Project Part 1

Introduction

The quality of life across different states in the US is a topic that attracts policymakers, researchers, economists, and the public. It impacts societal well-being and reflects how various factors contribute to the functioning of society. We are particularly interested in understanding how a combination of quantitative and qualitative factors shapes human experiences.

In a post-pandemic world where public health, economic recovery, and social equity are increasingly valued, studying how life expectancy, expenditures, and air quality impact quality of life is essential for shaping future policies (Woolf et al., 2018). By focusing on these variables, this study addresses disparities in state-level well-being and contributes to ongoing social and policy debates.

The rationale for selecting life expectancy, expenditures, and air quality as variables of interest is from their direct connection to well-being. Life expectancy serves as an indicator of health and access to resources (Saito et al., 2014). Economic expenditures reflect the financial stability, which can afford citizens better living standards (Chetty et al., 2016). Air quality is another crucial factor, as environmental conditions linked to health outcomes (Dockery et al., 1993). By investigating these variables, we aim to determine which factors have the most significant impact on quality of life.

This study focuses on three key research questions to explore these relationships: - Do states with higher average life expectancy exhibit a higher quality of life? - Do states with higher average annual expenditures per capita tend to have a better quality of life? - Do states with higher air quality indices correlate with higher quality of life?

Data Summary

This analysis utilizes multiple datasets from reputable sources to explore the relationship between various socio-economic and environmental factors and the quality of life across U.S. states. The air quality index data is collected by the US Environmental Protection Agency

(EPA) in 2023 at outdoor monitors. Air quality data is important as it gives people an idea of how safe and healthy it is to access the outdoor environment. The personal consumption expenditure (PCE) data is collected by the Bureau of Economic Analysis (BEA) in 2023. This is an annual governmental data set to release each year. The life expectancy data is released by the Center for Disease Control (CDC) and collected by National Center for Health Statistics. This data is an average of US state life tables from 2018 – 2021. Each year, the US & world news will release its "best states to live in" report, which ranks the quality of life each state provides to its residents. This index is affected by multiple factors, including healthcare, education, infrastructure, fiscal stability, crime and corrections, and natural environments.

Our sample includes observations for each U.S. state. We didn't make any data modifications. However, there are potential issues to consider. First, some datasets don't exactly fall into the same collection period (e.g., life expectancy is an average from 2018–2021, while other data points are specific to 2023). Nevertheless, given the relatively stable nature of the variables, we don't expect such minor discrepancies to produce any significant disparities. Additionally, the QOL index data are composites derived from multiple metrics. Thus there might be subjectivity in how these indices are ranked.

Exploratory Analysis

QOL Data Exploration

```
summary_data <- lifedata %>%
  summarise(`Mean QOL` = round(mean(QOL, na.rm = TRUE), 2),
    `SD QOL` = round(sd(QOL, na.rm = TRUE), 2))
print(summary_data)
```

```
##   Mean QOL SD QOL
## 1    51.97   5.75
```

