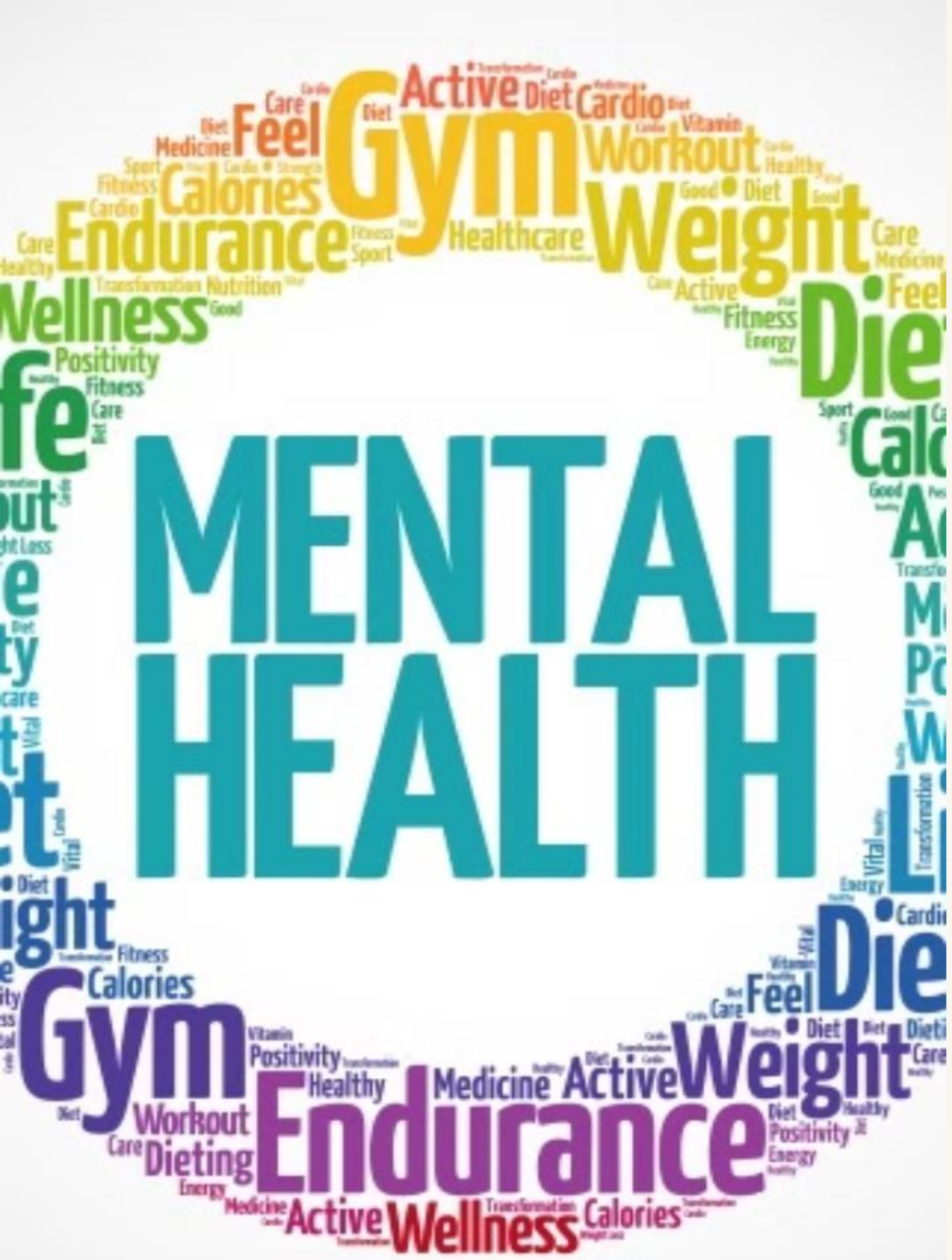


HealthWatch: A Data-Driven Approach to Mental Health Risk

HealthWatch is about leveraging the power of data to make informed decisions in mental healthcare. It's about understanding individual risks, optimizing intervention strategies, and personalizing mental health support.

Introduction

Background

The data used in this project is collected from the **U.S. Department of Health & Human Services** through their **Behavioral Risk Factor Surveillance System (BRFSS)**. This system gathers comprehensive health-related data on risk behaviors, chronic conditions, and preventive practices across the United States

Objective

- **Identify Health Disparities:** Analyze and highlight significant health disparities across different age groups, focusing on older adults′ challenges with mental health and cognitive decline.
- **Examine Correlations:** Investigate the strong relationship between frequent mental distress and cognitive issues to guide data-driven interventions.

