Investigating the Risk Factors of Cognitive Decline in Mexico

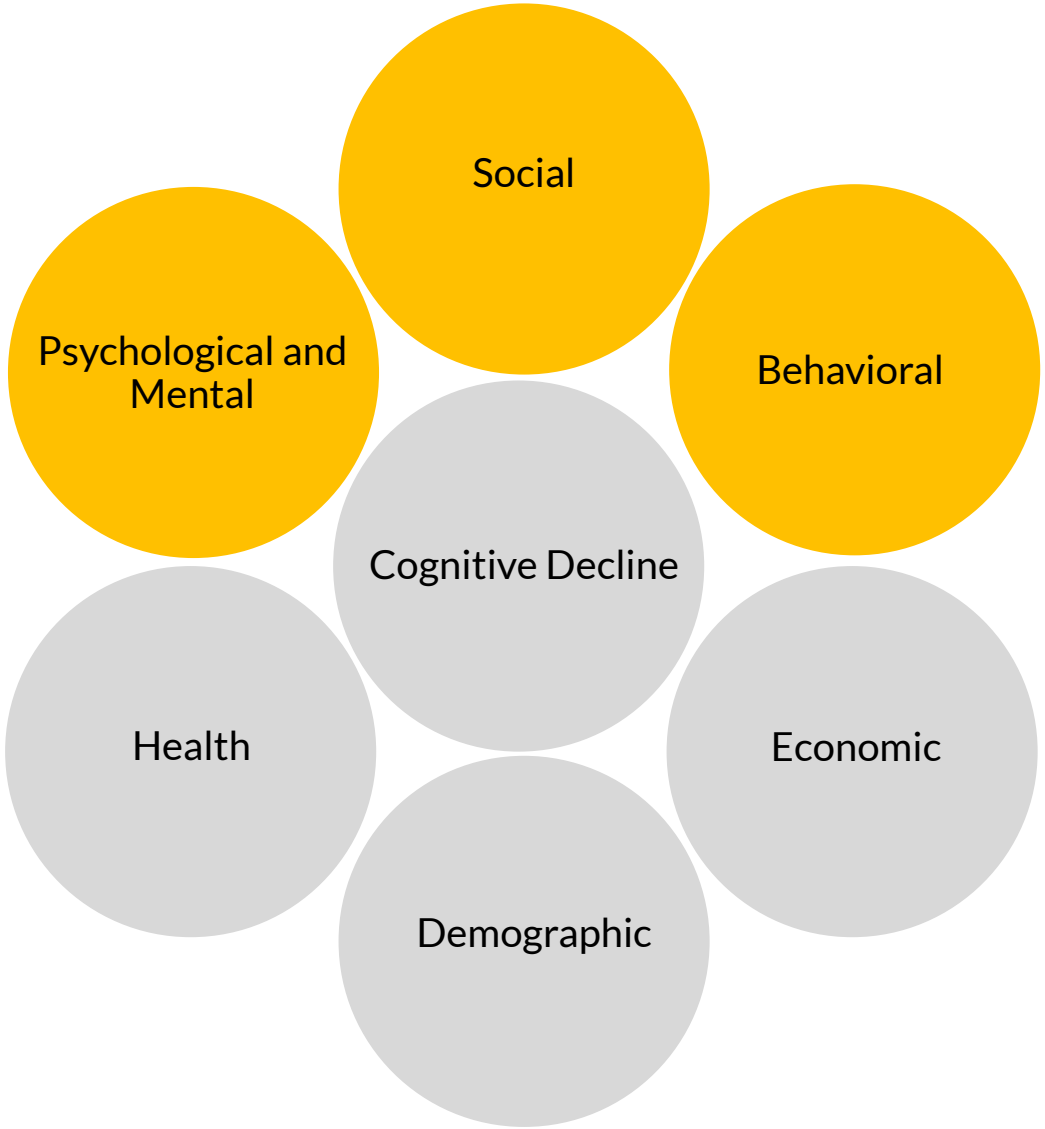
Python Exploratory Data Analysis

Gabriela Armenta

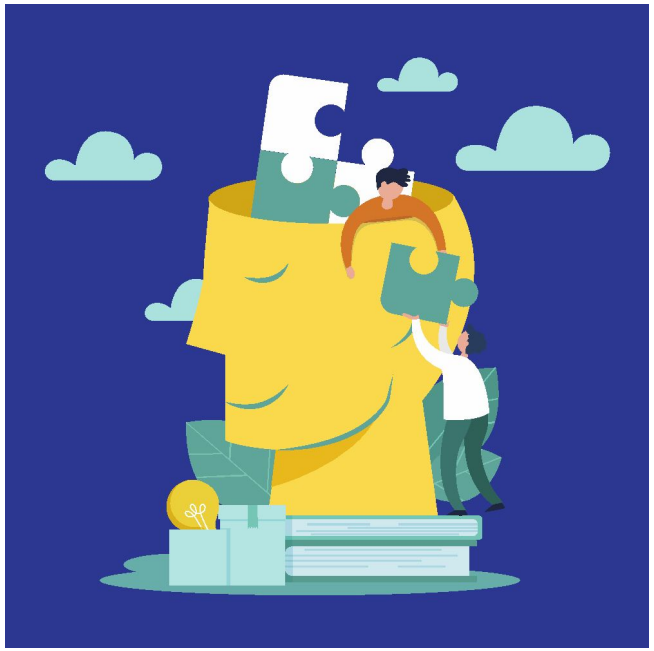
Outline

- **Objective**
- Background
- Data and methods
- Findings
- Conclusions
- Future work

Determine factors associated with cognitive decline for senior adults



Question

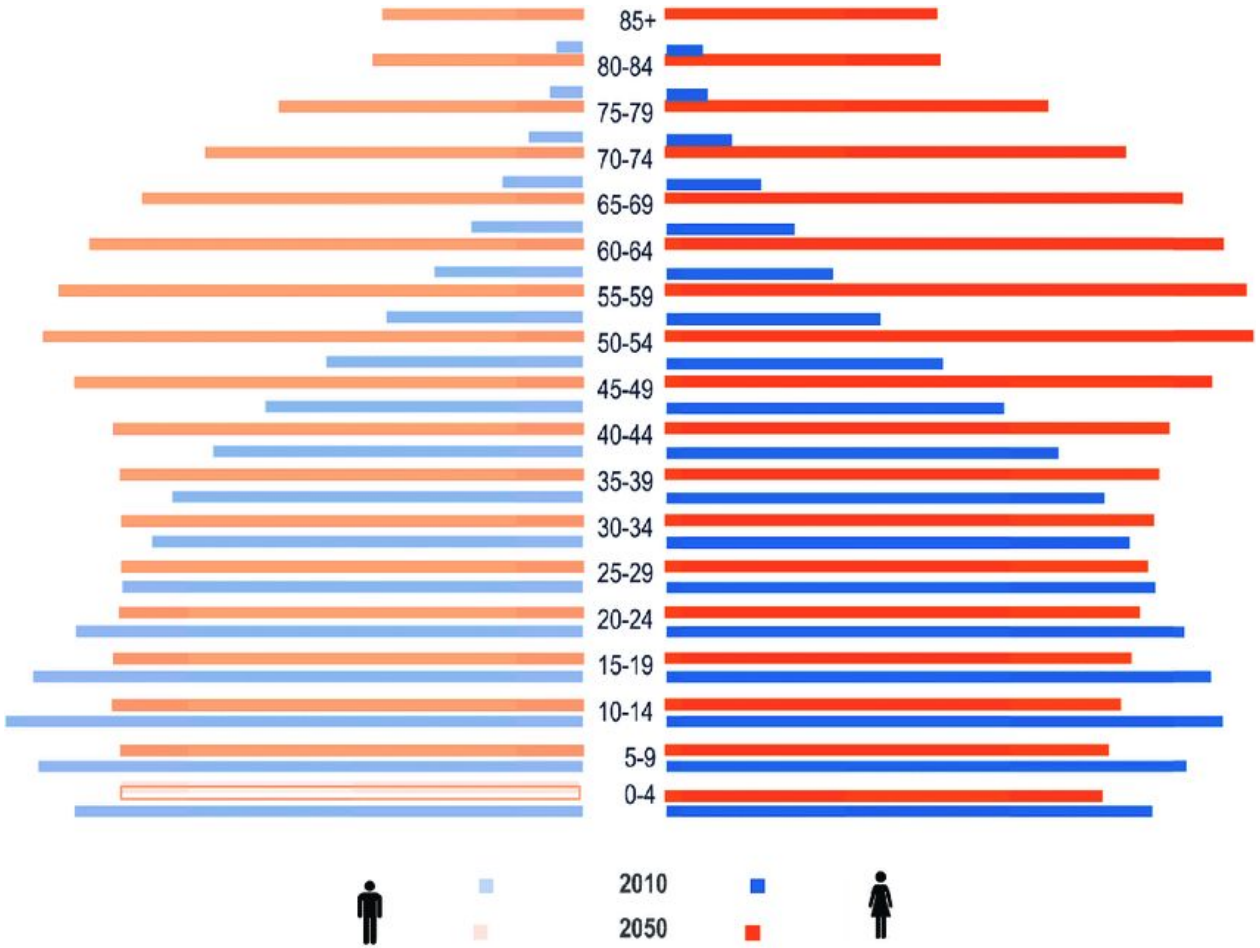


What are the factors associated with cognitive decline for people aged 50 years and older in Mexico?