Correlation Analysis between QOL and key variables

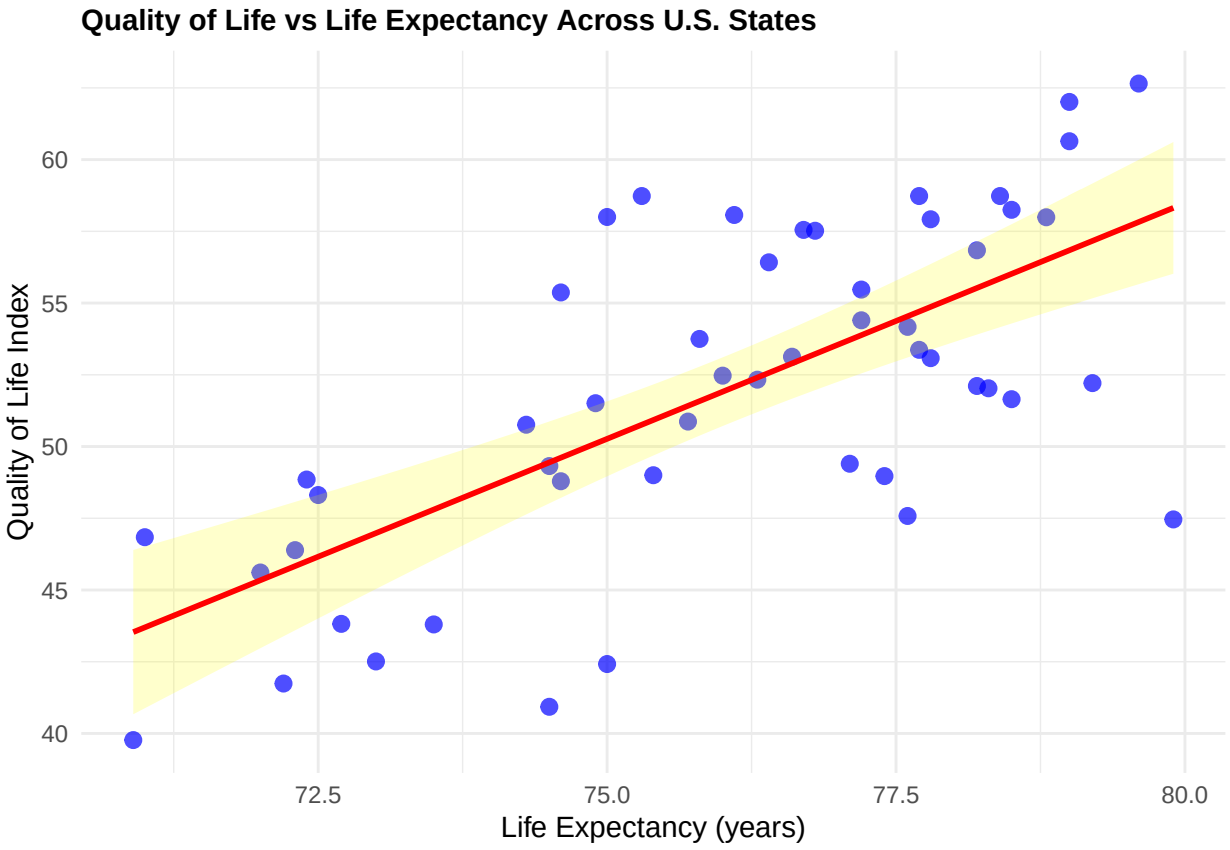
```
correlations <- lifedata %>%
  summarise(
    `QOL & Life Expectancy` = round(cor(QOL, LE, use = "complete.obs"), 3),
    `QOL & AQI` = round(cor(QOL, AQI, use = "complete.obs"), 3),
    `QOL & Expenditure` = round(cor(QOL, Expenditure, use = "complete.obs"), 3)
  )
print(correlations)
```

```
##   QOL & Life Expectancy QOL & AQI QOL & Expenditure
## 1                 0.681     -0.08             0.168
```

Scatterplot: QOL vs. Life Expectency

```
ggplot(lifedata, aes(x = LE, y = QOL)) +
  geom_point(color = "blue", size = 2.5, alpha = 0.7) +
  geom_smooth(method = "lm", se = TRUE, color = "red", fill = "yellow", alpha = 0.2) +
  labs(title = "Quality of Life vs Life Expectancy Across U.S. States",
       x = "Life Expectancy (years)",
       y = "Quality of Life Index") +
  theme_minimal() +
  theme(plot.title = element_text(size = 11, face = "bold"))
```

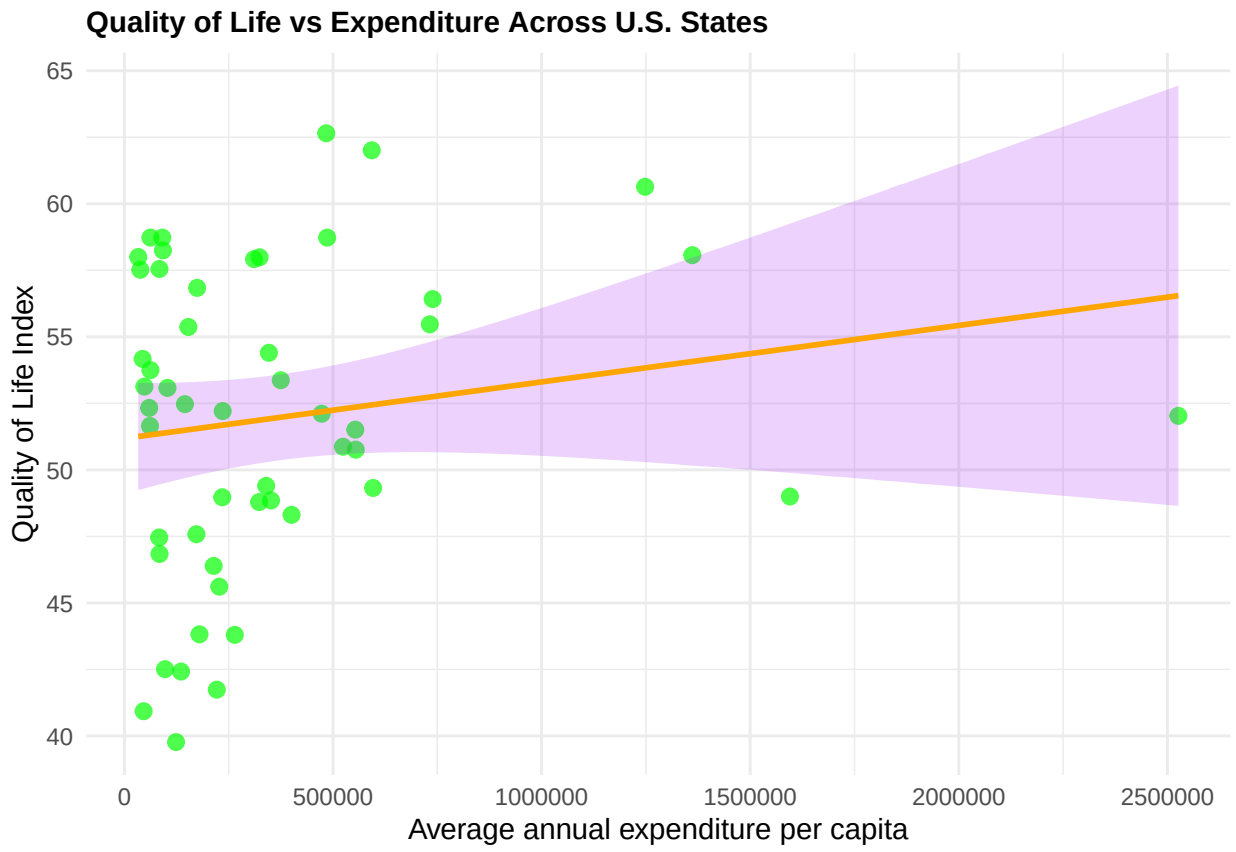
```
## `geom_smooth()` using formula = 'y ~ x'
```