Dataset Overview

1 Source

U.S. Department of Health & Human Services (BRFSS)

2 Key Features

Data segmented by age groups (50–64 years, 65+ years), race, and gender.

3 Geographic Scope

Includes data from all 50 states and U.S. territories.

Data Cleaning and Transformation

Stratification

Grouped data into meaningful categories by:

- Age: 50–64 years, 65+ years.
- Gender and Race/Ethnicity for demographic analysis.

Data Type Conversion

- Converted data fields to appropriate types for analysis (e.g., numeric for metrics, categorical for demographics).

Handling Missing Values

Addressed missing data by:

- Removing rows with excessive null values.
- Imputing missing values for key metrics using mean or median values.

Data Transformation

```
[1]: import pandas as pd

# Step 1: Load the dataset from CSV
file_path = 'Healthdata_final.csv'  # Replace with the path to your CSV file
data = pd.read_csv(file_path)

# Step 2: Display dataset information
print("Dataset Information:")
print(data.info())

# Step 3: Drop specified columns
columns_to_drop = [
    'RowId', 'Data_Value_Unit', 'DataValueTypeID', 'Data_Value_Type', 'Data_Value_Alt',
    'Low_Confidence_Limit', 'High_Confidence_Limit', 'Geolocation', 'ClassID', 'TopicID',
    'QuestionID', 'LocationID', 'StratificationCategory1', 'StratificationCategoryID1',
    'StratificationID1', 'StratificationCategoryID2', 'StratificationID2'
]
data = data.drop(columns=columns_to_drop, errors='ignore')

# Step 4: Rename columns
columns_rename = {
    'YearStart': 'Year_Start',
    'YearEnd': 'Year_End',
    'LocationAbbr': 'Location_Abbr',
    'LocationDesc': 'Location_Desc',
    'Class': 'Survey_Class',
    'Topic': 'Survey_Topic',
    'Question': 'Survey_Question',
    'Data_Value': 'Data_Value',
    'Stratification1': 'Age_Group',
    'StratificationCategory2': 'StratificationCategory2',
    'Stratification2': 'Stratification2'
}
data = data.rename(columns=columns_rename)
```

- **Column Dropping:** Unnecessary columns such as identifiers, confidence limits, and redundant stratifications were removed to streamline the dataset.
- **Column Renaming:** Columns were renamed for consistency and clarity, ensuring easier readability and analysis.

```
# Filter 1: Only_agegroup_data
only_agegroup_data = data[data['StratificationCategory2'].isnull()]

# Filter 2: Only_start2_data
only_start2_data = data[(data['StratificationCategory2'].notnull()) & (data['Age_Group'] == 'Overall')]

# Filter 3: agegroup_and_strat2_data
agegroup_and_strat2_data = data[(data['StratificationCategory2'].notnull()) & (data['Age_Group'] != 'Overall')]

# Display the resulting dataframes
print("Main Processed DataFrame:")
print(data.head())

print("\nOnly Age Group Data:")
print(only_agegroup_data.head())

print("\nOnly Stratification Category 2 Data:")
print(only_start2_data.head())

print("\nAge Group and Stratification Category 2 Data:")
print(agegroup_and_strat2_data.head())
```

- **Filtering by Categories:** The dataset was split into subsets based on conditions for `Stratification` and `Age_Group` to focus on specific analyses.
- **Data Frames** are created by applying **conditional filters** to the dataset to extract specific subsets of data that meet certain criteria.

Subset Creation

Three distinct subsets were created for "only Age Group," "Stratification Category 2," and a combination of both, for targeted exploration.

