Online Submission ID: 1326

Palettailor: Discriminable Colorization for Categorical Data

– Supplementary Material –

Category: Research

Paper Type: Algorithm/Technique

This supplemental material file provides additional experimental results for our submitted paper titled "Palettailor: Discriminable Colorization for Categorical Data."

Overview

We include materials for:

1. Details of Colorization method implementation
2. Parameters of Simulated Annealing
3. Partial Trials Used in User Study
4. Data Organization Details
5. User Study Setup
6. Pilot Study Details and Statistics
7. Formal Study Time Statistics
8. Results of Bar Charts and Line Charts

Details of Colorization method implementation. Due to limit information we got from the paper [1], we can only implement the method by setting the parameters empirically. There are multiple unknown parameters in Chen's method, including the size of the M local regions, the adjustable weight k, the constraint of color distance d and the user adjustable parameter L*. In our implementation, we set the local region to be 25×25 pixels to include more local region, set $k = 1000$ to ensure the minimum color distance is larger than the threshold $d = 100$, and set $L* = 50$. Then optimized the palette by Nelder and Mead method[2].

Parameters of Simulated Annealing. There are three parameters which influence the iteration times in the simulated annealing process, i.e., the cooling coefficient α, the initial temperature T, the end temperature T_{end}. We quantitatively compared these three parameters, see Figure 1.

Partial Trials Used in User Study. We have 180 trials(30 datasets $\times$ 6 conditions) in total for scatterplots and 150 trials(30 datasets $\times$ 5 conditions) for line charts, due to the limit space, we provide the partial representative trials used in the user study. See Figure 2 and Figure 3.

Data Organization Details. We organized the data following Latin Square, see Table.1 and Table.2 in Experiment 1, and it was similarly used in Experiment 2.

Table 1: Datasets Organization for the Discrimination Tasks: 30 scatterplots $\times$ 6 palettes.

	Palettailor	**Tableau Best**	**Tableau Worst**	**Colorgorical Best**	**Colorgorical Random**	**Colorization**
Dataset 1	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Dataset 2	Group 6	Group 1	Group 2	Group 3	Group 4	Group 5
Dataset 3	Group 5	Group 6	Group 1	Group 2	Group 3	Group 4
Dataset 4	Group 4	Group 5	Group 6	Group 1	Group 2	Group 3
Dataset 5	Group 3	Group 4	Group 5	Group 6	Group 1	Group 2
Dataset 6	Group 2	Group 3	Group 4	Group 5	Group 6	Group 1
...						