Outline

- Objective
- **Background**
- Data and methods
- Findings
- Conclusions
- Future work

Mexico's aging population will drive dementia cases



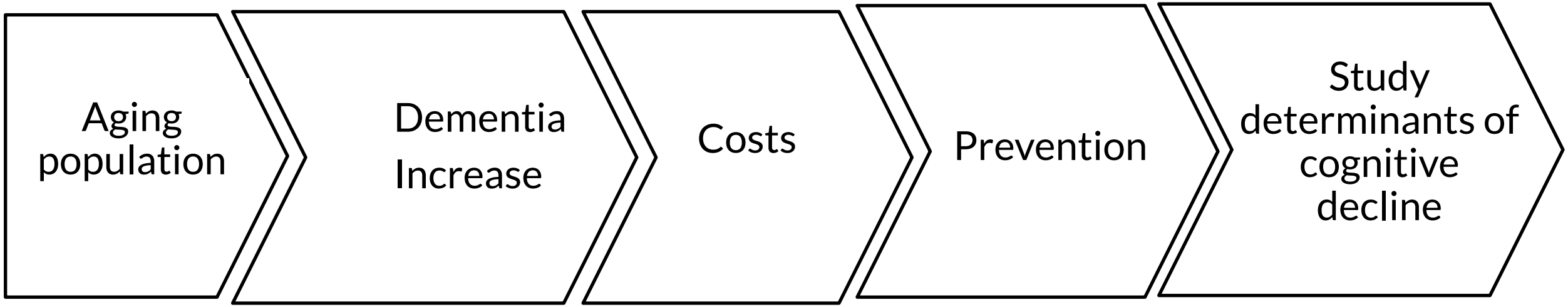
Mexico's population ages 60+ → 10% in 2017 → 25% in 2050

Low and Middle Income Country (LMIC) Focus

- 80+ percent of older people will be living in LMIC by 2050
- Currently, LMIC already have 60% of cases of dementia.

[1] WHO 2022; [2] Pew Research Center 2014

Cases will drive costs, increasing the need for prevention and studying the determinants of cognitive decline



People 60+
10% in 2017
25% in 2050

Increase ~150%
from 2020 to 2050

Up to 40%
undiagnosed

From diagnosed
60% Alzheimer's

40% specialized care
60% catastrophic
exp

40% can be
prevented or
delayed

In summary…

- Limited evidence exists from social, behavioral, psychological and mental factors for Latinos, Mexican Americans, and Mexicans.

- Most evidence focuses on economic, demographic and health-related factors.

- Important to expand understanding on unexplored factors that are could be relevant.

Outline

- Objective
- Background
- **Data and methods**
- Findings
- Conclusions
- Future work

Data

- Mexican Health and Aging Study
- Publicly available longitudinal survey for adults 50+ (partners younger than 50 years old)
- National representation
- Direct in-person interviews
- Years 2012, 2015, 2018

Methods

- **Language**: Python
- **Libraries:** Pandas, NumPy, SciPy, matplotlib, os, seaborn.
- **Two steps in data analysis:**
  - First, data preprocessing and feature engineering.
  - Second, descriptive statistics, data visualizations, basic statistical tests to identify statistically significant associations between variables.

Data preprocessing and feature engineering ~75% of time

Challenge	One example of change implemented
Data were inconsistent across waves	**Kept waves for years 2012, 2015, 2018** where data were consistent.
Samples for interviewers could have significant variations: ● partners of direct interviews could be younger than 50 years old ● proxies were interviewed	**Kept direct interviews** (dropped data from proxies) and **people ages 50+**
Features had names that were not interpretable without a dictionary	**Renamed features** to names more easily interpretable
Similar values had differing meaning for different features	**Recoded all missing values** (e.g., .d, .m, .i, .r, 888, 88, 8, 999, 99, 9) to pd.NA
Features contained both numeric and text values within the same row or column	**Feature values were converted into either numeric or string.**

Data preprocessing and feature engineering ~75% of time

Challenge	Example of change implemented
Higher values could indicate both more or less of a feature	**Higher values were recoded to indicate a greater presence of a feature (e.g.,** self-reported health score of 1 meant poor health and 5 meant excellent health).
Features required engineering	**Based on literature, features were turned into scales, indexes, or analyzed independently.** For instance, social support and conscientiousness.
Literature supported more than one feature transformation	**More than one scale or index was created.** For instance, BMI was engineered as indexes for both senior adults and general population, as well as binary categories.
Features could be analyzed as either continuous, categorical, or binary	**Feature variations were tested**, such as transforming continuous into categorical/binary features (e.g., BMI to underweight or obese), redefining categorical features into fewer/different categories (e.g., marital status to partner).
Highly correlated features	**One feature was selected among highly correlated ones**.

Output feature: Cognitive Status

Shows degree of cognitive decline

Takes four values

0. Normal decline
Gradual brain function loss, memory loss, and brain shrinking

1. Normal decline with instrumental impairment
Problems with medication intake, keeping appointments, finding things at home, using technology)

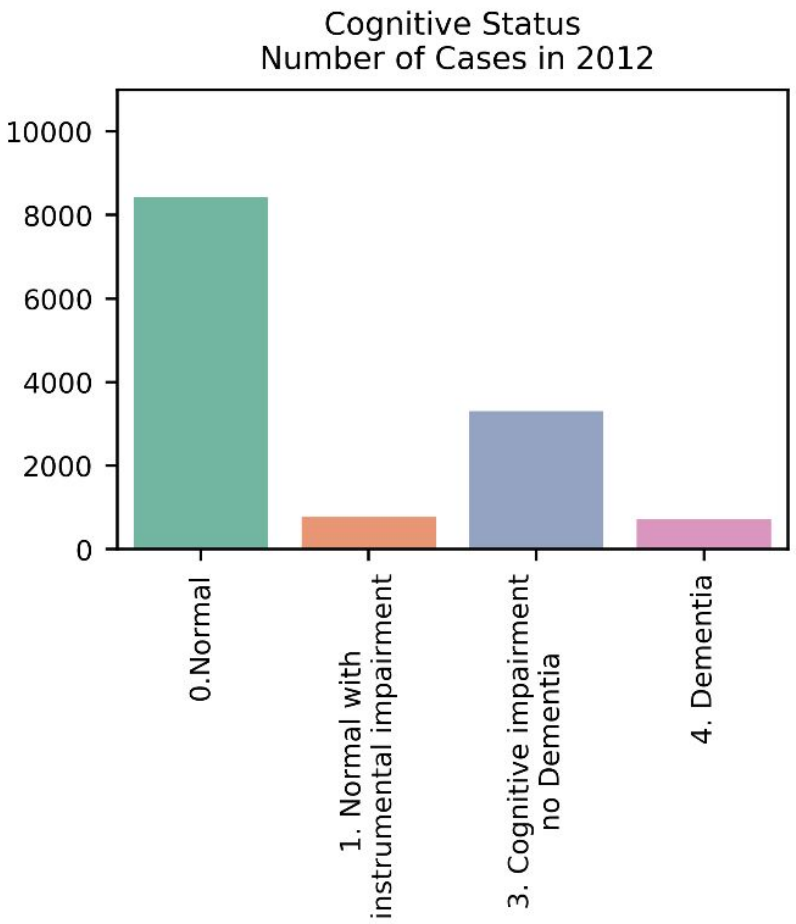
2. Cognitive impairment no dementia
Problems with memory, speech and ability to think clearly

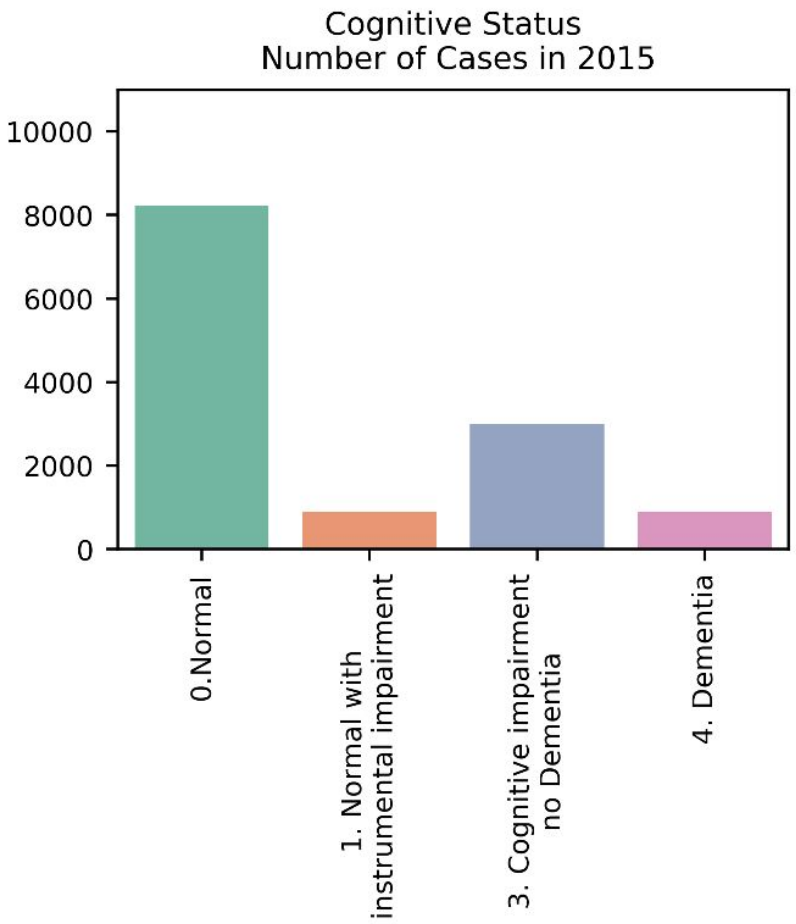
3. Dementia
Impaired memory, reasoning, mobility, communication, daily functions severely affected

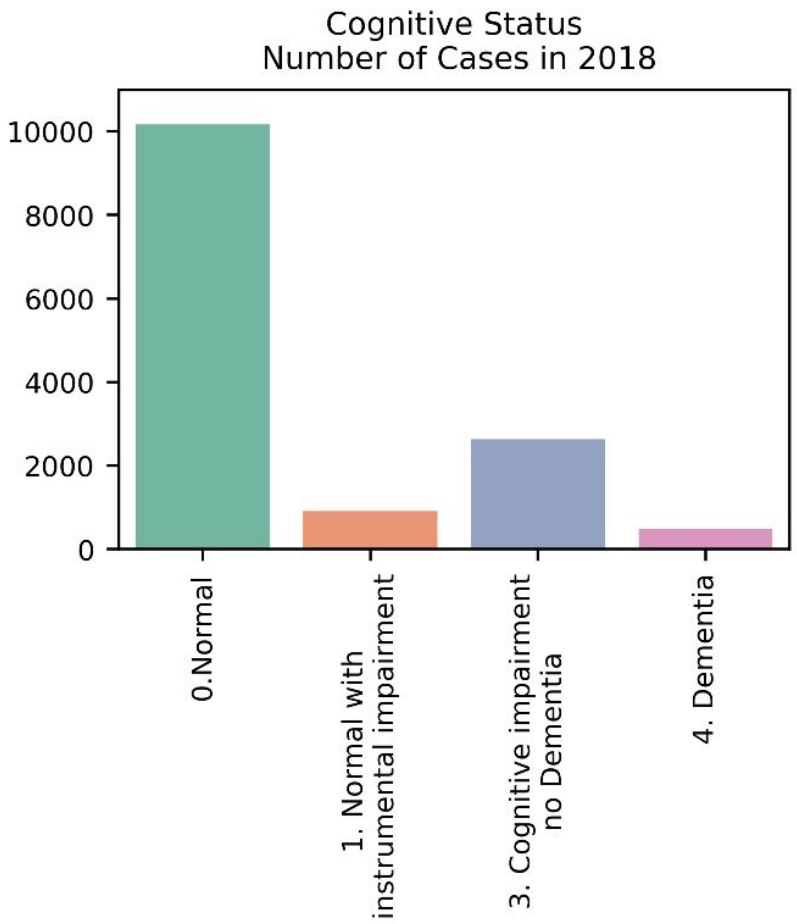
Outline

- Objective
- Background
- Data and methods
- **Findings**
- Conclusions
- Future work

