Multilingualism and Cognitive Ability

Oviya Adhan, Nory Arroyo, Helin Yilmaz

DATASCI 203

April 17, 2025

Agenda

- ✓ Introduction
- ✓ Dataset
- ✓ Models
- ✓ Assumptions
- ✓ Takeaways

Introduction

- Multilingualism is on the rise globally
- Prior studies show mixed cognitive outcomes
- Research Question: Is there a relationship between multilingualism and cognitive ability?

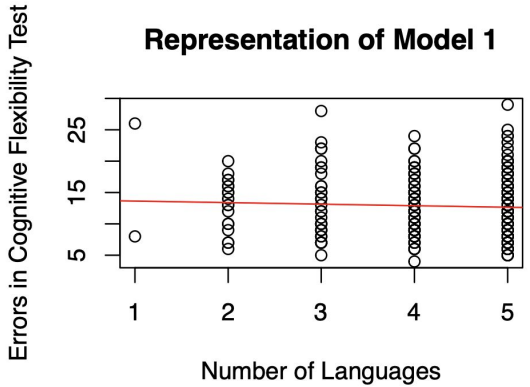
Dataset

- **Source:** University of Groningen (Netherlands)
- **Sample**: 387 older adults
- *Region is known for high multilingualism
- **Data**: 2 cognitive tests + language & lifestyle questionnaire
- **Key measures**:
  - *Multilingualism* = # of languages spoken (1–5)
  - *Cognitive ability* = errors on WCST (4–29)

Model 1

Errors WCST = Number of Languages

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.8815     1.3931   9.964   <2e-16 ***
number_lang  -0.2466     0.3194  -0.772    0.441
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```



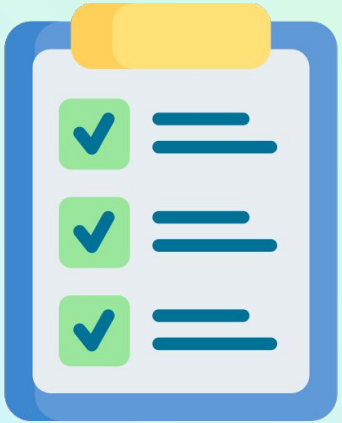
Model 2

Errors WCST = Number of Languages + Gender +
Age + Education Level + Income Level

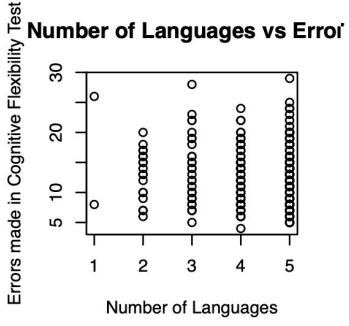
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.72152    4.71846   2.060  0.04042 *
number_lang -0.06183    0.31530  -0.196  0.84470
gender      -1.04945    0.62149  -1.689  0.09256 .
age          0.15528    0.05618   2.764  0.00615 **
education   -1.05116    0.30150  -3.486  0.00058 ***
income      -0.14817    0.24136  -0.614  0.53985
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

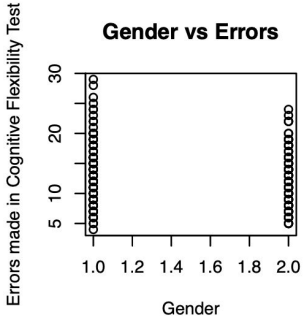
Assumptions

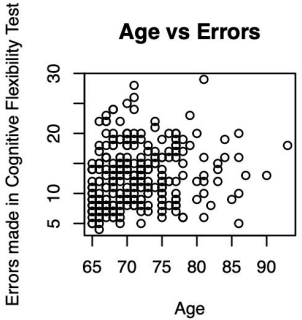
1. IID
2. Linear Conditional Expectation
3. No Perfect Collinearity
4. Homoscedasticity
5. Normally Distributed Errors

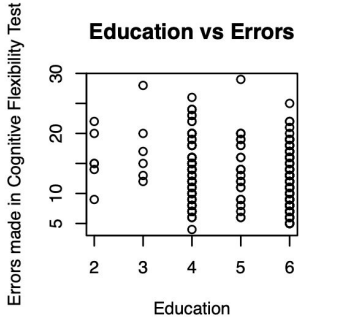


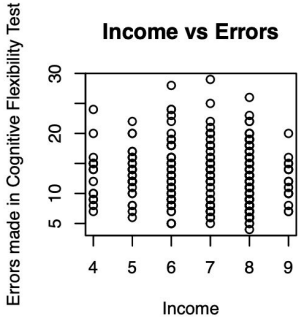
Independent and Identically Distributed (IID)