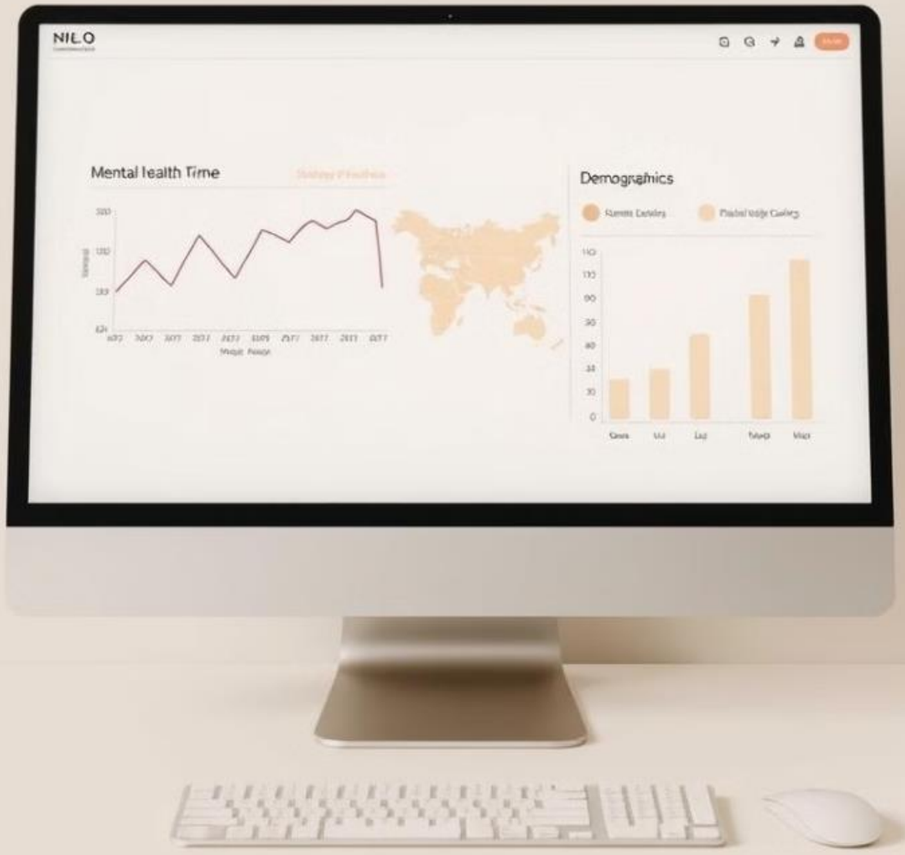
	Survey_Topic	Survey_Question	Data_Value
	Frequent mental distress	Percentage of older adults who are experiencin...	9.0
	Frequent mental distress	Percentage of older adults who are experiencin...	5.6
	Expect to provide care for someone in the next...	Percentage of older adults currently not provi...	14.5
	Obesity	Percentage of older adults who are currently o...	32.7
	Arthritis among older adults	Percentage of older adults ever told they have...	42.7

	Age_Group	StratificationCategory2	Stratification2
	65 years or older	Race/Ethnicity	Black, non-Hispanic
	65 years or older	Gender	Male
	50-64 years	Gender	Male
	65 years or older	Race/Ethnicity	Hispanic
	65 years or older	Race/Ethnicity	Hispanic

	Year_Start	Year_End	Location_Abbr	Location_Desc	Survey_Class
0	2022	2022	MD	Maryland	Mental Health
1	2022	2022	WI	Wisconsin	Mental Health
2	2022	2022	OK	Oklahoma	Mental Health
3	2022	2022	PA	Pennsylvania	Mental Health
4	2022	2022	OH	Ohio	Caregiving

	Survey_Topic	Survey_Question	Data_Value
0	Frequent mental distress	Percentage of older adults who are experiencin...	9.0
1	Frequent mental distress	Percentage of older adults who are experiencin...	5.6
2	Frequent mental distress	Percentage of older adults who are experiencin...	21.5
3	Frequent mental distress	Percentage of older adults who are experiencin...	10.0
4	Expect to provide care for someone in the next...	Percentage of older adults currently not provi...	14.5

	Age_Group	StratificationCategory2	Stratification2
0	65 years or older	Race/Ethnicity	Black, non-Hispanic
1	65 years or older	Gender	Male
2	Overall	Race/Ethnicity	Native Am/Alaskan Native
3	Overall	Race/Ethnicity	White, non-Hispanic
4	50-64 years	Gender	Male



Key Visualizations

Visualization 1

State-wise Comparison of Leisure Inactivity and Obesity.

