ITC6000

24 NOV2024

MODULE 4: ASSIGNMENT (TEAM DIGITAL PRESENTATION)

Co-Authored:

Nicholas Hooper

Yash Amit Singh

Amit Patel

Trusha Pradeep Sonawane

Ganya Prakash Reddy

Agenda

1. Team Introduction and Contribution Description
2. Research Focus and KPIs
3. Data Organization and Tools
4. KPI Analysis and Research Findings
5. Conclusion
6. Bibliography

Team and Contribution

- Nicholas Hooper
  - Topic and KPI Development; Database structure; KPI 1 Analysis; Conclusion
- Yash Amit Singh
  - KPI 2 Analysis and summary
- Trusha Pradeep Sonawane
  - KPI 3 Analysis and Future implications
- Ganya Prakash Reddy
  - KPI 4 Analysis
- Amit Patel

  KPI 5 Analysis and summary

Research Focus and Desired Outcomes

TOPIC

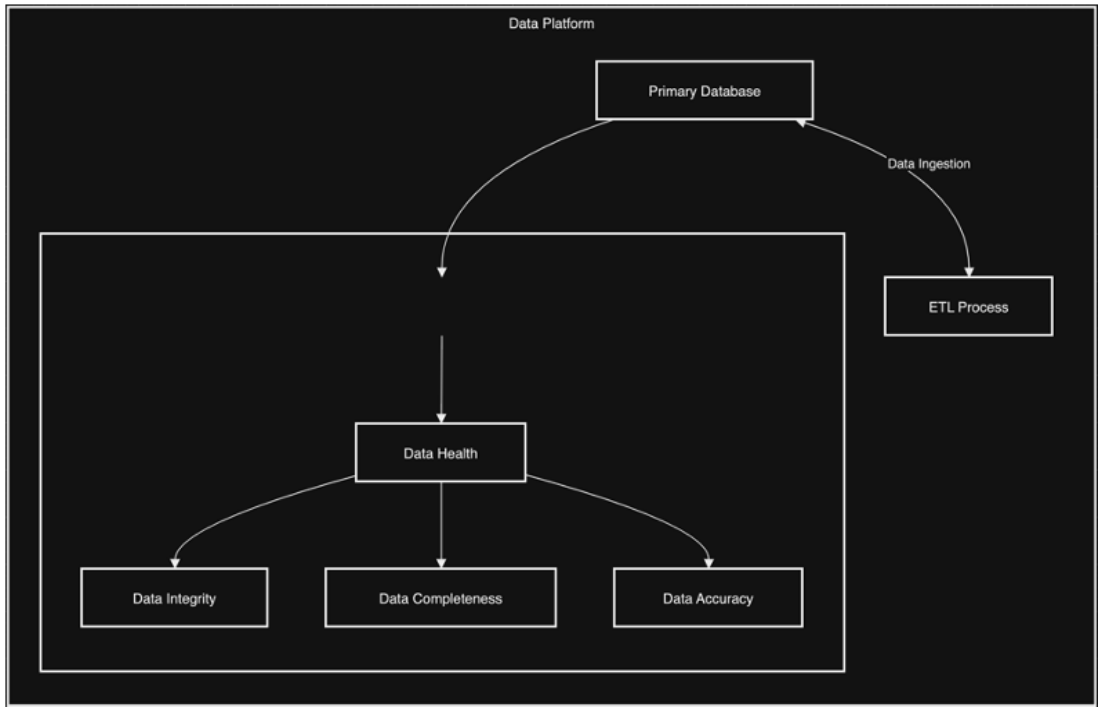
This project analyzes health-related data from the U.S. Department of Health & Human Services' *2023 Local Data for Better Health* report, encompassing state-level statistics from 2021-2022. The dataset includes 280,000 rows covering topics such as chronic conditions, preventive care, mental health, and social vulnerabilities, with values represented as percentages of prevalence. The goal is to develop key performance indicators (KPIs) to assess healthcare access, chronic disease trends, social determinants of health, and preventive screening uptake while utilizing database techniques such as data aggregation, anomaly detection, ER modeling, and visualization.

KPI'S

1. Social Vulnerability Index: Transportation, Housing, and Food Security
2. Percentage of Adults with No Access to Preventive Healthcare
3. State Trends in Obesity and Related Chronic Conditions
4. Mental Health and Social Isolation Trends
5. Preventive Screening Uptake: Mammography and Colorectal Cancer

Data Organization and Analytic Process

DATA-BASE STRUCTURE



METHODS AND TOOLS

- DB Browser
  - Cleaning and Importing
  - Primary Keys
  - Aggregation
  - Grouping
- R
  - Visualizations
- Draw.IO
  - ERD Visualization

KPI – 1

Social Vulnerability Index: Transportation, Housing, and Food Security

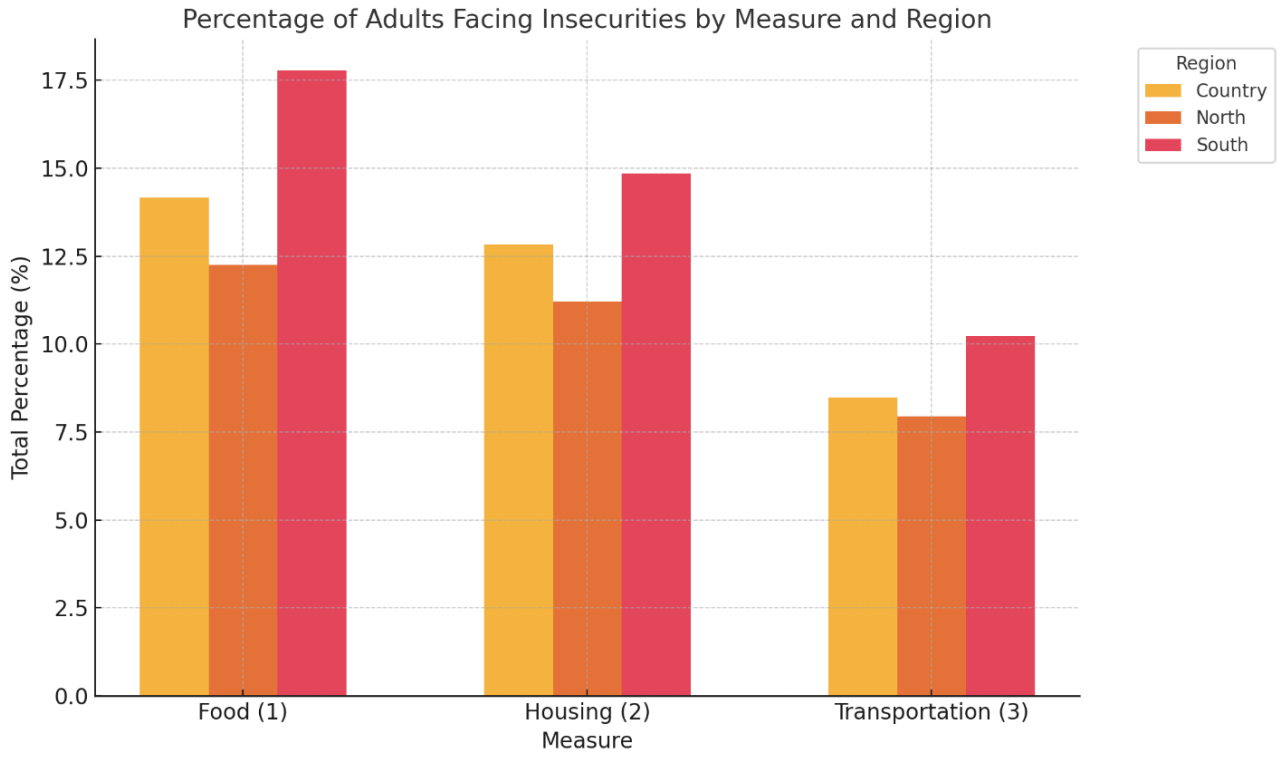
Purpose: Assess the prevalence of social factors that contribute to health inequalities.

Fields Needed:

- Lack of reliable transportation in the past 12 months among adults
- Housing insecurity in the past 12 months among adults
- Food insecurity in the past 12 months among adults

Components of Analysis:

- Aggregate data to calculate an index score by state and region.
- Use anomalies to identify areas for intervention.
- Determine if there is geographic significance.



KPI 1: Database Cleanup, Filtering, and Aggregation

```sql
CREATE TABLE social_vulnerability AS
-- Select relevant columns
WITH formatted_health_data AS (
  SELECT
    Year,
    StateDesc,
    Measure,
    Data_Value,
    TotalPop18plus
  FROM health_data
)

-- Isolate transportation data
, transportation AS (
  SELECT *
  FROM formatted_health_data
  WHERE measure = 'Lack of reliable transportation in the past 12 months among adults'
  AND Year = 2022
)

-- Isolate housing data
, housing AS (
  SELECT *
  FROM formatted_health_data
  WHERE measure = 'Housing insecurity in the past 12 months among adults'
  AND Year = 2022
)

-- Isolate food data
, food AS (
  SELECT *
  FROM formatted_health_data
  WHERE measure = 'Food insecurity in the past 12 months among adults'
  AND Year = 2022
)

-- Output the final result
SELECT * FROM transportation
UNION ALL
SELECT * FROM housing
UNION ALL
SELECT * FROM food;
```

1

The initial data set was cleaned and parsed down to the relevant components for measurement

2

The components were then grouped and aggregated

```sql
CREATE TABLE state_and_measure AS
SELECT
  StateDesc,
  Measure,
  avg(data_value) AS total_data_value,
  avg(TotalPop18plus) AS total_pop
FROM
  social_vulnerability
GROUP BY
  StateDesc, Measure
ORDER BY
  StateDesc, Measure;
```

3

An additional database was compounded to capture the desired characteristics

```sql
-- add a region column
ALTER TABLE state_and_measure
ADD region VARCHAR(10) ;

-- update the `region` column for states in the `north` column of North_South
UPDATE state_and_measure
SET region = 'North'
WHERE StateDesc IN (SELECT north FROM North_South WHERE north IS NOT NULL);

-- update the `region` column for states in the `south` column of North_South
UPDATE state_and_measure
SET region = 'South'
WHERE StateDesc IN (SELECT south FROM North_South WHERE south IS NOT NULL);
```

4

The compounded table was aggregated to get the final comparison for analysis

```sql
CREATE TABLE region_measure AS
SELECT
  region,
  Measure,
  avg(total_data_value) AS total_perc
FROM
  state_and_measure
GROUP BY
  region, Measure
ORDER BY
  region, Measure;
```

KPI – 1 Analysis

Social Vulnerability Index: Assess the prevalence of social factors that contribute to health inequalities.

12% of the unhealthy population report being affected by a lack of housing availability. Making it a meaning contributor.