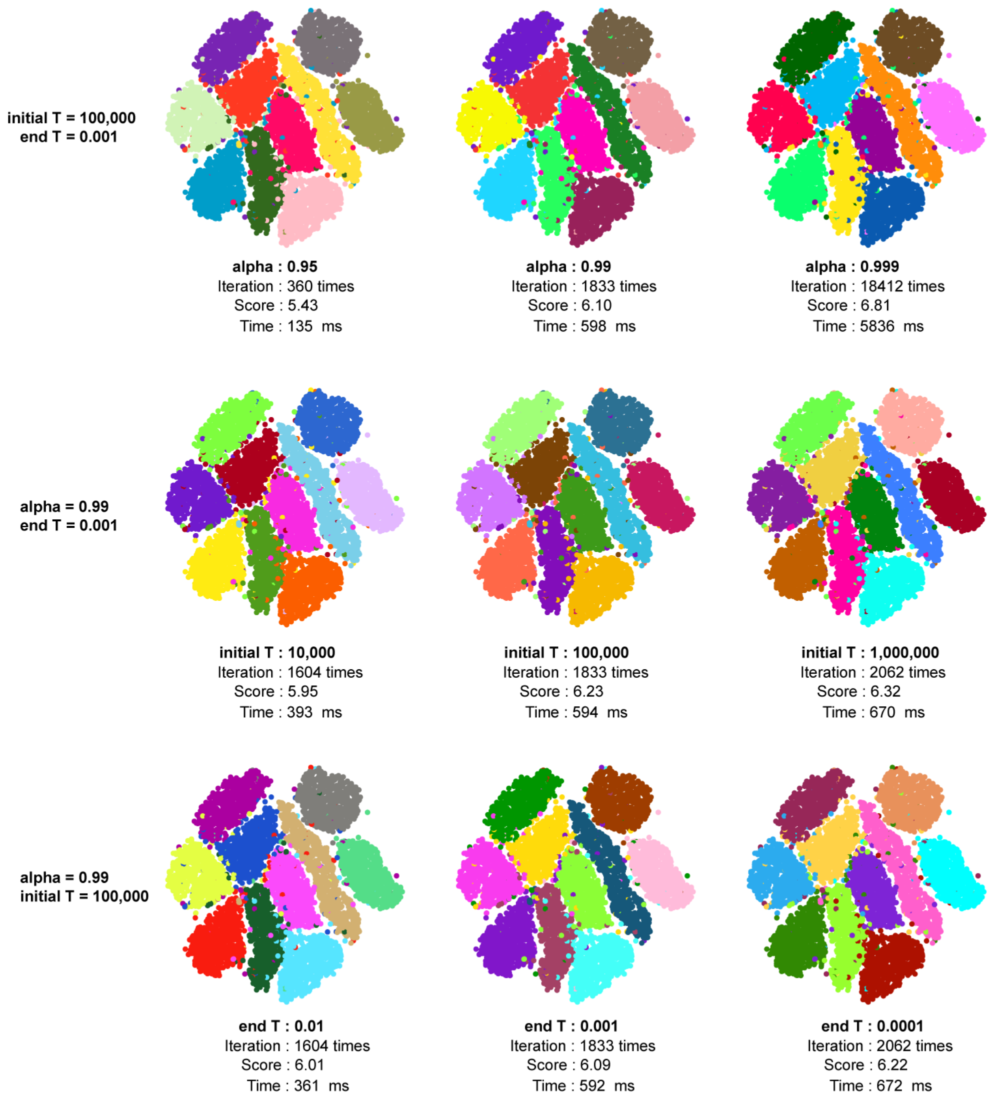

Figure 1: Comparison of different parameters. We can see that more iterations resulting in larger score and longer time.

Table 2: Datasets Organization for the Preference Task: 30 scatterplots × 5 pairs.

	Palettailor v.s. Tableau Best	Palettailor v.s. Tableau Worst	Palettailor v.s. Colorgorical Best	Palettailor v.s. Colorgorical Random	Palettailor v.s. Colorization
Dataset 1	Group 1	Group 2	Group 3	Group 4	Group 5
Dataset 2	Group 5	Group 1	Group 2	Group 3	Group 4
Dataset 3	Group 4	Group 5	Group 1	Group 2	Group 3
Dataset 4	Group 3	Group 4	Group 5	Group 1	Group 2
Dataset 5	Group 2	Group 3	Group 4	Group 5	Group 1
...					