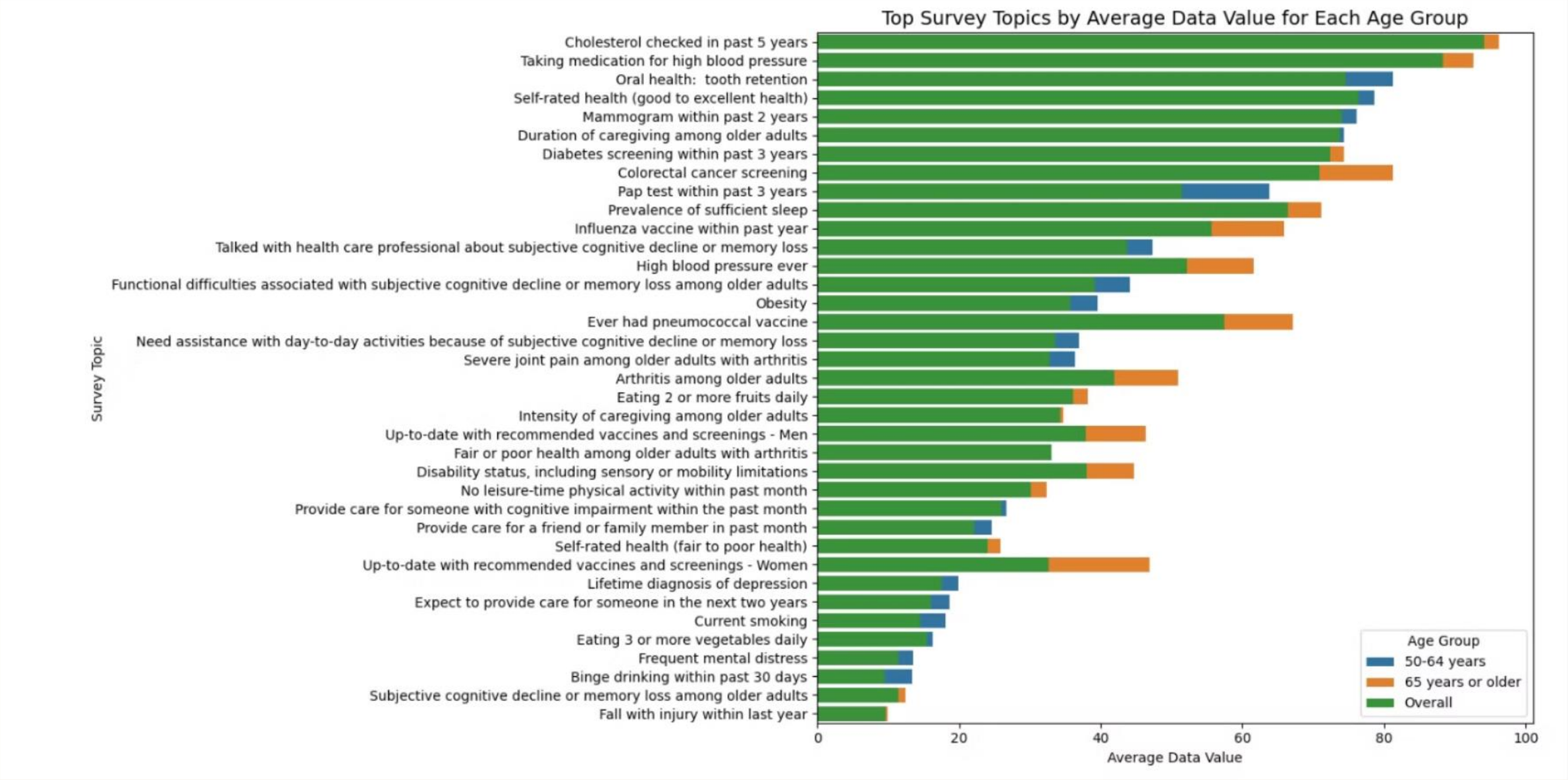
Visualization 2

Cognitive Decline Across Race/Ethnicity and Gender.

Visualization 3

Correlation Between Mental Distress and Cognitive Decline.

