U.S. Cost of Living Project Proposal

Group 2: Devin Fogerty, Demiko Birden, Spencer Thompson

We plan on using U.S cost of living data by county to analyze key factors such as housing, food, healthcare, transportation, childcare, as well as tax rates, to determine a comprehensive picture of cost of living in different counties within the state of Ohio, and if possible, throughout different regions of the U.S.

Questions to answer:

* Do metro areas or rural areas have higher cost of living?
* Which counties in each state have the highest/lowest cost of living?
* Do families with children have significantly higher costs of living, and how significant are childcare costs compared to total income?
* What percentage of total family income goes towards housing in each county?
* What percentage of total family income goes towards food, and do metro or rural counties spend more?
* Do counties outside of metro areas have higher transportation costs?
* Are healthcare costs higher for families with children?
* Do metro areas spend more or less than rural areas on healthcare?
* How significant of an effect do taxes have on cost of living, and what percentage of total income goes straight to taxes?
* What regions have the highest/lowest cost of living?

Datasets:

<https://www.kaggle.com/datasets/asaniczka/us-cost-of-living-dataset-3171-counties/>

https://www.geoapify.com/

Rough breakdown of tasks:

* Import CSV data and group by state and by county
* Clean dataset and fix formatting of columns (# of family members/children)
* Analyze effects of metro/rural and # of family members
* Compare costs for each county
* Create bar plots for # of family members versus each cost of living category
* Create cost of living plots on a state-by-state basis
* Map total cost of living for each county onto state/U.S. map using Geoapify