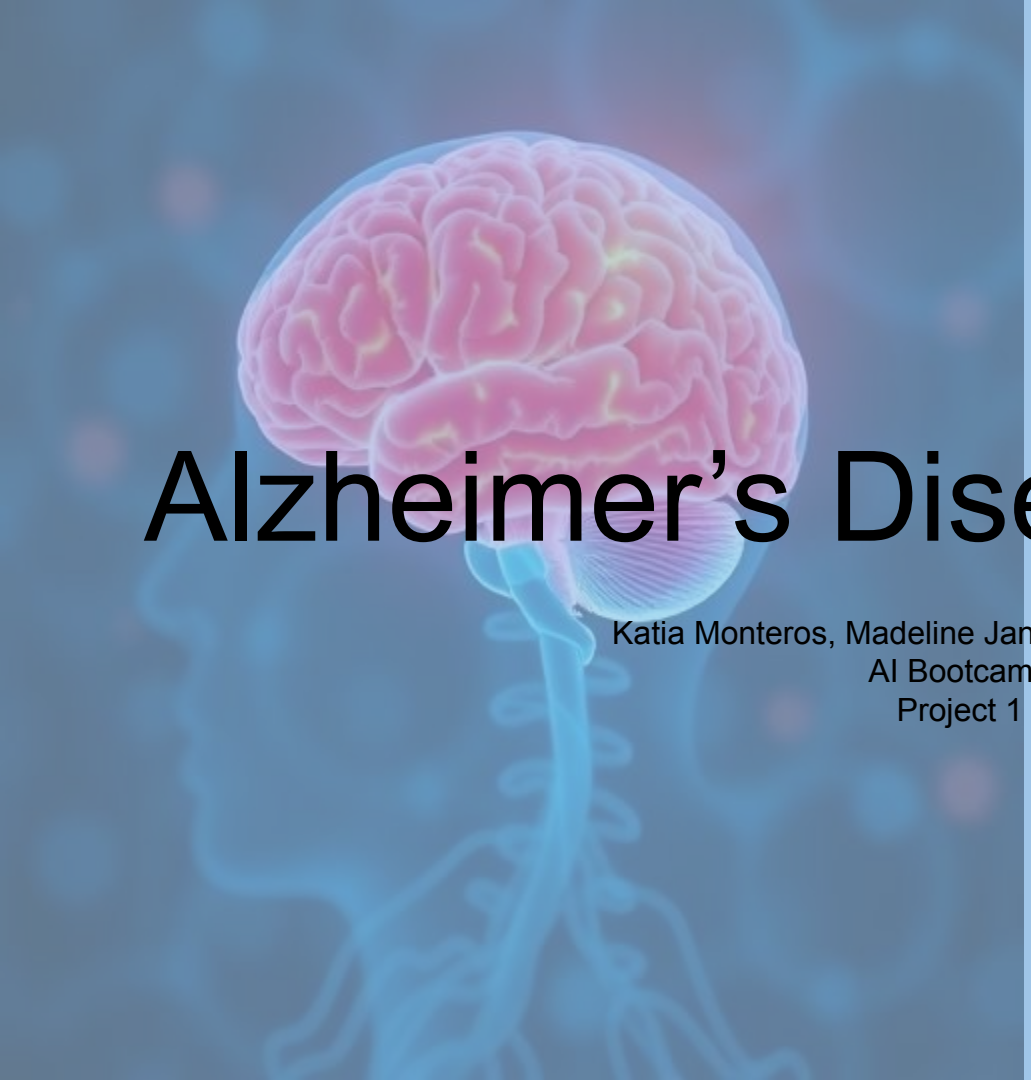


Alzheimer's Disease Analysis

Katia Monteros, Madeline Jankowski, James Kim
AI Bootcamp
Project 1

Dementia and Alzheimer's Overview:

- ➔ Dementia is an umbrella term used to describe different brain disorders that affect memory, thinking, behavior and emotion
- ➔ Alzheimer's disease is a progressive brain disorder and the most common cause of Dementia
  - ◆ The buildup of a protein called 'beta-amyloid' outside of brain cells and tangled strands of a protein called 'tau' inside these cells are key features of the condition. This is often linked to the death of brain cells and harm to brain tissue. In addition, there can be inflammation and shrinkage of the brain tissue
- ➔ There is a spectrum of Alzheimer's, ranging from asymptomatic to severe

Dementia and Alzheimer’s Overview:

- ➔ In the US alone, there is nearly 7 million people living with Alzheimer's
- ➔ The medical community is starting to understand more about the disease, but there is currently no known prevention or cure
- ➔ Early detection is critical as it allows for better care

Project Overview

- This project focuses on the analysis of an Alzheimer's dataset to identify trends in demographics, lifestyle factors, and health-related risks associated with the disease.
- In the healthcare industry, there is active research studying this disease and a focus on preventative care as an aging population.

Problem Statement

- The cause of Alzheimer's is still largely unknown
- There is a growing elderly population - in 2022, almost 58 million people were over 65, and all of the Baby Boomer population will be over 65 by 2030.
- Almost 62% of healthcare practitioners worldwide incorrectly identify dementia as a normal part of aging

Hypothesis

Before analyzing the dataset, we hoped to see patterns to answer:

- ➔ Which health conditions are most related to Alzheimer’s?
- ➔ What lifestyle factors factors contribute most to Alzheimer’s?
- ➔ What can we glean from the numbers of patients with symptoms versus those who are diagnosed?

Initial Thoughts

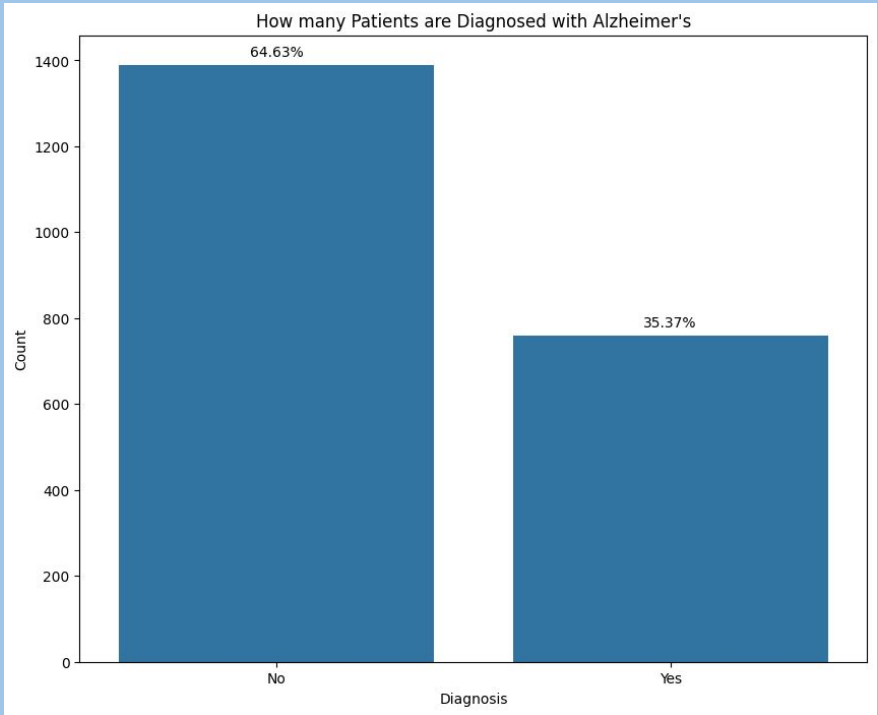
- ➔ Even though there is no definitive cause for why people get Alzheimer's, research suggests that factors like demographics, lifestyle, medical history, and cognitive function play a role in Alzheimer’s.