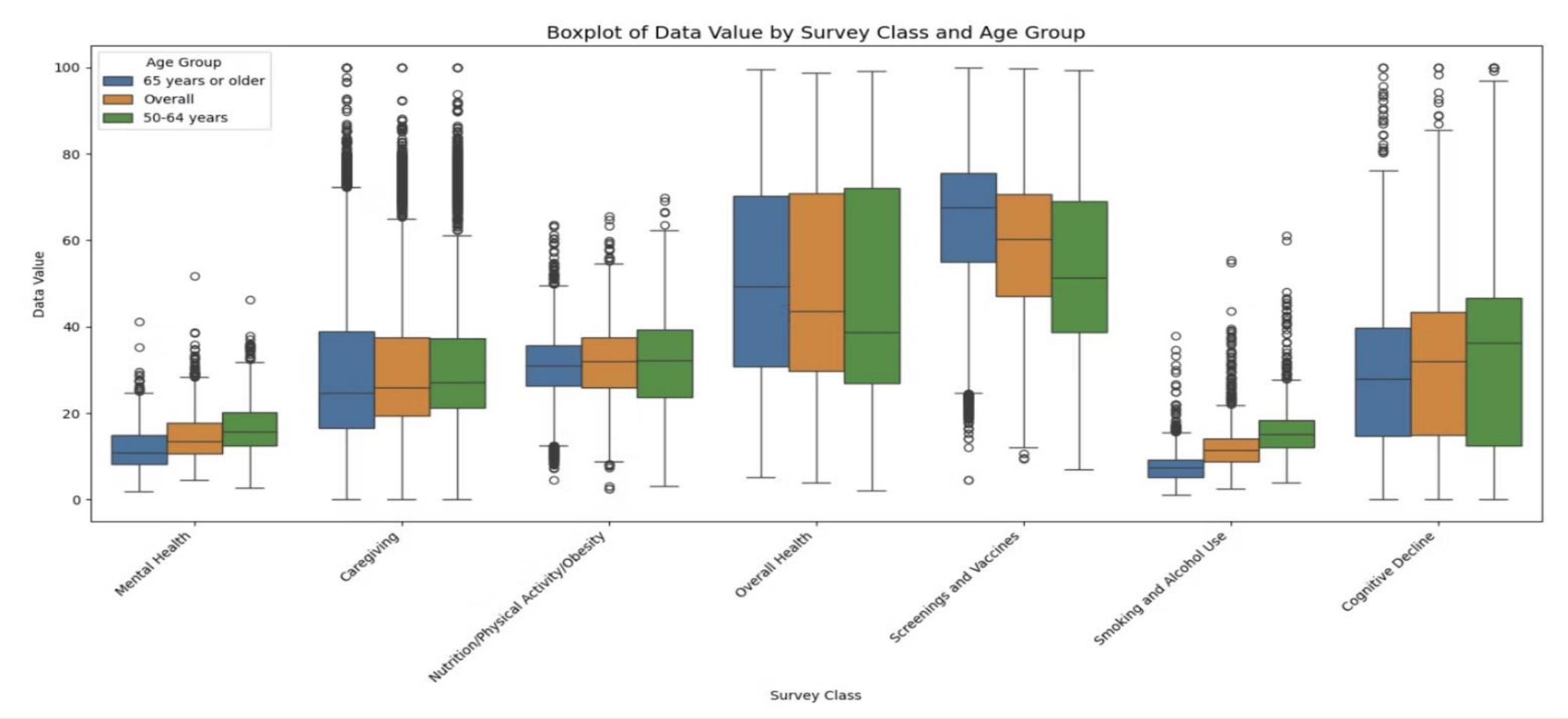
Top Survey Topics by Average Data Value Across Age Groups