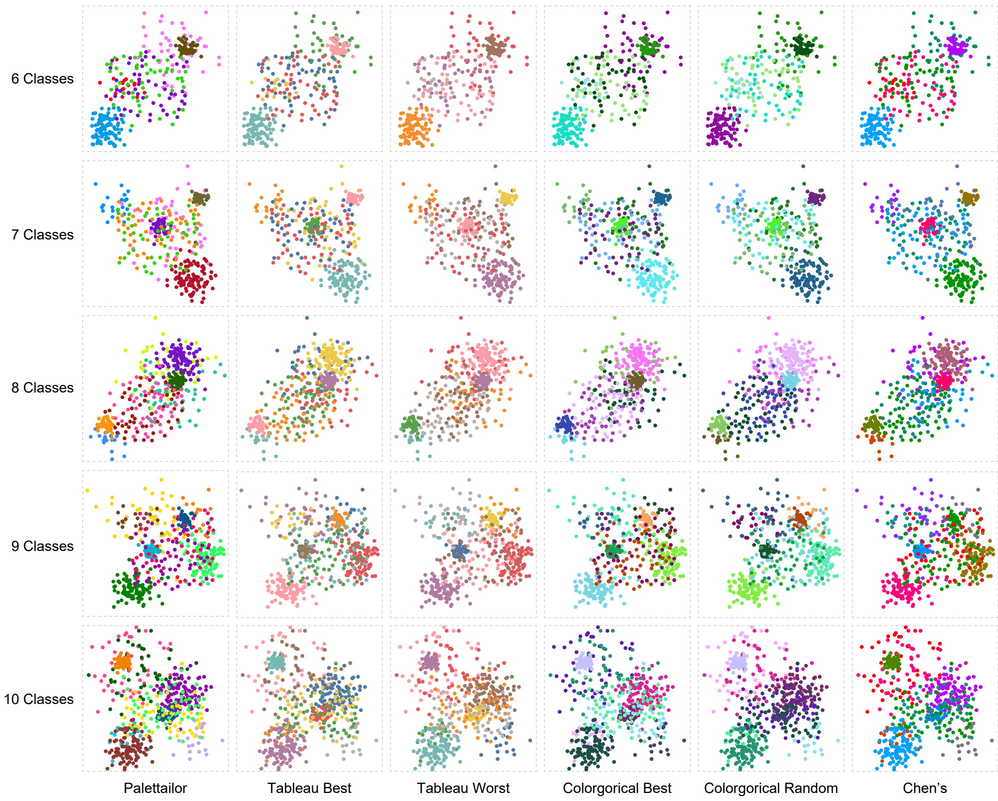

Figure 2: Partial representative trials of different palette size for scatterplot.

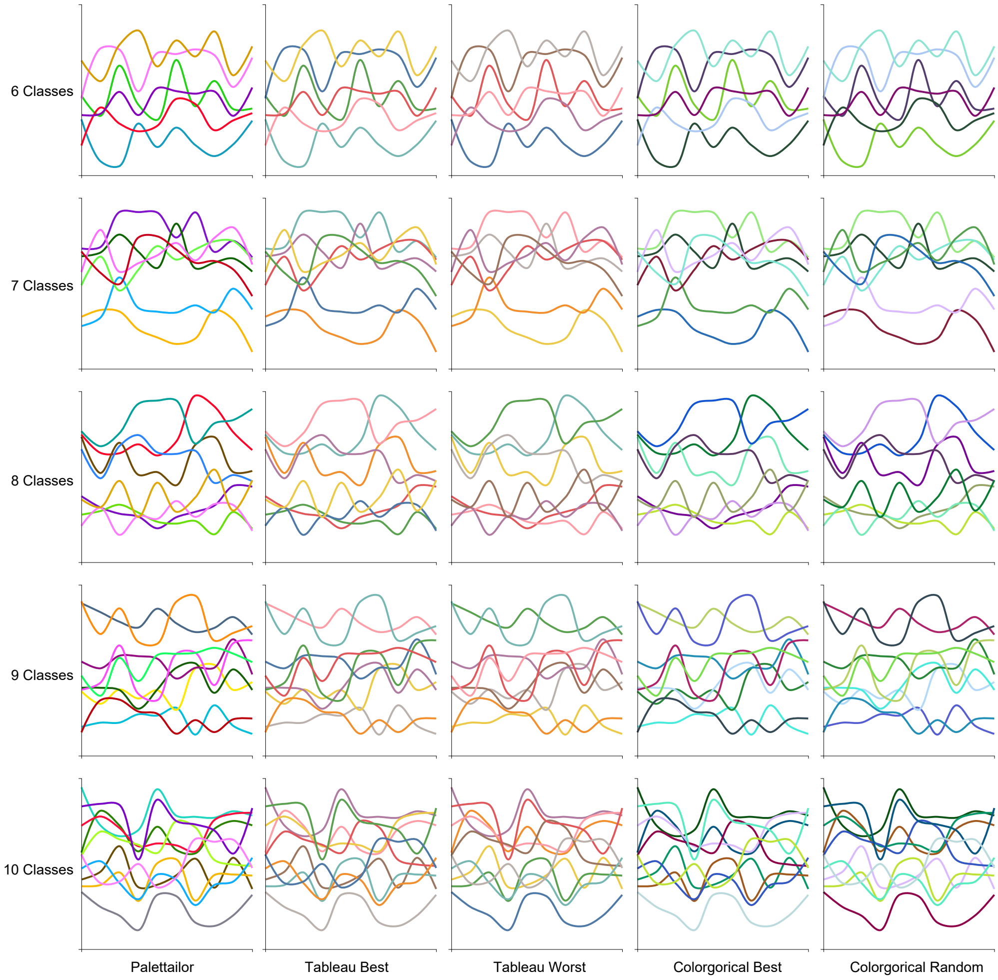

Figure 3: Partial representative trials of different palette size for line chart.