Data Set Analysis

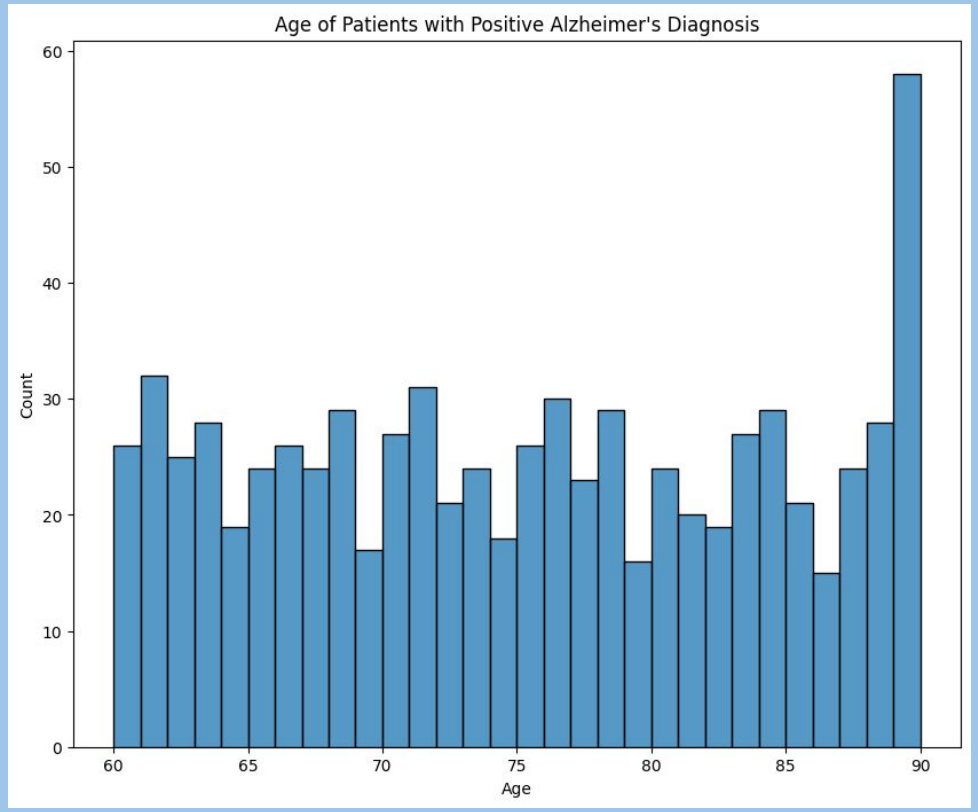
The dataset consists of 2,149 patients and 35 columns. It covers a range of factors that may contribute to Alzheimer's Disease, including demographic details, lifestyle factors, medical history, and reported symptoms.

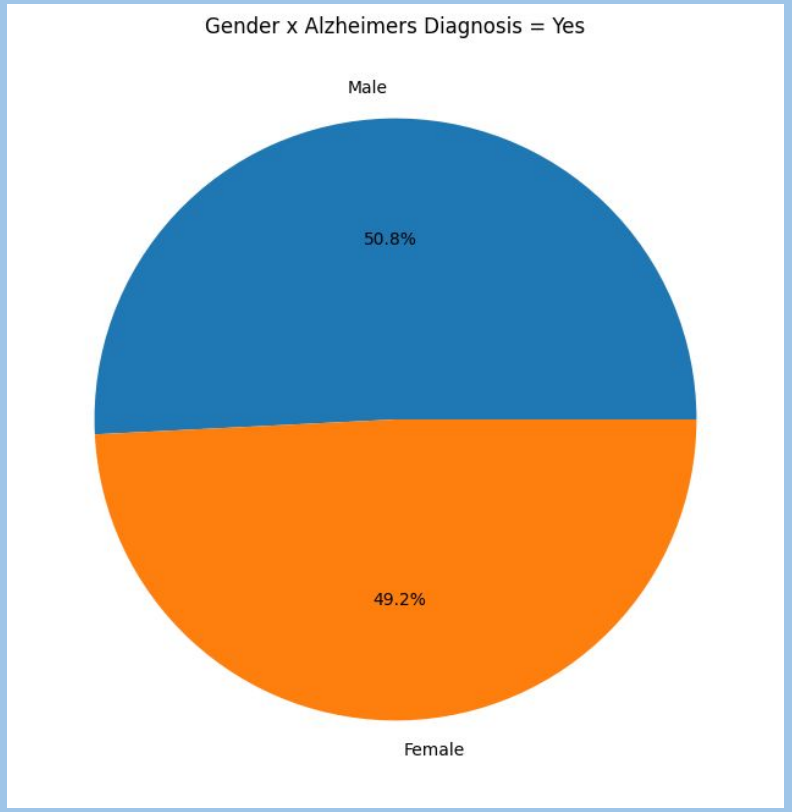
Data Collection & Cleanup:

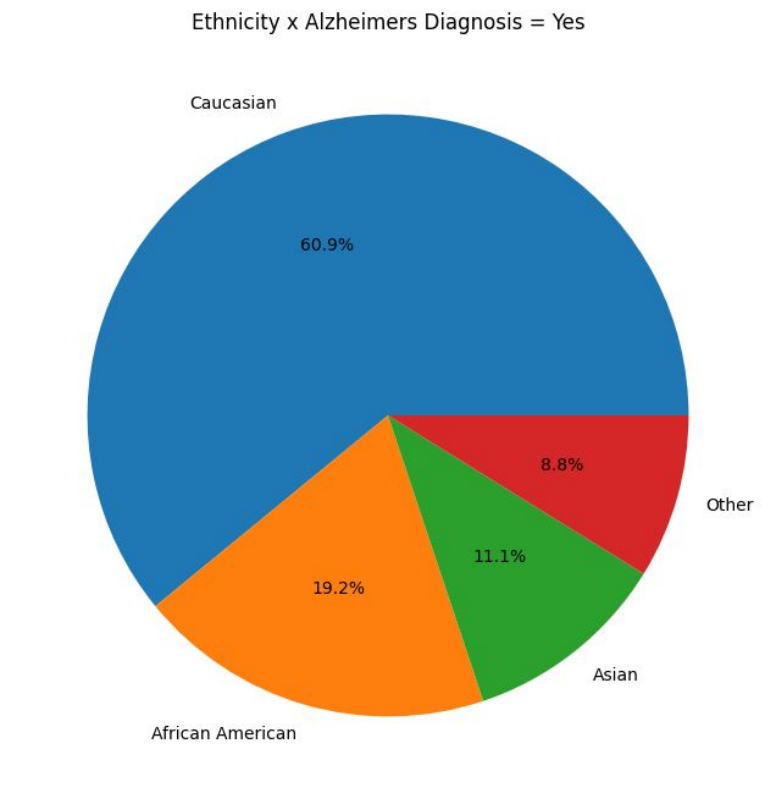
- ➔ The data was collected from Kaggle, a publicly available data set platform
- ➔ The data cleanup consisted of:
  - ◆ Initial data inspection
  - ◆ Missing data
  - ◆ Irrelevant data
  - ◆ Data formatting
  - ◆ Separating data into subcategories
  - ◆ Finding sum and mean
  - ◆ Filtered and grouped based on diagnosis



Demographics



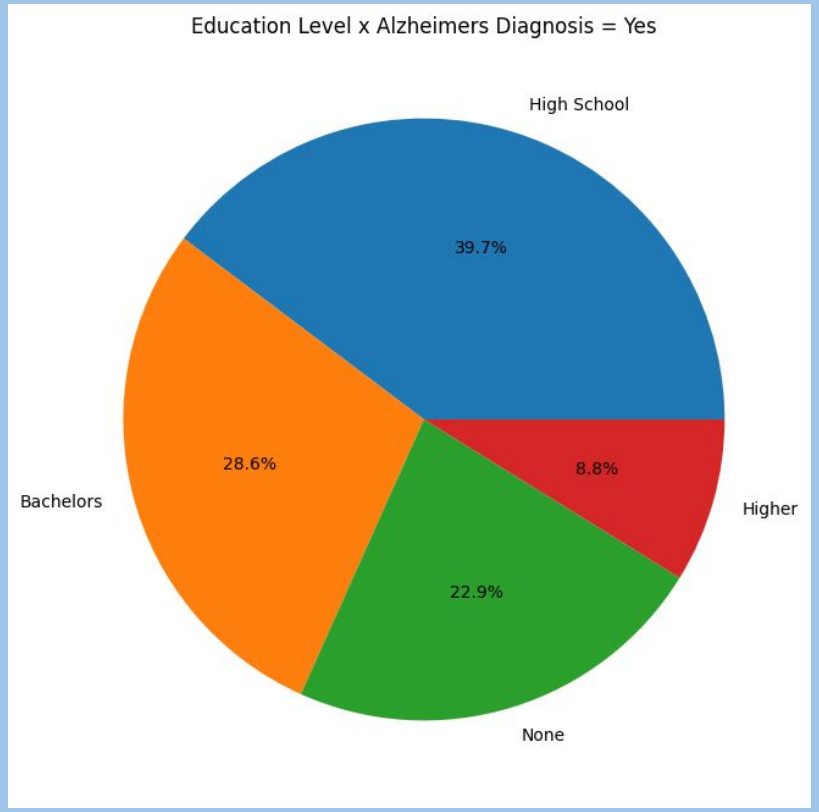




Ethnicity

- **Caucasian:** 60.9%
- **African American:** 19.2%
- **Asian:** 11.1%
- **Other:** 8.8%

<u>Note:</u> Dependent on where the study was taken like country, location of hospital, etc.



