Timothy Abramov

Data Mining

1st Draft

**“THE” Place to Live**

*\*length is approximate length in text only - no graphs/tables/figures assumed*

**Introduction(0.5 - 1 page): What did you/others do? Why did you do it?**

*In this study I want to explore some ways a person might approach the question of choosing where to live and how to evaluate it based on your preferences. And obviously, this time I’ll be the one who will dictate the preferences in this study =)*

*\*Not how I plan to write my introduction, but it seems like will be able to do a better job once I actually get to some results in my study.*

**Methods(1-2 pages): How did you do it?**

1. Describe the data:

*Don't think it will be a big section.*

*Will include the datasets that I used in*

*addition of any preprocessing I might*

*needed to do on them.*

1. Describe algorithms/processing of the data:

*Decision tree will be used to categorize the data into good/not good. After that, a clustering algorithm will be applied to each of the categories, most likely the k-means algorithm, to find 3 clusters of regions: less favorable, neutral options, and more favorable regions. Thus I will use k = 3 for the k-means.*

**Results(0.5 - 1 page): What did you find?**

*Haven’t found anything yet, but expect to have a binary-classified (either good or not good region for me) data set in the end, and having the data grouped into clusters to see what regions are more/less favorable in each category (from good/not good categories) to further refine my findings.*

*\*Most graphs/maps I expect to reside in this section.*

**Discussion(0.5 - 1 page): What does it all mean?**

*Well, we will see. Hopefully, I will be able to draw some conclusions about the best/worst and anything in the between regions for my future consideration in life. Not much in terms of contribution to the general society, except for maybe my methodology of getting to my conclusions.*

*\*Now, as I think about it, to make my paper less egocentric, I might make it more focused on the process of finding/measuring how different regions suit a person and how it can be analyzed, compared to my individual quest of finding ideal places for myself.*

**Sources(1 page)**

*Where I've got the datasets from and any supplement material that was used to derive with my work.*