User Study Setup.

We provide the details setup of the user study in our paper. There are three tasks in Experiment 1, including Counting Task, Comparison Task and Preference Task. In each task, we first show the participants an Instruction that describe the task, see Figure 4 top row. Followed by the Instruction, we show an example of how to do the trials, this part we called User Guide. Then we provide two training trials in the Counting and Comparison tasks, there are no training part in the Preference task, see Figure 4 third row. After training, we asked the users do the trials as accurately as possible, see Figure 4 bottom row. Then the formal study is starting. Experiment 2 is similar, see Figure 5.

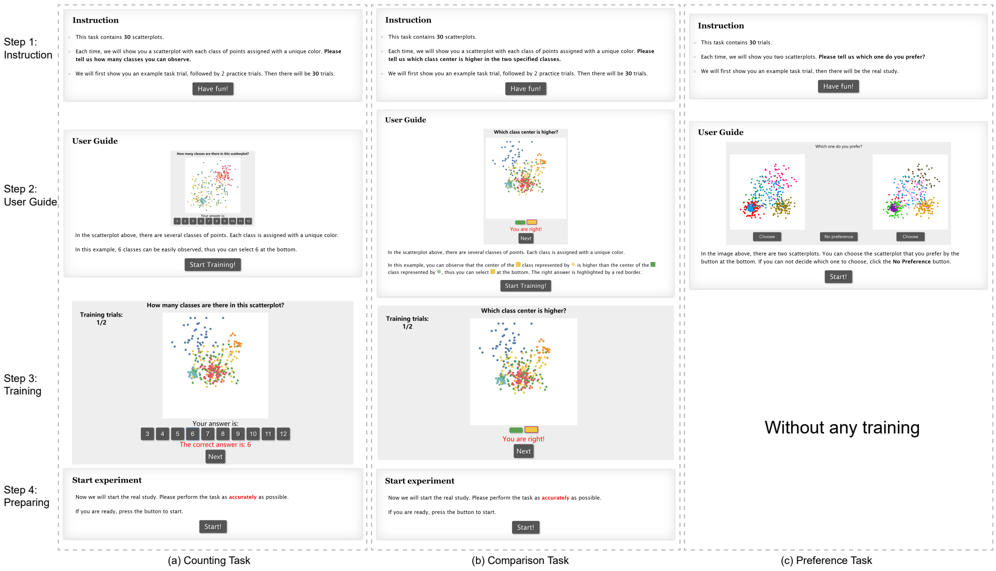

Figure 4: Pipeline of the user study for Scatterplot.

Table 3: Experiment details for pilot study.

	Task 1	**Task 2**	**Task 3**
User 1	Group 1	Group 2	*Group 1*
User 2	Group 2	Group 3	*Group 2*
User 3	Group 3	Group 4	*Group 3*
User 4	Group 4	Group 5	*Group 4*
User 5	Group 5	Group 6	*Group 5*
User 6	Group 6	Group 1	*Group 1*

Pilot Study Details and Statistics. We provide the detailed results of the pilot study in our paper. For Experiment 1, we recruited 6 students in our school and each student carried out all the three tasks with different data, see Table.3. Due to the COVID-19, we recruited 15 people(11 for discrimination task and 4 for preference task) through the Amazon Mechanical Turk to do our tasks for Experiment 2, see Table.4.

The statistics results can be seen in Figure 6. As shown in Figure 6(c), Palettailor's results are comparable with Tableau Best. This is different with the formal study results. One possible reason is that the pilot study users are all familiar with visualization and they have done all three tasks, thus they might prefer results with better discrimination. While in formal study, users just do the preference task with no training and they have different background.

Formal Study Time Statistics. We plotted the time consuming in the formal study, see Figure.7

Results of Bar Charts and Line Charts. We showed Palettailor's results and the Tableau results for the **bar chart** and **line chart** extensions, see Figure 8 and Figure 9.