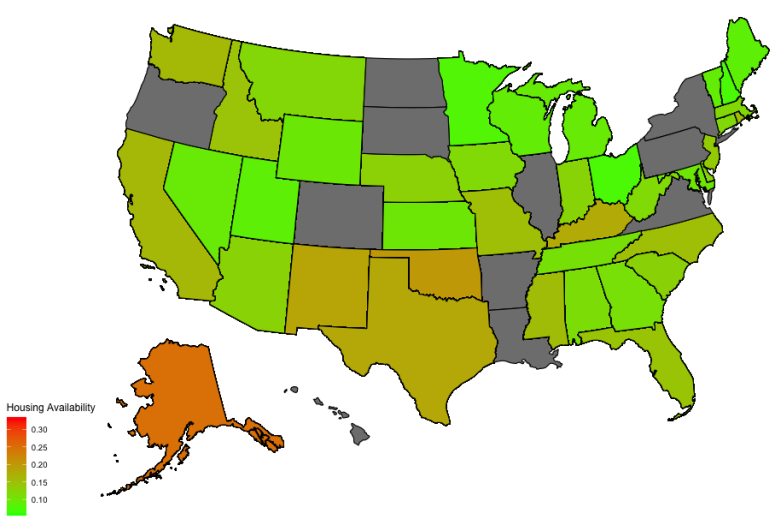
1% of the unhealthy population report being affected by a lack of transportation, indicating it has little effect on the measure.

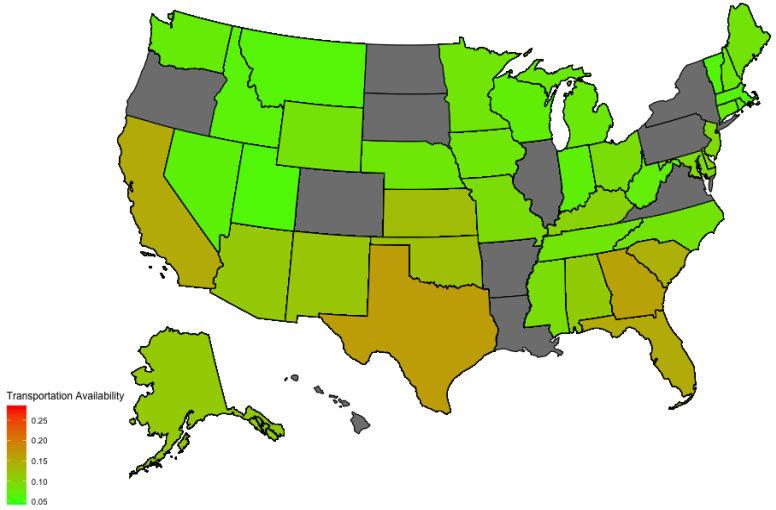
15% of the unhealthy population report being affected by a lack of food availability. Making it a meaning contributor.

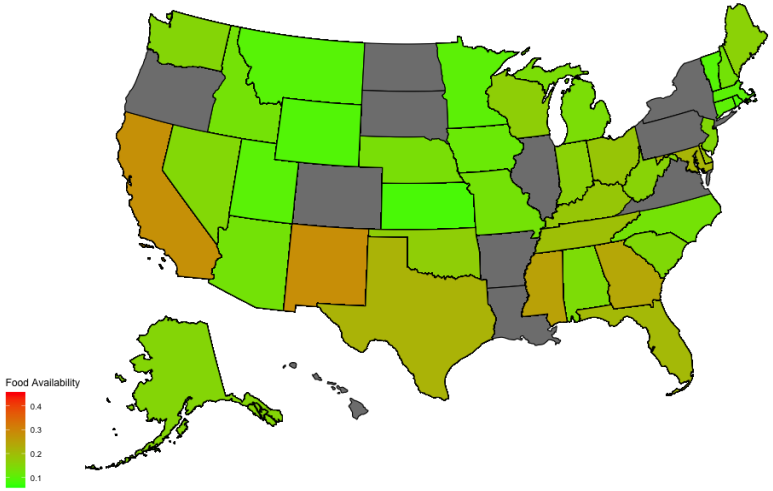
Housing

Transportation

Food







Social vulnerabilities effect 7% of the total population and account for 4% of the 'unhealthy' population. The South is disproportionately effected by social factors by 38%

KPI – 2

Percentage of Adults with No Access to Preventive Healthcare

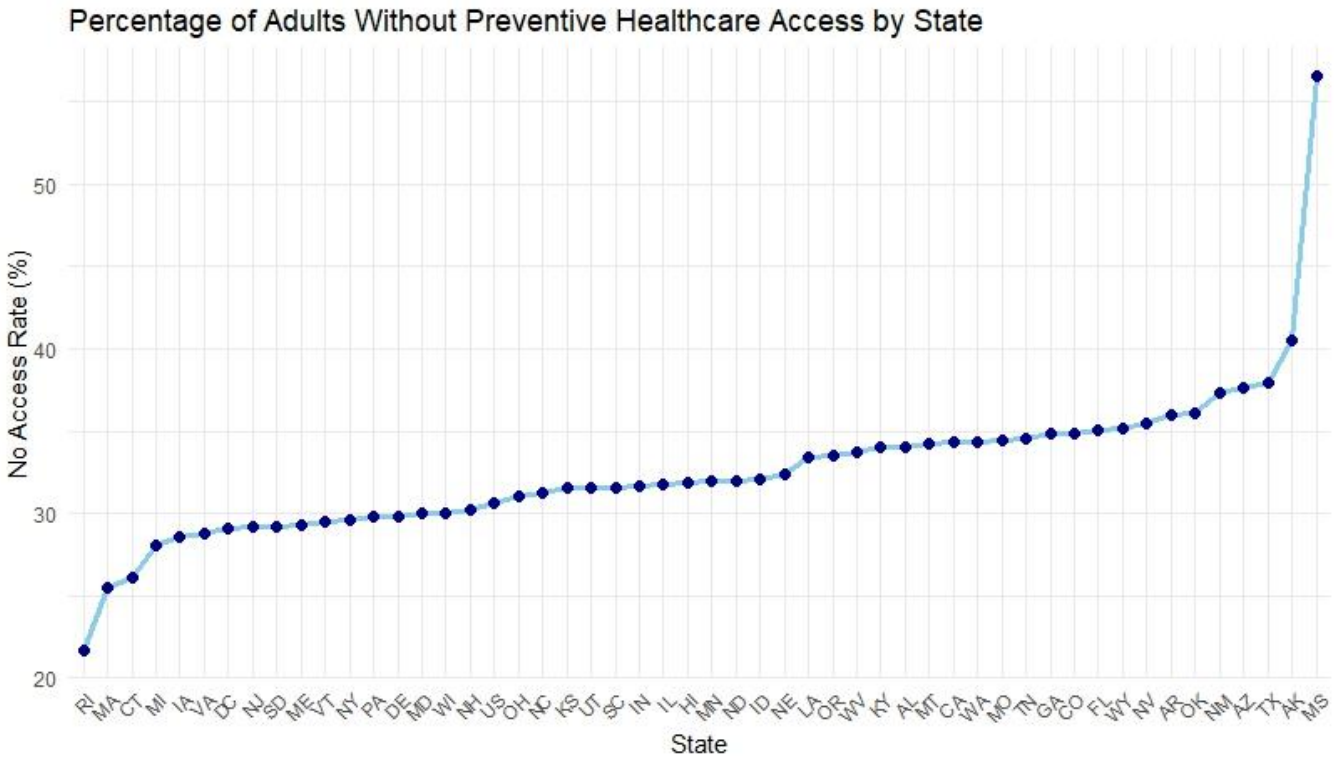
Purpose: Measure disparities in preventive healthcare access, such as routine checkups or dental visits.

Fields Needed:

- Visits to doctor for routine checkup within the past year among adults
- Visited dentist or dental clinic in the past year among adults

Components of Analysis:

- Aggregate by state and year to compare across regions and time.
- Chart trends or anomalies in access.
- Highlight states or years with notably low percentages.



KPI 2: Database Cleanup, Filtering, and Aggregation

```r
usa <- read_csv("PLACES__Local_Data_for_Better_Health__County_Data_2024_release.csv")

kpi2_measures <- c("Visits to doctor for routine checkup within the past year among adults",
                   "Visited dentist or dental clinic in the past year among adults")

kpi2 <- usa %>%
  filter(Measure %in% kpi2_measures) %>%
  group_by(StateAbbr,StateDesc, Measure) %>%
  summarise(Avg_Value = mean(Data_Value, na.rm = TRUE), .groups = "drop") %>%
  pivot_wider(names_from = Measure, values_from = Avg_Value) %>%
  mutate(
    No_Access_Rate = 100 - rowMeans(across(where(is.numeric)), na.rm = TRUE)
  )

ggplot(kpi2, aes(x = reorder(StateAbbr, No_Access_Rate), y = No_Access_Rate, group = 1)) +
  geom_line(color = "skyblue", size = 1.2) +
  geom_point(color = "darkblue", size = 2) +
  labs(
    title = "Percentage of Adults Without Preventive Healthcare Access by State",
    x = "State",
    y = "No Access Rate (%)"
  ) +
  theme_minimal()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size = 8))
```

1

The code reads data, filters for specific measures, calculates average values, pivots the data, computes the no access rate, and creates a plot.

```r
routine_disparities <- kpi2 %>%
  arrange(`Visits to doctor for routine checkup within the past year among adults`) %>%
  slice_head(n = 5)

ggplot(routine_disparities, aes(x = reorder(StateDesc, `Visits to doctor for routine che
                                y = `Visits to doctor for routine checkup within the pas
  geom_col(fill = "orange", color = "black", width = 0.7) +
  labs(
    title = "States with the Lowest Routine Checkup Rates",
    x = "State",
    y = "Routine Checkup Rate (%)"
  ) +
  theme_minimal()
```

2

Identifies and orders states with lowest routine checkup rates, and creates a plot.

```r
dental_disparities <- kpi2 %>%
  arrange(`Visited dentist or dental clinic in the past year among adults`) %>%
  slice_head(n = 5)

ggplot(dental_disparities, aes(x = reorder(StateDesc, `Visited dentist or dental
                               y = `Visited dentist or dental clinic in the past
  geom_col(fill = "orange", color = "black", width = 0.7) +
  labs(
    title = "States with the Lowest Dental Visit Rates",
    x = "State",
    y = "Dental Visit Rate (%)"
  ) +
  theme_minimal()
```

3

Identifies and orders states with lowest dental visit rates, and creates a plot.

```r
kpi2_usheatmap <- kpi2 %>%
  select(
    state = StateAbbr,
    avg =  No_Access_Rate,
    routine = `Visits to doctor for routine checkup within the past year among adults`,
    dental = `Visited dentist or dental clinic in the past year among adults`
  )

plot_usmap(data = kpi2_usheatmap, values = "avg", regions = "states") +
  scale_fill_gradient(low = "lemonchiffon", high = "navy", name = "No Access Rate (%)") +
  labs(
    title = "Heatmap of Preventive Healthcare Access (KPI 2)",
    subtitle = "Percentage of Adults Without Preventive Healthcare Access"
  ) +
  theme_minimal()

plot_usmap(data = kpi2_usheatmap, values = "routine", regions = "states") +
  scale_fill_gradient(low = "lemonchiffon", high = "navy", name = "No Access Rate (%)") +
  labs(
    title = "Heatmap of Preventive Healthcare Access (KPI 2)",
    subtitle = "Percentage of States with the Lowest Routine Checkup Rates among Adults"
  ) +
  theme_minimal()

plot_usmap(data = kpi2_usheatmap, values = "dental", regions = "states") +
  scale_fill_gradient(low = "lemonchiffon", high = "navy", name = "No Access Rate (%)") +
  labs(
    title = "Heatmap of Preventive Healthcare Access (KPI 2)",
    subtitle = "Percentage of States with the Lowest Dental Visit Rates among Adu
  ) +
  theme_minimal()
```

