PREDICTING ALZHEIMERS' DISEASE PROGRESSIONS USING MACHINE LEARNING

Normal   Early AD   Advanced AD
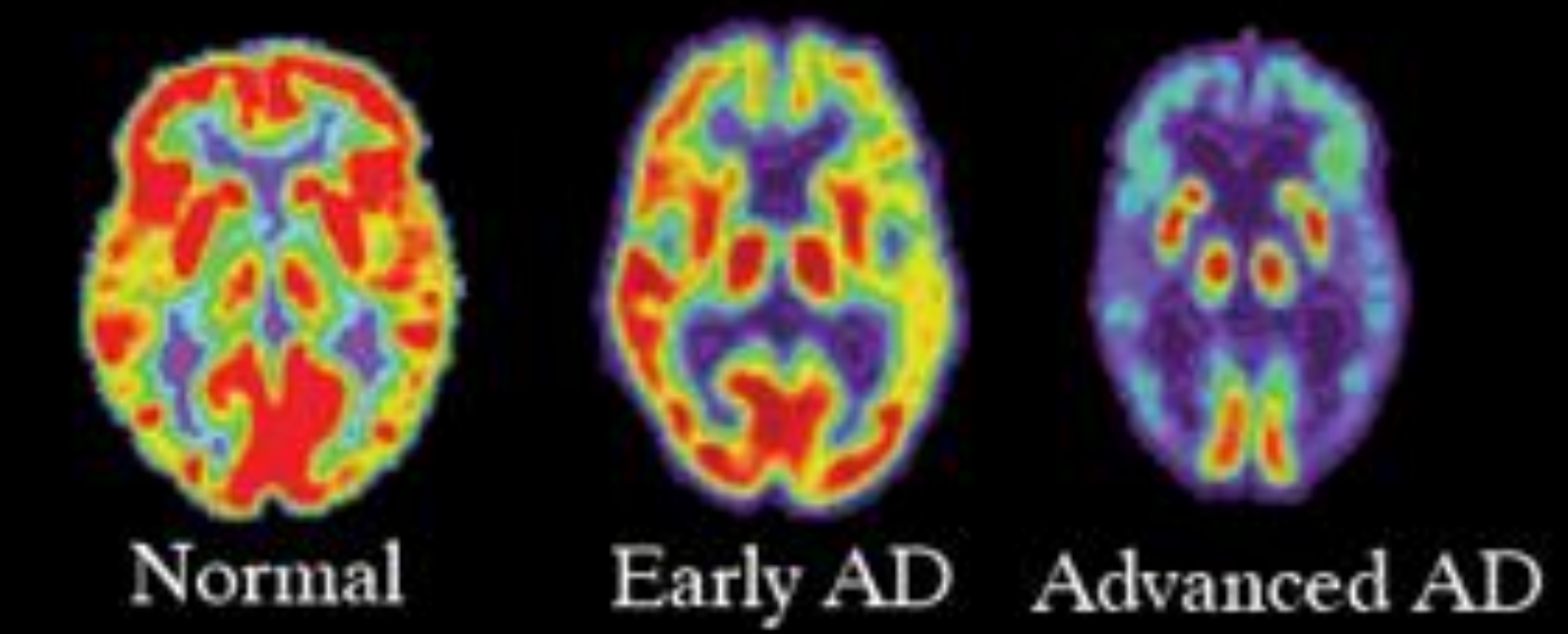
VIPUL SATONE, RACHNEET KAUR | GUIDED BY: FARAZ FAGHRI, PROF. ROY CAMPBELL | UNIVERSITY OF ILLINOIS AT URBANA CHAMPAIGN

Motivation

Describing Alzheimer's Disease through clinical tests and predicting it 's progression in early stages.

Objective

Our goal is to predict progressions in Alzheimer's Disease (AD) through only clinical features using machine learning techniques.

AD being an irreversible neurodegenerative disease, affecting an estimated 5.3 million Americans with no discovered preventive cure, the study would help doctors subduing it's effects with better information.

Dataset

- ADNI dataset: 800 subjects from 50 sites from the US and Canada.
- Some clinical features: MOCA, MODACH, NEUROBAT etc.
- Labels present: CN - Normal Aging, SMC - Significant Memory Concern, EMCI - Early Mild Cognitive Impairment, MCI - Mild Cognitive Impairment, LMCI – Late MCI, AD - Alzheimer's disease
- Visits: sc (Baseline), m06 (6-month), m12 (12 months) and so on.

