Childcare Costs

Joshua Ewer

2025/08/05

<h1 style="text-align:center">Summary of Analysis</h1>

Findings

      The dataset contains child care costs across the United States, by county, for the years 2008 to 2018. Child care costs are primarily reported in the MFCC[Infant|Todder|Preschool] column, which represents the median weekly cost for child care in those respective categories of children. Findings include:

- Infant child care is the most expensive category

- Costs have consistently risen in all categories throughout the entire dataset. As an example, the average cost of infant care rose from $105 to $125 a week from 2008-2018, which is a nearly 20% increase.

- There is a very wide degree of cost difference between the lower and higher states. As an example, the highest state (California) has infant care costs over 300% higher than the lowest (Mississippi).

Assumptions

- The weekly cost is per child and does not take into account any federal, state, or local taxes.

- I could not find any explanation of what the columns that have a "flag" prefix mean, so I assumed they were not useful for this exploration.

Items that still need clarification

- There were some states that did not have as much data available (e.g. South Carolina and Alaska only had data starting in 2014). Is the process for collecting data in each area the same, and can we trust the values are correct?

<h2 style="text-align:center">Desired Message</h2>

In general, the cost of child care is increasing, for all types of child care. Infant care cost is the highest, but all costs are increasing. Some areas have a disproportionate rise in child care costs, and families with infants will share an increasingly large portion of the cost. If these prices are per child, large families will find child care to be a burdensome expense. This is not necessarily a problem that individual families can fix, it's up to policy makers and politicians to advocate for change to reduce this burden on growing families.

Target Audiences

- A non-profit group that can provide care in specific areas to combat the rising cost of child care.

- An elected official or a policy maker that can influence government subsidies or programs

- Parents or families doing research on where to start families

Mediums

- A narrative article or blog post that explains the problem in simple terms, and can show how every day families are impacted. The data story for this medium focused on using income vs cost to explain areas that have a disproportionately higher cost of childcare burden. The call to action includes looking at states that have a lower burden for care.

- An infographic, one-page overview of areas of higher- or lower- cost, which could be used to quickly highlight outliers in terms of child care cost. The data story for this medium focuses on giving readers suggestions on alternate places to move if childcare burden is a driving factor for relocation.

- A formal presentation, perhaps to a CEO or a public official. Should include concrete references to data. For this medium, I roleplayed that the presentation is being done by a

lobbyist trying to convince a state senator in a state with a high cost of living to invest in publicly-funded childcare.

Design Decisions

For the first two mediums, I focused on simple, consistent visualizations that helped to support the data story. Since the target audiences for those groups did not likely have high data literacy, I focused on the narrative and let the charts/graphs support the story. For the last medium, the presentation focused on more granular charts and included figures and numbers that would appeal to someone with higher data literacy and is focused on specific values instead of general trends. I intended to make the last medium interactive (e.g. a Tableau dashboard) but ran out of time and switched to a power point presentation because that did fit the target audience better.

Lessons Learned

One lesson that I learned is that it's important not to enter into analysis with a bias or assumption. I had assumed that large states on the coast (e.g. California and New York) would have higher costs, that the cost would even out because of higher wages. It turns out that, when income is factored in, those states have a significantly higher burden on childcare cost. I assumed the opposite of southern states (expected lower costs, but lower wages mean similar burden) but it turned out that southern states *do* have a lower burden of healthcare costs.

I definitely enjoyed building the visualizations in Python/Seaborn the most. The combination of data wrangling and exploration tools along with simple and flexible visualizations made it possible to quickly experiment with different visualizations to see what helped explain the story best. I also learned that I do *not* enjoy working with PowerBI or Tableau. I tried using both as an opportunity to learn something new and spent an unfortunate

amount of time doing anything meaningful. I wanted to do an interactive dashboard with either

but time didn't allow me to complete that with either tool.