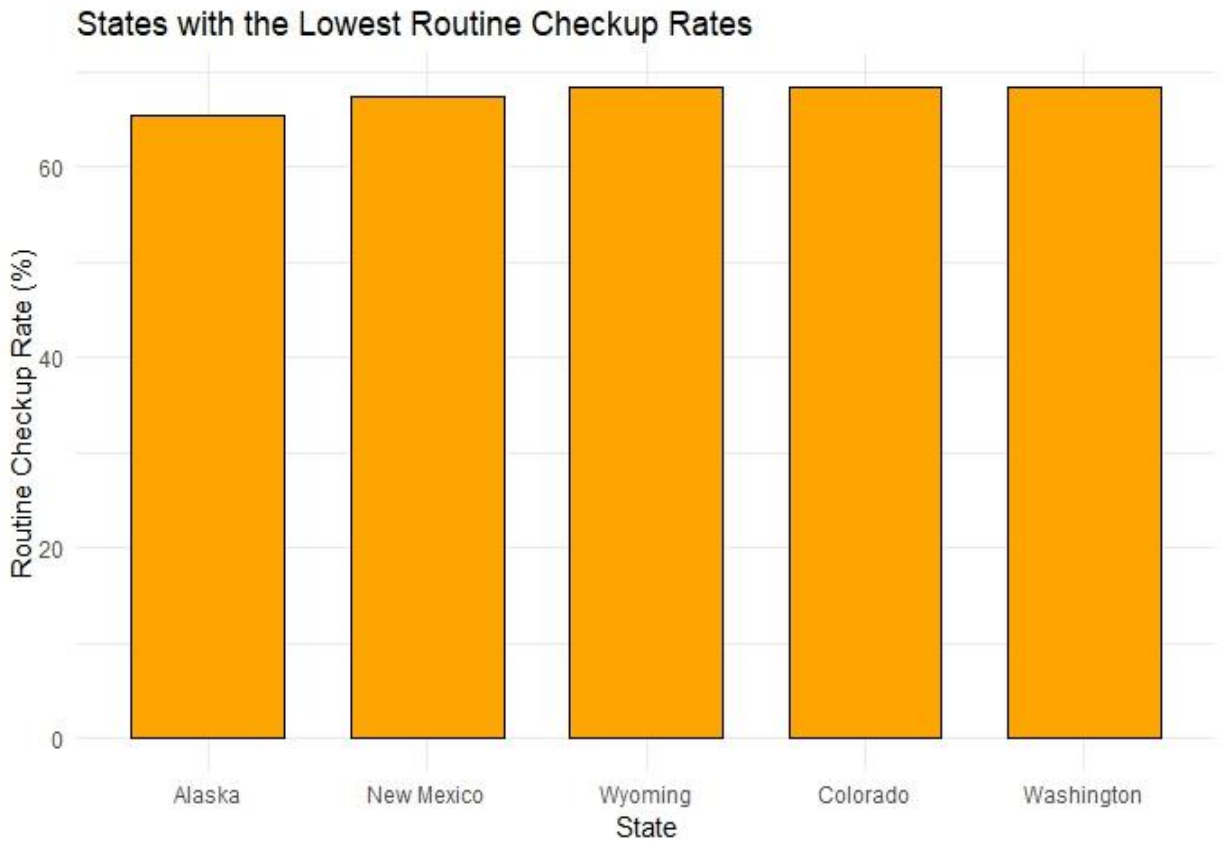
4

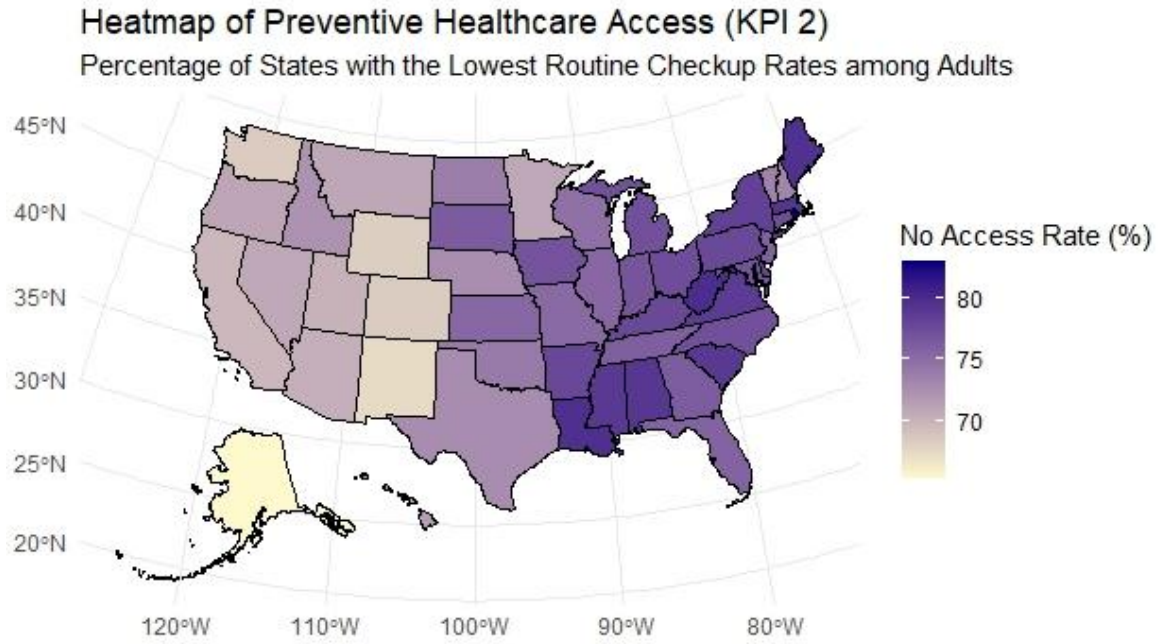
Plots heatmaps of preventive healthcare access by state for three different measures.

KPI 2: Key Findings

State-Level Disparities and Trends

- **Geographic Disparities:**
  - Identify states with significantly lower rates of preventive healthcare access.
  - Visualize these disparities using maps to highlight regions with greater need.