Descriptive Statistics
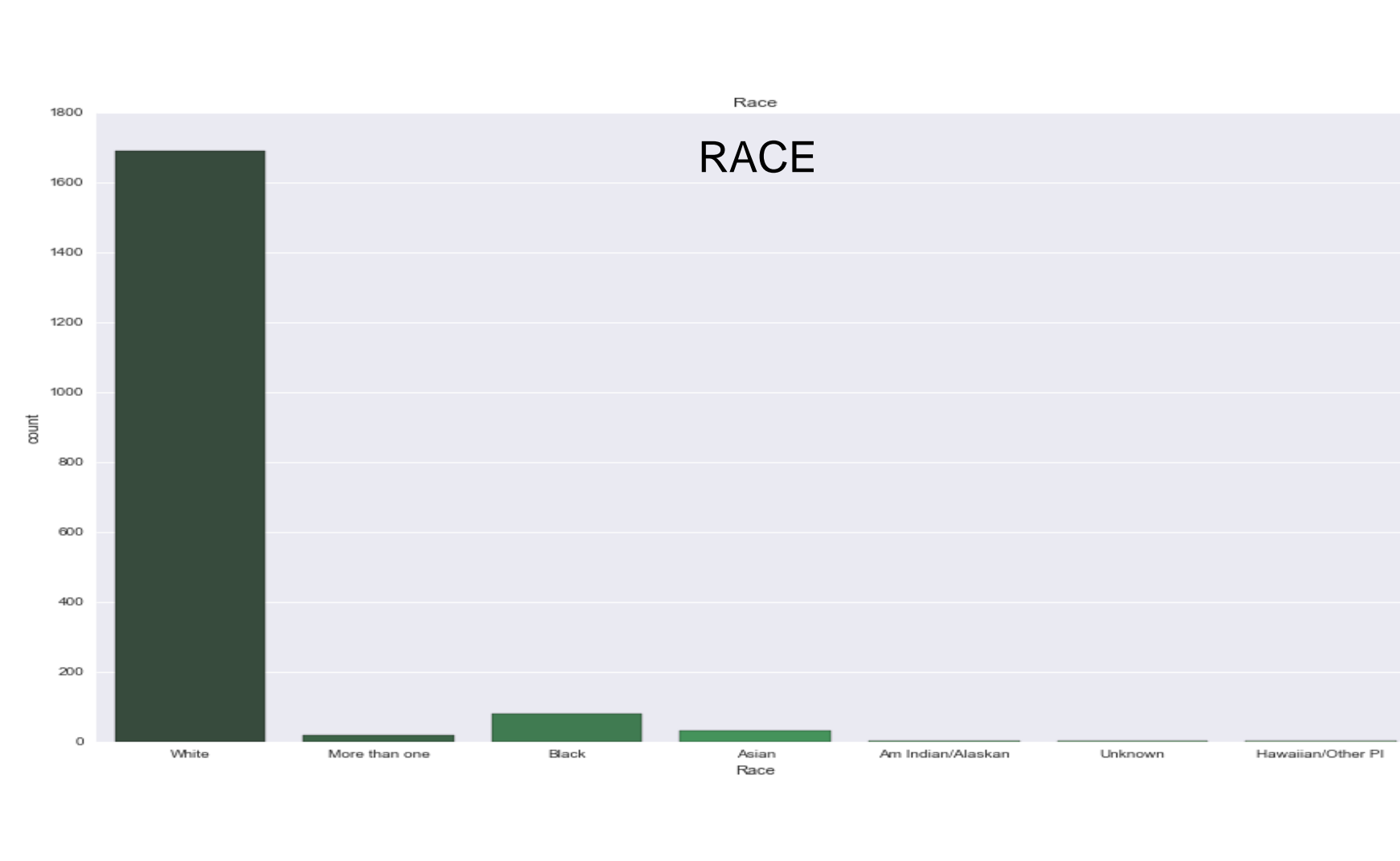