Distribution of cognitive status by year







Total 14,366

Total 13,868

Total 15,385

Cognitive status converted to binary feature

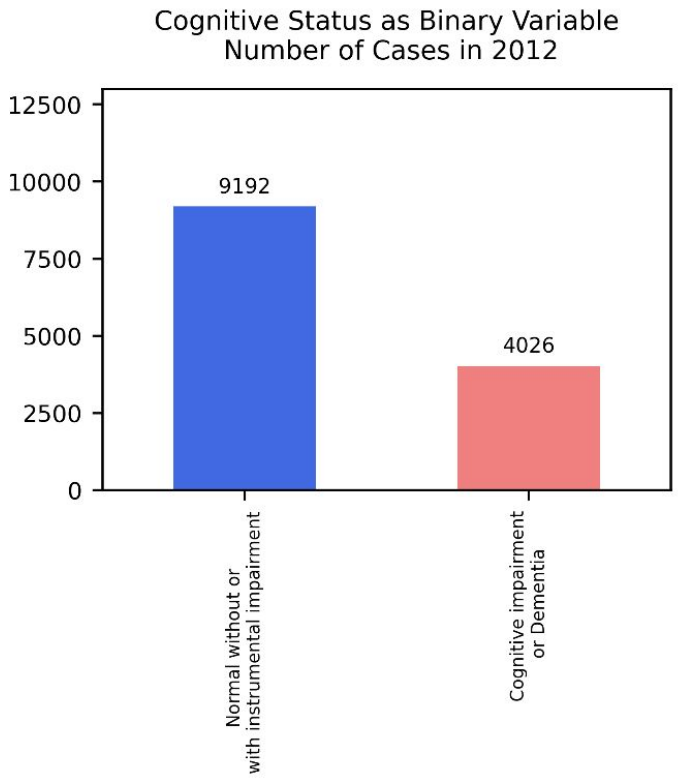
Cognitive status:

0. Normal
1. Normal with instrumental impairment

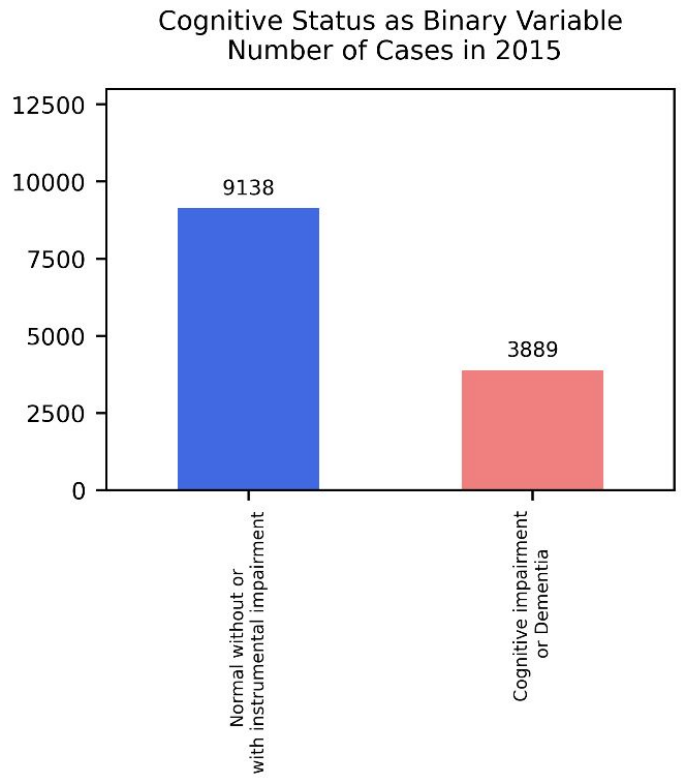
Blue

3. Cognitive impairment no dementia
4. Dementia

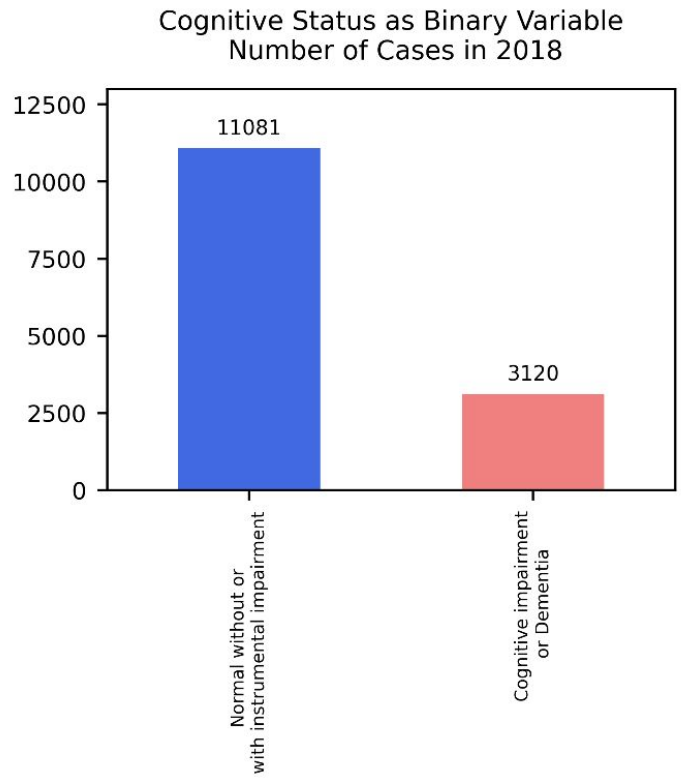
Pink



Total 14,366



Total 13,868

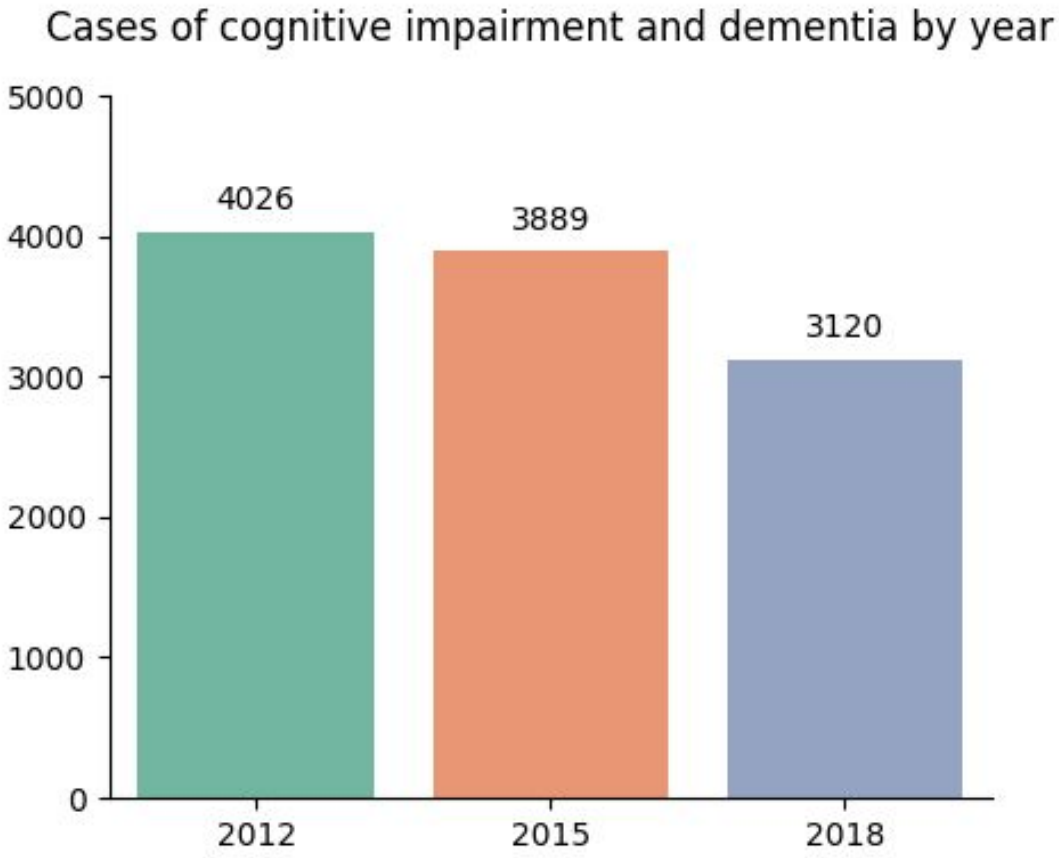


Total 15,385

Cases of cognitive impairment and dementia decreased

Cognitive status (degree of cognitive decline):