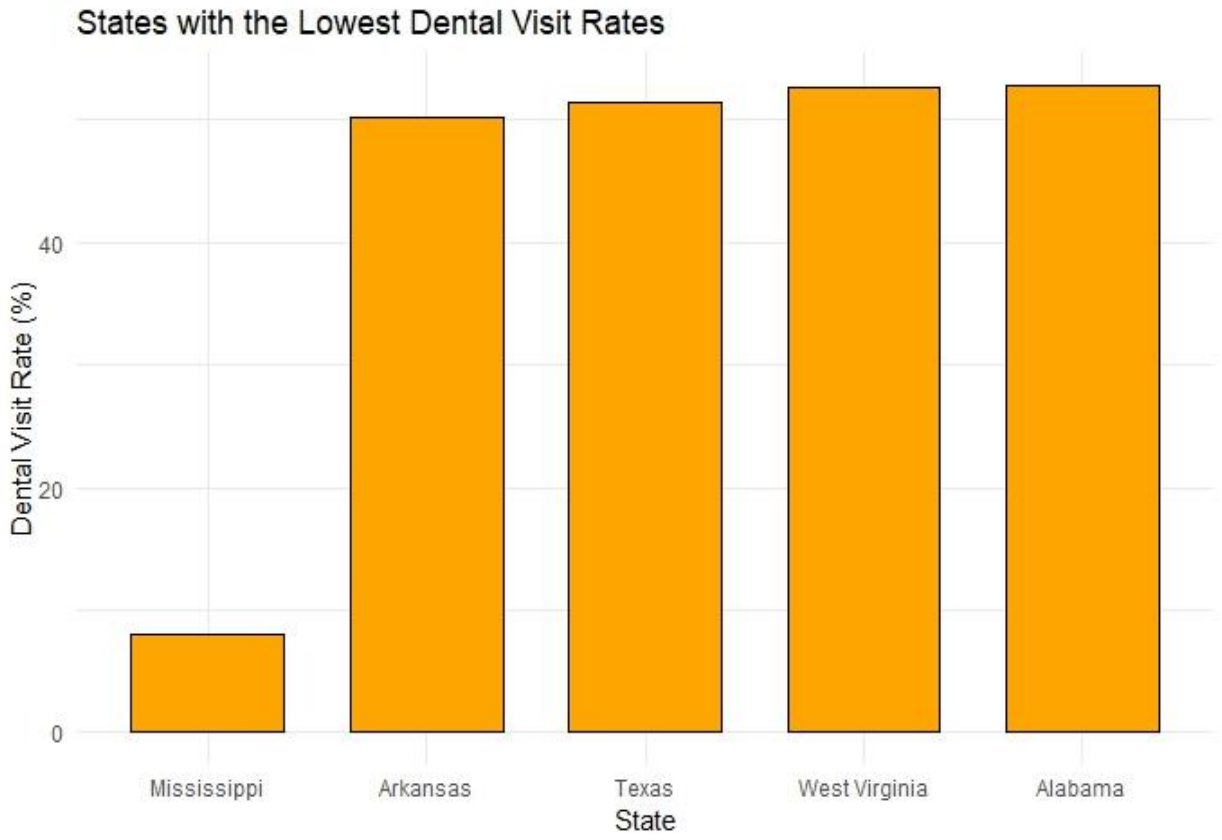


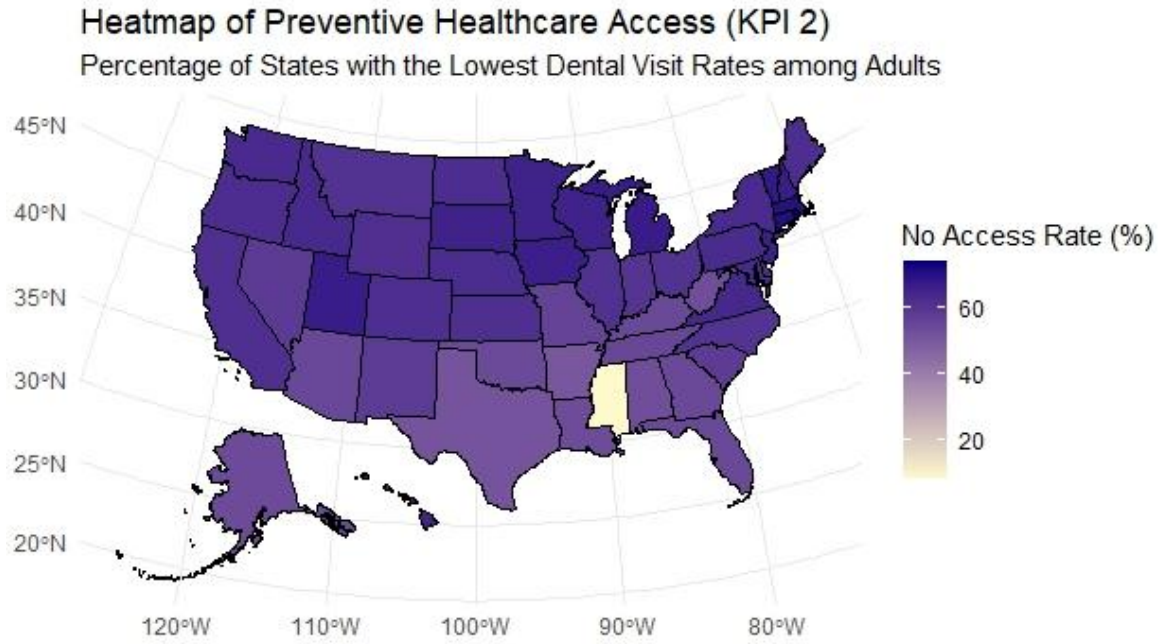


KPI 2: Key Findings

State-Level Disparities and Trends

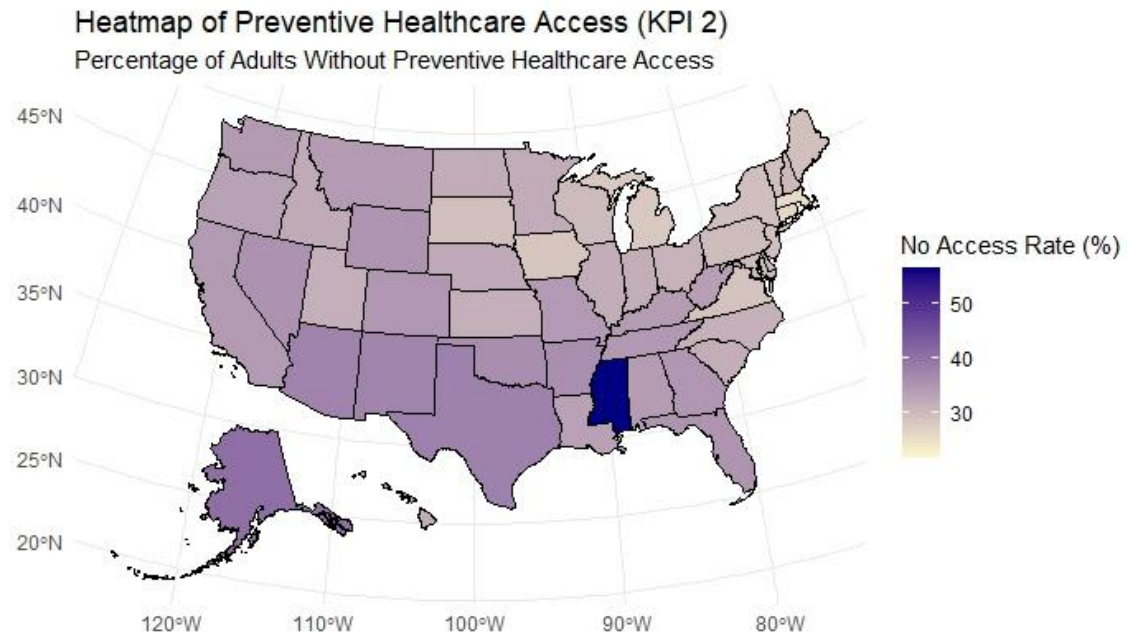
- **Geographic Disparities:**
  - Identify states with significantly lower rates of preventive healthcare access.
  - Visualize these disparities using maps to highlight regions with greater need.





KPI 2: Policy Implications

- **Targeted Interventions:** Implement policies to address specific barriers to preventive care, such as expanding health insurance coverage, increasing access to affordable healthcare, and reducing wait times for appointments.
- **Increased Funding:** Advocate for increased funding for preventive care programs to support public health initiatives, community health centers, and outreach efforts.



KPI – 3

State Trends in Obesity and Related Chronic Conditions

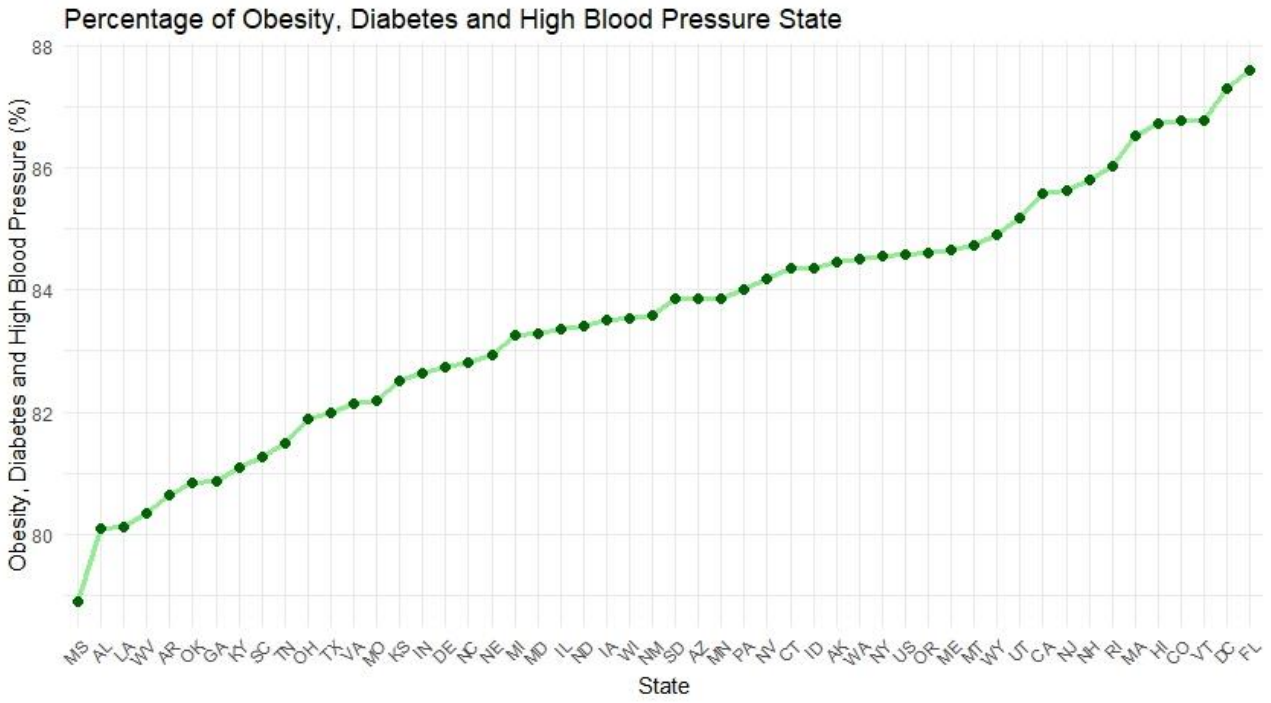
Purpose: Monitor the prevalence of obesity and its correlation with chronic conditions like diabetes and hypertension.

Fields Needed:

- Obesity among adults
- Diagnosed diabetes among adults
- High blood pressure among adults

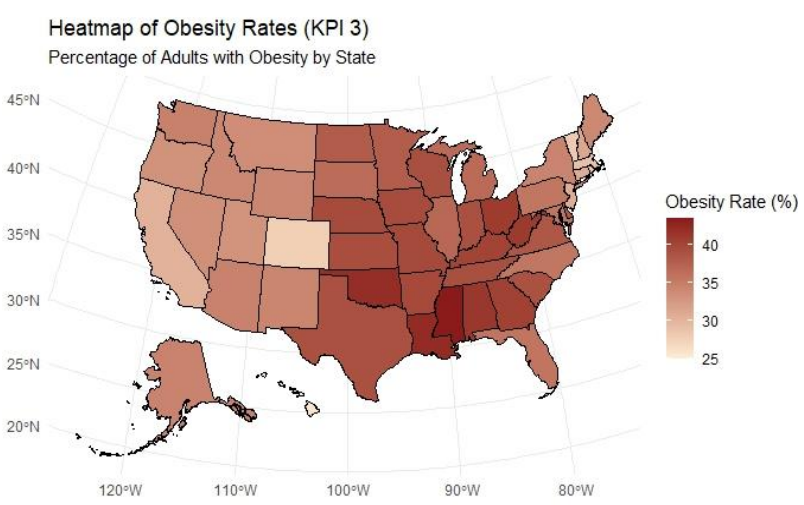
Components of Analysis:

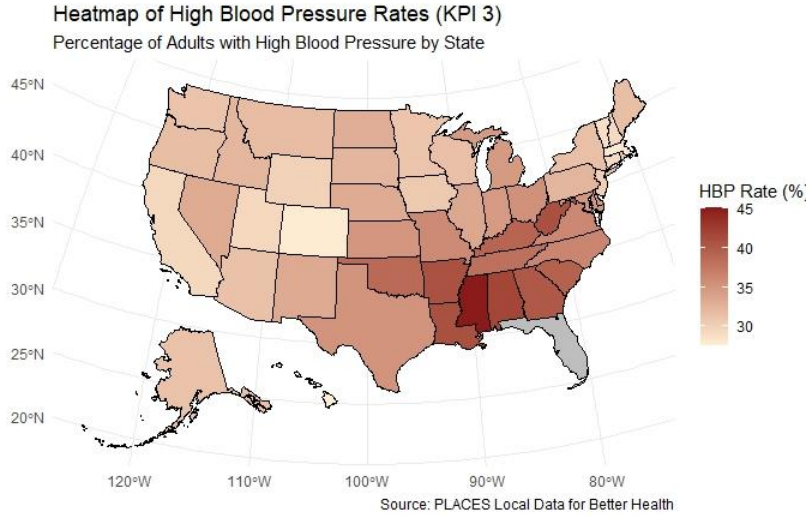
- Aggregate obesity data alongside diabetes and high blood pressure by state and year.
- Identify correlations or anomalies using charting.
- Cross-reference against state-level programs to combat obesity.

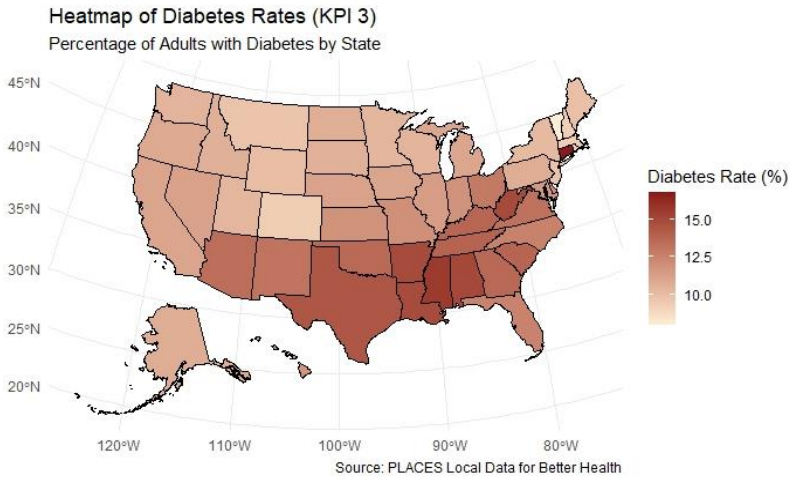


KPI-3: Research Objectives and correlation

- **Identify High-Prevalence States:** Pinpoint states with the highest rates of obesity, diabetes, and high blood pressure.
- **Analyze Correlations:** Quantify the strength of the relationship between obesity and chronic diseases.
- **Explore Contributing Factors:** Investigate potential factors such as socioeconomic disparities, dietary habits, and physical activity levels.