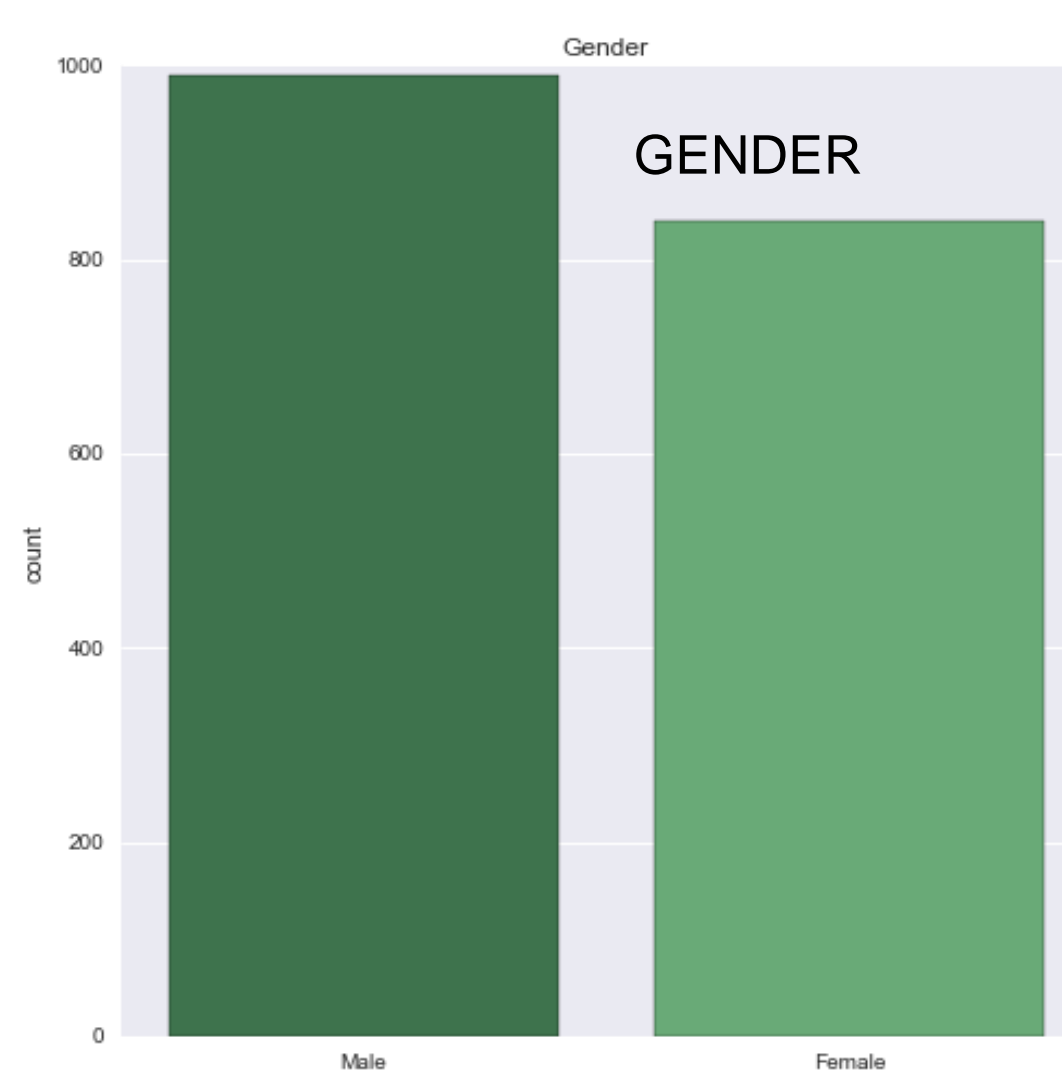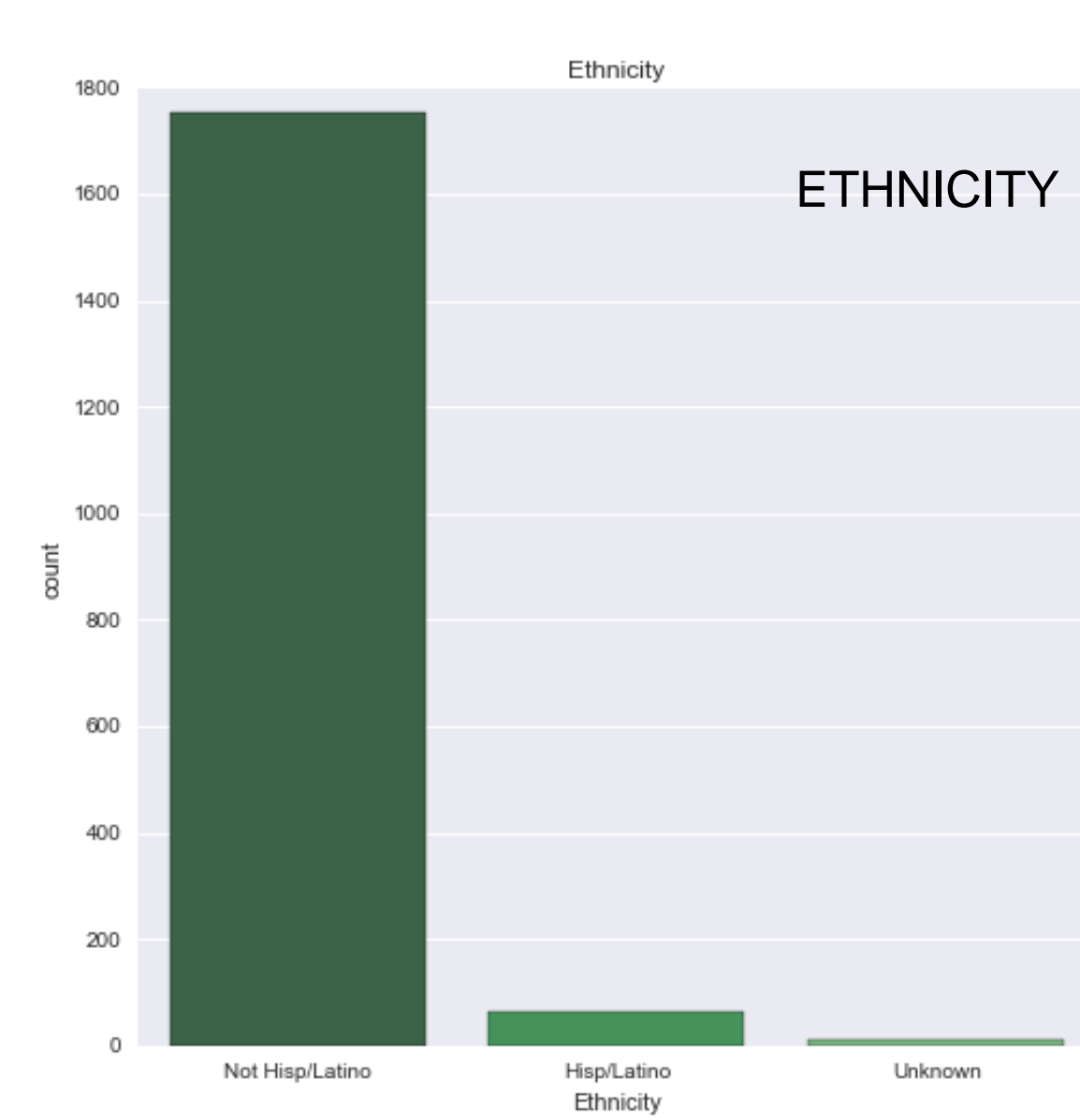