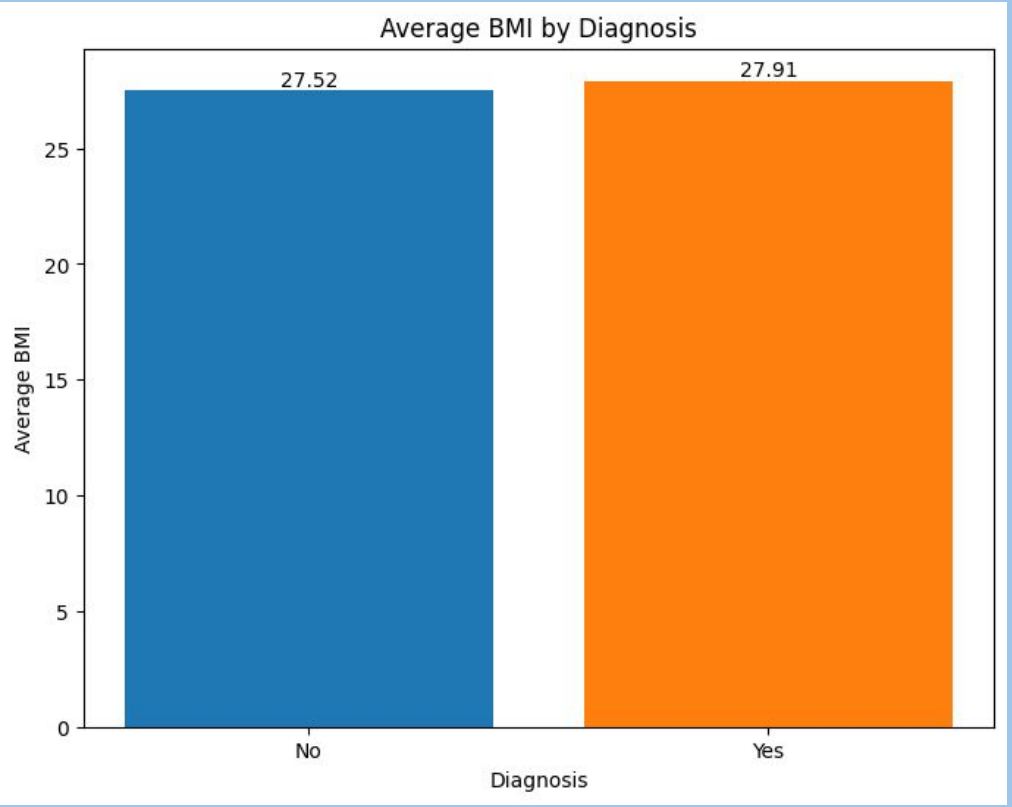
Education Level

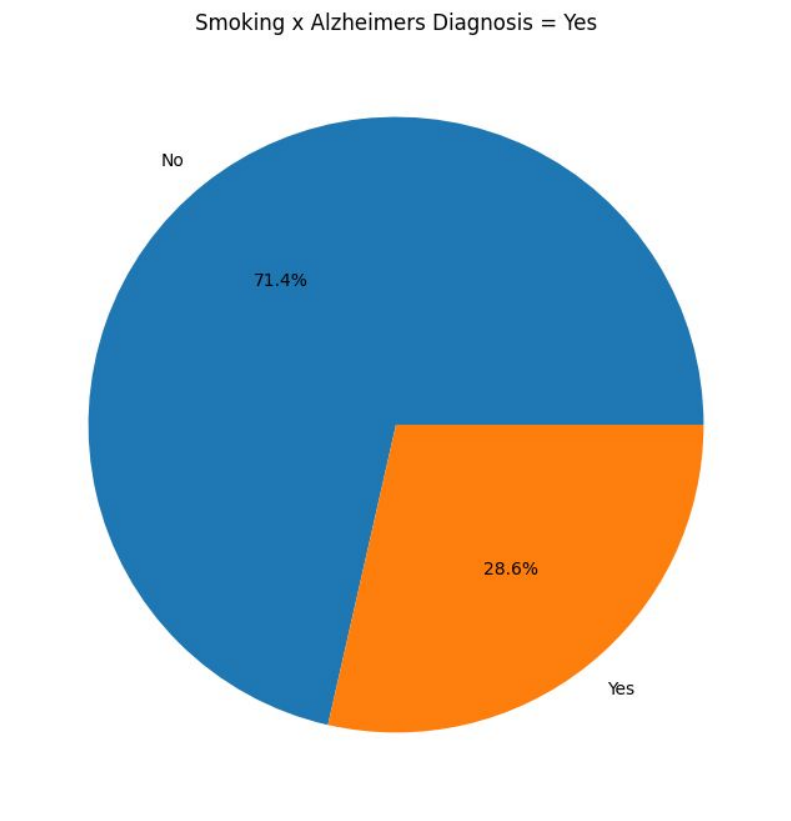
- ➔ **None:** 22.9%
- ➔ **High School:** 39.7%
- ➔ **Bachelors:** 28.6%
- ➔ **Higher:** 8.8%

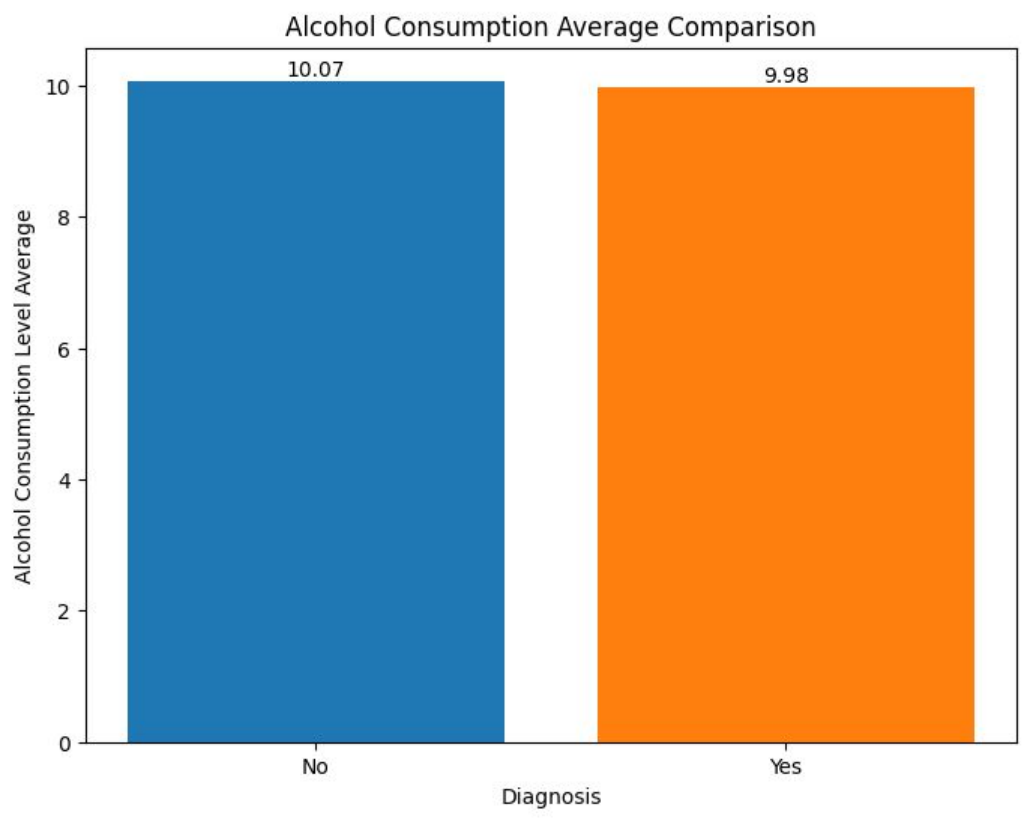
<u>Note:</u> Dependent on where the study was taken like country, location of hospital, etc.

