. How can we define a visual cluster on a graph without the geographical dimension?

Our definition of the concept of visual cluster is more general than what is specified in the Ratwani's Framework of Visual Cluster. More importantly, we introduce the concept of visual grouping order. We propose that a visual group has a visual order in it. Visual layers, visual clusters, and visual elements form a visual grouping order (from higher to lower). A visual layer consists of one or more visual clusters, which, in turn, consists of one or more visual elements. In the following, we use a column chart (Figure 4) to illustrate our definitions.

Visual Element: A visual element is the atomic or most basic building block for a graph. It can be a dot, a line, an area or polygon, or a label (e.g., text, number, or currency). For example, as shown in Figure 4, each column is a visual element, and each text label is also a visual element. In the same logic, each line, slice, or dot is a visual element on line charts, pie charts, and scatter plots respectively.

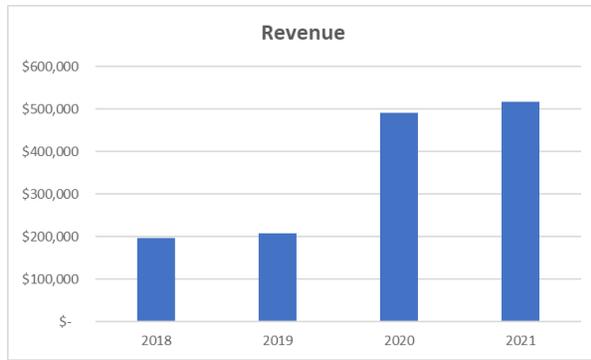


Figure 4: A column chart

Visual Cluster: A visual cluster consists of one or more visual elements that form a clear boundary. When the viewer’s eyes move from one visual element to another visual element belonging to the same cluster, the eyes’ movement does not get interrupted. In other words, the eyes’ movement flow is smooth. On Figure 4, all four columns can be viewed as a visual cluster. Alternatively, we can regard each column and its X axis label as a cluster. As a result, there are four clusters on this chart.

Visual Layer: A visual layer consists of one or more visual clusters. When two or more visual clusters are plotted on the same 2D plane and all these clusters do not overlap with one another, they are counted as one visual layer. Otherwise, they are counted as multiple visual layers. For example, Figure 3A has three layers, and Figure 3B has only one layer. Comparing the two charts, most people would agree that Figure 3B is easier to understand.

For some types of charts, such as clustered column chart, the number of visual layers depends on how we count the visual clusters. For example, if we consider the two columns for each year in Figure 5 as a cluster, then there are four visual clusters which are on one visual layer. However, if we count all blue columns as one visual cluster, and all orange ones as another visual cluster, then there are two layers. One layer for the blue cluster, another layer for the orange. This is where we borrowed the concept of “layer” from ArcGIS, which is a leading software for GIS graphs.

When ArcGIS draws/plots a map, it draws by layers. First, it draws the base layer, which is a base map. Then it draws feature layers. There is only one base map layer but could be many feature layers. For example, all highways on a map are plotted in one round (layer), all parks on a map are plotted in another round (layer). With

each feature forms a layer, the feature would have a name/label or legend associated it. This is why ArcGIS specifies that a feature layer is more or less equivalent to a legend on a paper map. Our definition of visual layer is based on, but different from ArcGIS’ feature layer. For example, in Figure 3B, there are three feature layers, because there are three legends (each type of flower is a legend). But because three clusters on these feature layers do not overlap, we count these three feature layers as one visual layer. In other words, feature layer is about how the software would draw the graph, while visual layer is about how users view the graph. In all three figures (Figures 3A, 3B, and 5), the base layer is the X axis and Y axis. In the next section, we explain how the number of visual clusters and visual layers affect graph complexity.