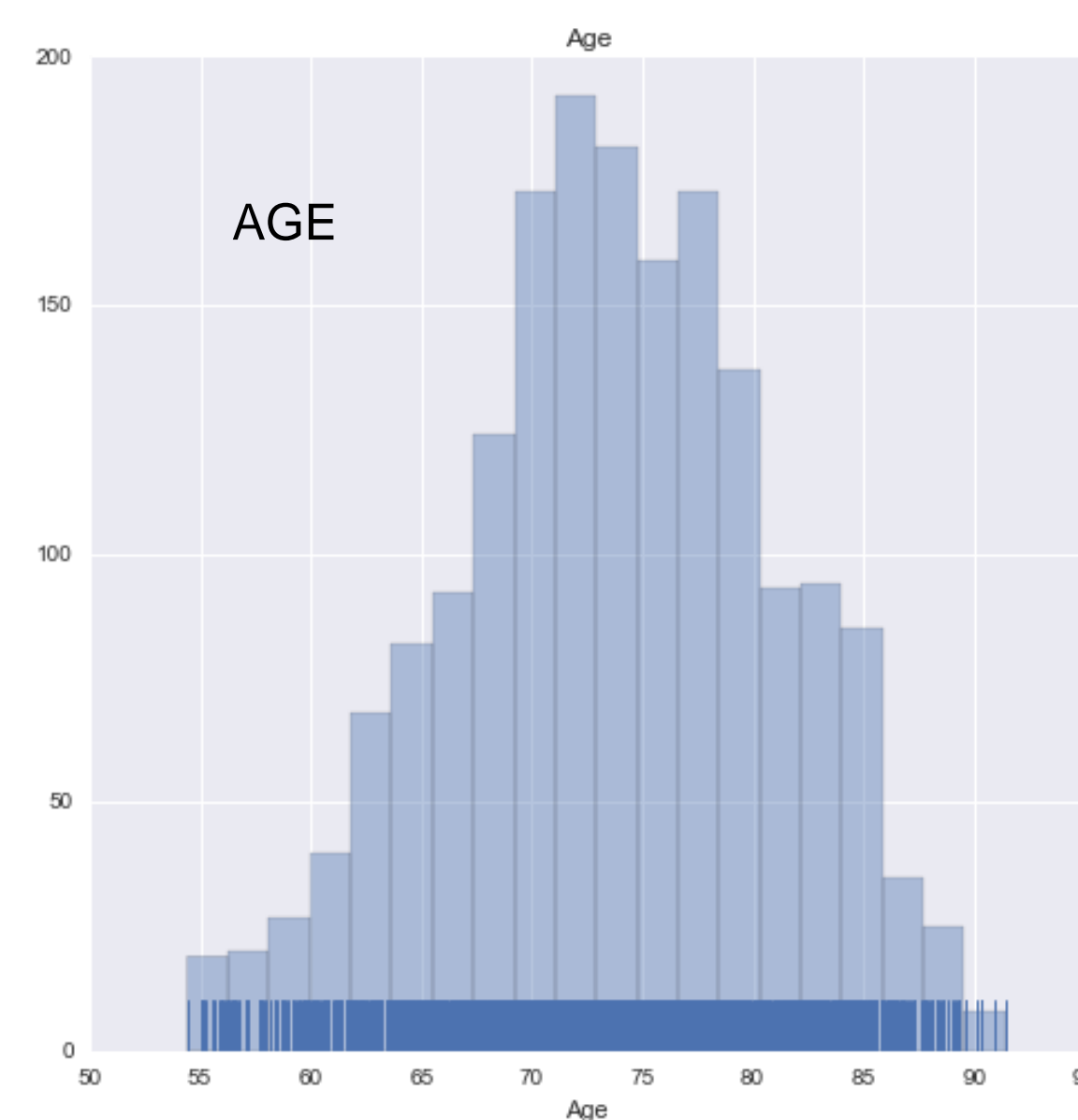
Procedure