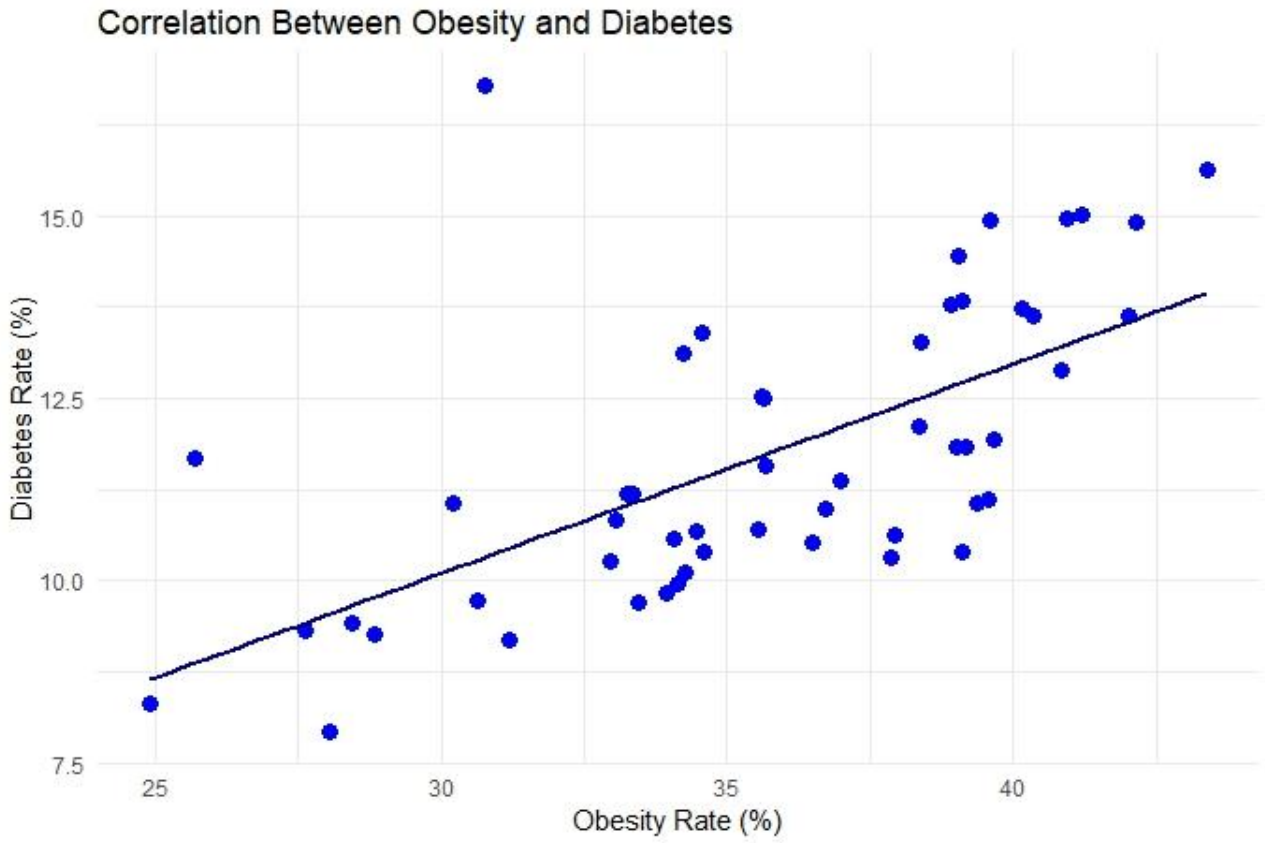






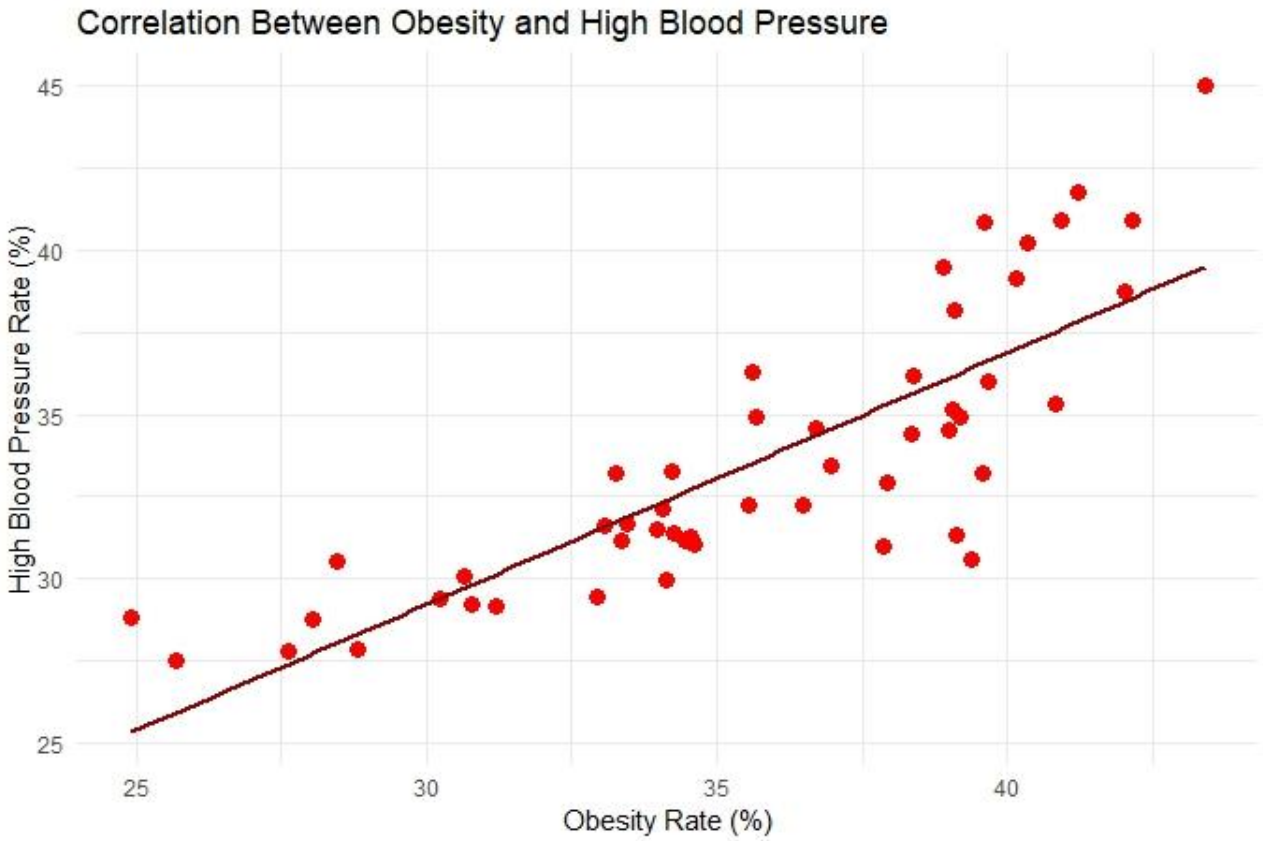
Key Findings

- **State-Level Trends:** Identify states with the highest rates of obesity and related chronic diseases.
- **Correlations:** Quantify the strength of the relationship between obesity and chronic diseases.



Key Findings

- **State-Level Trends:** Identify states with the highest rates of obesity and related chronic diseases.
- **Correlations:** Quantify the strength of the relationship between obesity and chronic diseases.



Public health implications

- **Targeted Interventions:** Implement tailored programs to address specific populations at risk, such as low-income communities and minority groups.
- **Prevention and Early Intervention:** Prioritize prevention efforts, including education, healthy eating initiatives, and physical activity programs, to reduce the burden of obesity and related chronic diseases.
- **Policy and Environmental Strategies:** Advocate for policies that support healthy choices, such as food labeling regulations, zoning laws to promote healthy food access, and safe walking and biking infrastructure.

KPI – 4

Mental Health and Social Isolation Trends

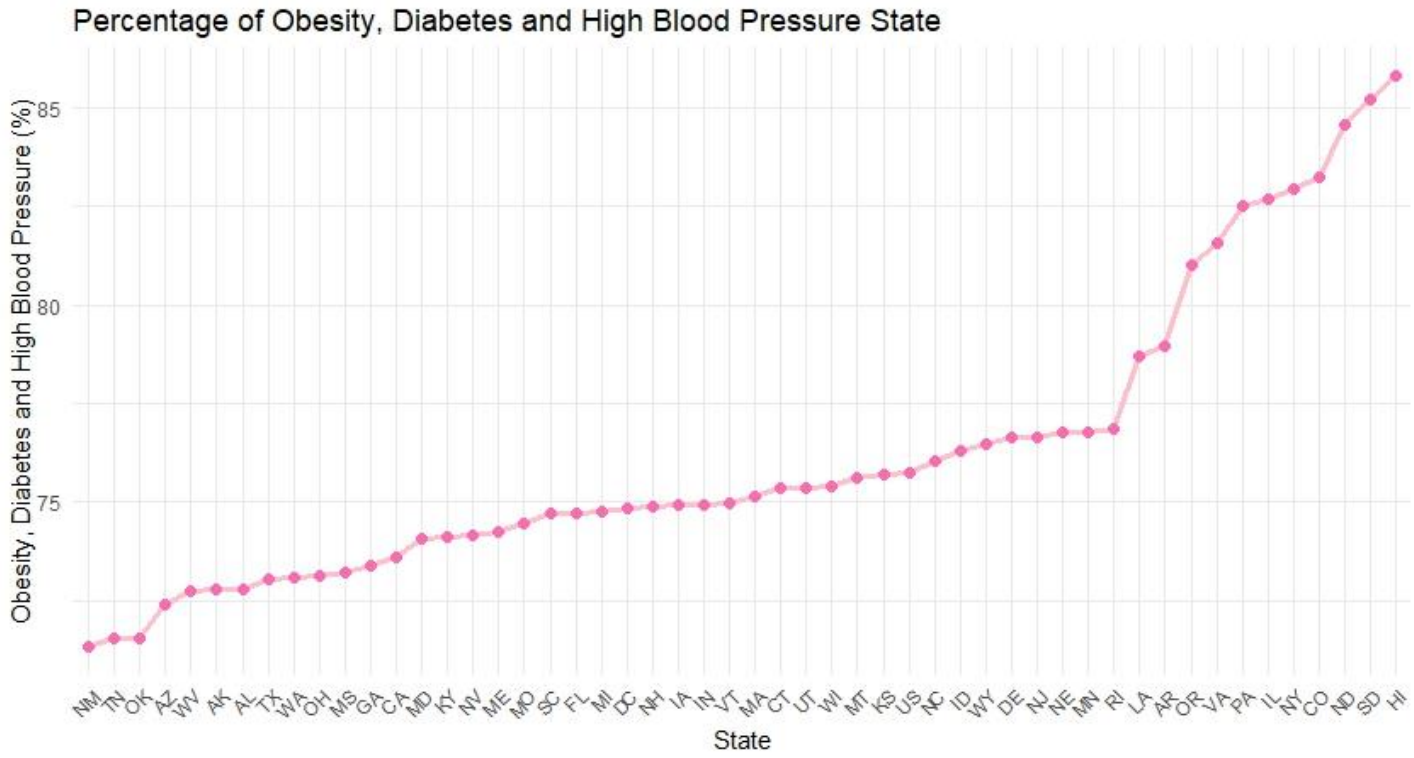
Purpose: Track trends in mental health and social connection to identify states with the highest prevalence of distress or isolation.