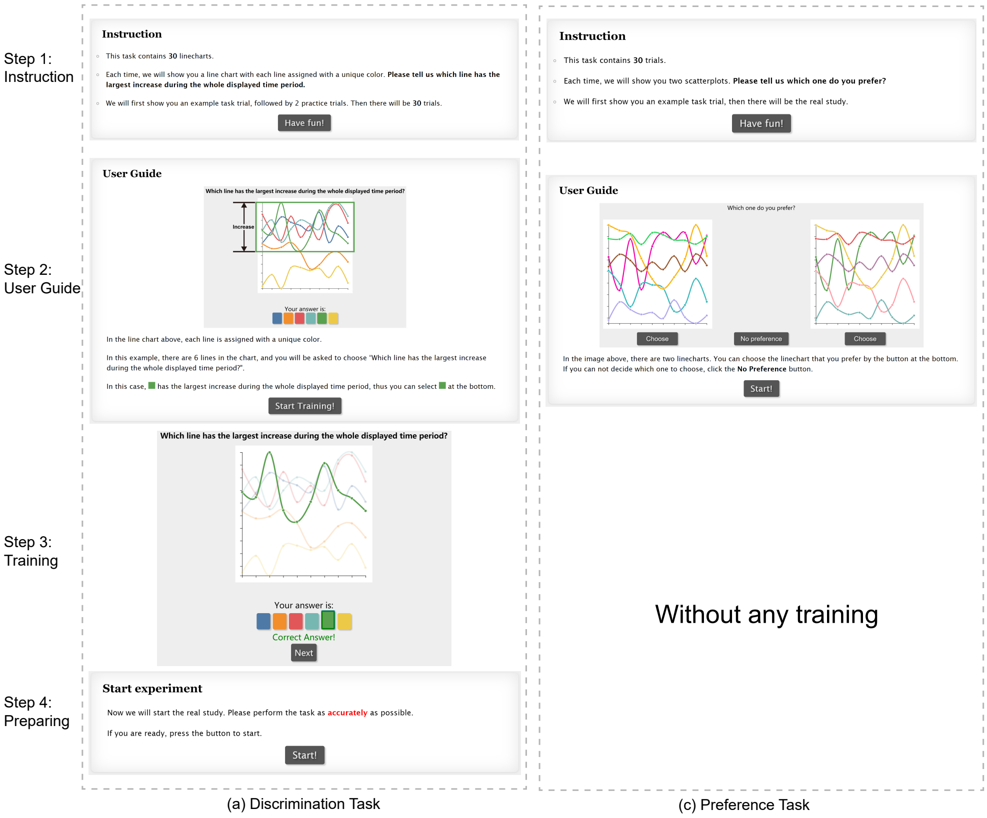

Figure 5: Pipeline of the user study for Line Chart.

Table 4: Participants recruited for the pilot study. G# specifies a participant group number.

Task	Participants	G1	G2	G3	G4	G5
Discrimination	11	3	2	2	2	2
Preference	4	1	1	1	1	NA