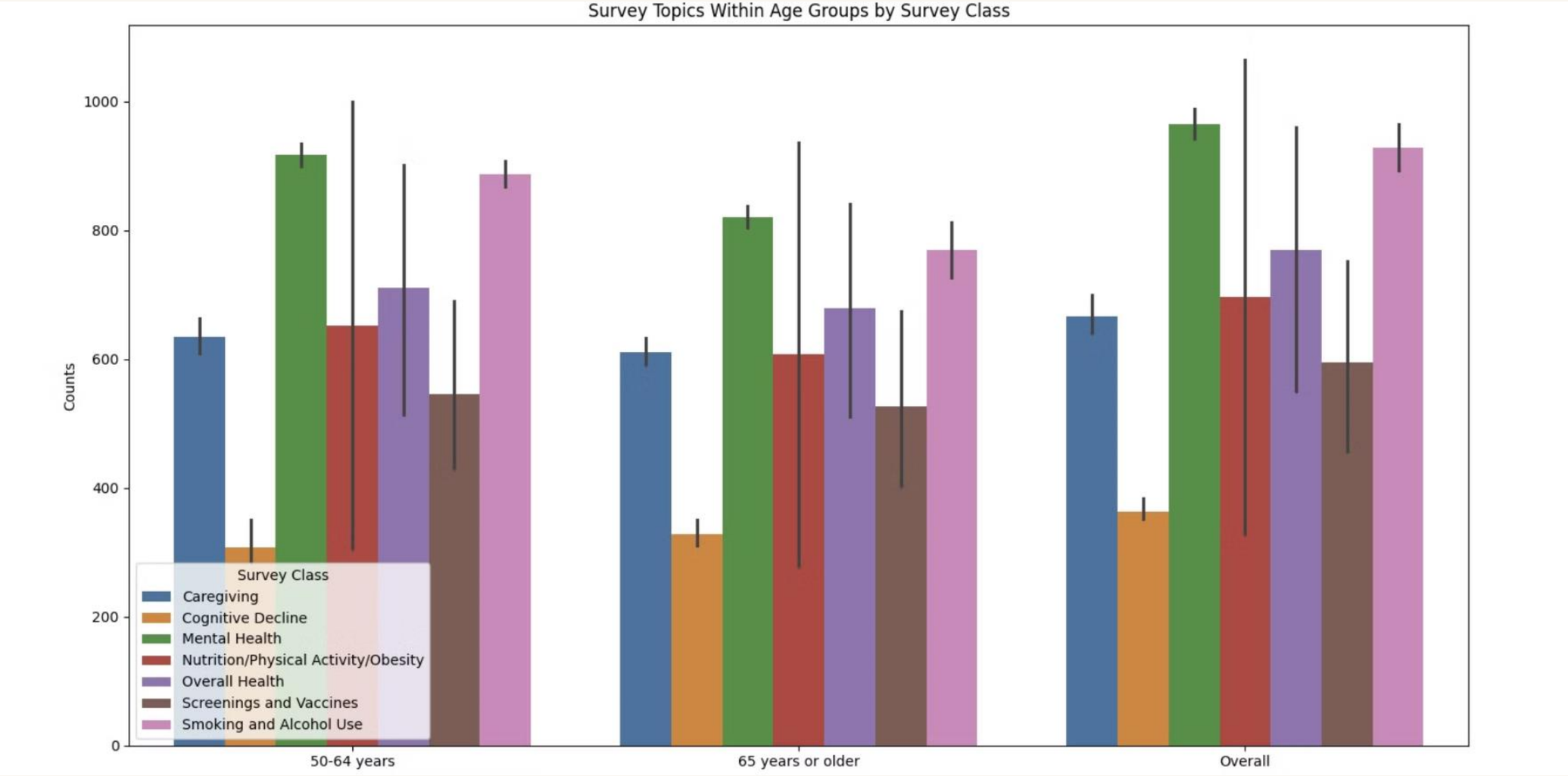
Findings

Comparison of Health Indicators Across Age Groups and Survey Categories:

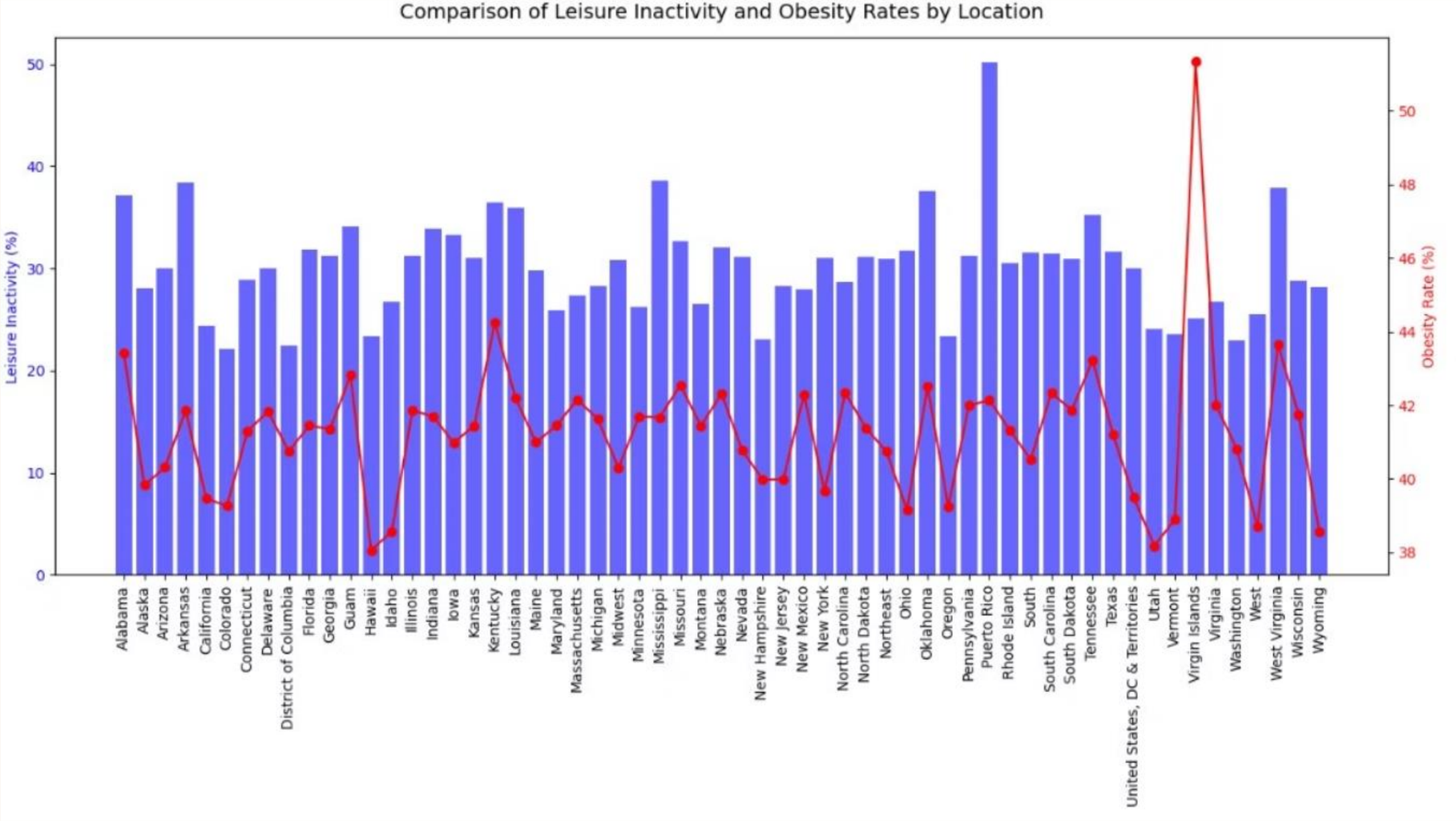
- The 50–64 age group exhibits relatively high values in the mental health quartile compared to other categories. This could be attributed to higher work-related pressures often experienced by individuals in this age group.
- Cognitive decline shows a high median and broader variability for the 50-64 age group, which aligns with expectations of increased cognitive health issues in older populations.