Scatterplot: QOL vs Expenditure

```
ggplot(lifedata, aes(x = Expenditure, y = QOL)) +
  geom_point(color = "green", size = 2.5, alpha = 0.7) +
  geom_smooth(method = "lm", se = TRUE, color = "orange", fill = "purple", alpha = 0.2)
  labs(title = "Quality of Life vs Expenditure Across U.S. States",
       x = "Average annual expenditure per capita",
       y = "Quality of Life Index") +
  theme_minimal() +
  theme(plot.title = element_text(size = 11, face = "bold"))
```

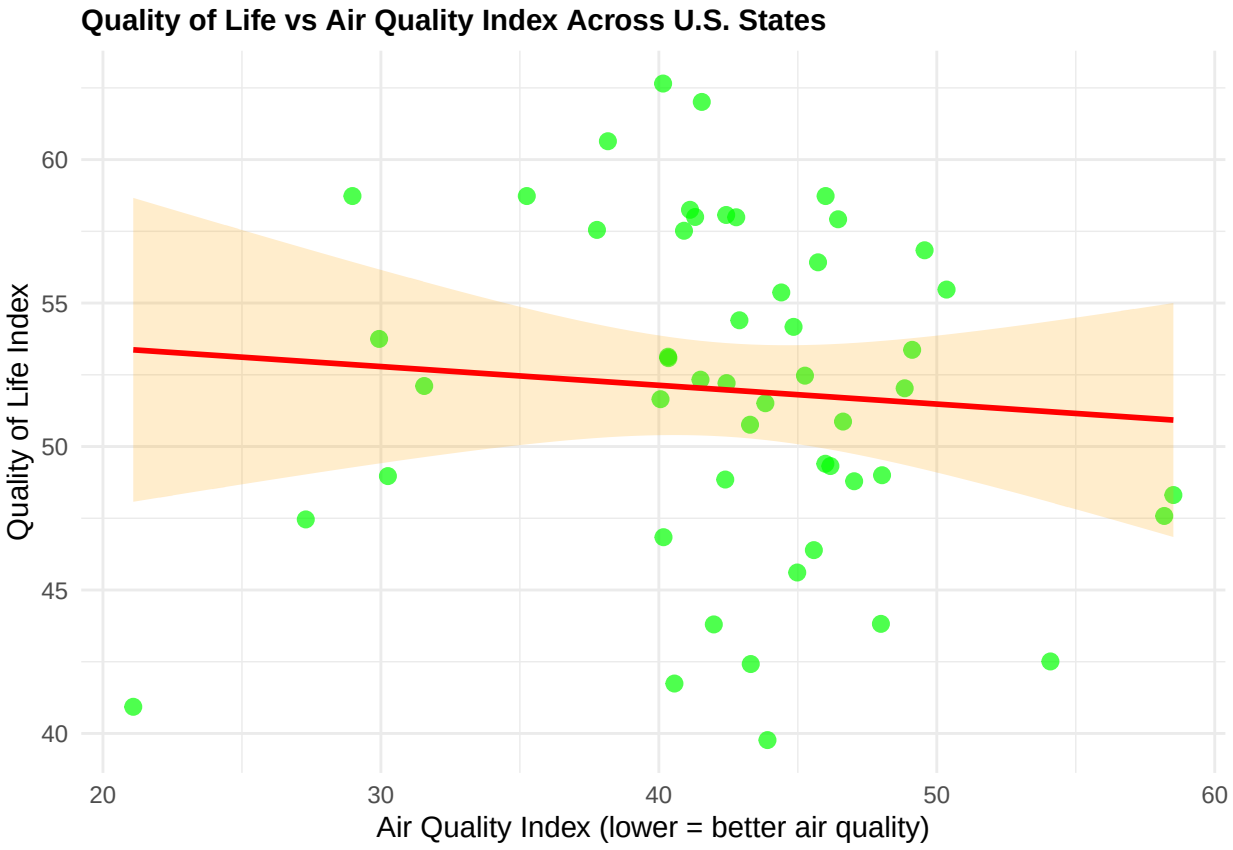
```
## `geom_smooth()` using formula = 'y ~ x'
```