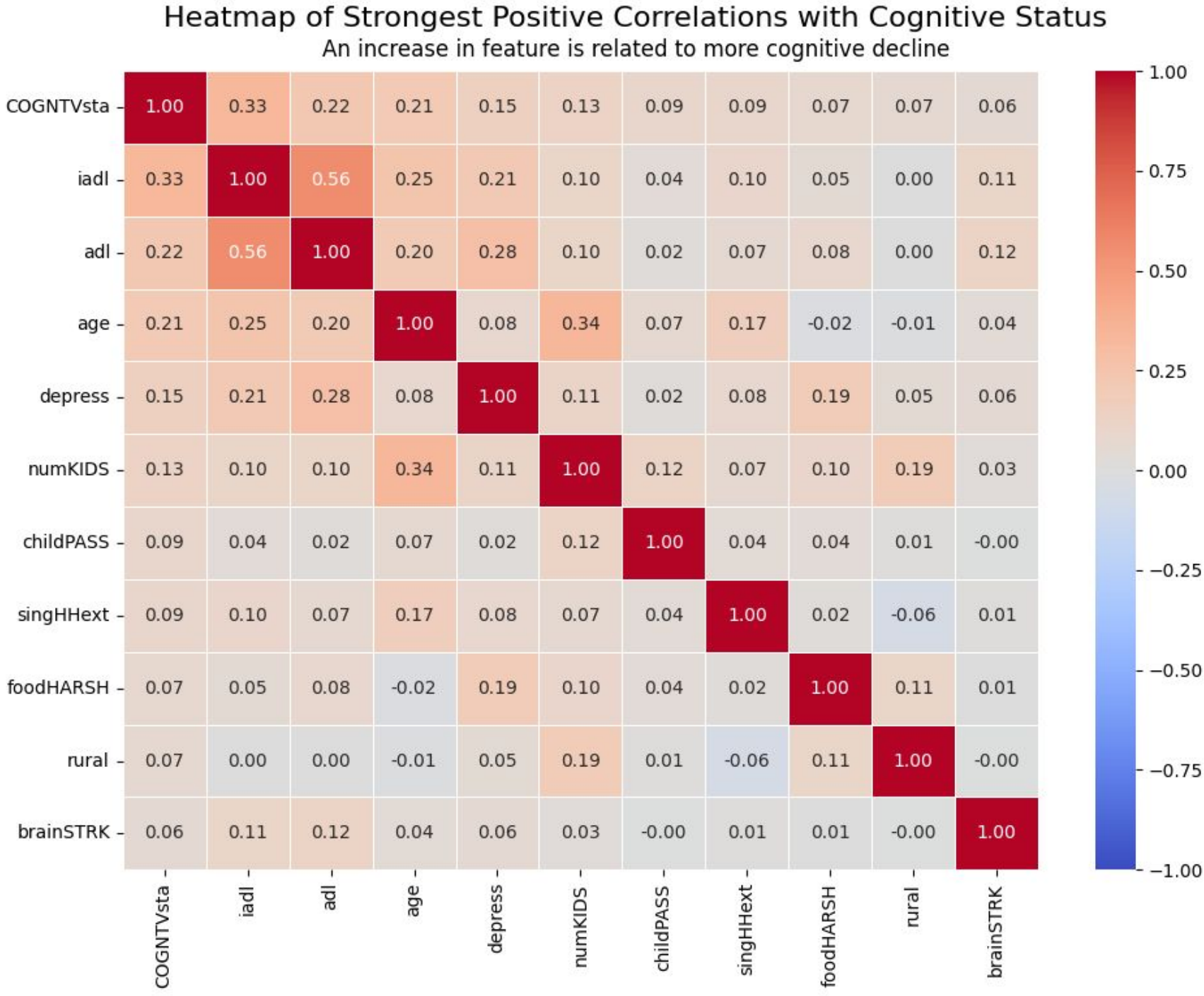
0. Normal
1. Normal with instrumental impairment
2. **Cognitive impairment no dementia**
3. **Dementia**



- We observe a negative trend in cases of cognitive impairment no dementia and dementia over time.
- May be due to people leaving the sample (e.g., passing away).

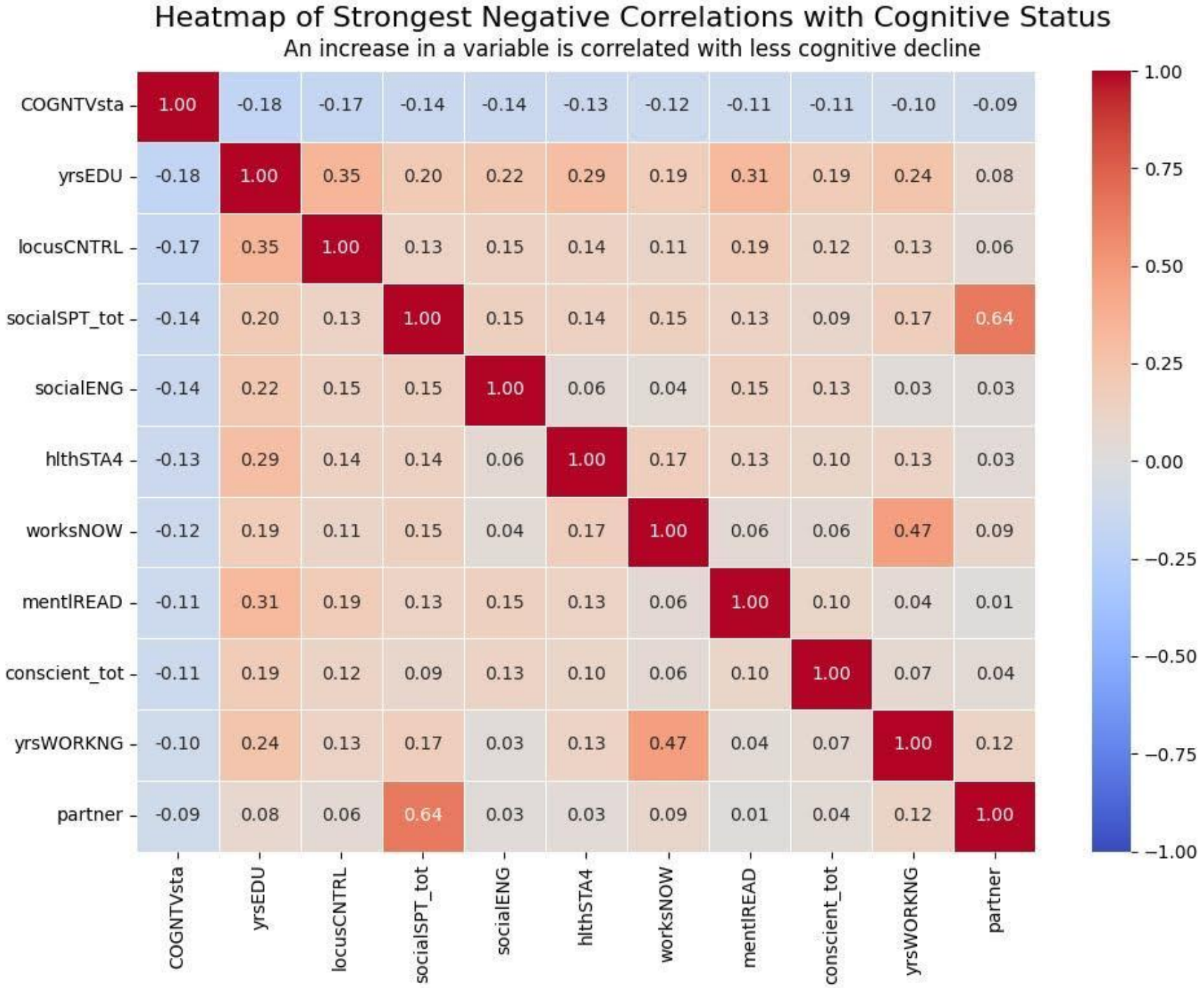
Top 10 features for which a higher presence is related to more cognitive decline

- Difficulties with Instrumental Activities of Daily Living
- Difficulties with Activities of Daily Living
- Years of age
- Depression
- Having more children
- Trauma experience -child passed away
- Being single in extended household
- Food hardship due to lack of money
- Living in rural community
- Experiencing a brain stroke

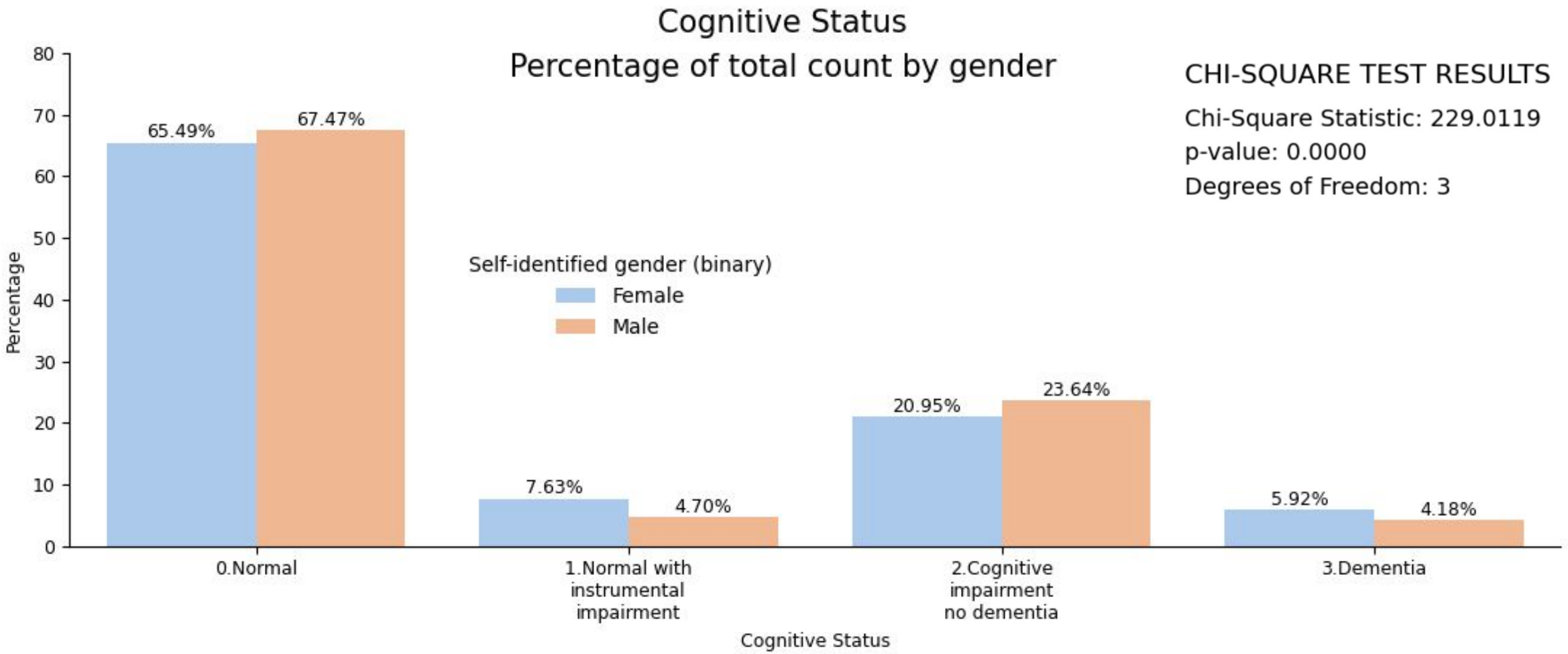


Top 10 features for for which a higher presence is related to less cognitive decline

- Years of education
- Internal locus of control
- Social support received
- Social engagement
- Better health status
- Currently working
- Reading activities
- Conscientiousness
- Years actively working
- Having a partner--also positively related to social support



Gender relationship to cognitive status



- Gender did not show up as one of the top correlations with cognitive status.
- The Chi-Square test shows a statistically significant relationship between cognitive status and gender, which justifies deeper investigation into the nature of this relationship. Plus literature also supports.