Preprocessing

1. Data Cleaning:
   I. Reduced to 21 relevant datasets
   II. Visits: Up to 12 months
   III. Eliminated features and entries with considerable missing values
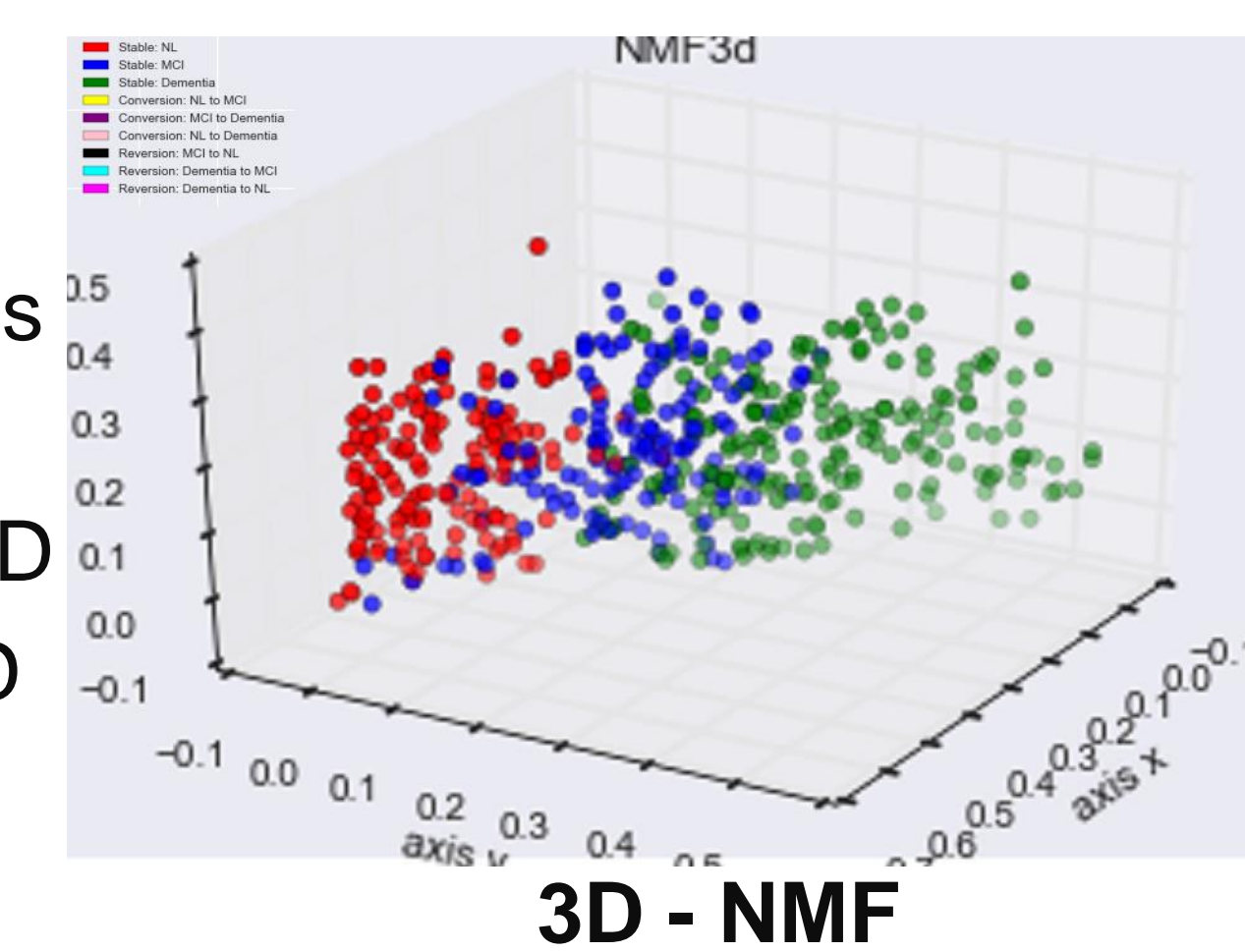   IV. Filled missing values using linear interpolation