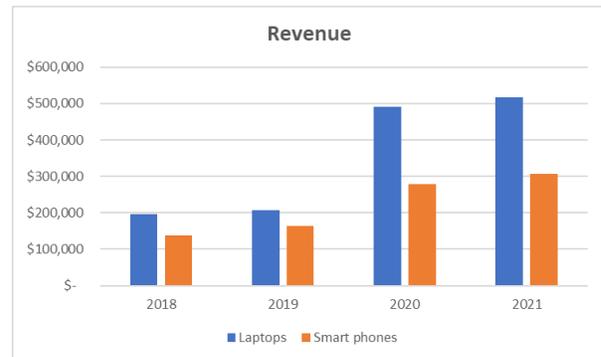


Figure 5: Column chart with two measures

3. RESEARCH MODEL

The Model of Visual Clustering and Visual Layering (Figure 6) is a novel model of graph complexity for the following three reasons. First, we invent the concept of visual layer, which is related to, but different from the concept of feature layer used by ArcGIS map plotting practice. Second, we propose the concept of visual cluster for non-map related data by borrowing the concept of visual boundaries defined by Ratwani and his colleagues (2008). Third, we identify visual element, visual cluster, and visual layer as three key components on a graph and establish the order of these components that forms visual perception.

Our model indicates that visual layers consist of visual clusters, and the number of visual layers and visual clusters determines the graph complexity. That is, increasing the number of visual layers and/or the number of visual clusters increases graph complexity because both visual layers and visual clusters require processing cycles in the brain. The graph complexity affects

a graph reader's comprehension performance, such as speed, accuracy and quality of decision making.

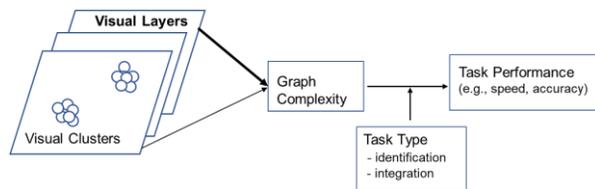


Figure 6: Model of Visual Clustering and Visual Layering

The model also specifies task type as a moderating factor. For graph comprehension, there are two major task types: identification and integration. The question “what is the revenue for year 2021” is an example of identification task. Integration tasks can be further classified into comparison tasks (e.g., which year has the least sales) and pattern recognition tasks (e.g., what is the sales trend in the last four years). Generally speaking, identification tasks are easier to solve compared to integration tasks.

In the proposed model (Figure 6), the concept “Visual Layers” and the associated arrow are bolded because we speculate that the number of visual layers plays a more important role than the number of visual clusters in terms of determining the graph complexity. It is conceivable that the more visual layers the more eye movement interruptions during the processing cycle. However, we decide not to test whether the number of visual layers play a more important role in determining the graph complexity. We will leave it for the next study.

We formulate the following hypotheses:

H1a: For identification tasks with the same number of visual layers, there is no significant difference of processing time between users who read a chart with more visual clusters than users who read a chart with fewer visual clusters.

H1b: For integration tasks with the same number of visual layers, users who read a chart with more visual clusters will take more time to complete the task than users who read a chart with fewer visual clusters.

H2a: For identification tasks with the same number of visual clusters, there is no significant difference of processing time between users who read a chart with more visual layers and users who read a chart with fewer visual layers.

H2b: For integration tasks with the same number of visual clusters, users who read a chart with more visual layers will take more time to complete the task than users who read a chart with fewer visual layers.

4. EXPERIMENT DESIGN AND PILOT STUDY

We did a pilot study in a business undergraduate class in summer 2022. Participating students would earn 1% extra credit for the course. Neither co-author taught the class. The survey itself did not capture students' identities. The class instructor knows which student participated in this study but has no access to the survey results. Only the co-authors of this paper have access to the survey results which are completely anonymous. We designed 12 charts from simple imaginary business datasets. Each chart was associated with two question. All students would answer the first 12 questions. For the remaining six questions, students were randomly assigned to the control group or the treatment group. As a result, each student answered 18 questions total. Both control and treatment groups were given the same tasks and charts with the same number of data points and feature layers. The difference between the two groups was that the treatment group viewed the charts with fewer visual layers due to color grouping effect, as shown in Figure 7A and 7B as examples. Please see the Appendix for a complete list of the experiment tasks.

Students participated in the experiment and filled out the surveys on Qualtrics. The study procedure is listed as follows:

Step 1: Login to the university survey administering system (SONA) (for participation mark earning).

Step 2: Launch the Qualtrics survey link, read study information, and agree to participate.

Step 3: Fill out pre-task survey (including demographic information, such as age, gender, and familiarity with charts).

Step 4: Answer four questions for color-blind detection.

Step 5: Do the experiment tasks, with one chart and the associated question per page. We used “last click” to measure the time each participant used to complete each task/question. “Last click” recorded the number of seconds that the user spent to click the answer after the page was

loaded and right before the page was submitted.