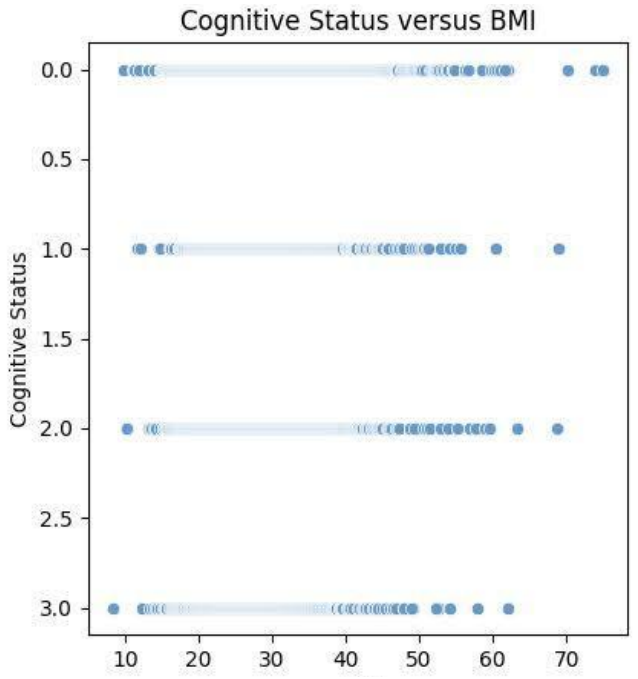
BMI relationship may not be linear

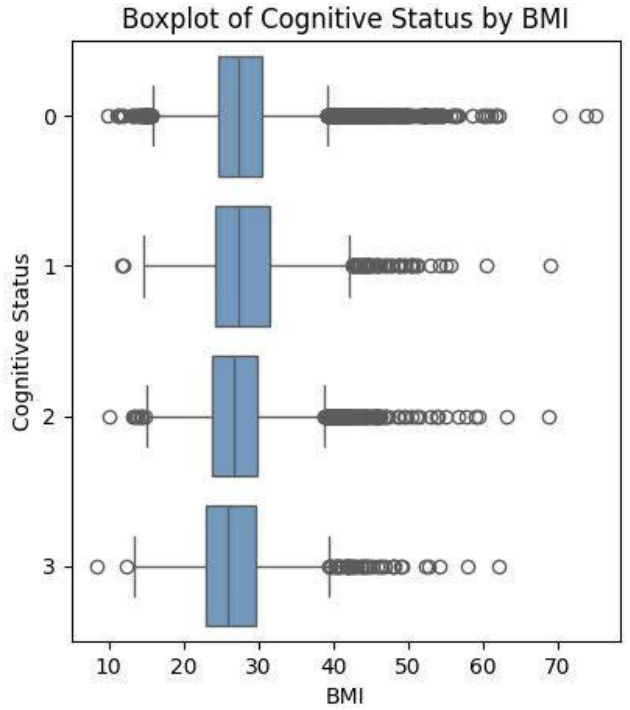
BMI did not show up in heatmap as top correlated feature.

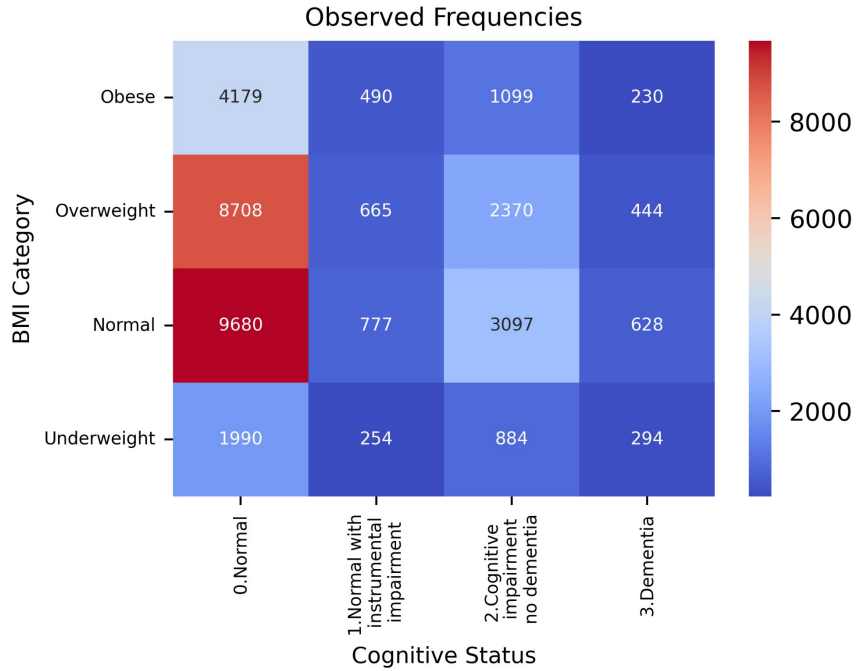
However, Chi-Square test shows a statistically significant relationship with cognitive decline, and the box plot seems to show a negative relationship.

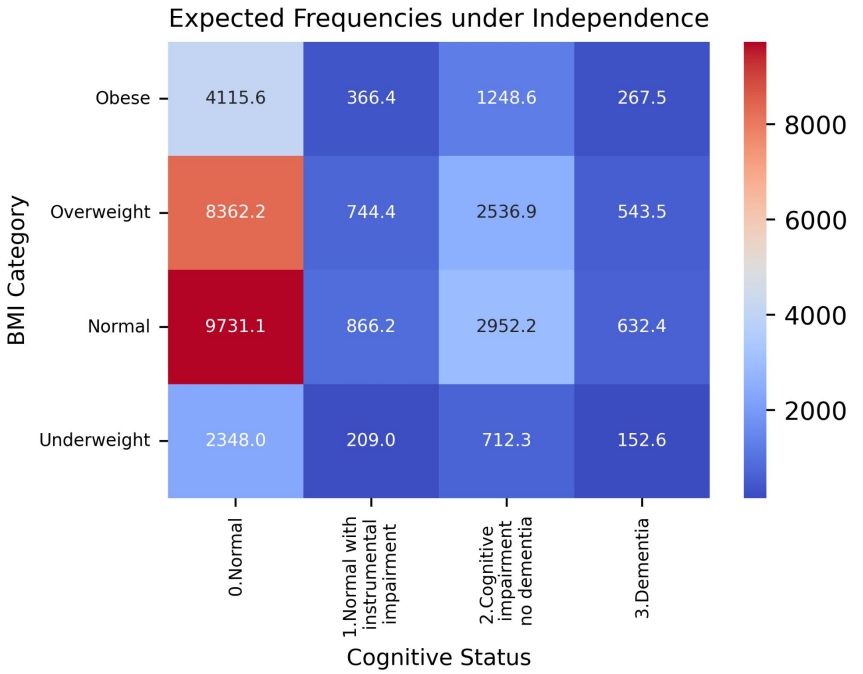
CHI-SQUARE TEST RESULTS

Chi-Square Statistic: 371.0780
p-value: 0.0000
Degrees of Freedom: 9









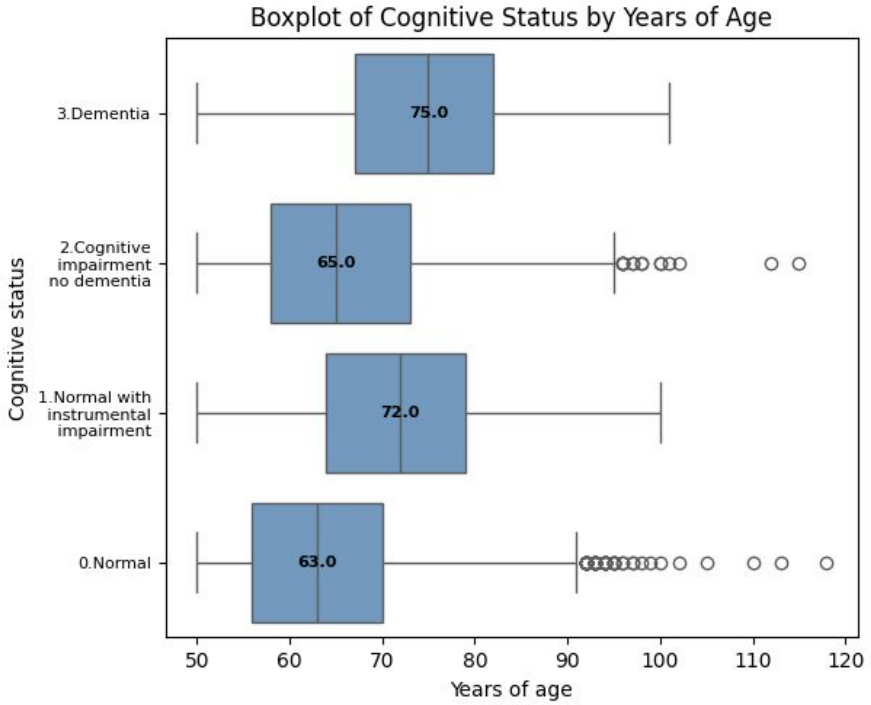
Relationships with age, social support, locus of control, and education have expected signs