2. Normalization
   I. Z-score: $z = \frac{x - \mu}{\sigma}$
   II. Min-Max: $X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$
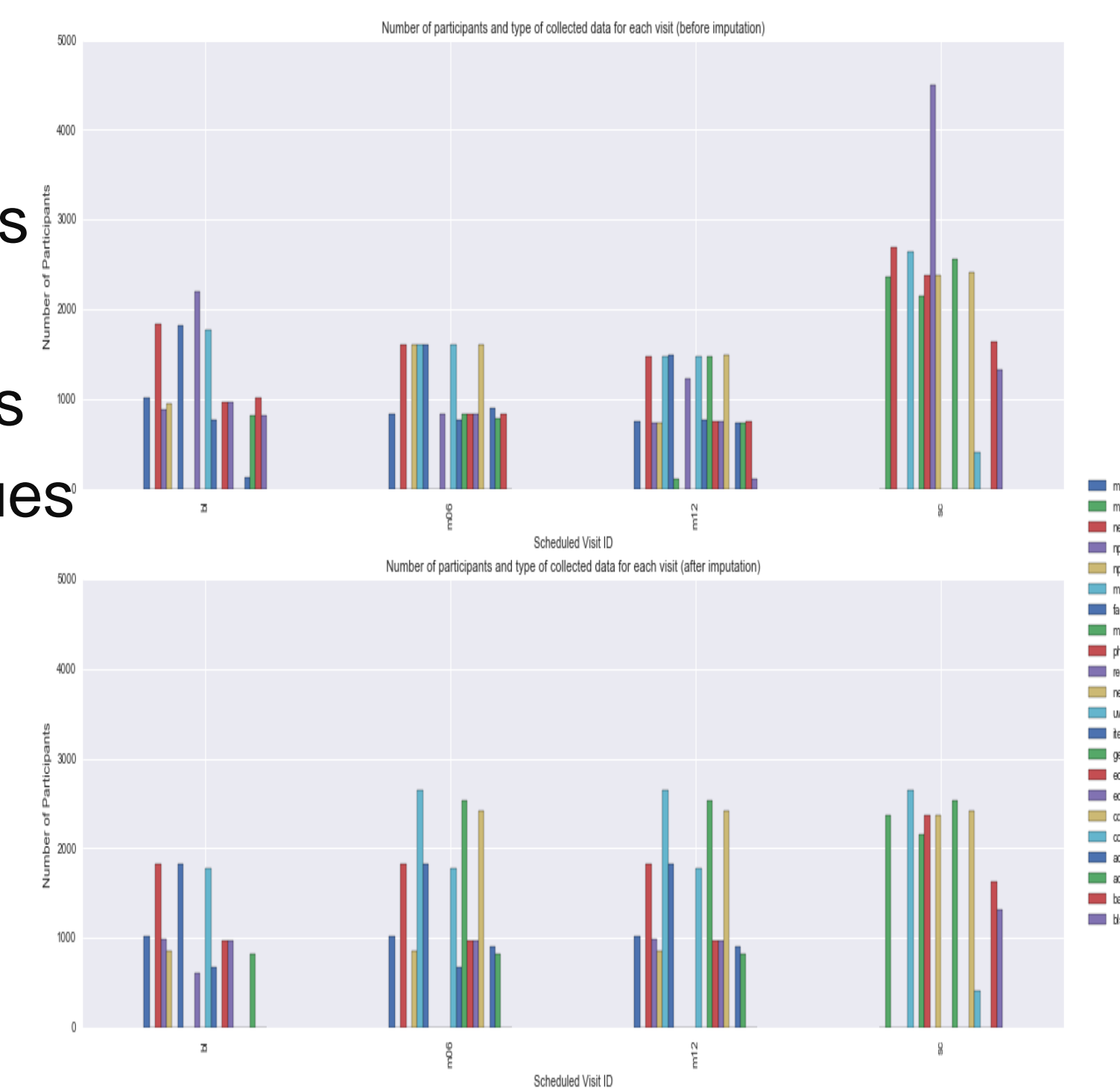
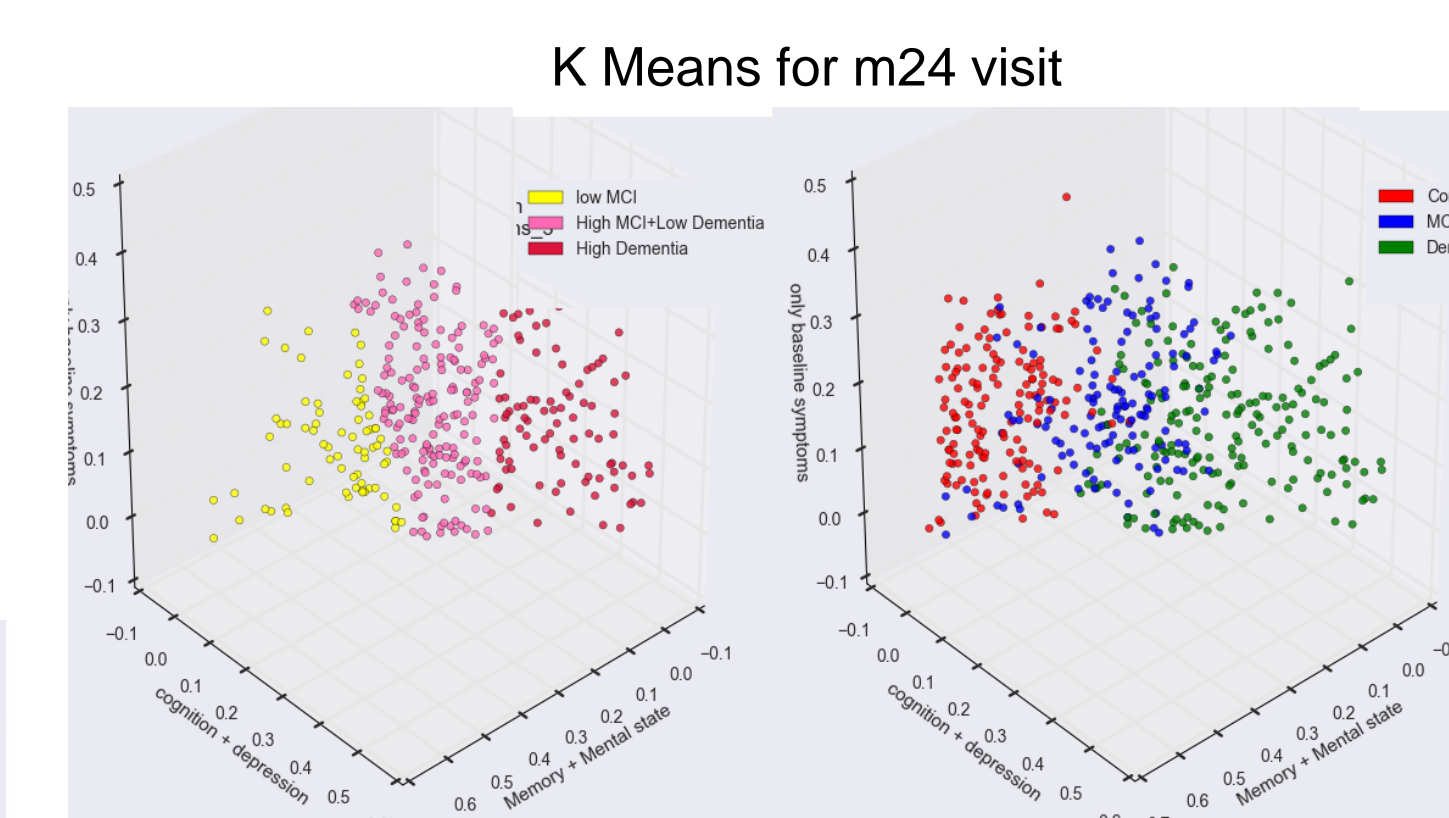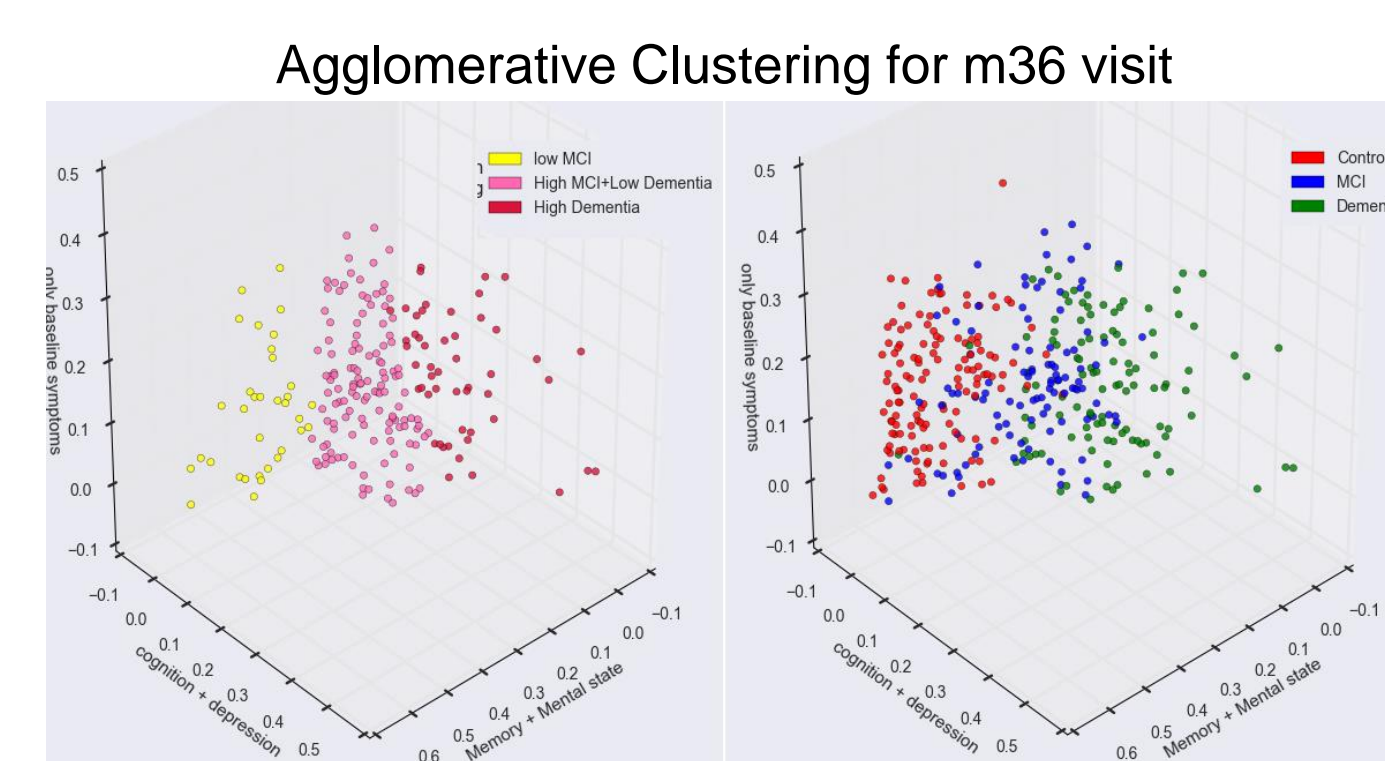### Dimensionality Reduction

