Sara Banai

Introduction to My Project

The Human Early Learning Partnership (HELP) is an interdisciplinary research network at the University of British Columbia. Operating out of the School of Population and Public Health, they bring together a variety of scientific perspectives on early childhood development (Human Early Learning Partnership). Their research is done with the approach that long-term health and well-being is most dependent on the early years of childhood development.

One of HELP's methods for addressing this is the Early Development Instrument (EDI), which is a method of assessing how children rank on certain scales that contribute to long term success. The data is gathered through a questionnaire completed by kindergarten teaches throughout British Columbia. Data is collected every year, but not every school completes it every year depending on the number of students they have to ensure there is equal representation for small and large schools alike. The responses are grouped together every two to three years to create a wave, and there is one dataset for each wave (Human Early Learning Partnership 2016).

The five scales which are measured by this questionnaire are as follows: Physical Health & Well-Being, Social Competence, Emotional Maturity, Language & Cognitive Development, Communication Skills & General Knowledge. The scales are measured as percent vulnerability on a scale of 0-100% based on how far the scores are below a threshold. Vulnerability refers to the level of risk that children have of experiencing challenges throughout school and in the future without additional help (Human Early Learning Partnership 2016). The percent vulnerable for each scale is determined based on subscales that are specific to each scale. Communication Skills &

General Knowledge does not have subscales and therefore is not included in this project. The subscales are standardized because the nature of the scales is such that on average students will do better on some than others (e.g. Basic Literacy versus Advanced Literacy). The scores are standardized to the provincial average at Wave 2 for each subscale, which is zero, and any score higher than zero is above that average and lower than zero is below the average. The standardized score allows the subscales to be compared between each other and over time (Human Early Learning Partnership 2017).

EDI scale data has been available to the public in the past, and each school district has their own report including maps of each of the scales. EDI subscale data has recently been released to the public, and it has not yet been formally mapped. The purpose of my project was to map HELP EDI subscale data from Vancouver Wave 6 in a way that is easy to understand and could be used as the standard for how to map all the subscales across the province for all the waves. I chose to map Vancouver because it is the part of the province I am most familiar with, which means I am more aware of the context and implications of the data that I am mapping than I would be for a part of the province with which I am completely unfamiliar. Additionally, Vancouver data is relatively diverse, making the process of displaying the data more complex and the final product more interesting than would be for a school district that is more uniform. Although the concept for my design is meant to be one that could be applied across all school districts at all waves, some of my data analysis choices were specific to Vancouver Wave 6 and would not necessarily be the best method for other districts and waves. For the purposes of this project, I decided it was more appropriate to do an

analysis that was specific to this data. However, the design and cartographic choices are not only applicable to this context and could be applied to the other subscale datasets.

Geospatial Data Visualization Pipeline

The data used for my project was acquired and parsed by HELP. It was gathered through the EDI questionnaire, and the data is divided by neighborhood, scale and subscale, and wave for all neighborhoods and waves in the province. I filtered the data by removing everything except the Vancouver School District neighborhoods at Wave 6. I separated the scale and subscale data, because the scale data did not need to be mined, as it had already been mined and represented by HELP in their previous reports, and I intended to follow their method.

Next, to mine the data, I had to discern patterns in the data to determine the appropriate classification method and class sizes. All the subscales needed to have the same classification so that the maps could be easily compared between each other. I decided to use Equal Interval classification because it is easy for the audience to interpret, and it matches the classification method for the scale data, which is also equal interval. I looked at the highest and lowest scores of the data and found that all the scores fell within +/- 0.6, except three scores that were below -0.6. Because the subscale data diverges from a meaningful average (zero), I found the standard deviation to determine the class sizes. The standard deviation was around 0.2, although 0 was not actually the average for this wave because the scores are standardized to the average at Wave 2, but it was still close enough to 0 that I decided classes of 0.2

diverging from 0 would be appropriate. Additionally, this class size would create six classes, which is not too many that it overwhelms the map user nor too little that it does not show the diversity in the data. Given that this process of data mining was specific to the Vancouver Wave 6 data, I would not recommend applying it to other waves and school districts when that subscale data is mapped. If all the subscale data is mapped, I would recommend still using equal intervals, but creating classes based on the Wave 2 data for the entire province because that is how the data was originally standardized.

To represent the data, I used ArcMap to create choropleth maps for each scale and subscale. The scale maps followed the same classification method and color scheme as HELP had previously created. For the subscales, I classified each subscale using Manual Interval, so that I could use the intervals I had predetermined, and used a red-to-blue diverging color scheme selected from ColorBrewer to represent the diverging data. I avoided a red/green color scheme to avoid issues of red/green color-blindness. Additionally, I used the same color scheme for all subscales so the maps could be easily compared. Because there were three data points that did not fall within +/- 0.6, I labelled those on the map so as not to misrepresent the data.