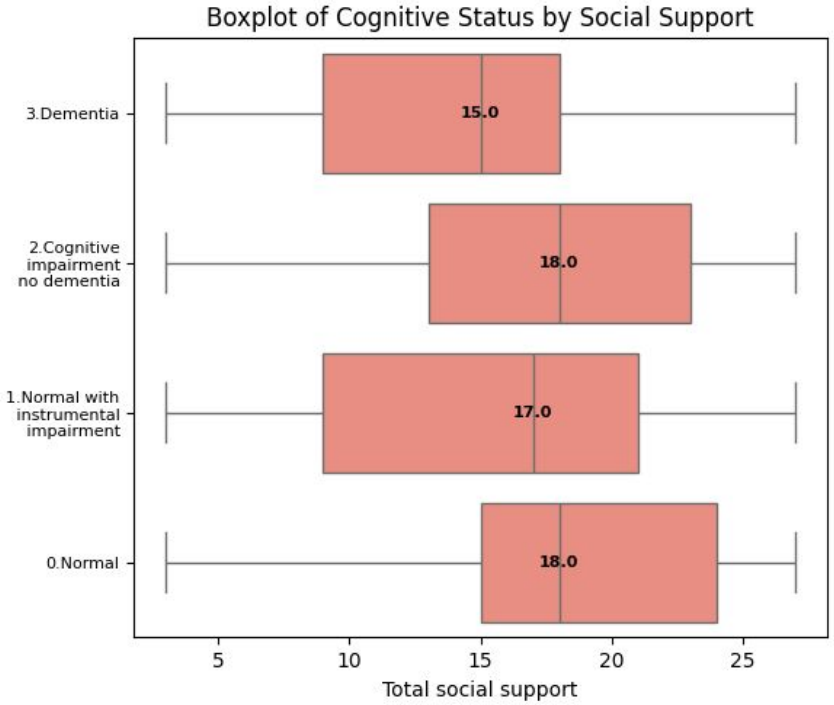
As the median years of age increases, cognitive decline increases.

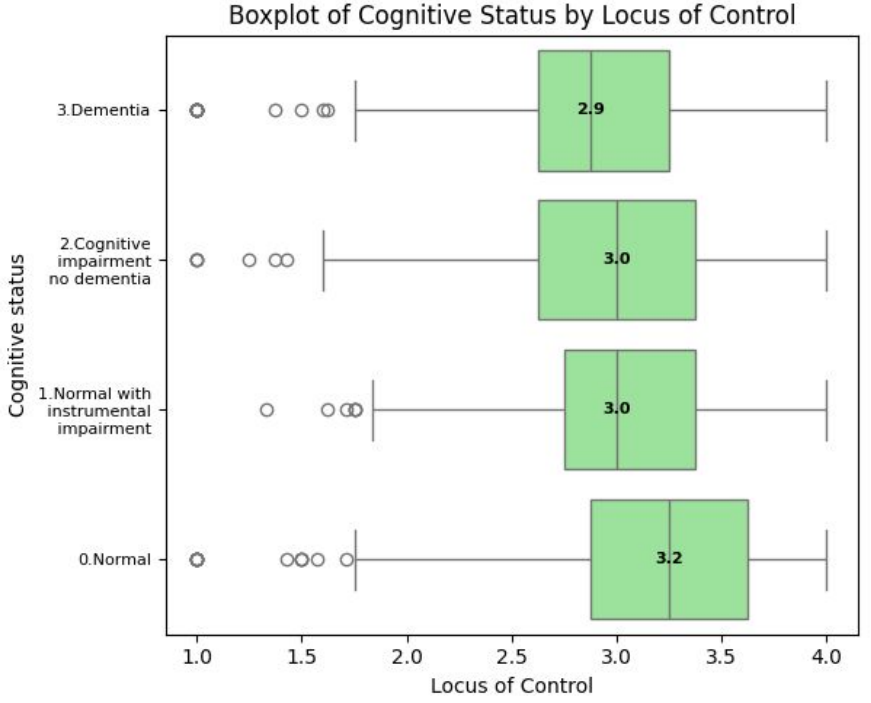
As social support received increases, cognitive decline decreases.

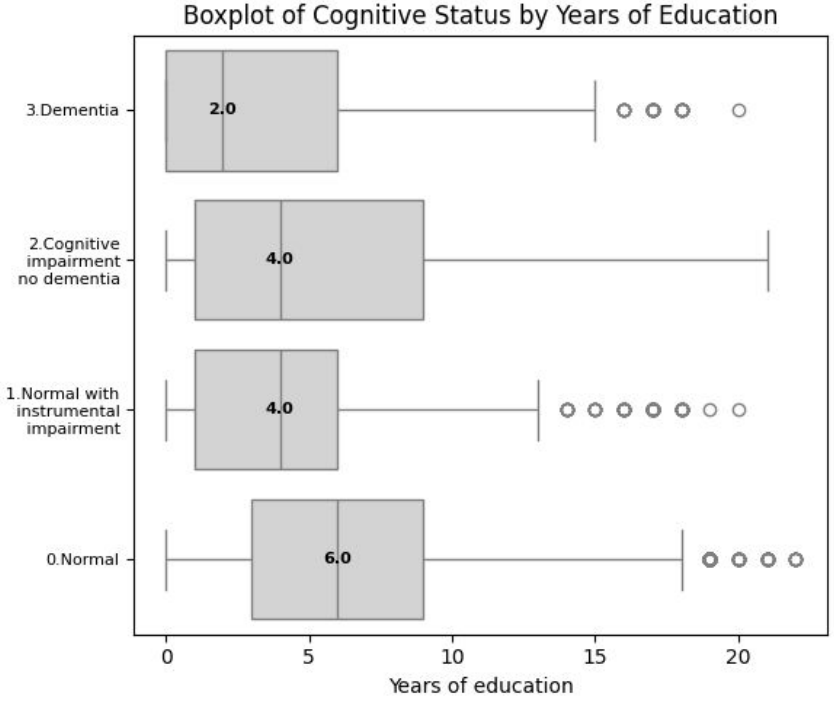
As internal locus of control increases, cognitive decline decreases.

As years of education increases, cognitive decline decreases.

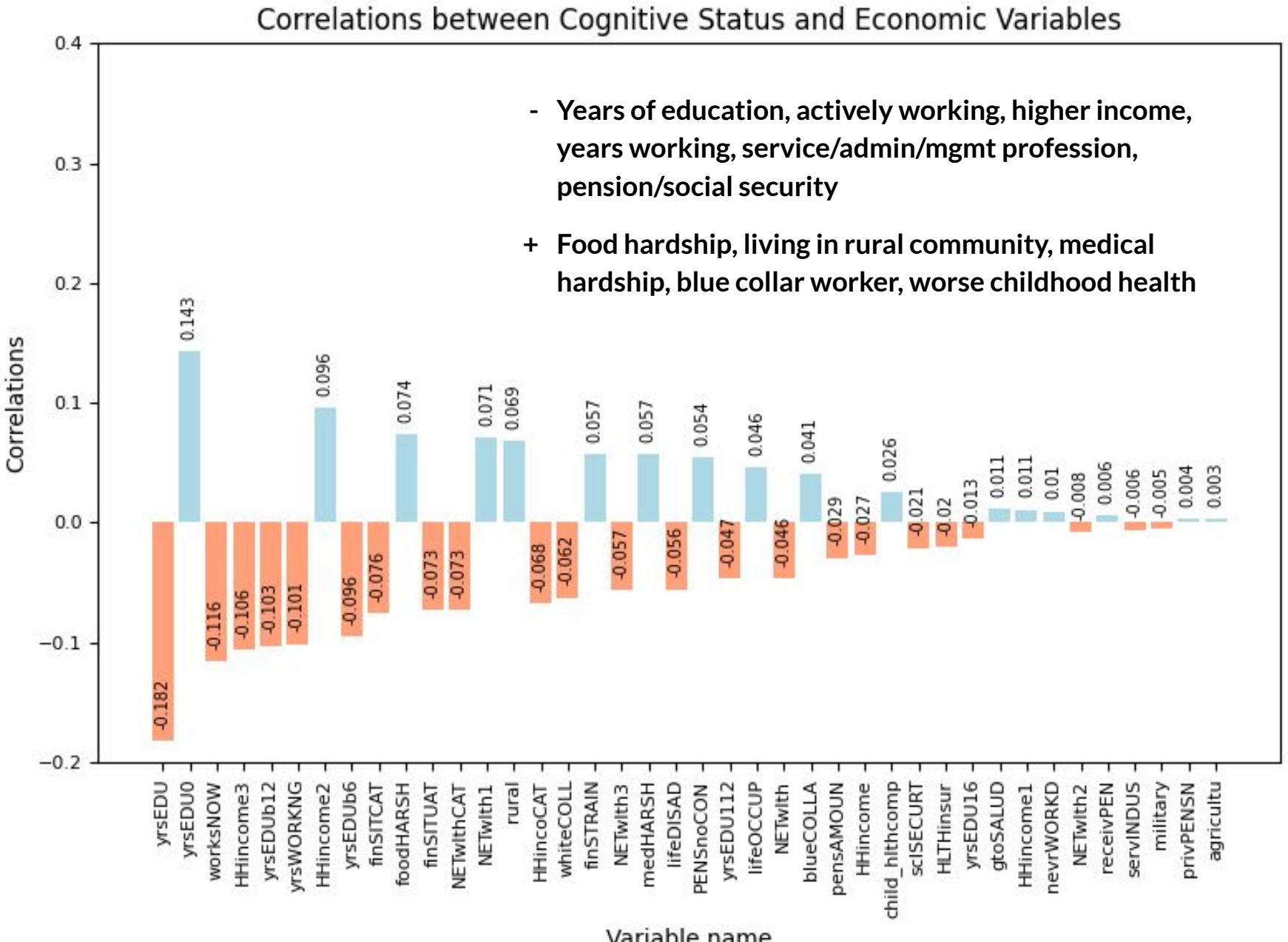








Correlations with economic factors



- **Years of education, actively working, higher income, years working, service/admin/mgmt profession, pension/social security**