Scatterplot: QOL vs AQI

```
ggplot(lifedata, aes(x = AQI, y = QOL)) +
  geom_point(color = "green", size = 2.5, alpha = 0.7) +
  geom_smooth(method = "lm", se = TRUE, color = "red", fill = "orange", alpha = 0.2) +
  labs(title = "Quality of Life vs Air Quality Index Across U.S. States",
       x = "Air Quality Index (lower = better air quality)",
       y = "Quality of Life Index") +
  theme_minimal() +
  theme(plot.title = element_text(size = 11, face = "bold"))
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Conclusions

The data exploration reveals compelling patterns regarding the factors associated with quality of life across U.S. states. Several key findings emerge from our analysis.

The average quality of life index across all states is 51.97 with a standard deviation of 5.75, suggesting moderate variation in quality of life measures.

Life expectancy shows a strong positive correlation with quality of life (0.681), suggesting states with longer-lived populations tend to have higher quality of life scores. The correlation between QOL and Air Quality Index is near zero (-0.08), indicating little linear relationship between these variables as measured. QOL shows a low positive correlation with Expenditure (0.168), suggesting that the amount of money spent for an individual does not have a strong correlation with their quallity of life.

The scatterplot between QOL and Life Expectency illustrates the strong positive relationship between them. States where residents live longer tend to have substantially higher QOL scores. The trend line shows a clear upward pattern. The tight clustering around the trend line suggests this relationship is consistent across most states. The scatterplot between QOL

and Expenditure illustrates the weak positive relationship between expenditure and quality of life. States with higher expenditure do not have significantly higher QOL compared to the sates with lower expenditure. In the scatterplot between QOL and AQI, the trend line is nearly flat, indicating minimal linear relationship between these variables. States with low AQI and those with high AQI show similar ranges of QOL scores.

Our analysis provides foundational insights that can inform targeted policy recommendations aimed at improving quality of life by addressing critical factors such as health, environmental factors, and equitable access to resources.

References

Background

1. Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The Association between Income and Life Expectancy in the United States, 2001-2014. JAMA, 315(16), 1750. https://doi.org/10.1001/jama.2016.4226
2. Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., & Speizer, F. E. (1993). An Association between Air Pollution and Mortality in Six U.S. Cities. New England Journal of Medicine, 329(24), 1753–1759. https://doi.org/10.1056/nejm199312093292401
3. Saito, Y., Robine, J., & Crimmins, E. M. (2014). The methods and materials of health expectancy. HAL (Le Centre Pour La Communication Scientifique Directe), 30(3), 209–223. https://doi.org/10.3233/sji-140840
4. Woolf, S. H., Chapman, D. A., Buchanich, J. M., Bobby, K. J., Zimmerman, E. B., & Blackburn, S. M. (2018). Changes in midlife death rates across racial and ethnic groups in the United States: systematic analysis of vital statistics. BMJ, k3096. https://doi.org/10.1136/bmj.k3096

Data

1. USAFacts. (2023, April 20). Which states have the best and Worst Air Quality? https://usafacts.org/articles/which-states-have-the-best-and-worst-air-quality/
2. News release. Personal Consumption Expenditures by State, 2023 | U.S. Bureau of Economic Analysis (BEA). (n.d.). https://www.bea.gov/news/2024/personal-consumption-expenditures-state-2023

3. Centers for Disease Control and Prevention. (2024, August 21). Life expectancy at birth by State. Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/pressroom/sosmap/life_expectancy/life_expectancy.htm
4. Quality of life by state 2024. (n.d.). https://worldpopulationreview.com/state-rankings/quality-of-life-by-state