Step 6. A “thank you” message displays at the end of the survey to confirm the completion of the study.

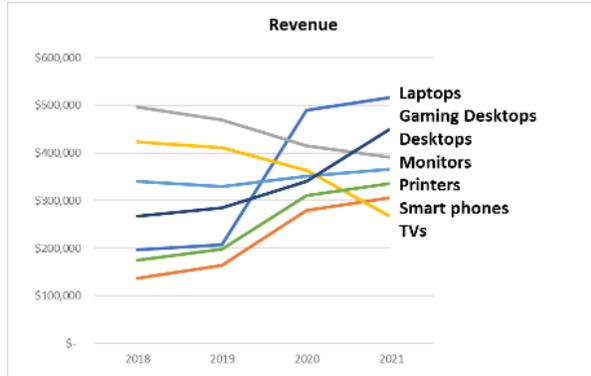


Figure 7A, Control group A line chart with 7 feature layers & 7 visual layers

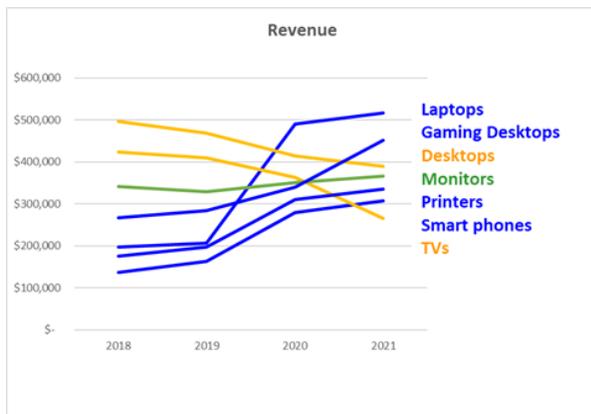


Figure 7B: Treatment group A line chart with 7 feature layers and 3 visual layers

There were 26 students in the class, and 23 participated in this study. Twelve students were in the control group and 11 students in the treatment group. Out of the 23 students, 20 were in age group 18-29 and three were in age group 40-49 (two in the control group and one in the treatment group). Nine students were male, and 14 students were female. Eighteen students rated themselves as being moderately familiar or very familiar with using charts. One student rated him/herself as being extremely familiar with using charts. The four remaining students were slightly familiar with charts. A comparison between the students who were extremely familiar with using charts and the others shows that the former did not perform significantly different from the latter. All students had experiences of creating charts in Excel, which was

the software that was used to create charts for the experiment.

Charts	Accuracy	Time spent (in seconds)	P value hypothesis supported?
Pie charts: 3 vs. 7 slices	96% vs. 96%	9.0 vs. 16.3	P = 0.003 No
Stacked column charts: 3 vs. 7 slices	70% vs. 83%	22.2 vs. 21.4	P = 0.44 Yes

Table 1: Summary of H1a Testing

There were four questions to test whether participants were color blind. All students correctly answered two of these questions. For the other two questions, two and three students chose the wrong answers respectively. Since no student chose more than one incorrect answer, we included all participants in the analysis. For Q1 to Q12, we used paired t test. For Q13 to Q18 (see the Appendix for questions), we used t test with two samples assuming unequal variances. The following tables summarize the results.

Charts	Accuracy	Time spent (in seconds)	P value hypothesis supported?
Clustered column chart: 3 layers vs. 7 layers	91% vs. 91%	11.9 vs. 17.7	P < 0.01 no
Line chart: 3 layers vs. 7 layers	36% vs. 67%	42.9 vs. 29	P = 0.14 yes
Scatter plot: 1 layer vs. 3 layers	91% vs. 92%	19.3 vs. 22.2	P = 0.42 yes
Scatter plot: 1 layer vs. 3 layers	82% vs. 92%	19.7 vs. 16.1	P = 0.50 yes
Scatter plot: 1 vs. 7 layers	91% vs. 100%	38.0 vs. 27.9	P = 0.27 yes
Scatter plot: 1 vs. 7 layers	82% vs. 83%	22.75 vs. 22.74	P = 0.50 yes

Table 2: Summary of H2a Testing

Tables 1 and 2 summarize the test results of hypotheses H1a and H2a, both are about identification tasks.

Tables 3 and 4 summarize the test results of hypotheses H1b and H2b, both are about integration tasks.