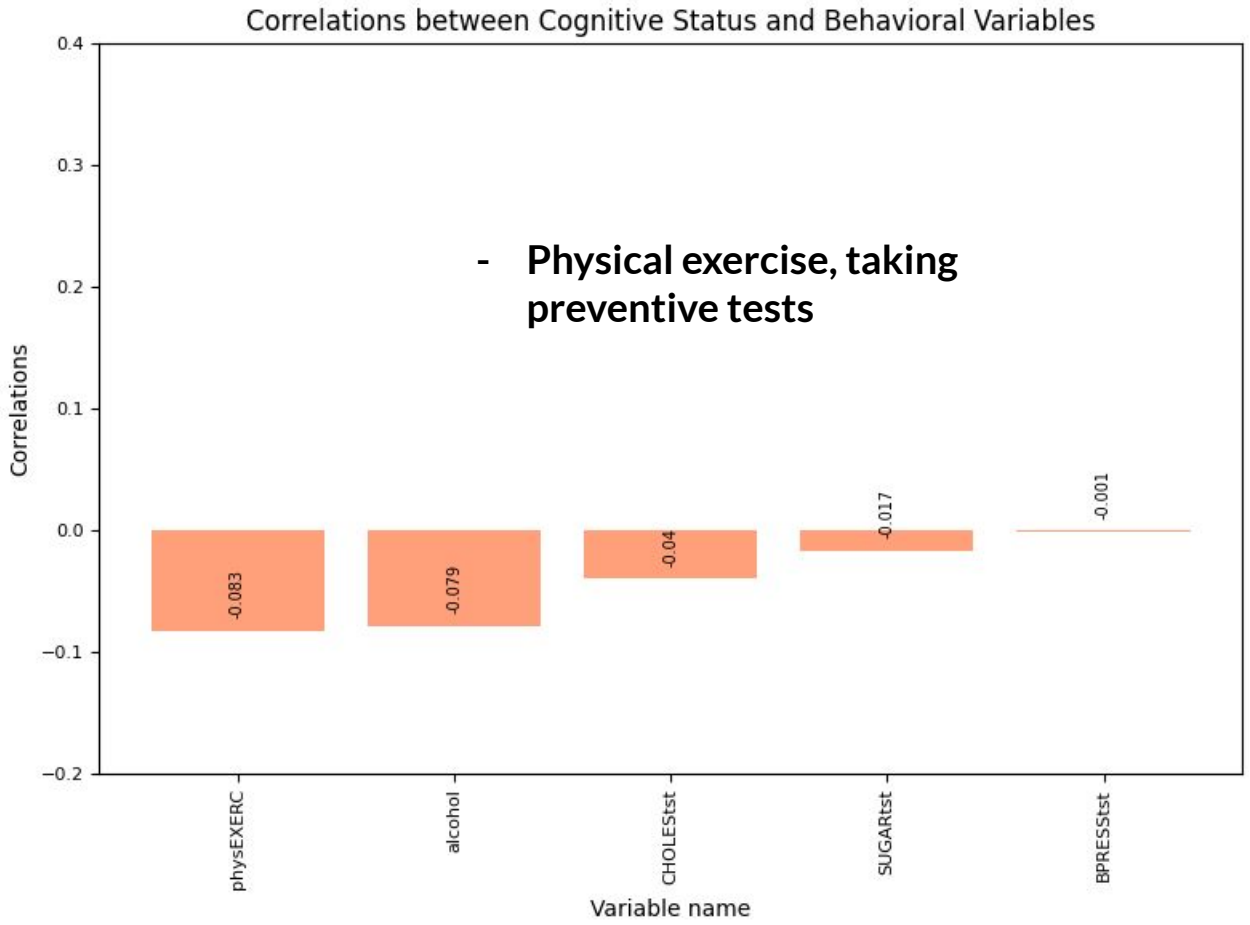
+ **Food hardship, living in rural community, medical hardship, blue collar worker, worse childhood health**

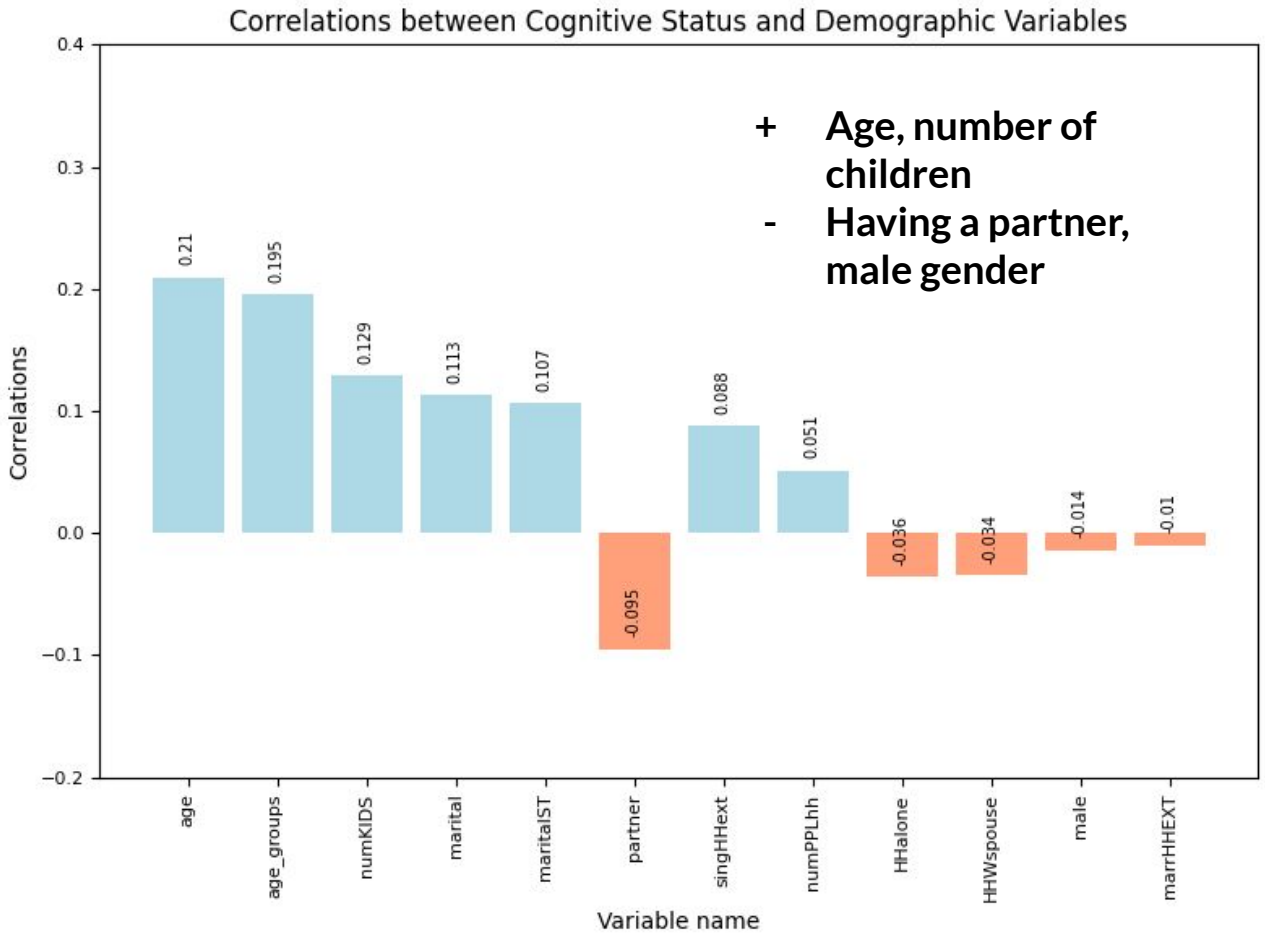
Features had expected sign thus supporting literature

Positive correlation (blue): an increase in feature is associated with more cognitive decline.

Negative correlation (orange): an increase in feature is associated with less cognitive decline.

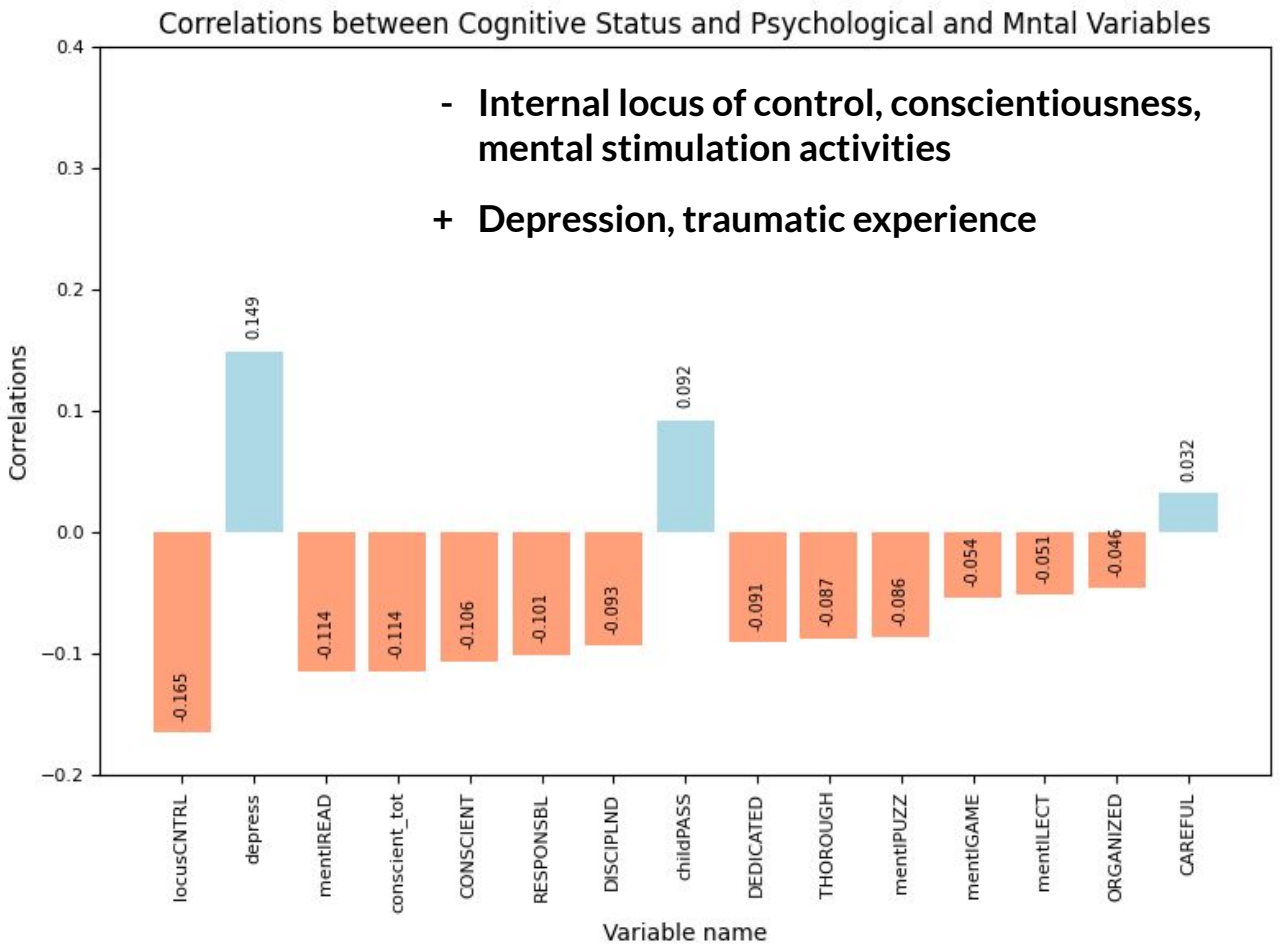
Correlations with behavioral and demographic factors