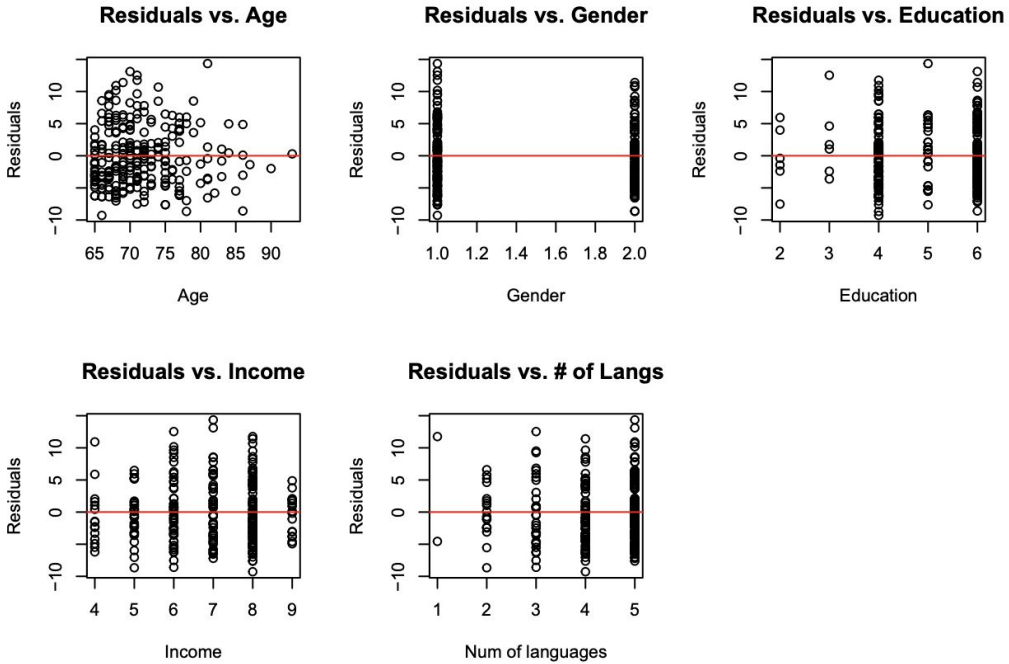




Considerations:

- No strong clusters
- Data Collection:
  - Limited generalizability
  - Independent recruitment and testing