Charts	Accuracy	Time to complete task (in seconds)	P value hypothesis supported?
Pie charts: 3 vs. 7 slices	96% vs. 78%	8.0 vs. 17.5	P < 0.01 yes
Stacked column charts: 3 vs. 7 slices	78% vs. 96%	13.9 vs. 22	P = 0.03 yes

Table 3: Summary of H1b Testing

Charts	Accuracy	Time spent (in seconds)	P value hypothesis supported?
Clustered column chart: 3 vs. 7 layers	96% vs. 83%	17.9 vs. 16.6	P = 0.21 no
Line chart: 3 vs. 7 layers	36% vs. 17%	32.3 vs. 50.3	P = 0.09 no

Table 4: Summary of H2b Testing

5. RESULT DISCUSSION AND STUDY LIMITATIONS

In this section, we discuss the results of the hypothesis testing (Table 5) followed by the study limitations.

Hypothesis	# of tasks	Questions	# of tasks supporting hypothesis
H1a	2	Q1 vs Q3 Q5 vs. Q7	1
H2a	6	Q9 vs. Q11 Q14,15,16, 17, 18: control vs. treatment	5
H1b	2	Q2 vs. Q4 Q6 vs. Q8	2
H2b	2	Q13, control vs. treatment	0

Table 5: Summary of the hypothesis testing

H1a and H2a: For identification tasks, the number of visual layers or visual clusters did not affect users' speed of processing the visuals (six out of eight tasks supporting the hypothesis). Both hypotheses were partially supported.

H1b and H2b: For integration tasks, two tasks supported the hypothesis, however, the other two tasks failed to support the hypothesis. Therefore, H1b and H2b received partial support. We will

need to design more tasks to test H1b and H2b in the future studies.

The study results need to be interpreted with the following limitations in mind. First, the number of participants was small. There were only 23 participants in this pilot study. Previous studies in this area had more participants. Some visual studies had at least 50 participants for each condition (e.g., Zhu & Watts, 2009; Walden, et. al. 2014). The problem with small sample size is that there is not enough statistical power to detect the difference, even if there is one.

The second limitation is that the participants did the study online at the time and place of their choice. As a result, there were outliers for each of the experiment tasks. We excluded the outliers in the t tests, but the variances for several tasks were still large. As we did not have the chance to observe participants while they were performing the tasks, we were not able to tell whether or not the participants were distracted or interrupted during the tasks (e.g., replying a text message).

The third limitation is that there were not enough integration tasks in the experiment. However, as a pilot study, this experiment served the purpose of informing a better design for the future studies.

6. CONCLUSIONS AND FUTURE STUDIES

Graph designers should understand and be able to estimate graph complexity, because they need to make sure that readers can process the charts quickly and accurately. As demonstrated by the CHI conference expenses graph (Figure 2), inappropriate design prevents designers from conveying the insights from the data. In this paper, we propose a novel research model for graph complexity based on the observation that the sheer number of data points may not be the most important factor to determine graph complexity. Instead, how these data points are displayed or plotted on a graph may play a more important role in determining graph complexity.

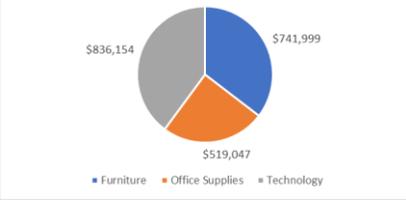
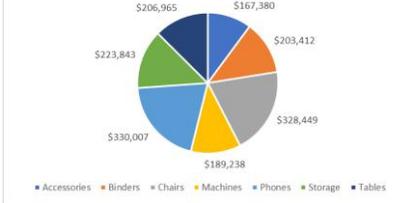
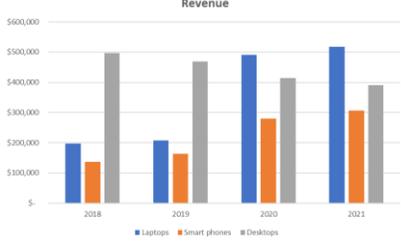
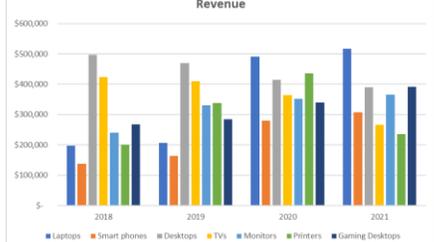
In this model, we hypothesize that 1) readers who conduct easy task such as simple search or identification may perform equally well using graphs with more or fewer visual clusters or layers; and 2) readers who conduct integration tasks such as comparison or pattern recognition will take more time to process graphs with more clusters or layers than to process graphs with fewer clusters or layers. Our pilot study partially supported hypothesis 1 but the results for hypothesis 2 were mixed. In our next study, we will design an experiment that addresses the