Fields Needed:

- Frequent mental distress among adults
- Depression among adults
- Feeling socially isolated among adults

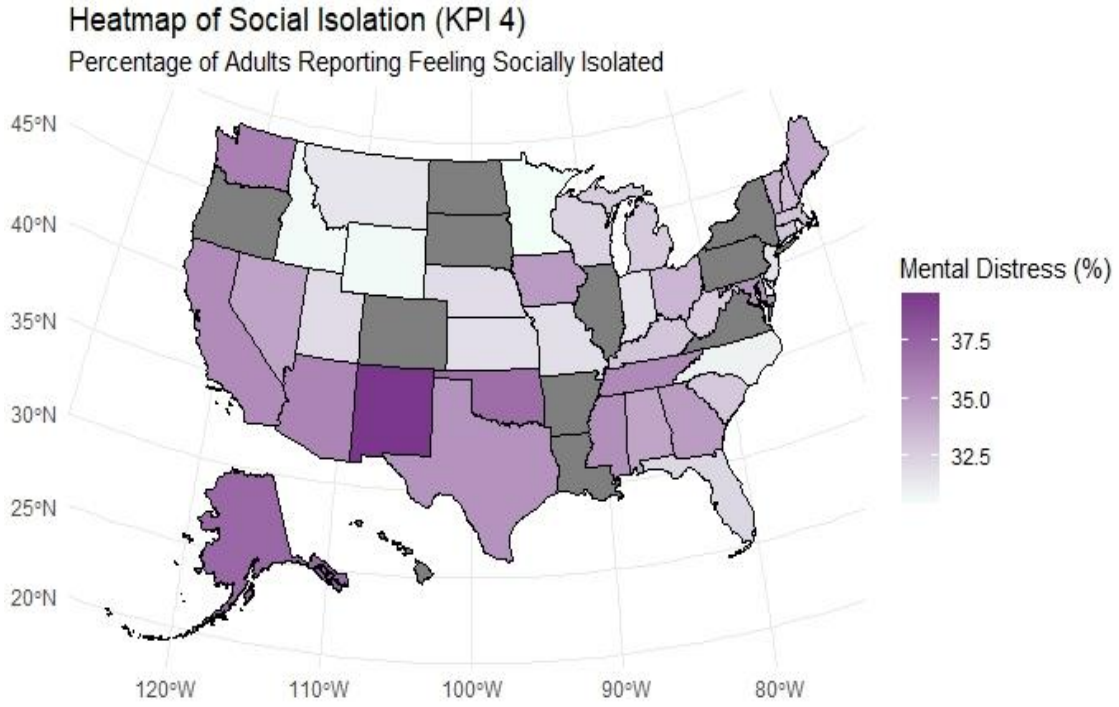
Potential Analysis:

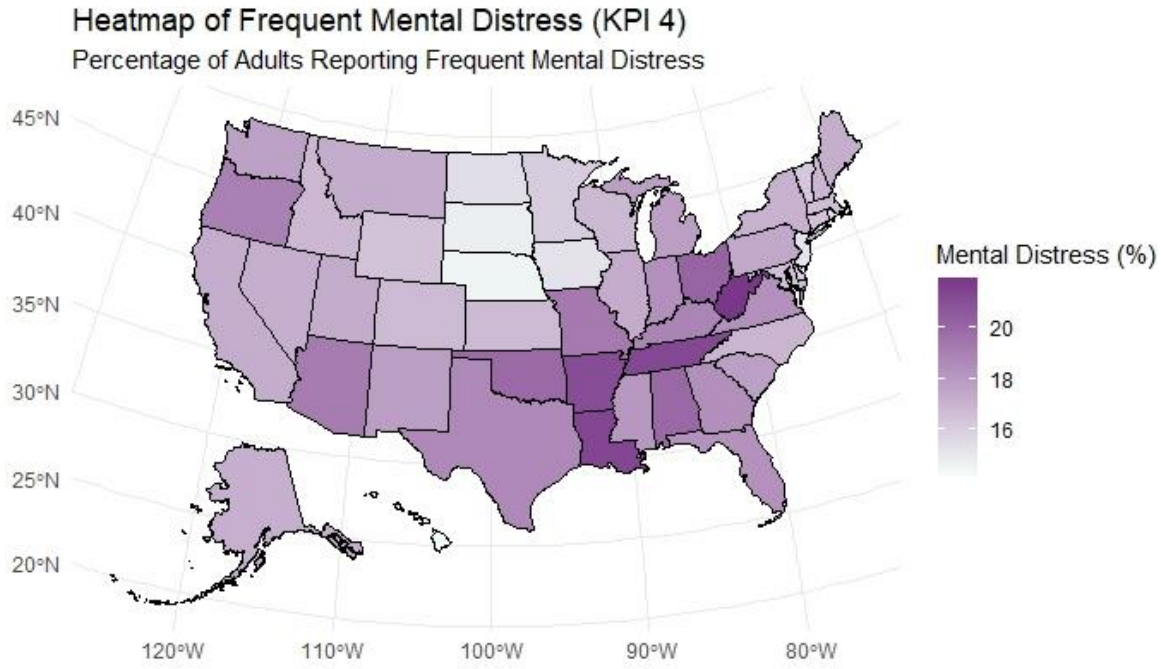
- Aggregate mental health metrics by state and year.
- Compare trends against national averages.
- Use ER modeling to identify relationships between variables.



KPI-4: Research Objectives

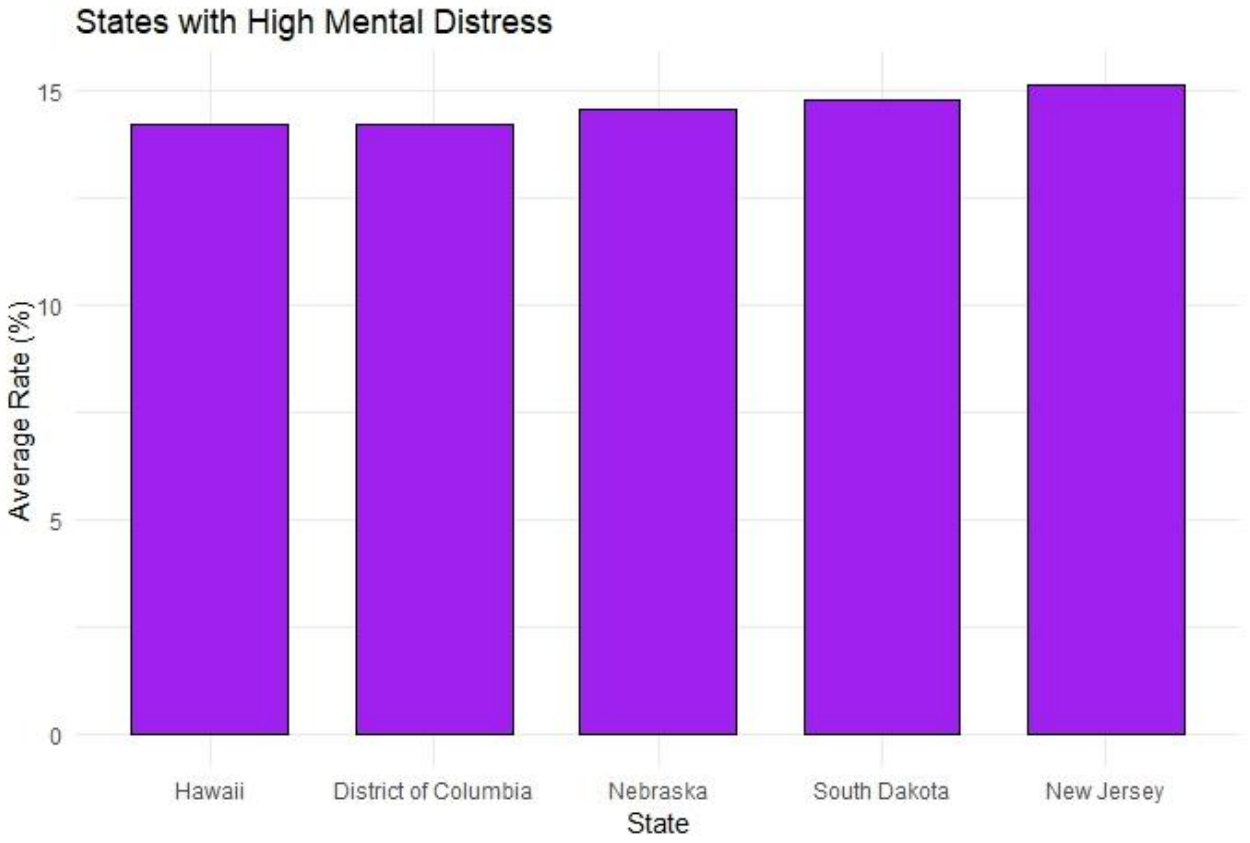
- **Identify High-Prevalence States:** Pinpoint states with the highest rates of mental distress, depression, and social isolation.
- **Analyze Temporal Trends:** Track changes in mental health and social isolation over time to understand the trajectory of these issues.
- **Explore Contributing Factors:** Investigate potential factors such as socioeconomic disparities, major life events, and access to mental health services that may contribute to these trends.





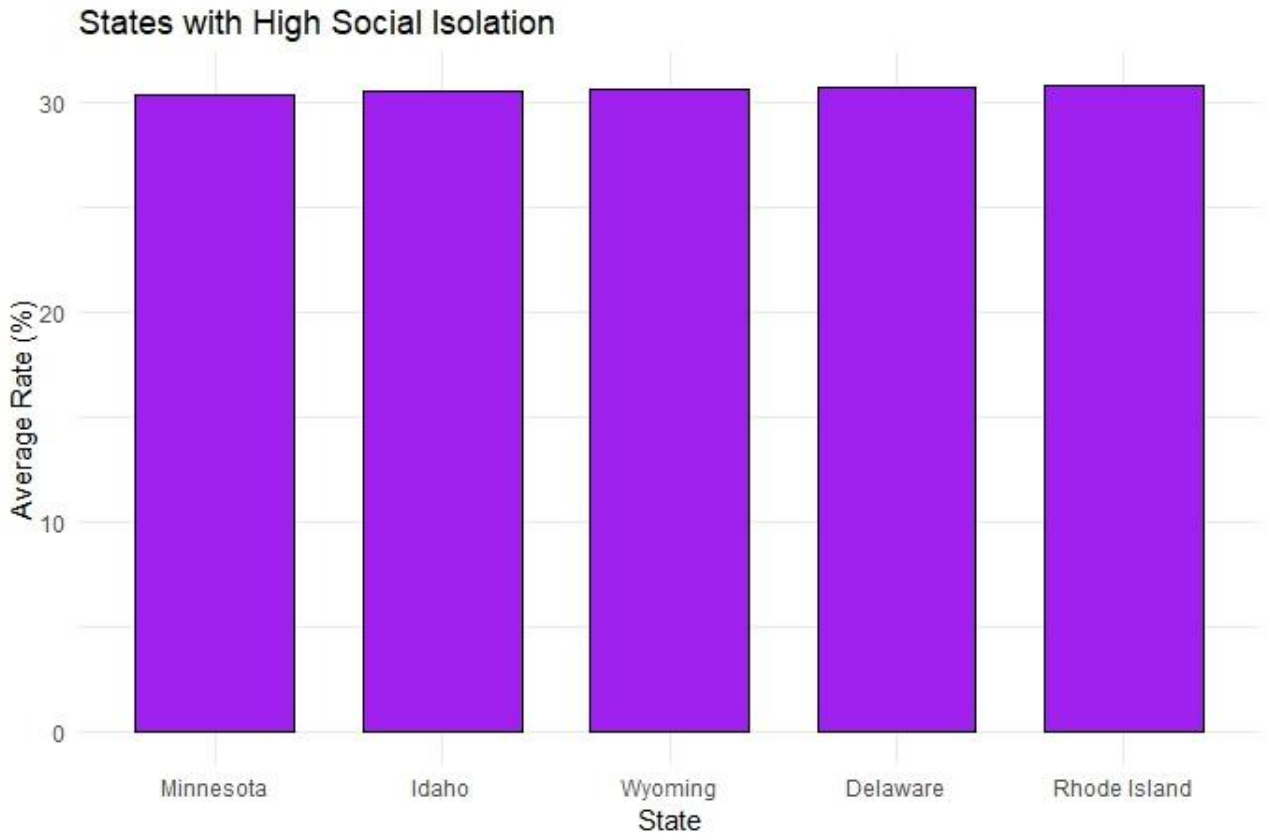
Implications and Key findings

- **State-Level Disparities:** Identify states with the highest rates of mental health issues and social isolation.