348 common observations
Projections on m24 and m36 month visits
- Principal Component Analysis: 2D, 3D
- Independent Component Analysis: 2D, 3D
- Non negative matrix factorization: 2D, 3D

3D - NMF

Clustering

1. K-means
2. Gaussian Mixture Models
3. Agglomerative Clustering

K Means for m24 visit

Agglomerative Clustering for m36 visit

Observations and Results

- 3D NMF best projects the stages of progression in Alzheimer's namely Healthy Cohort, Mild Cognitive Impairment (MCI) and Dementia for m24 and m36 visits using the data until 12 months.

- Best Clusters representing stages of Progression are obtained using K-means for m24 and Agglomerative Clustering for m36.

- The axis of most relevant and informative features represent memory, cognition and baseline symptoms.

Conclusions

- Prediction of future Alzheimer's progressive stage using the past
- Extracted features responsible for different progression rates

Impact

- Additional information about Alzheimer's progression velocity
- Predictive tests allow early detection and characterization of distinct disease subgroups based on clinical heterogeneity

- Helps doctors subduing effects with better stage information
- 9.2 million fewer cases of the disease in 2050 (Refer [3])
- Improved clinical trial design
- Individualized clinical care

Future Directions

- Explore features causing progressions in both Parkinson's and Alzheimer's disease.
- Correlate CSF biospecimen with Alzheimer's progression space.
- Identify biomarkers predicting progressions to customize treatments.
- Replicate study on other datasets say AddNeuroMed and ItalianADNI.

References

- Predicting onset, progression, and clinical subtypes of Parkinson disease using machine learning, Faraz Faghri et. al., Brain

- Cognitive variability—A marker for incident MCI and AD: An analysis for the Alzheimer's Disease Neuroimaging Initiative, Anderson, Eric D., et al., Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring

- The magnitude of dementia occurrence in the world." Alzheimer Disease &Associated Disorders, Wimo, Anders, Bengt Winblad, Hedda Aguero-Torres, and Eva von Strauss.