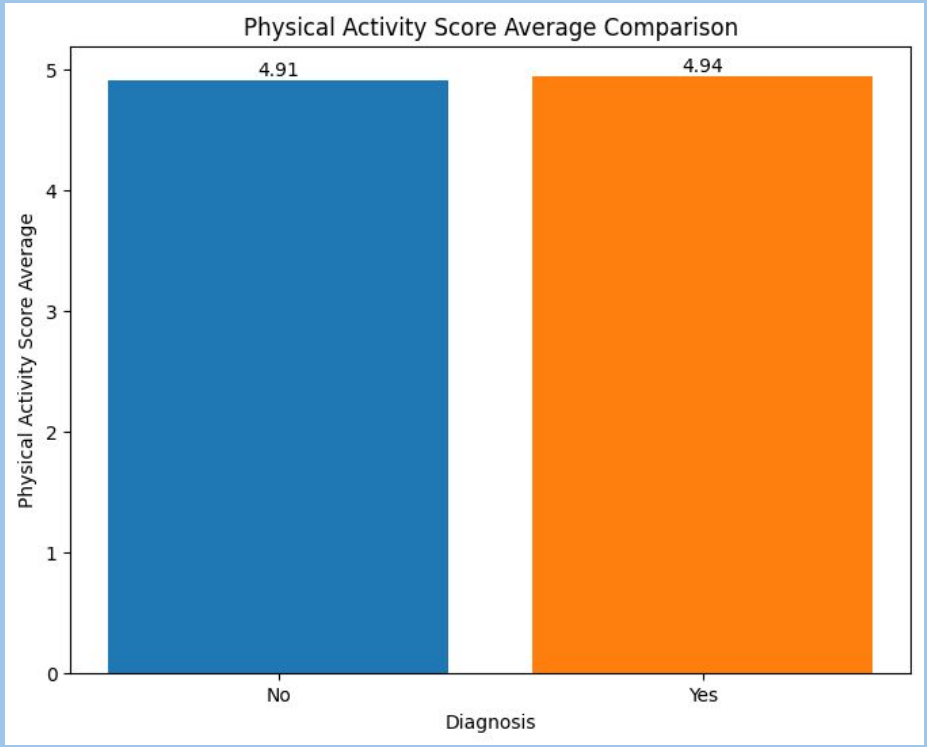
Lifestyle

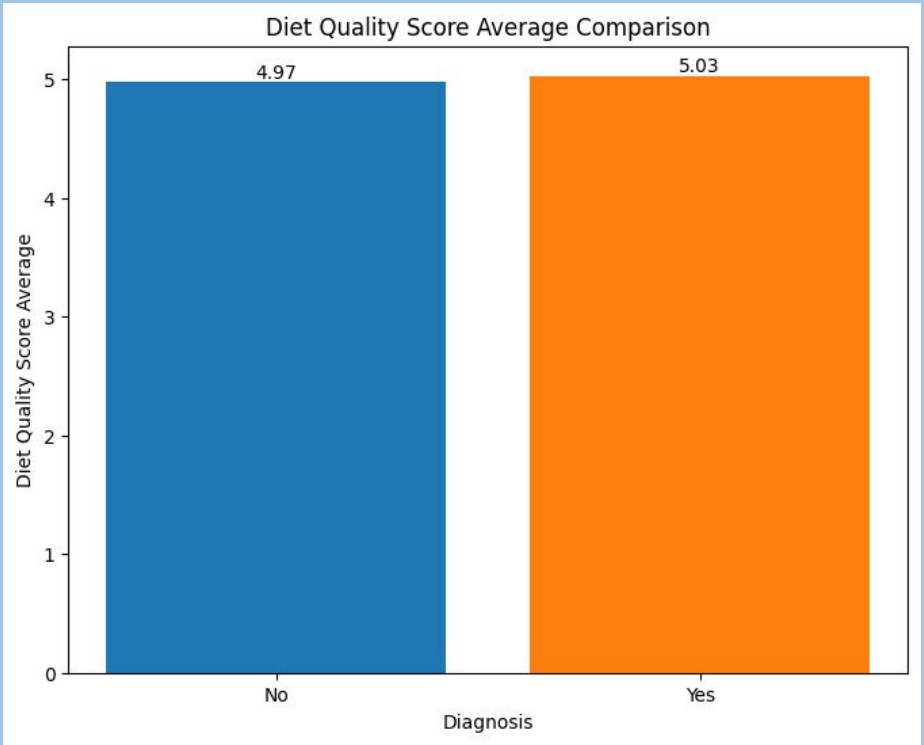
- ➔ Average BMI
- ➔ Alcohol Consumption
- ➔ Smoking (Y/N)
- ➔ Physical Activity Score
- ➔ Diet Quality Score
- ➔ Sleep Quality Score




Smoking x Alzheimers Diagnosis = Yes
No
71.4%
28.6%
Yes


Alcohol Consumption Average Comparison
10.07
9.98
Alcohol Consumption Level Average
0
2
4
6
8
10
No
Yes
Diagnosis


Physical Activity Score Average Comparison
4.91
4.94
Physical Activity Score Average
0
1
2
3
4
5
No
Yes
Diagnosis

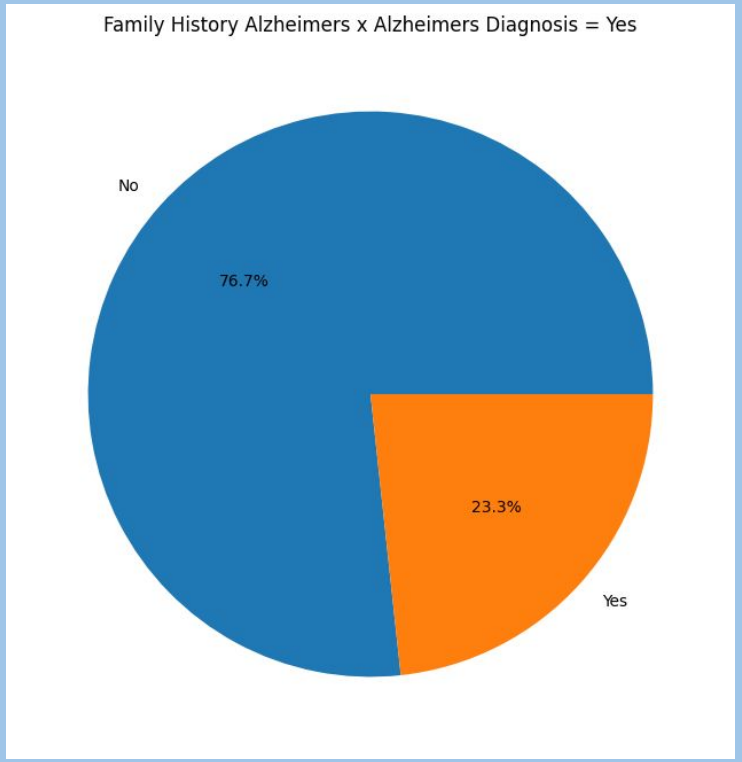

Diet Quality Score Average Comparison
4.97
5.03
Diet Quality Score Average
0
1
2
3
4
5
No
Yes
Diagnosis

Lifestyle Conclusion

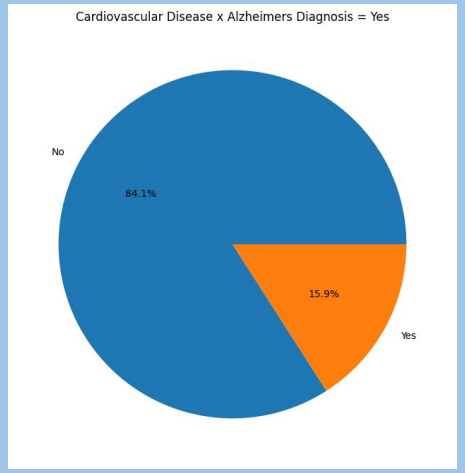
- ➔ All lifestyle factors showed similar or equal results between the 2 groups
- ➔ Sleep quality score was the only factor that showed a small difference
- ➔ Not consistent to what we have researched or read in other sources.

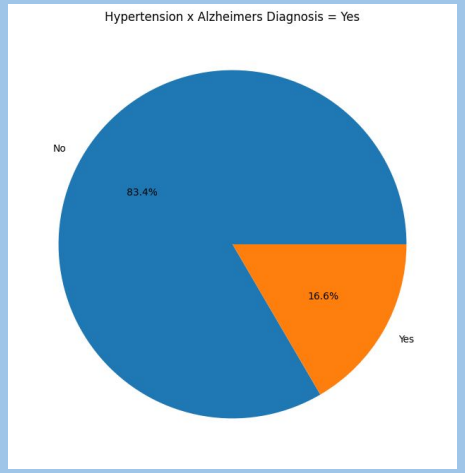
Other Source: A population-based study of 1,500 people on aging and dementia in Finland found that those who engaged in leisure-time physical activity at least twice a week through middle age, carried half the risk of developing dementia and a 60 percent lower risk of developing Alzheimer's than did those who remained sedentary in their midlife.