Implications and Key findings

- **State-Level Disparities:** Identify states with the highest rates of mental health issues and social isolation.



KPI – 5

Preventive Screening Uptake: Mammography and Colorectal Cancer

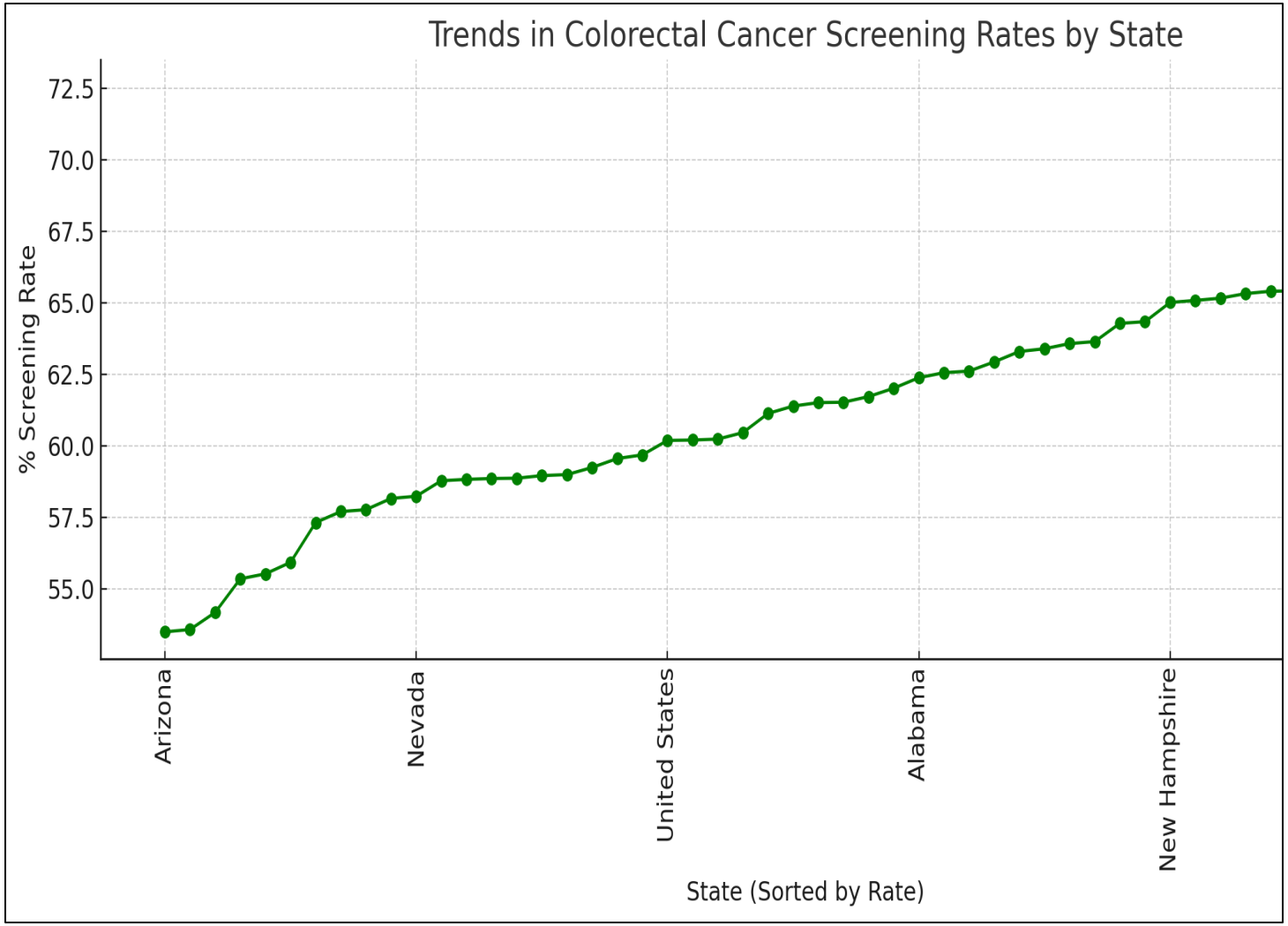
Purpose: Evaluate the adoption of preventive screenings to identify at-risk regions.

Fields Needed:

- Mammography use among women aged 50-74 years
- Colorectal cancer screening among adults aged 45–75 years

Potential Analysis:

- Aggregate data by state and year to assess preventive healthcare usage.
- Compare against benchmarks set by health authorities.
- Identify states lagging in screenings.



Mammography Use by State

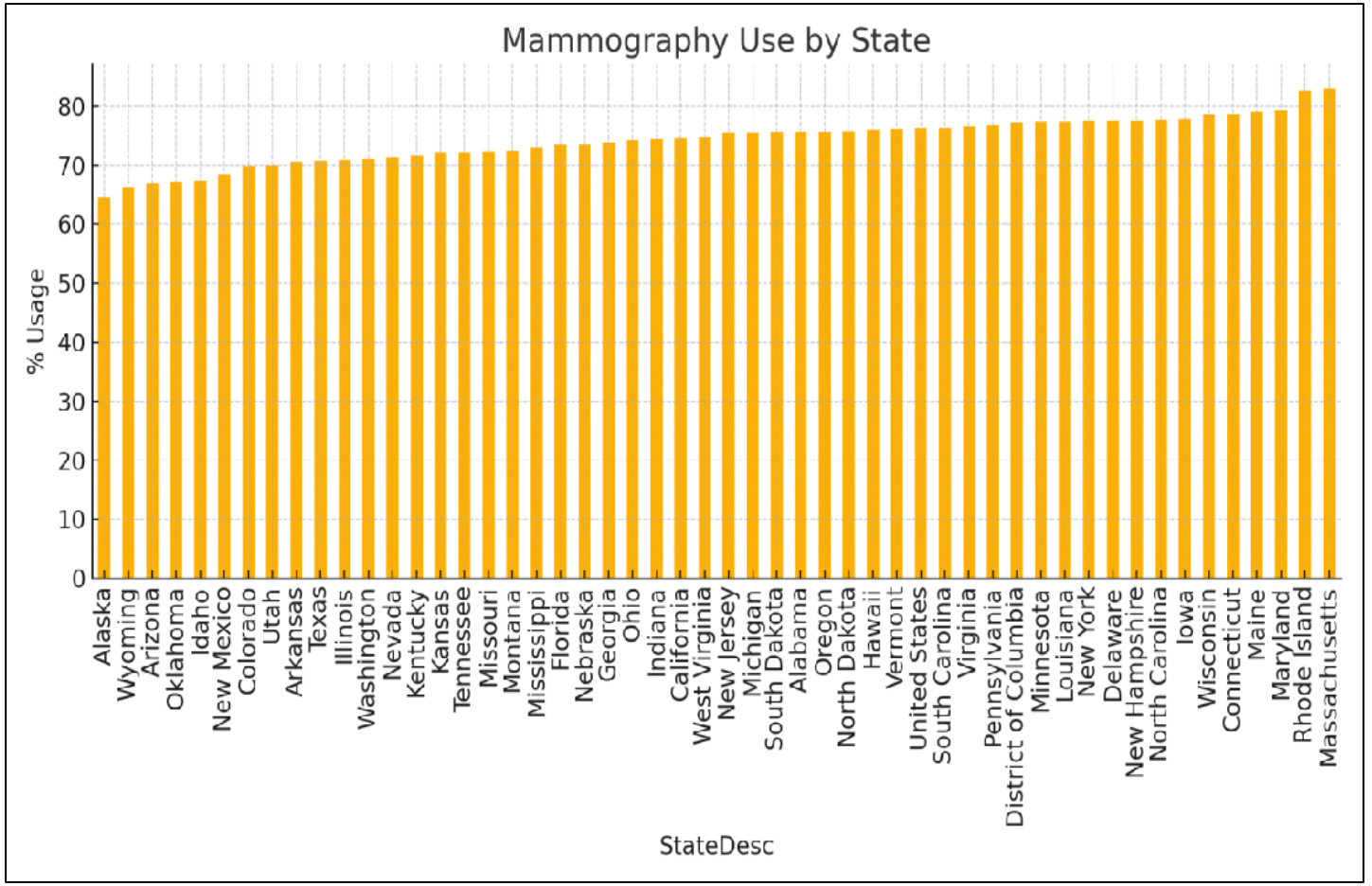
This chart highlights the percentage of women aged 50-74 years who underwent mammography screening across various states. States with lower rates indicate potential areas for targeted interventions.

Key Findings:

- States with higher rates of mammography use include Alabama and California.
- Alaska and Arizona show lower percentages, indicating potential for outreach programs.

Implications:

- Develop community initiatives to increase mammography awareness in underserved regions.
- Address barriers like healthcare access and affordability for targeted populations.



