Xin Tang

Oct/2024

Data Presentation and Visualization

DSC640-T301

**Term Project ML4**

**Write Up**

**Summary of Analysis:**

**Findings:**

1. In general, the home care and center care costs increase at a comparable pace. So, looking into one will represent the trend of the other.
2. In general, the center care costs are more expensive than family care costs.
3. In general, when income grows, the childcare costs also grow.
4. In some regions, the childcare costs grow much faster than income growth, which caused the affordability issue.
5. In City of Austin, there are abnormal drop of center based childcare cost in 2007.
6. By searching the news in 2017, there is a policy change related to childcare center.

**Assumptions:**

1. I assumein the region/area of higher income, the labor cost will also be higher, therefore, the childcare cost will also be higher. vice versa for lower income regions.
2. Without special cause, the income will grow year after year, the same as the childcare cost.
3. if the childcare cost is affordable, females will be more willing to go out for work, otherwise, they will be force to stay at home to avoid expensive center based childcare.
4. If there is a sudden change against the trend, there must be some special cause triggered the abnormal behavior
5. If the event happened at the same time and associated with same topic, they are correlated.

**Items need to be clarified**

1. If there are any other factors in 2017 which contributed to the sudden drop of Austin area center based childcare cost.
2. If there is another case of childcare cost reverse trend, which can be used to investigate the cause.

**Direction of Story/Message you want to get across**

The focus of the study is on childcare costs. Without an in-depth study of childcare cost increases and the contributing factors, it is hard to draw a story at national level, as well as compare things at state level. So, I would like to focus on a region I am familiar with and use data analysis to find some interesting facts to build up a story.

To build a story, I started with the national level average, confirming a common assumption that income and cost are both at growth. Then I used a commonly concerned topic: childcare cost affordability to make my story interesting and related to my audience. Next, I set up the stage of making changes, aka, election time.

Now the study starts to make sense to my audience; the interesting findings of reverse trend in 2007 at Austin area will draw audience attention, I can come to the point of call of action.

**Customer/Audience**: My audience is leaders from some nonprofit groups who are curious why this bond is needed. They can provide additional voice through their events. My audience also includes the general voters, who can vote yes to this bond.

**Medium**:

1. **PowerPoint slides**: My target audience are leaders from non-profit groups, who are not commercial or market people but are looking for ideas to support this bond. I will not show too much detail data tables or sort. Instead, I will primarily use charts with a line of summary to lay out my story. I am thinking of using a trend chart to compare income growth vs childcare cost growth, childcare cost growth between different states or counties etc. the type of chart could be bar chart, line chart, distribution chart etc.

2. **Power BI Dashboard**: I will create a Power BI dashboard of KPI cards to show some information. I plan to ask a web designer to put it on the election website, which will be used to promote this bond proposal. People can have a summary right from the dashboard.

3. **Poster**: To get voters’ support, I will design a poster with a clear call for action. The poster will be put outside of the vote station. This will bring voters’ attention to the bond

proposal: therefore, voters will potentially support the bond.

**Design decision**:

For PowerPoint, I want to make charts as simple as possible. I also place a sentence at the most visual place of the page to summarize the chart. I used font change and color change to draw audience attention.

For dashboard, to prepare for broader interests in the childcare topics, I also list the female labor participate rate to prove my assumption that female will work if the center care is affordable.

For the poster, instead of using many words, I introduced a picture of mother and baby and a chart, hopefully draw voters’ attention and get their buy in on the actions.

**Ethical Consideration**:

* I checked the data, which is clean, so I did not need to remove/adjust anything. I used Excel to do some grouping since power BI is not strong in data transformation.
* The data did not cover the years when Austin grows most, so I googled out the post 2018 data to show the insane increase of the childcare cost.
* My analysis will be used for election champion, so I do see a huge legal or ethical risk/consequence. Fortunately, this is school practice work so I can place it in a simplified situation and do not worry about the risk. However, it is still a good idea to call out the source of the data to avoid any liability issue.
* I trusted the data since the data resource is from a government website.
* Before the analysis is posted to social media, I will add discriminator to protect myself.

**Lesson Learnt**:

* Even this is practice and will not be used for real. I treated this as a real- life task and enjoyed linking the data analysis result to a real-life scenario and a real-life call of action.
* On the other hand, if it is for real, I will be facing a lot of questions and challenges from opponents. My analysis will need to be much more in depth and broader to build a robust case. I also need to seek review and approval from the subject experts before publishing.