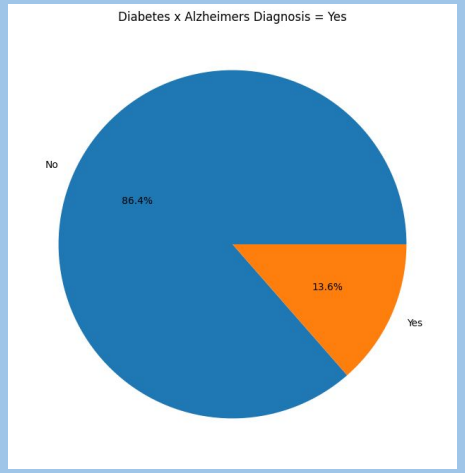
Medical History



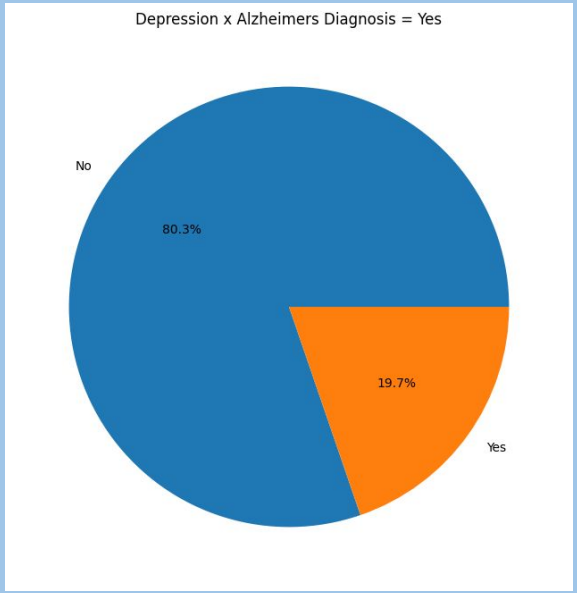
- ➔ Family History of Alzheimer's
- ➔ Cardiovascular Disease
- ➔ Diabetes
- ➔ Depression
- ➔ Head Injury
- ➔ Hypertension

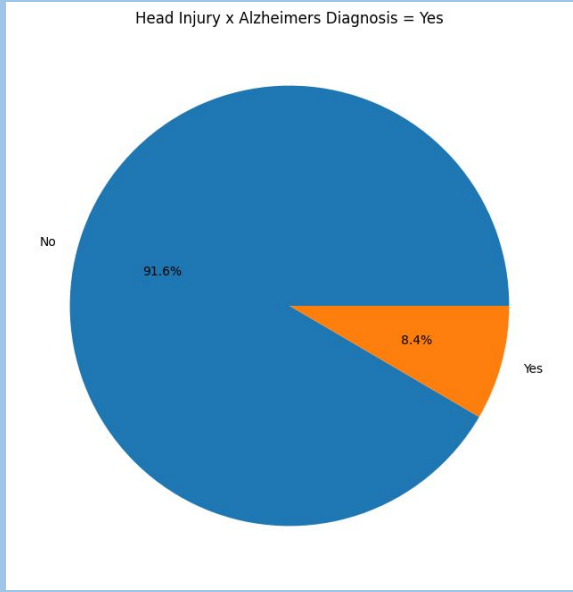






- ➔ Family History of Alzheimer’s
- ➔ Cardiovascular Disease
- ➔ Diabetes
- ➔ Hypertension
- ➔ Head Injury
- ➔ Depression





- ➔ Family History of Alzheimer’s
- ➔ Cardiovascular Disease
- ➔ Diabetes
- ➔ Hypertension
- ➔ Head Injury
- ➔ Depression

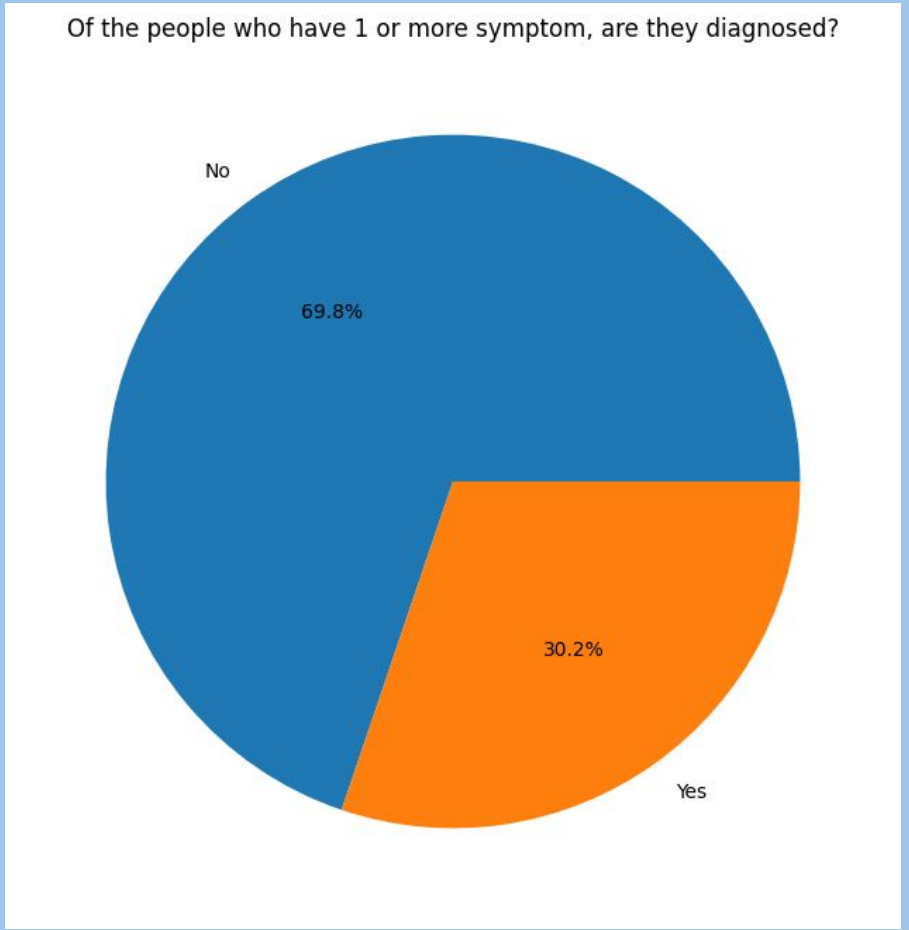
Medical History Conclusion

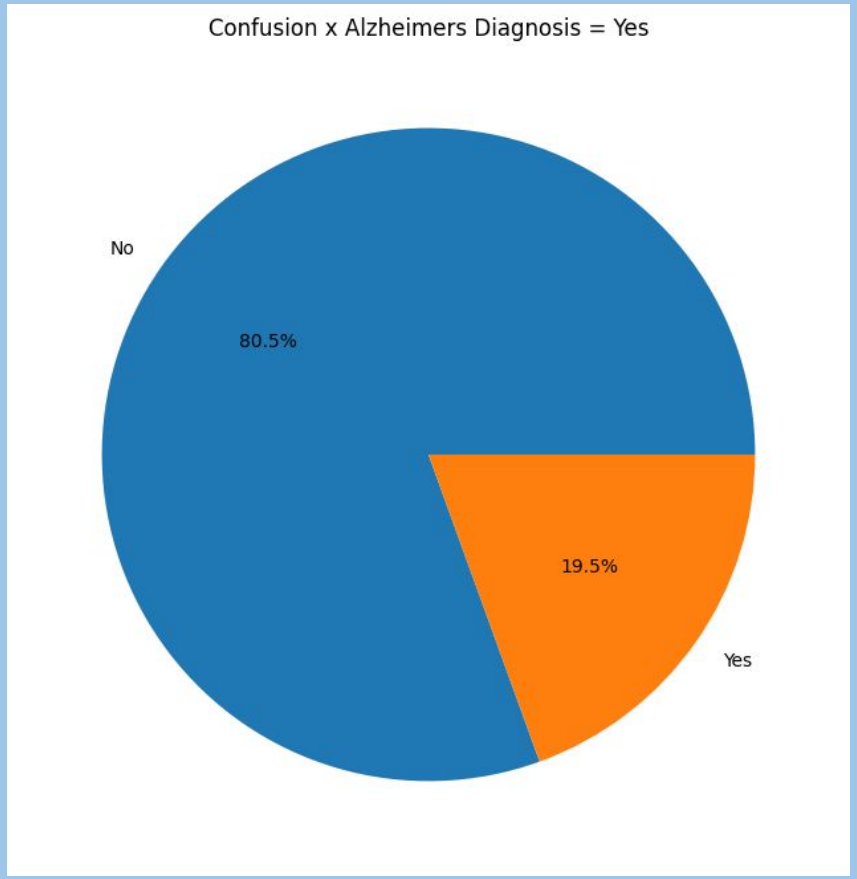
- ➔ There was minimal correlation between medical history and diagnosis of Alzheimer's in our dataset
- ➔ Research has shown these comorbidities impact the risk of Alzheimer's:
  - ◆ Family history/genetic - people with a different for of the APOE gene increases the risk of Alzheimer's
  - ◆ Head trauma - especially over the age of 50 and the number of injuries
  - ◆ Hypertension
  - ◆ Diabetes

