Measuring Quality of Life Across U.S. States Using Salary, Housing, and Cost of Living Data

Team: Jun Kim, Eddy Kim

1. Dataset Description

We will evaluate the *Quality of Life (QoL)* across four U.S. states—**Utah, Texas, California, and New York**—using three datasets publicly available:

(1) U.S. Census Bureau (American Community Survey, ACS): Median Household Income (B19013_001E), Median Gross Rent (B25064_001E), Median Home Value (B25077_001E), Rent Burden (≥30% income spent on rent; B25070 group), Owner Cost Burden (≥30% income spent on housing; B25091 group)

(2) Bureau of Labor Statistics (BLS): Regional CPI (BLS API) for real purchasing power adjustment

(3) Tax Foundation: State/Local Tax Burden (Tax Foundation) for disposable income adjustments

* (1) & (2): API, (3): Excel files

2. Python Package: qol_analyzer

We will design a modular Python package enabling automated QoL analysis.

- **Data Retrieval & Cleaning:** Query Census and BLS APIs for selected states, Normalize and merge datasets into a unified DataFrame
- **Feature Engineering**
  - Compute Real Purchasing Power (RPP) = Income ÷ CPI
  - Compute Housing Burden Index = Housing Cost ÷ Income
  - Create a composite QoL Score using standardized z-scores
- **Visualize** income vs. rent/home value differences

Package structure:

(1) **data_fetch.py:** fetch_census_data(), fetch_bls_cpi(), fetch_tax_data();
(2) **data_clean.py:** standardize_state_names(), merge_datasets(), handle_missing_values()
(3) **feature_eng.py:** calculate_real_purchasing_power(), calculate_housing_burden_index(), compute_composite_qol_score()
(4) **analysis.py:** run_regression_analysis(), generate_summary_statistics()
(5) **visualize.py:** plot_income_vs_housing(), create_comparative_dashboard()

Workflow:
Automated API retrieval → merge via FIPS codes → calculate Real Purchasing Power & Housing Burden Index → regression to determine if housing or CPI contributes more to QoL → visualizations comparing states

3. Challenges and Mitigation

Challenge	Mitigation
Missing/Inconsistent data	Use FIPS codes; standardize to same year; document missing values
CPI not in Census	Integrate BLS CPI; create conversion factors
Different units/scales	Automated z-score normalization; unit conversion utilities
API rate limits	Response caching; retry logic; register for API keys
5-week timeline	Limit to 4 states; weekly milestones; prioritize core functions

4. Team Roles

- **Eddy Kim:** API extraction, feature engineering, unit testing, error handling
- **Jun Kim:** Statistical analysis, visualization, regression modeling, documentation
- **Both:** Integration testing, package structure, presentation

5. Evaluation Metrics

- **Technical:** > 90% test coverage; API calls succeed for all states; R^2 > 0.7 for regression; cross-machine compatibility
- **Quality:** QoL scores show logical relationships; clear visualizations; reproducible calculations

6. Replicability

- **Environment:** requirements.txt, setup.py, pytest
- **Data:** .env (gitignored), sample data for offline testing
- **Version Control:** GitHub with clear commits

7. Ethical Considerations

- **Privacy:** Only aggregated, anonymized state-level data; no PII; compiles with agency terms
- **Interpretation:** Avoid ranking states as "best/worst"; acknowledge QoL is multidimensional; state averages mask within-state variation
- **Transparency:** Publish code/methodology on Github; document assumptions; enable verification
- **Bias Awareness:** Census may undercount transient populations; CPI reflects urban consumers; clearly state limitations

8. Interest to Classmates

- **Relatable:** Post-graduation location decisions; concrete financial planning (e.g., '$80K in CA vs. TX').
- **Transferable skills:** API automation, data merging, regression analysis, package development, visualization
- **Extensible:** Framework allows extension to other states, international comparisons, or other domains (healthcare, education).

9. References

- U.S. Census Bureau (2024). ACS 5-Year Estimates 2023. census.gov/programs-surveys/acs/data.html
- Bureau of Labor Statistics (2025). CPI Data. bls.gov/cpi/
- Tax Foundation (2024). State/Local Tax Burden Rankings 2022. taxfoundation.org/data/all/state/tax-burden-by-state-2022/