**Notes**

* Using RData files to preserve column descriptions. They would be lost in excel
* Several Typos in spreadsheet
  + HBAC for a couple sheets
  + Wrong file name for a few
  + Year 11 independent variable name changes
* Missing Data:
  + Y3CogVit SAS file
  + “faminc” in Year5ClnVis, removed row from spreadsheet
* In read process, PH data just included in year1. To be joined later

**Longitudinal Data**

* All data organized in year subdirectories
* Includes PH data in year 1

**Other Patient Data**

* All data not in yearly subdirectories
* psych, and DNA
* Calling it “Patient Data”

**Medical Data**

* Just using “MIFCod”, which has the same info as “MIF”, but formatted differently

Source:

<https://medlineplus.gov/genetics/gene/apoe/#conditions>

# Filters

* Removed years that were missing *both* dementia data and work status in the same year