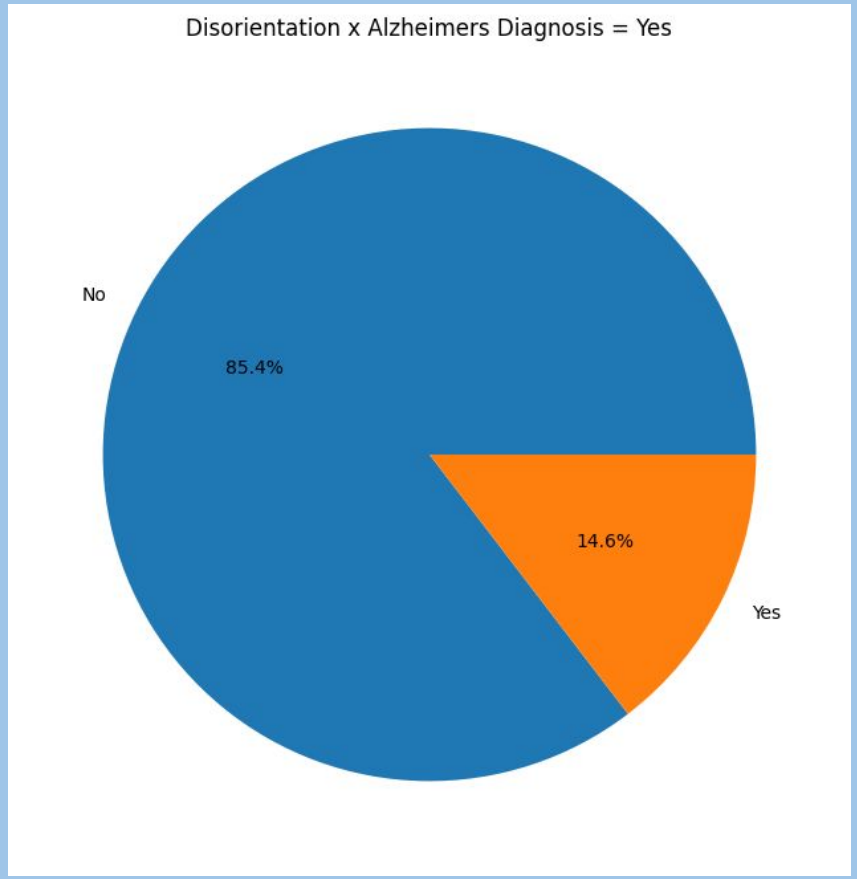
https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/symptoms-causes/syc-20350447

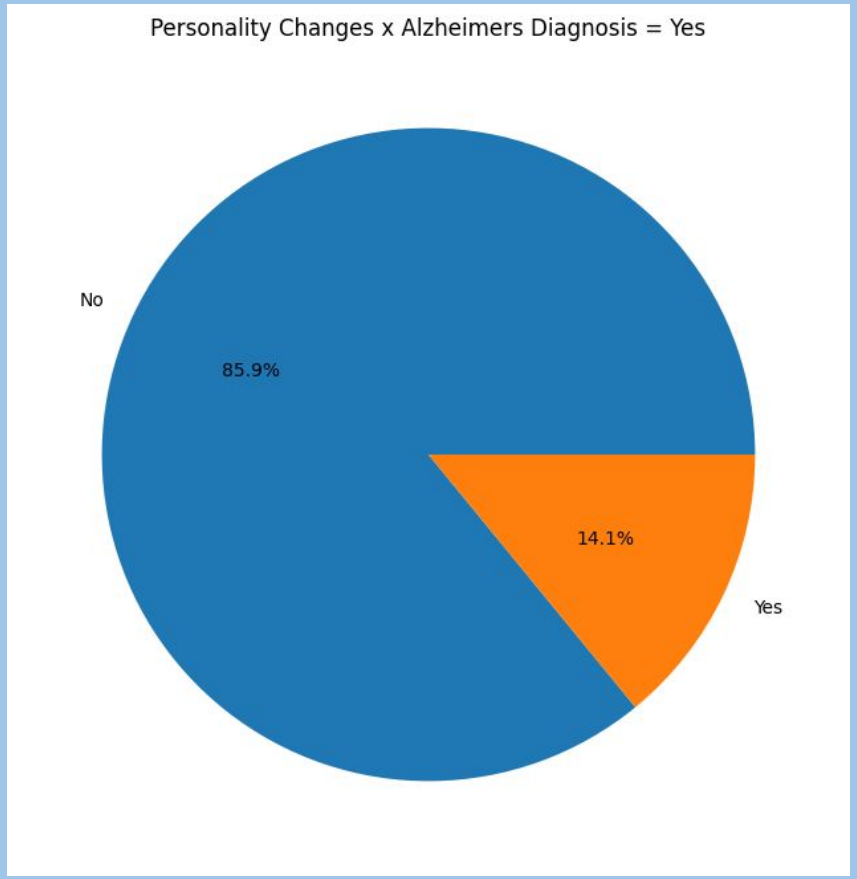
Symptoms

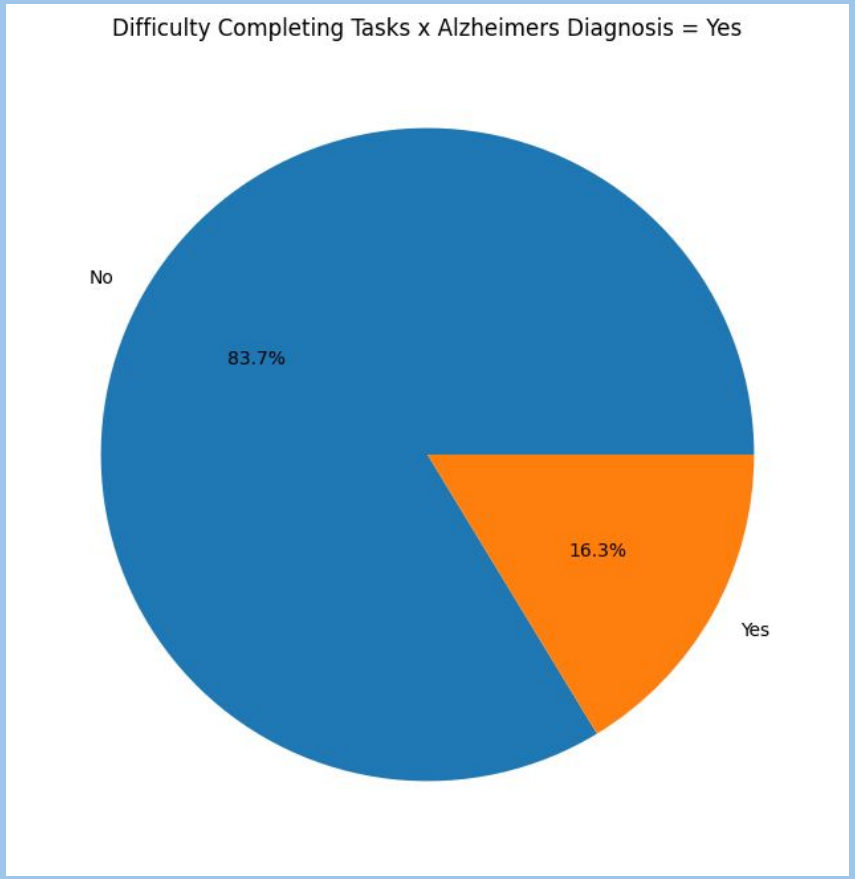
- ➔ Confusion
- ➔ Disorientation
- ➔ Personality Changes
- ➔ Difficulty Completing Tasks
- ➔ Forgetfulness
- ➔ Memory Complaints
- ➔ Behavioral Problems