Linear Conditional Expectation

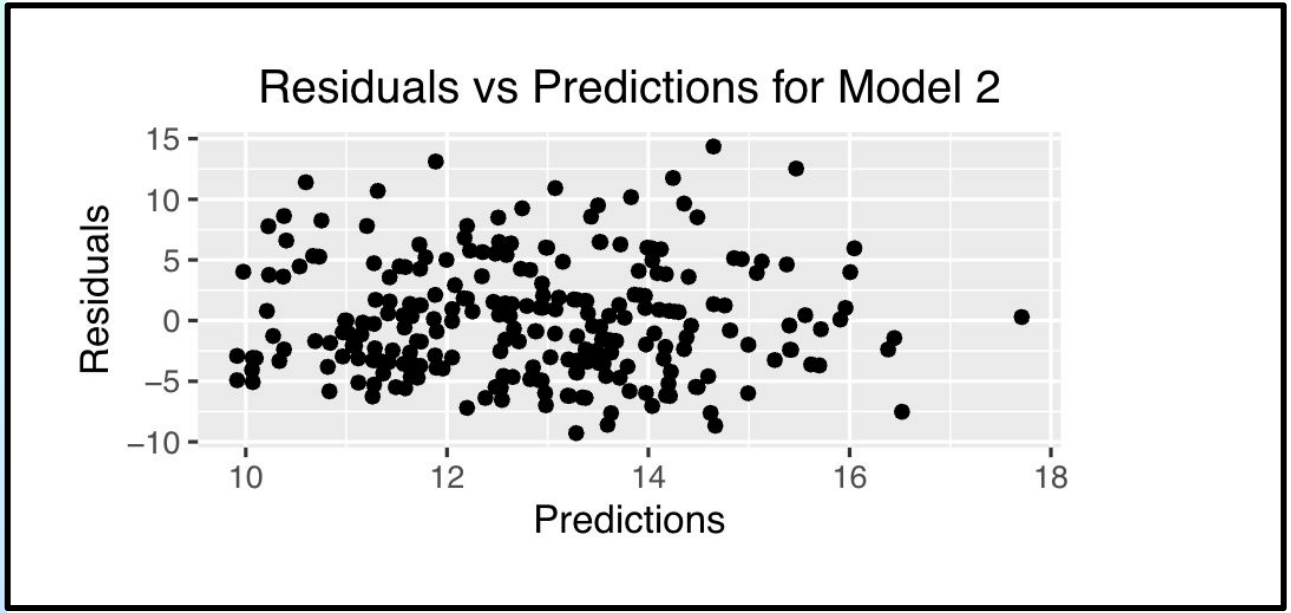


No Perfect Collinearity

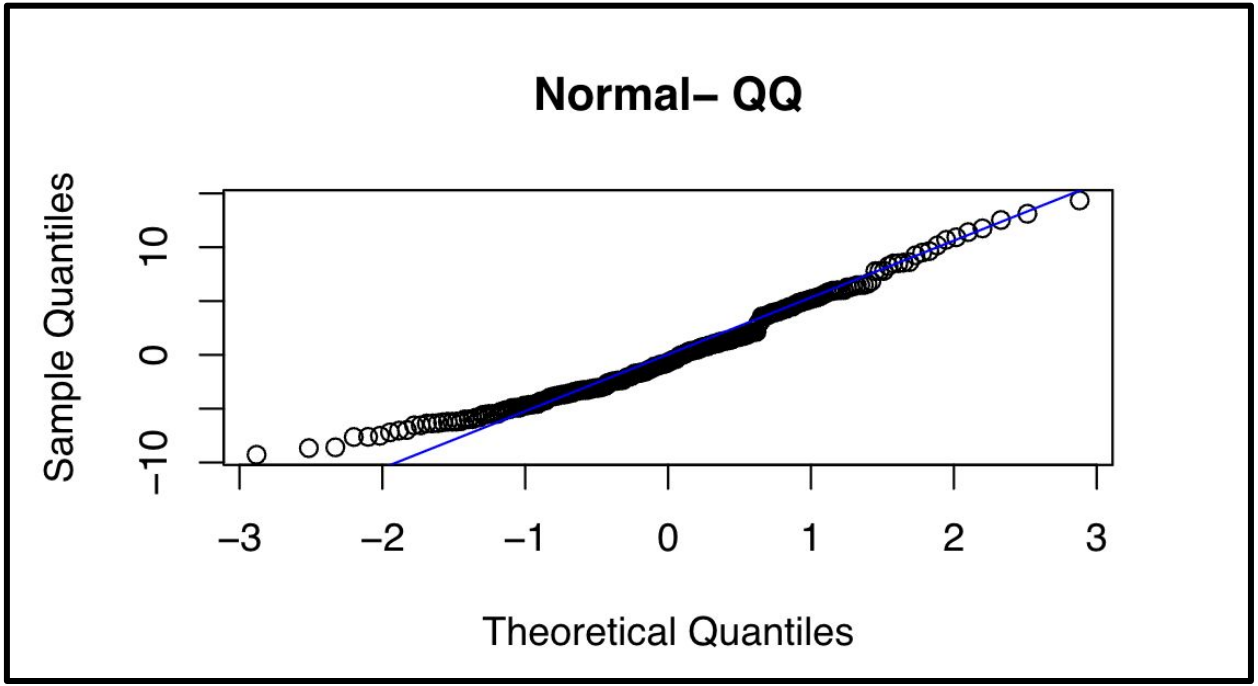
Variance Inflation Factor (VIF)

Number of languages	Gender	Age	Education	Income
1.058058	1.062037	1.004101	1.16781	1.181965

Homoskedastic Conditional Variance



Normally Distributed Errors



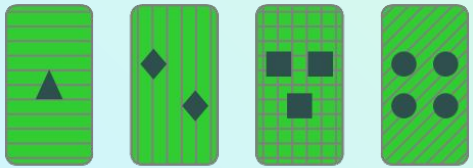
Findings

- Statistically Significant:
  - Age
  - Education Level
- Practically Significant:
  - Education Level
- No Significance:
  - Number of Languages
  - Gender
  - Income Level

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.72152    4.71846   2.060  0.04042 *
number_lang -0.06183    0.31530  -0.196  0.84470
gender      -1.04945    0.62149  -1.689  0.09256 .
age          0.15528    0.05618   2.764  0.00615 **
education   -1.05116    0.30150  -3.486  0.00058 ***
income      -0.14817    0.24136  -0.614  0.53985
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Challenges + Reflections

- Y variable is discrete → logistic regression ideal
- Limited and localized sample size.
- Further studies could analyze other behavior related to multilingualism.



WCST card game!

Thank You!