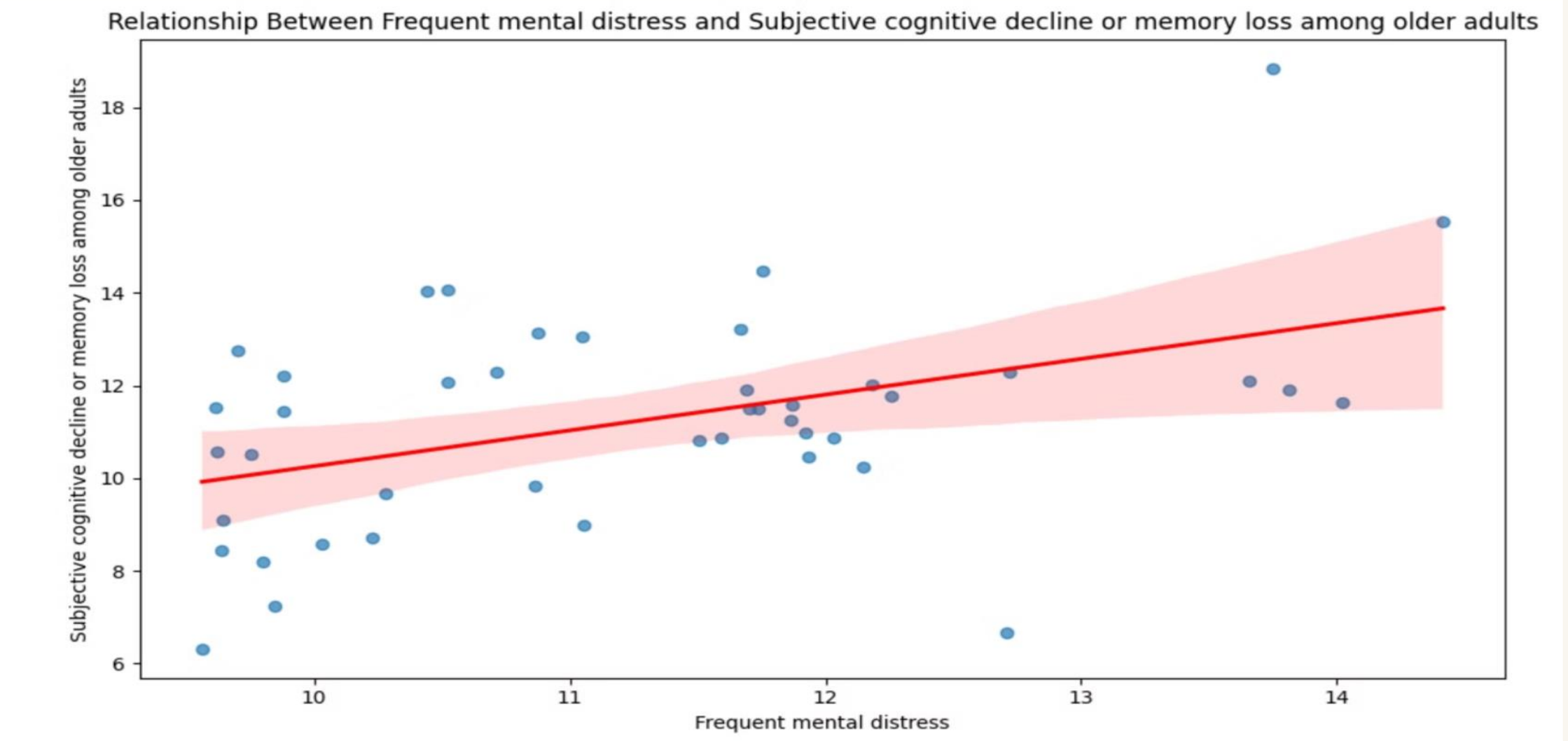
Exploring Obesity as a Central Health Concern Across Age Groups