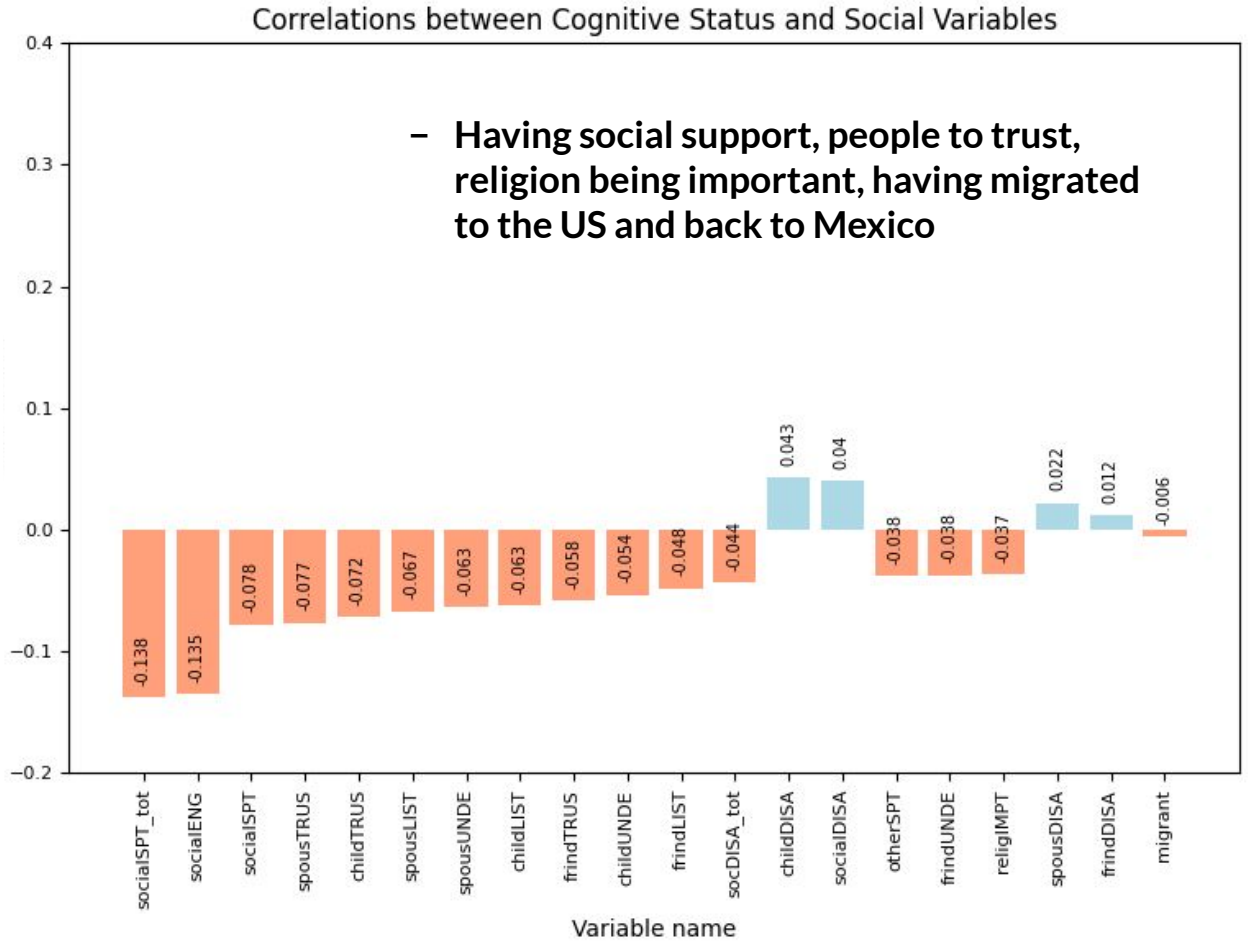




- Very low correlations between behavioral factors and cognitive status, higher correlations with demographic factors.
- Unexpected finding for alcohol consumption--though small number.

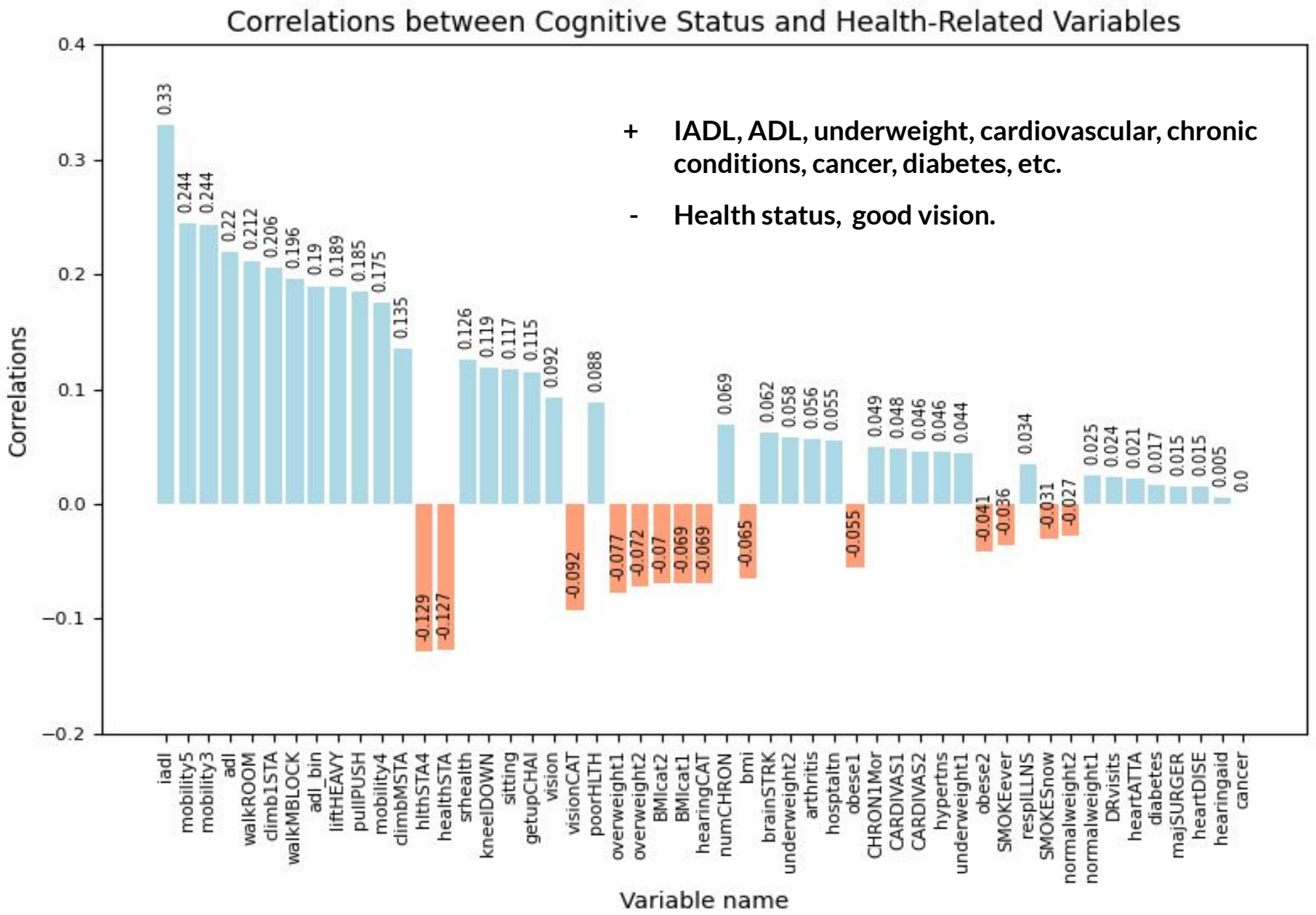
Correlations with social and psychological and mental factors





Social support, and mental and psychological features had the expected sign.

Correlations with health-related factors



Higher correlations with health-related variables than with other groups of factors

Most correlations with expected sign. Unexpected sign for smoking.

Outline

- Objective
- Background
- Data and methods
- Findings
- **Conclusions**
- Future work

Conclusions

- Social, psychological and mental factors emerged as related to cognitive decline, some correlations were comparable in magnitude or even superior to those of health-related variables.
    - Internal locus of control, conscientiousness, depression, mental stimulation and social engagement were all correlated to cognitive decline
- Most correlated factors:
    - Difficulties with Instrumental Activities of Daily Living, difficulties with Activities of Daily Living, age, depression, years of education, internal locus of control, social support, social engagement, self-reported health status, and number of children.
- Many of these are modifiable risk factors.
- Some relationships may not to be linear.
- More evidence is needed to shed light on the actual relationships and on directionality.

Outline

- Objective
- Background
- Data and methods
- Findings
- Conclusions
- **Future work**

Future work

- Using machine learning algorithms to predict cognitive decline.
- Potentially, take advantage of the longitudinal nature of the data.
- Conduct deeper investigation into the nature of relationships that may not be linear (e.g., BMI, gender).
- Potentially add more years of data and explore other features (e.g., technology use, pets at home).