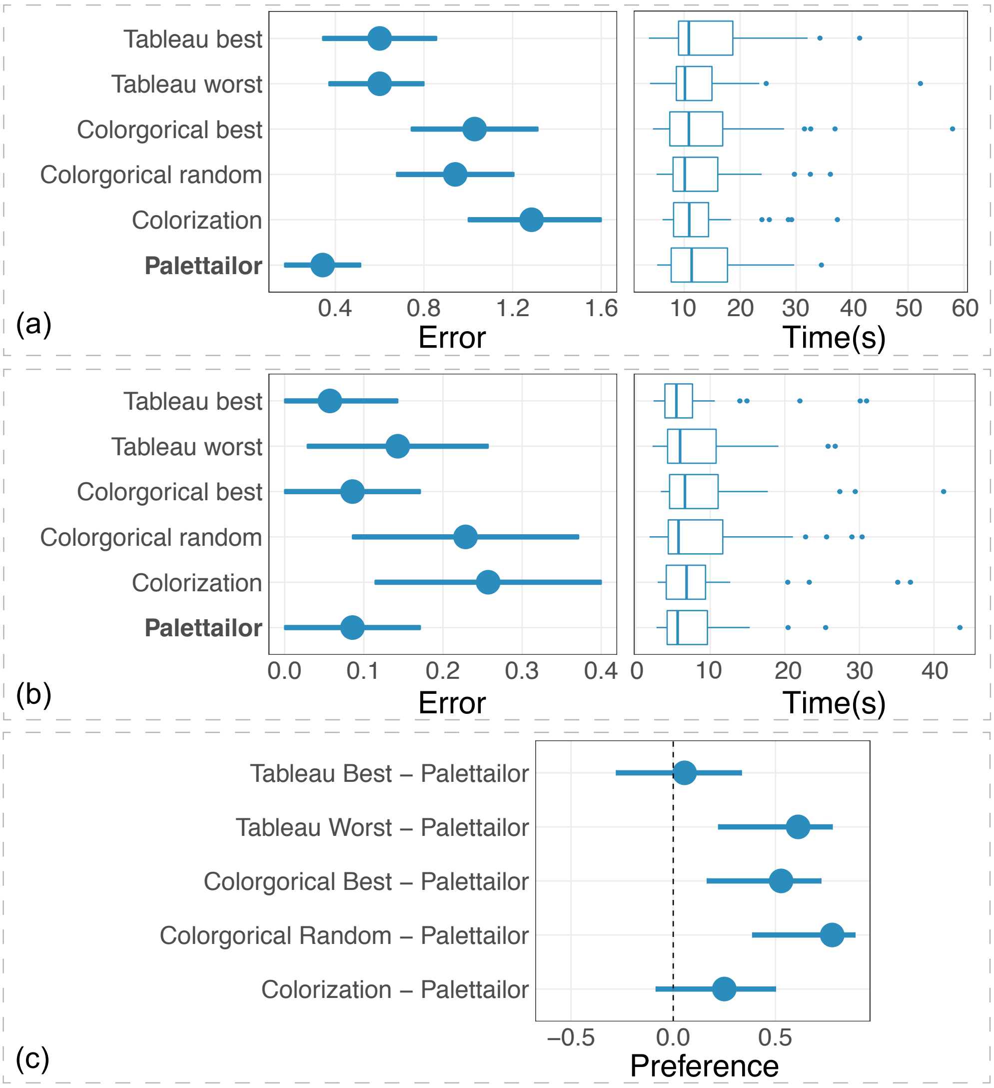

Figure 6: Pilot study results. (a) Pilot study results for the counting task of scatterplots; (b) Pilot study results for the comparison task of scatterplots. (c) Preference between Palettailor and other conditions.

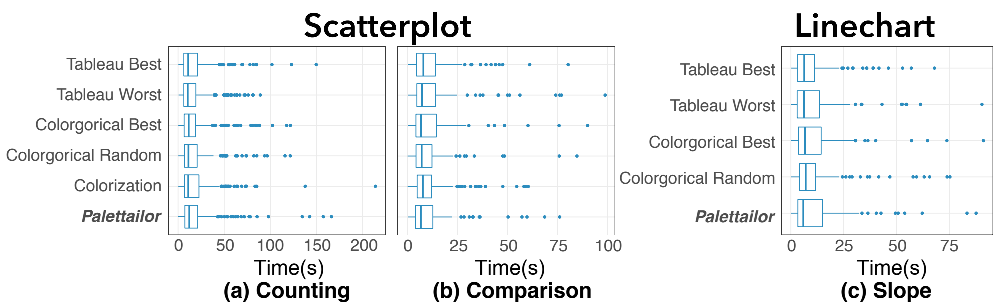

Figure 7: Time boxplot for three discrimination tasks.