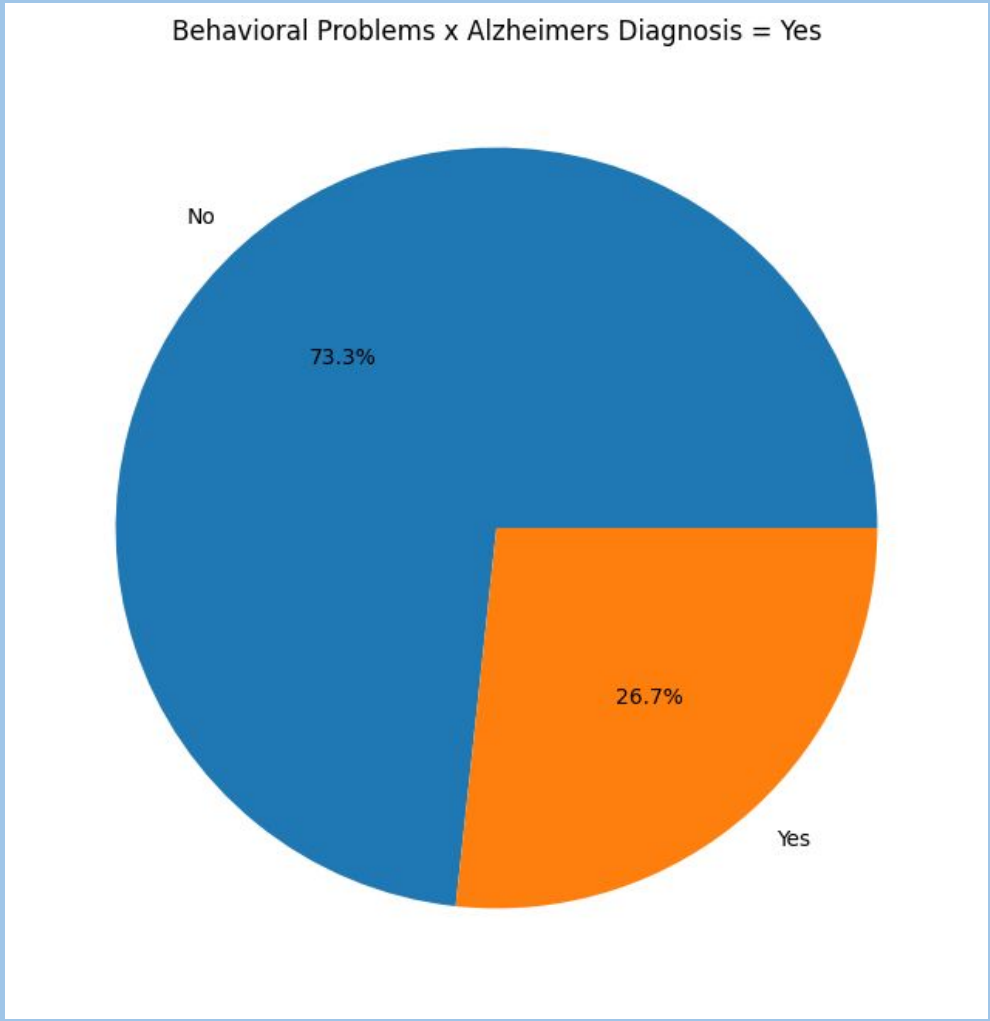



Confusion x Alzheimers Diagnosis = Yes
No
80.5%
19.5%
Yes


Disorientation x Alzheimers Diagnosis = Yes
No
85.4%
14.6%
Yes


Personality Changes x Alzheimers Diagnosis = Yes
No
85.9%
14.1%
Yes

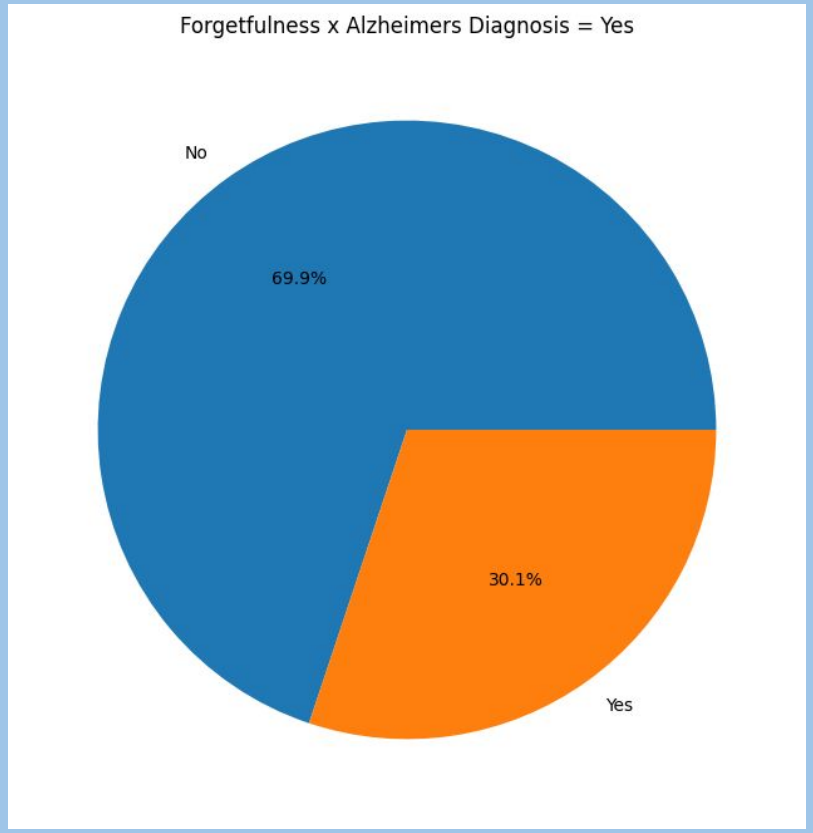

Difficulty Completing Tasks x Alzheimers Diagnosis = Yes
No
83.7%
16.3%
Yes

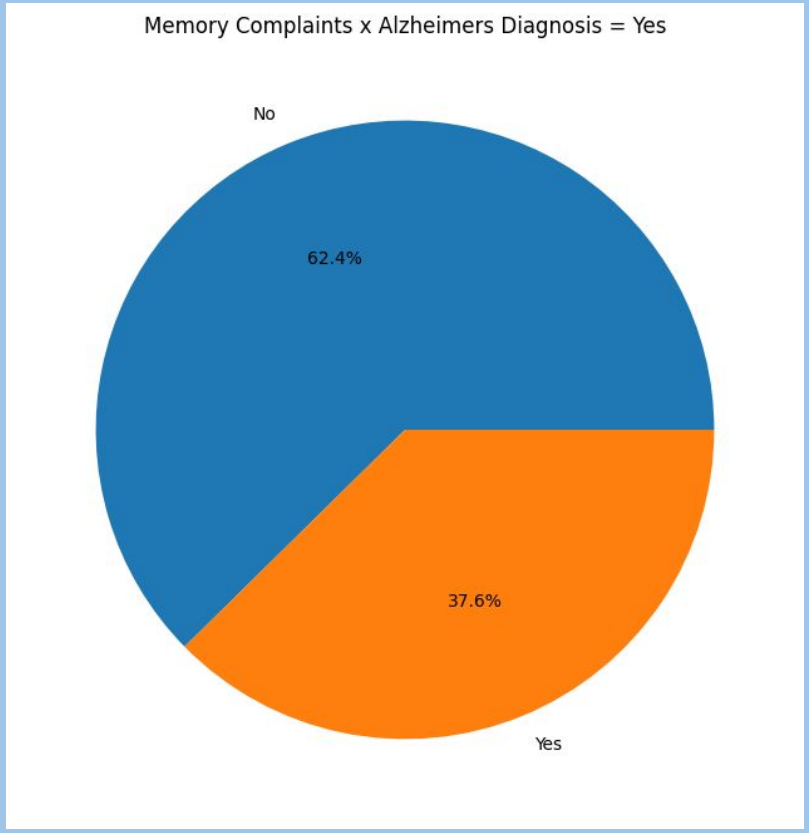


When comparing the three different symptoms:

- ➔ Personality Changes
- ➔ Difficulty Completing Tasks
- ➔ Behavioral Problems

Personality Changes and DCT show very similar results while Behavioral Problems have a 10-12% increase of people who had BP symptoms vs the other two.





➔ Forgetfulness and Memory Complaints appear similar, but there is a 7.5% increase with memory complaints among those who have Alzheimer's.

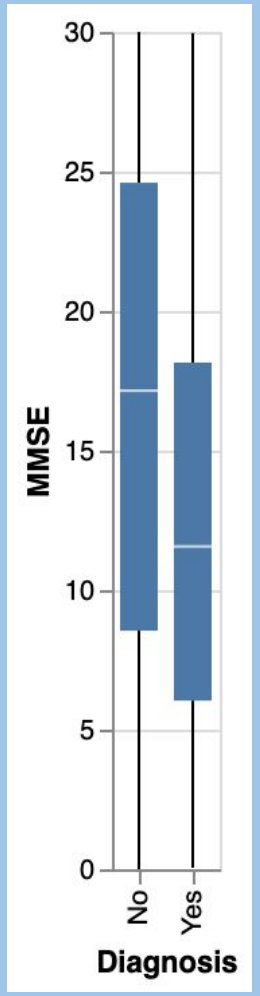
Symptoms Conclusion

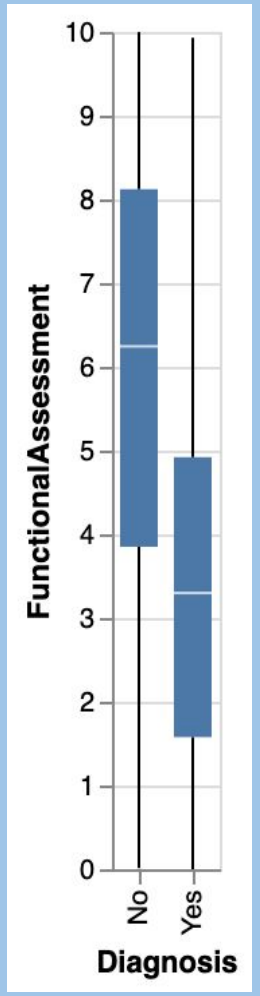
- ➔ Only compared the data that had all the patients that were diagnosed with Alzheimer's Disease
- ➔ There were way more patients that did not have the symptom that was being measured.
- ➔ When comparing these results to something like lifestyle there are at least noticeable changes in percentages
- ➔ Still not consistent with with other sources and findings say about the main symptoms an individual can have.