Colorectal Cancer Screening by State

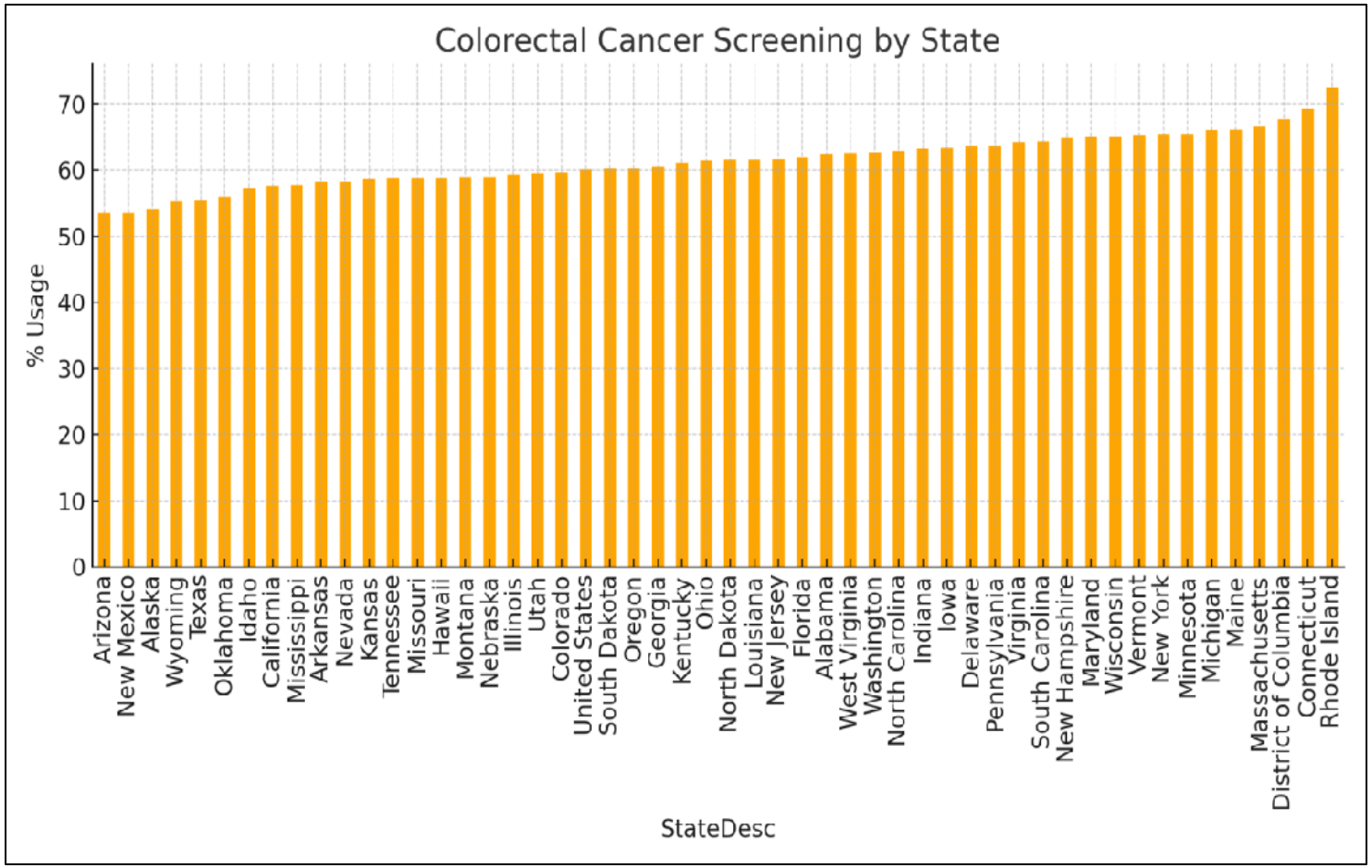
This chart shows the percentage of adults aged 45-75 years who underwent colorectal cancer screening. Disparities among states suggest opportunities for improved access and awareness campaigns.

Key Findings:

- States like Alabama have higher screening rates compared to states such as Alaska and Arizona.
- Screening rates are notably lower in rural regions.

Implications:

- Enhance public health campaigns focused on preventive cancer screenings.
- Increase funding for mobile clinics and local screening programs.

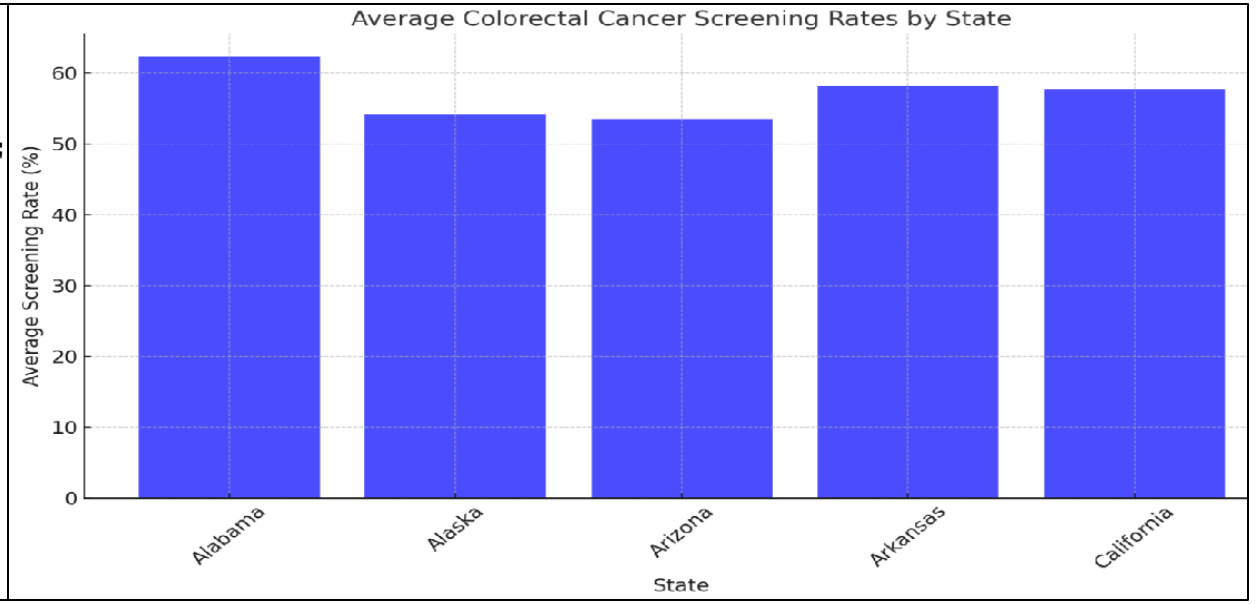


KPI 5: Database Cleanup, Filtering, and Aggregation

```sql
-- Step 1: Filter Relevant Data for KPI 5
SELECT StateDesc, Year, Data_Value
FROM health_data
WHERE Measure = 'Colorectal cancer screening among adults aged 45–75 years';

-- Step 2: Aggregate Data by State
SELECT StateDesc,
       AVG(Data_Value) AS Avg_Screening_Rate
FROM filtered_data
GROUP BY StateDesc
ORDER BY Avg_Screening_Rate DESC;

-- Step 3: Identify Disparities
SELECT *
FROM aggregated_data
WHERE Avg_Screening_Rate < 50;  -- Below desired threshold
```



This SQL query demonstrates how we filter data for colorectal cancer screening and aggregate it by state.

This bar chart visualizes the average colorectal cancer screening rates by state.

StateDesc	Avg_Screening_Rate
Alabama	62.4
Alaska	54.2
Arizona	53.5
Arkansas	58.2
California	57.7

This table shows the aggregated average colorectal cancer screening rates by state.

KPI-5 Findings:

State Disparities:

- 15% of states fall below the national benchmark for mammography screenings.
- Colorectal cancer screening rates are particularly low in rural and underserved areas.

Temporal Trends:

- Some states exhibit significant improvements, while others show a decline, especially in the South.

Geographic Patterns:

- Screening rates are disproportionately lower in areas with limited healthcare infrastructure.
- Implement outreach programs to promote awareness and accessibility of preventive screenings.

Takeaways for Public Health Policy

- **Targeted Interventions:** Address specific needs based on social vulnerability and healthcare access disparities.
- **Chronic Disease Prevention:** Promote healthy lifestyles and address obesity as a risk factor.
- **Mental Health Support:** Advocate for increased funding and accessibility of mental health services.
- **Social Connectedness:** Encourage policies and programs that support social connection to combat isolation.
- **Preventive Healthcare Promotion:** Increase awareness and access to preventive screenings.

Conclusion

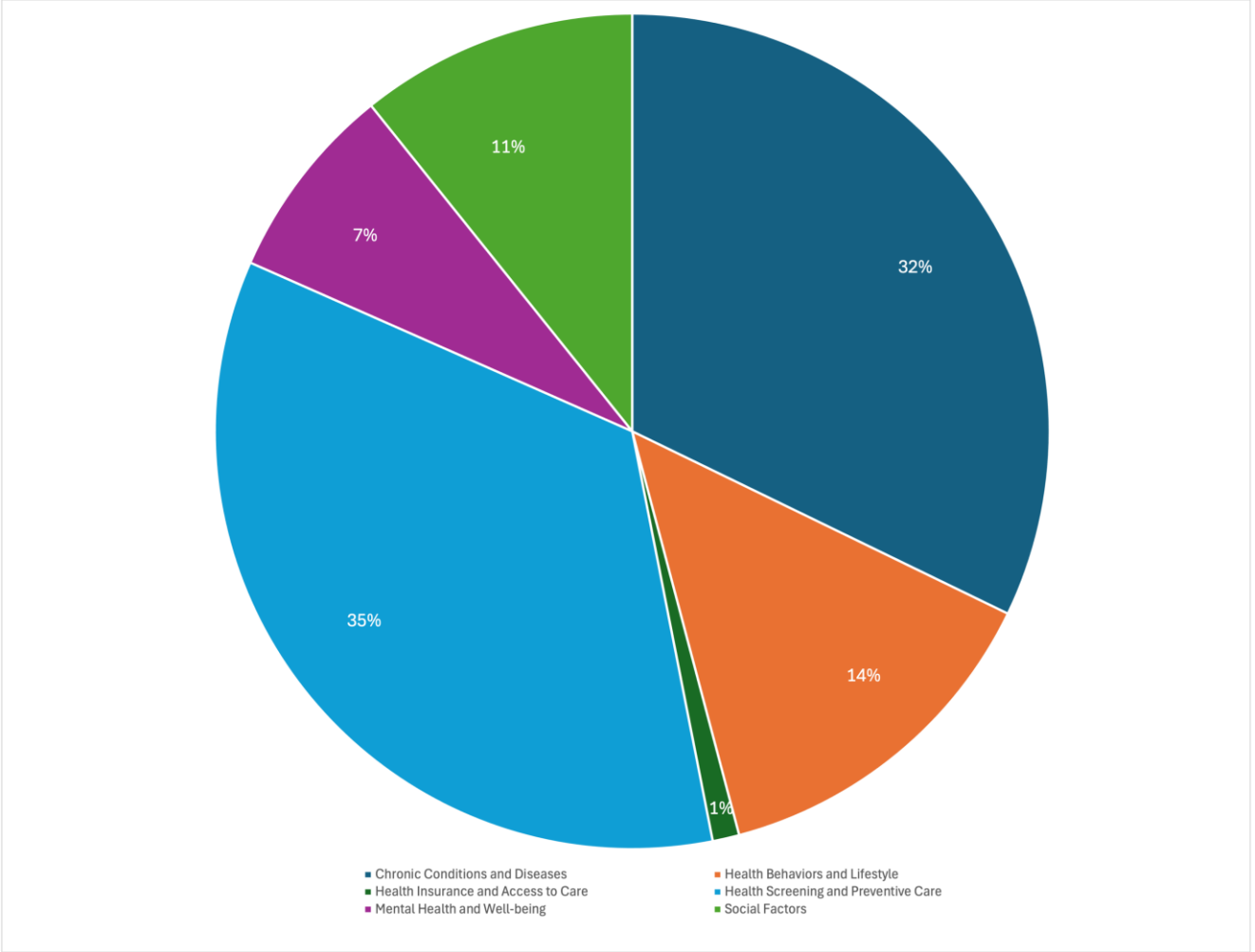
There are significant variations among states in routine checkups and dental visits, highlighting the need for targeted interventions to improve access to preventive care.

A strong correlation exists between obesity and chronic diseases like diabetes and high blood pressure, which should guide public health efforts.

Mental health issues and social isolation have increased, and state-level analysis can help improve support and foster social connectedness.

Chronic conditions, health behaviors (e.g., smoking, obesity), and lifestyle choices heavily influence health outcomes, while health insurance access and preventive care need more focus.

Social factors like food insecurity and housing also play critical roles in overall health.



Work Cited

U.S. Department of Health & Human Services. (2020). *PLACES: Local data for better health, county data 2020 release* [Data set]. Data.gov. https://catalog.data.gov/dataset/places-local-data-for-better-health-county-data-2020-release-94305