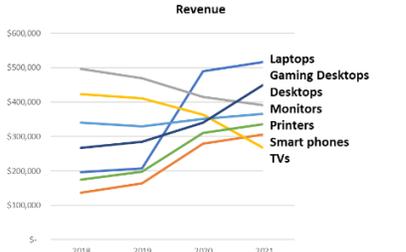
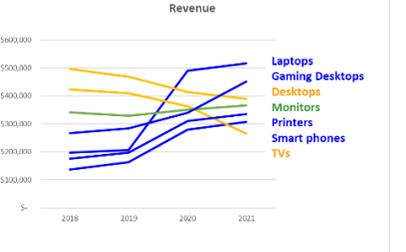
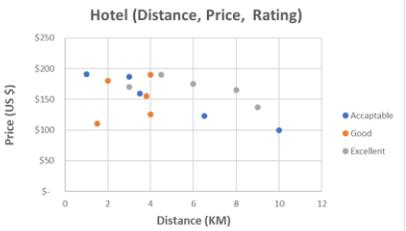
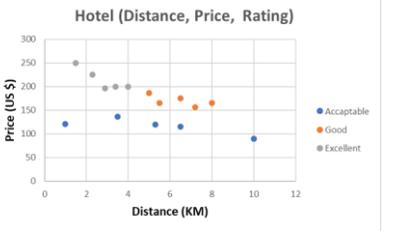
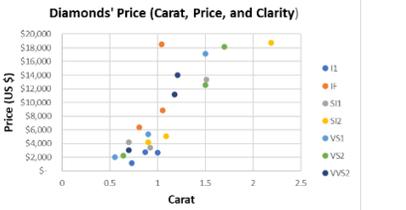
limitations of this study. In addition, we will test whether the number of visual layers play a more important role than the number of clusters in terms of determining graph complexity.

7. REFERENCES

- Arnheim, R. (1969). *Visual Thinking*. University of California Press.
- Bera, P., Soffer, P., & Parsons, J. (2019). Using eye tracking to expose cognitive processes in understanding conceptual models. *MIS quarterly*, 43(4), 1105-1126.
- Cairo, A. (2012). *Functional Art: An Introduction to Information Graphics and Visualization*, New Riders (a division of Pearson Education), USA.
- Cao, J. (2015). Gestalt Principles for Designers – Applying Visual Psychology to Modern Day Design. Retrieved on Sept. 16, 2022 from <https://blog.teamtreehouse.com/gestalt-principles-designers-applying-visual-psychology-modern-day-design>,
- Chen, K.T, Dwyer, T., Bach, B., & Marriott, K. (2021). Rotate or Wrap? Interactive Visualizations of Cyclical Data on Cylindrical or Toroidal Topologies, *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1-1
DOI:10.1109/TVCG.2021.3114693
- “Feature Layers” (2022). Retrieved on Aug. 25, 2022 from ArcGIS Website, <https://doc.arcgis.com/en/arcgis-online/reference/feature-layers.htm>
- Gordon, I. E. (1989). *Theories of Visual Perception*. John Wiley & Sons.
- “Graph vs Graphics - What's the difference?” Retrieved on Aug. 25, 2022. <https://wikidiff.com/graph/graphics>,
- Hensley, J. (2016a), Improve your Designs with the Principles of Similarity and Proximity (Part 1), *Smashing Magazine* (an online web magazine). Retrieved on Sept. 16, 2022 from: <https://www.smashingmagazine.com/2016/05/improve-your-designs-with-principles-similarity-proximity-part-1/>
- Hensley, J. (2016b), Improve Your Designs With The Principles Of Closure and Figure-Ground (Part 2), *Smashing Magazine*. Retrieved on Sept. 16, 2022 from: <https://www.smashingmagazine.com/2016/05/improve-your-designs-with-the-principles-of-closure-and-figure-ground-part-2/>.
- Keogh, A., Johnston, W., Ashton, M., Sett, N., Mullan, R., Donnelly, S. & Caulfield, B. (2020). It’s Not as Simple as Just Looking at One Chart: A Qualitative Study Exploring Clinician’s Opinions on Various Visualization Strategies to Represent Longitudinal Actigraphy Data. *Digital Biomarkers*, 4(1), 87-99.
- Knaflic, C.N. (2015), *Storytelling with Data: A Data Visualization Guide for Business Professionals*, pp. 74 – 80.
- Ratwani, R. M., Trafton, J. G., & Boehm-Davis, D. A. (2008). Thinking graphically: Connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology: Applied*, 14(1), 36-49.
- Talbot, J., Setlur, V., & Anand, A. (2014). Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2152-2160.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & Vo der Heydt, R. (2012), A Century of Gestalt Psychology in Visual Perception I. Perceptual Grouping and Figure-Ground Organization, *Psychol Bull.*, Nov. 138(6): pp.1172-1217, doi: 10.1037/a0029333. Retrieved on Sept. 16, 2022 from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3482144/>
- Walden, E.A., Cogo, G.S., Lucas, D.J., Moradiabadi, E., & Safi, R. (2018). Neural Correlates of Multidimensional Visualizations: An fMRI Comparison of Bubble and Three-Dimensional Surface Graphs Using Evolutionary Theory. *MIS Quarterly*, 42(4), 1097-1116.
DOI: 10.25300/MISQ/2018/14419
- Ware, C. (2008). *Visual Thinking for Design. The Morgan Kaufmann Series in Interactive Technologies*. Elsevier.
- Zhu, B., & Watts, S. A. (2009) Visualization of Network Concepts: The Impact of Working Memory Capacity Differences. *Information Systems Research* 21(2):327-344. <https://doi.org/10.1287/isre.1080.0215>