Once I had created the basic maps in ArcMap, I moved to Illustrator to refine the design. Following Tufte's Fundamental Principles of Analytical Design, I wanted to make subscale maps for one scale easily comparable (Tufte, 2006). My community partner mentioned that they often print out the maps to lay them out on the table when they are looking at the data. Based on this, I decided to fit my design on a regular 8.5x11" paper with small multiples of each scale and its subscales on one page. I added a grey box around the scale map to make it lower on the visual hierarchy and added a drop

shadow to the subscale maps to raise them on the visual hierarchy as well as to mitigate the island effect.

My final design closely follows Tufte's principle of efficiency, which states that designs should minimize the amount of ink on a page that does not display the data (Cairo, 2013). I chose to follow this principle for a couple reasons. First, I did not want to distract from the actual message of the data or confuse the reader by putting lots of information on one page. Second, extra design features that do not directly represent data are generally used to make the information intriguing and draw people's attention to the design. However, due to the sensitivity and complexity of this data, I did not want to create a design that would draw just anyone's attention because they may not take the time to properly understand and interpret the data. By keeping the design simple, and somewhat un-intriguing, it is more likely that the people who look at the maps will be those who are specifically interesting in understanding this data. However, I did not follow Tufte's principle to its full extent as I did still include some design features that enhance visual hierarchy and make the design appealing because I believe there is a balance between minimalism and creative design that the efficiency principle does not allow.

The final maps are intended to accompany an EDI Subscale Report, which has already been produced by HELP. This report includes lots of in-depth information about the subscales including what each subscale means and charts of the data across all waves. Therefore, I did not include details about the data on the map to avoid repetition and messiness from trying to fit too much on one page. I included an introduction page just to give a brief overview of how to read the data, but I do not give much explanation

of what the data means because that would be too complex to fit on one page and is already explained in the report.

Reflection on Working with a Community Partner

Working with a community partner was a positive experience. Often for the cartography and GIS labs and projects that we do in Geography, students are given the data and we have no idea where it comes from or what it means, and the maps that we produce are only for ourselves and are not meant to be used for any practical purpose. Working with a community partner was a completely different experience. It was really interesting to meet some of the people who work directly with the data I used because I was able to ask them questions and learn about where the data came from and why it is important. I also liked that Emilia and Jeremy gave me the choice to play with the data and decide for myself how what I wanted to focus on but also were willing to give suggestions and tell me what they were hoping for. Ultimately, I chose to map the subscale data because it was recently released, and Jeremy and Emilia said that they are looking for a way to map it widespread. Even though I know there is a good chance that my design for mapping the subscale data will not become the final format, I liked knowing that there was a chance that what I produce could actually be used in the real world. Moreover, if they do decide to use some part of my design for their final format, it would be really exciting to know that I contributed to that.

One of the most valuable things I learned from working with HELP, is that not all mapping these days is about GIS analysis. I really enjoyed that the process of mapping this data was about how to make it most understandable to the viewer and how to give

the most unbiased perspective of the data. When I was making my design choices, I tried to avoid designs that would stigmatize certain areas because this is sensitive data and could affect the public's perception of education systems in certain neighborhoods. For example, I opted not to use a green/red color scheme for the data because I did not want to make the associated of certain areas being "good" and other areas "bad" just because of how they scored on certain exams. This was a good learning experience for me because it forced me to question the implications and ethics of how I displayed the data, which is so much of what we learn about it cartography and GIS classes in Geography.

Jeremy and Emilia were great partners to work with. They were always willing to meet and answer questions and encouraged me to explore my own creativity. I feel very lucky to have worked with them. Additionally, I liked that I had the opportunity to learn about HELP because I think the work they do is really important in ensuring the success of children later in life, yet I would likely never had heard of them if not for this project. Furthermore, it was eye-opening to see what other people who studied Geography at UBC have gone on to do with their careers and has made me consider more seriously if this is something I would like to do in my future. It was reassuring to see first-hand that cartographers are still relevant and needed.

References

Cairo, A. (2013). *The functional art: An introduction to information graphics and visualization.* New Riders.

Fry, B. (2008). *Visualizing Data.* Sebastopol, CA: O'Reilly Media, Inc.

Human Early Learning Partnership. *The human early learning partnership.* Retrieved from: http://earlylearning.ubc.ca/about/

Human Early Learning Partnership. (October 2016). *Wave 6 community profile, Vancouver (SD 39).* Retrieved from: http://earlylearning.ubc.ca/media/edi_w6_communityprofiles/edi_w6_communityprofile_sd_39.pdf

Human Early Learning Partnership. (September 2017). *EDI subscales community profile, Vancouver (SD 39).* Retrieved from: http://earlylearning.ubc.ca/media/subscales_wave_6_sept_2017/w6_edi_subscale_community_profile_-_39_vancouver.pdf

Tufte, E. R. (2006). *Beautiful Evidence.* Graphics Press LLC.