**Capstone Project Proposal**

**Alex Putz**

**Context**

Data relating to patient health status or the delivery of healthcare is routinely collected from a range of sources and is increasingly being used in clinical decision-making. Organizations are leveraging EHR data, patient registries, and mobile device information to better understand trends and outcomes, leading to improved care delivery.

By leveraging data and technology, the health care sector and its social care partners have an opportunity to improve the efficiency, effectiveness, and sustainability of efforts that address health-related social needs as a regular component of health care delivery.

**Data Set**

https://www.kaggle.com/datasets/shashwatwork/dementia-prediction-dataset

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Name | Description | Data Type | Non-null Count | Type of Variable |
| Age | This is an integer indicating the age of the participant | int64 | 1338 | Continuous |
| Hand | Dominant Hand | Object | 1338 | Discrete |
| EDUC | Years of education | Int64 | 1338 | Discrete |
| SES | Socio-Economic Status with 1 being low SES and 5 being the high SES | object | 1338 | Categorical |
| eTIV | Estimated Total Intra-Cranial Volume- The estimated volume of the patient’s brain structures | Float64 | 1338 | continuous |
| nWBV | Normalized Whole Brain Volumes - whole-brain and regional volumetric measures to be normalized for head size | float64 | 1338 | Continuous |
| CDR | Clinical Dementia Rating  0- No dementia  0.5- Mild  1- Mild  2- Moderate  3- Severe | Int64 | 1338 | Categorical |
| ASF | Amplitude/Stimulus Intensity Function  A measure of the amplitude of someone’s brainwaves with stimuli. | Int64 | 1338 | Continuous |

**Questions**

Do people with higher SES have a lower likelihood of developing Dementia compared to those who have a lower SES?

Does age influence the probability of dementia?

Does the education play a role in probability of dementia?

Does estimated total intracranial volume play a role in the likelihood of developing Dementia?

**Objective Hypothesis**

1. **Does lower nWBV increase the risk of dementia:**

H0: There is no difference in CDR based on nWBV.

Ha: There is a difference in CDR between nWBV

1. **Prove (or disprove) a difference in likelihood of dementia diagnosis by ASF.**

H0: There is no difference between higher and lower ASF.

Ha: There is a difference between higher and lower ASF.

1. **Does SES relate to likelihood of dementia?**

H0: There is no difference across SES

Ha: There is a difference across SES.

1. **Does education level differ the likelihood of developing dementia**

H0: No difference across education levels

Ha: Difference across education levels

1. **Does age influence the probability of dementia?**

H0: No difference across age.

Ha: Difference across age.

1. **Does MMSE score predict the likelihood of Dementia?**

H0: No based on MMSE

Ha: Difference based on high vs low MMSE

**Possible Impact of Findings**

Provide frameworks for identifying at risk population

Identifying variables to be explored more deeply in future research.