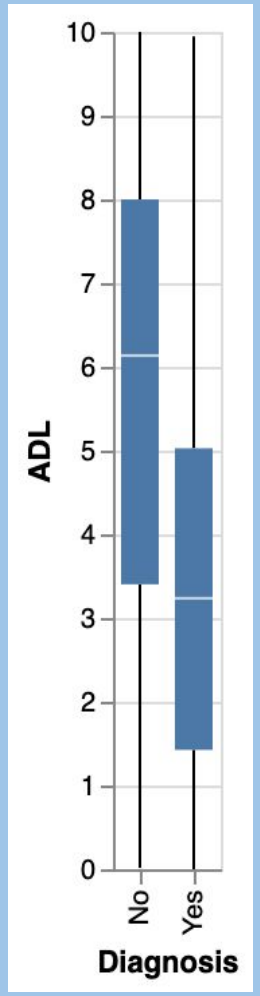
Other Source Summary: There are different symptoms (can be similar) when it comes to the severity of the patient's Alzheimer's condition: mild, moderate, and severe. The consistent ones are memory loss, difficult to organize and complete tasks, inability to learn new things, getting lost, etc.

Cognitive Assessments

- ➔ Cognitive Assessments were the strongest findings in the dataset analysis.
- ➔ We recommend that individuals over the age of 60 undergo regular Cognitive Assessments as a preventative measure to catch early signs of cognitive decline.







Importance of Administering Cognitive Assessments

Early Detection

- ➔ Alzheimer's symptoms are subtle and not immediately noticeable
- ➔ Cognitive Assessments can detect the disease before other symptoms show
- ➔ Patients and families can prepare to make decisions about treatment and care

Awareness

- ➔ Patients and caregivers can understand the severity of the decline
- ➔ Many symptoms are dismissed as normal signs of aging, but tests provide further evidence of the disease

Tracking

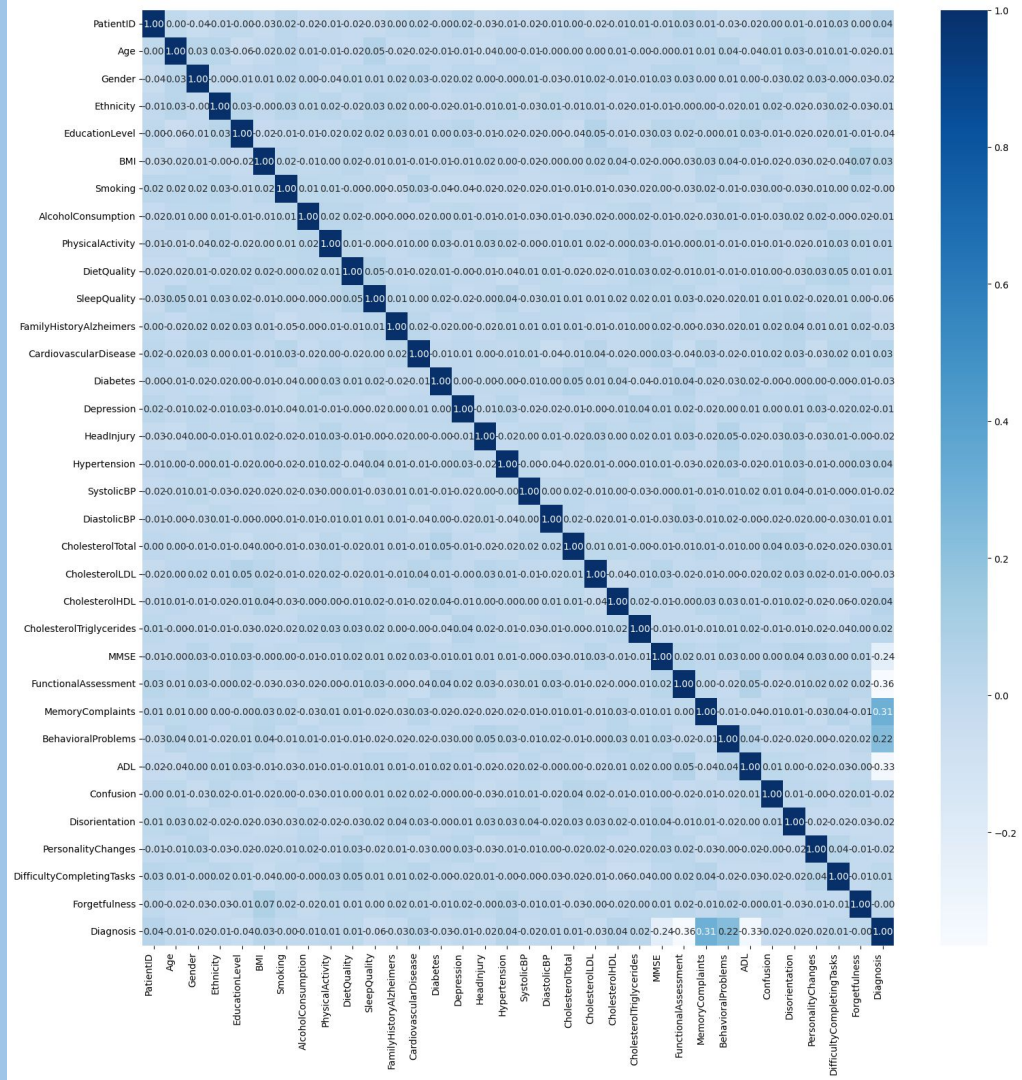
- ➔ Track the progression of Alzheimer's over time through repeated testing
- ➔ Treatment plans and care strategies can be adjusted for a more personalized experience

Clinical Measurements

- ➔ There were clinical measurements of blood pressure readings (systolic and diastolic) and readings of cholesterol measurements
- ➔ There was no correlation between clinical measurements in our dataset and the diagnosis of Alzheimer's
- ➔ In a future study, the A1C, a measure of blood sugar, would have been helpful to see a trend for diabetic patients with Alzheimer's

Heat Map

- This Heat Map helps us quickly identify strong relationships between factors and which ones were most influential in predicting Alzheimer's disease.
- This Heat Map shows the Cognitive Assessments to be the strongest correlations with Alzheimer's diagnosis.



Limitations

- Demographic Restriction and Sampling Bias:
  - Age: 60-90
  - Ethnicity: Location of data collection
  - Education Level: Location of data collection
- Inconsistent Data Entries:
  - Some Blood Pressure (Systolic BP and Diastolic BP) entries were reversed or entered wrong
- The data is not prospective
  - Alzheimer's is a progressive disease and tracking patients over time could provide better insights over time
- The dataset did not represent the same trends seen in Alzheimer's literature

Potential Next Steps

- ➔ Use of a second dataset with similar criteria with a bigger patient population
  - ◆ Possibility of seeing more trends in data with more diversity of patients
- ➔ Training a machine learning algorithm on the “diagnosis = yes” patients with this dataset and then using another data set to validate and see if trends could be seen
- ➔ Conducting a prospective trial to establish a time series dataset with the same values with each visit to see how patients do over time

Conclusions

Project aim: to uncover trends in an array of factors that could contribute to the risk of Alzheimer's in a dataset of 2159 patients

Findings:

- ➔ No significant correlation of any one factor to the presence of Alzheimer's
- ➔ Our hypothesis was that poor lifestyle factors would show the most correlation
- ➔ This suggests that the complexity of Alzheimer's etiology might require a broader or different set of data to uncover meaningful insights.

Next steps:

- ➔ Use other datasets to bolster patient populations and broaden dataset
- ➔ Use machine learning to develop algorithms to predict Alzheimer's using the "diagnosis = yes" dataset as training and other datasets for validation