Appendix A
Visual Experiment Tasks
 (answers are not listed due to space limit)

Graph	Questions
<p align="center">2021 Revenue by Product Category</p>  <p>Q1. What is the revenue for Office Supplies in 2021? (H1a) Q2. Which category generated most revenue in 2021? (H1b)</p>	<p align="center">2021 Revenue by Product Sub-Category</p>  <p>Q3. What was the revenue for Tables in 2021? (H1a) Q4. Which sub-category generated the second least revenue in 2021? (H1b)</p>
<p align="center">2021 Revenue</p>  <p>Q5. What is the revenue for Seattle in Quarter 4? (H1a) Q6. Which city has the lowest revenue for a single quarter? (H1b)</p>	<p align="center">2021 Revenue</p>  <p>Q7. What is the revenue for Austin in Quarter 1? (H1a) Q8. Which city has the lowest revenue for a single quarter? (H1b)</p>
<p align="center">Revenue</p>  <p>Q9. What was the revenue for smart phones in 2020? (H2a) Q10. The largest difference between the revenue of Desktops and Smart phones appeared in which year? (H2b)</p>	<p align="center">Revenue</p>  <p>Q11. What was the revenue for Printers in 2021? (H2a) Q12. The largest difference between the revenue of Printers and Gaming Desktops appeared in which year? (H2b)</p>

 <p>Q13-control. Which product generated the second least revenues in 2021? (H2b) Q14-control. The revenue of which products increased over the years? (This is a multi select question, which means you need to select all applicable alternatives)(H2a)</p>	 <p>Q13-treatment. Which product generated the second least revenues in 2021? (H2b) Q14-treatment. The revenue of which products increased over the years? (This is a multi select question, which means you need to select all applicable alternatives) (H2a)</p>
 <p>Q15-Control. What is the distance range for hotels with rating of "Excellent"? (H2a) Q16-Control. What is the price range for hotels with rating of "Acceptable"? (H2a)</p>	 <p>Q15-Treatment. What is the distance range for hotels with rating of "Excellent"? (H2a) Q16-Treatment. What is the price range for hotels with rating of "Acceptable"? (H2a)</p>
 <p>Q17-Control. What is the price range for diamonds with clarity of "VS2"? (H2a) Q18-Control. What is the Carat range for diamonds with Clarity of "SI1"? (H2a)</p>	 <p>Q17-Treatment. What is the price range for diamonds with clarity of "VS2"? (H2a) Q18-Treatment. What is the Carat range for diamonds with Clarity of "SI1"? (H2a)</p>