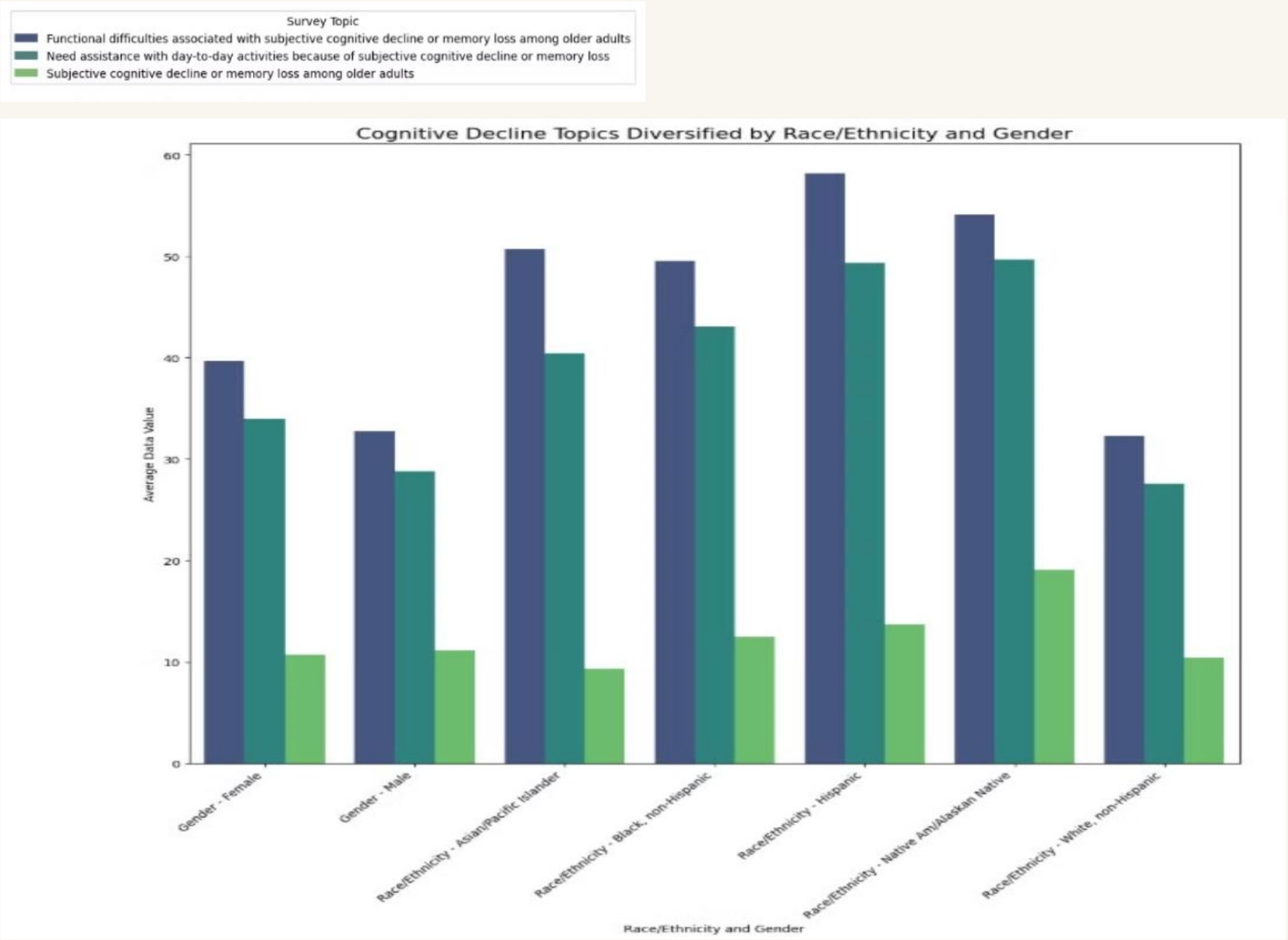
Comparison of Leisure Inactivity and Obesity Rates Across U.S. States



Relationship Between Frequent Mental Distress and Subjective Cognitive Decline Among Older Adults



Cognitive Decline Topics Analyzed by Race/Ethnicity and Gender Among Older Adults



Conclusion

- States with higher leisure inactivity percentages often exhibit higher obesity rates, emphasizing the influence of sedentary lifestyles on obesity trends.
- The strong correlation between frequent mental distress and cognitive issues calls for immediate action and targeted interventions.
- This study highlights significant health disparities across age groups, with older adults facing notable challenges in mental health and cognitive decline.

Questions?

Thank you for your time! We welcome any questions.