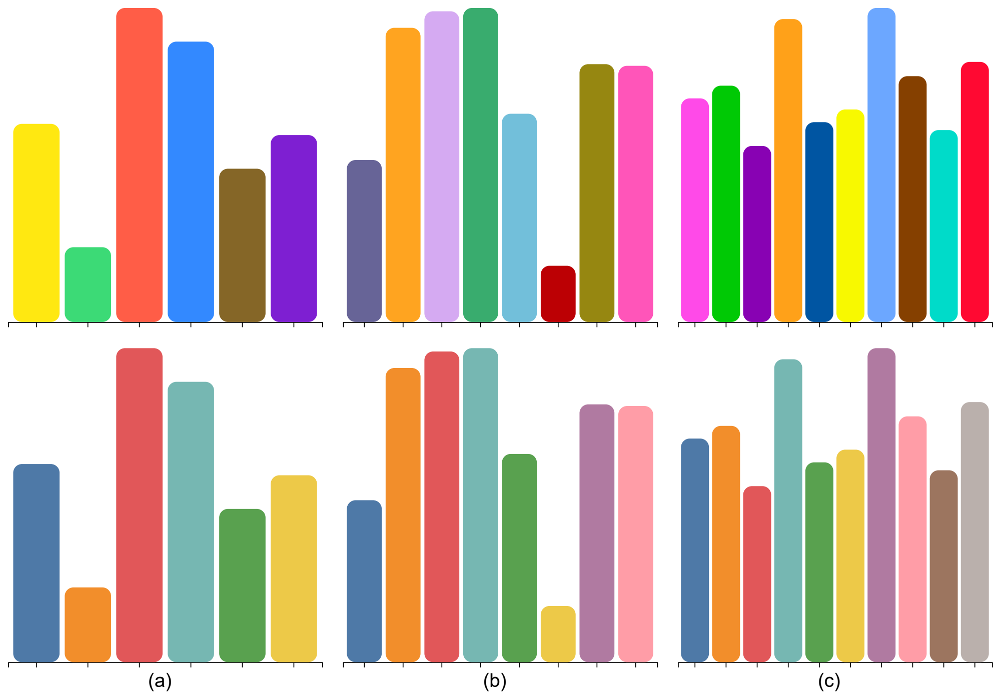

Figure 8: The results of bar chart. Top row is our results and bottom row is Tableau results. (a) 6 classes; (b) 8 classes; (c) 10 classes.

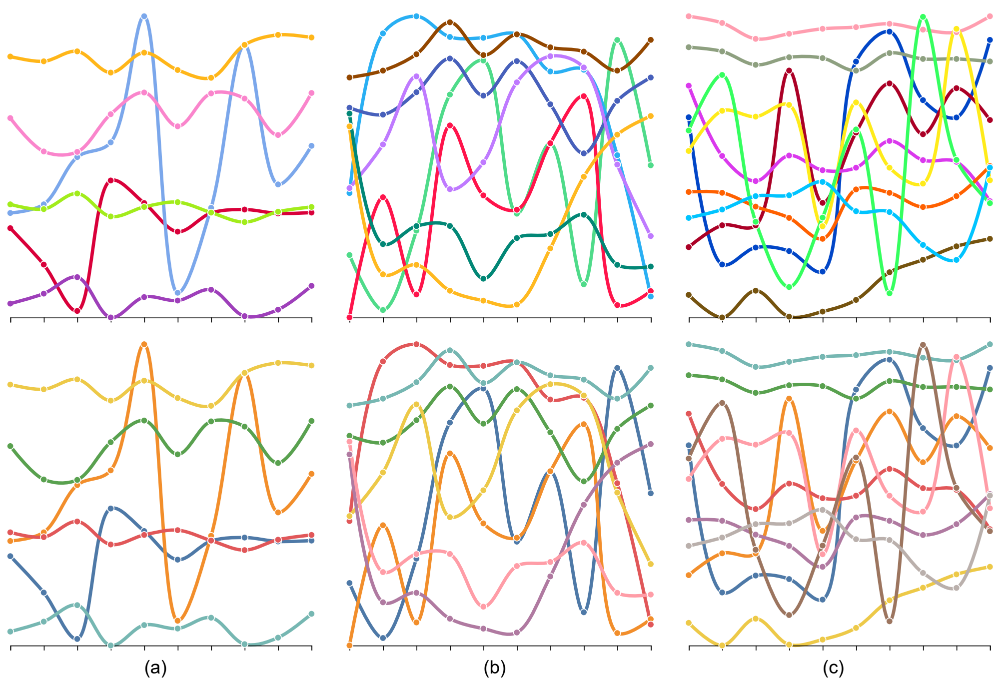

Figure 9: The results of line chart. Top row is our results and bottom row is Tableau results. (a) 6 classes; (b) 8 classes; (c) 10 classes.

References

[1] H. Chen, W. Chen, H. Mei, Z. Liu, K. Zhou, W. Chen, W. Gu, and K.-L. Ma. Visual abstraction and exploration of multi-class scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1683–1692, 2014.

[2] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1):112–147, 1998.