Data Visualization Final Project

Introduction

Aesthetics plays an integral role in promoting personal well-being. While individuals may not be consciously aware of their choices, they intrinsically prefer a setting where they can function efficiently. Previous research showed that people have a preference for natural over artificial environments. This aesthetic preference has been demonstrated to be strongly associated with nature's potential restorative effects in the research. For example, previous research has shown salubrious effects after engaging with natural environment, such as improving memory, attention and mood. In modern times, however, increasing man-made architect and construction have alienated individuals from the natural environment. Therefore, it is important to gain a better understanding of people's aesthetic preference of the environment in order to maintain a crucial engagement with nature and get potential benefits from it.

As driving has become a daily part of our everyday life, the field has yet, to date, studied people's aesthetic preferences with regards to roadside environment. Past studies on scenic beauty suggested that people prefer the properties of nature. They reported that people prefer semantic features of nature, such as trees, water, and flowers, and that removal of built-up features like billboards could increase landscape appreciation. A recent study quantified natural scenes by using low-level visual features (i.e., edge, hue, saturation, brightness, standard deviation of hue, standard deviation of saturation, etc.). In their study, participants were asked to rate the naturalness of the shown images and their likeness towards them. Results showed that low-level visual features significantly predicted people's preference towards the images as well as the naturalness ratings of the images.

In the current study, we aim to investigate people's aesthetic preference by using constructed highway environments with constantly changing surroundings. We generated simulation videos with the input of Geographical Information System (GIS) data, images and 3D models. The videos contain controlled environmental characteristics of a landscape, which are convenient for examining what specific design elements that affect scenic preference. Participant were asked to rate their likeness towards the videos while watching them throughout the experiment.

With this study, we hope to have a better understanding of people's aesthetic preferences in terms of both low-level visual features and semantic features. This information could potentially be helpful for urban design and landscape planning in constructing environment that optimizes subjective well-being.

Methods

This visualization aims at presenting research results in a more understandable and perceivable way for lay audience. I chose to use flexdashboard to present a research paper because the audience can choose which part of the paper they want to read. Besides, dashboard allows different layouts of the page and works well with plotly. Dashboard works very similar to a research paper where I could organize the texts into different sections but with more interactions with the audience.

In order to investigate how aesthetic preference was influenced under different environmental elements, I plotted the preference ratings graph and colored with different semantic features using ggplotly. Audience can check out for preference ratings, time, and the specific feature. The interactive graph is better than a static graph in a way that related statistics will show up as hovering over the graph. This plot was done by using geom_col where every bar

representing the height of the preference rating at every half second, and the bars were filled in different color according to which feature that the specific scene contained. For the color choices, I basically followed the common perceptions of those elements, for example, tree was colored in green, and lake was colored in blue. I also coupled the colors of a single feature and its combination with another feature, for example, lake was colored in blue, and lake with tree was colored in darker blue; billboard was colored in light yellow, and billboard with tree was colored in earthy yellow. I think in this way readers will be less likely to get confused when they are try to figure out which color represents which feature.

However, it was quite hard to insert images with arrows and couple with the previous plot. Thus, in the average ratings plot below, I just inserted an image where the ratings plot was coupled with snapshots from the video. In this plot, audience could clearly see that preference ratings largely dropped with the presence of billboard, whereas lake tended to boost the ratings. Trees sometimes contributed positively to the preference, and sometimes negatively impacted the preference. Especially in video Mountain 1, preference ratings did not correspond closely with the change in tree prevalence.

For the visualization of regression analyses, I did not choose the traditional regression table to present my results, instead I chose to use a dot-whisker plot. The purpose of visualization is to speed up information processing. Compared to reading numbers from a regression table and interpreting the results, this dot-whisker plot could easily tell the audience which a positive predictor is and which a negative predictor is. I also put two brackets and added significance to the predictors markers in the y-axis so that it is clear to see the two big categories of the predictors and how significant they are predicting the preference ratings. Taking some useful comments from the peer review, I plotted the bar graph below in order to compare how different

predictors performed across different models. By combining the coefficients from three models together, it could clearly be seen that there were some consistencies across the models, for example, more saturation and less brightness and variations in saturation positively contributed to the preference ratings; billboard negatively impacted the preference, whereas lake positively impacted the preference.

For the discussion, I also did a traditional correlation heatmap with all the features. I chose to only show the upper part of the map because the correlation map is usually symmetric. I think the color did a good job presenting the contrast between negative values and positive values and high correlations and low correlations, so I just used the default colors.

Results/discussion

Through my visualization of my research paper, audience could learn the conclusion that low-level visual features and semantic features significantly predict aesthetic preference. Billboard negatively impacts the preference whereas lake positively impacts the preference, and trees' influence vary in different scene types.

I think my visualization is quite truthful because the plots were drawn based on real data and I did not distort the scale when I visualize the results. Second, I think my visualization is functional in some ways than the traditional presenting method in a research paper. For instance, using dot-whisker plot to show regression results will easily tell people where positive or negative predictors are and how different predictors contribute to predict the results. The average ratings plot is also very functional that the height of the feature-shading area fluctuates with the change in actual ratings. The audience could not only see the changing trend of preference ratings, but also know what semantic features might contribute to that change. Third, in terms of beauty, I think the colors of my visualization are acceptable though might not be very beautiful. I

chose the colors in their light yet distinguishable versions so that they could be easily recognized but not have sharp contrasts with each other. Finally, as for insights and enlightenment, my visualization is insightful in a way that audience can learn some results of my research through the plots. But it might be something very enlightening, because there is no surprising findings in the results. The results are generally consistent with previous literature and people's intuitions.