Example 4b: General Linear Models with Multiple Fixed Effects of Multiple Predictors
(complete syntax, data, and output available for STATA, R, and SAS electronically)

The models for this example come from Hoffman (2015) chapter 2. Using this sample of 550 older adults (which was simulated based on real data), we will examine the extent to which cognition (as measured by the *information test*, a measure of crystallized general intelligence) can be predicted from quantitative age (centered at 85 years), quantitative grip strength (centered at 9 pounds per square inch), binary sex (with men as the reference group), and subsequent three-category dementia diagnosis (none = 1, future = 2, and current = 3, with the none group=1 as the reference). We will first examine the bivariate relationship of each predictor with the cognition outcome (via Pearson correlations if possible or with general linear models as needed), followed by their effects in sequential models, building up to a combined model with additive effects of each predictor.

STATA Syntax for Importing and Preparing Data for Analysis:

```
// Paste in the folder address where "Example4b_Data.xlsx" is saved between " "
cd "C:\Dropbox\24_PSQF6243\PSQF6243_Example4b"
// Using UIowa virtual desktop instead
//cd "\\Client\C:\Dropbox\24_PSQF6243\PSQF6243_Example4b"

// IMPORT Example4b.xlsx data from working directory using exact file name
// To change all variable names to lowercase, remove "case(preserve")
clear // Clear before means close any open data
import excel "Example4b_Data.xlsx", case(preserve) firstrow sheet("Example4b") clear
// Clear after means re-import if it already exists (if need to start over)

// Center quantitative predictors near their means
gen age85 = age - 85
gen grip9 = grip - 9
// Create 2 indicator-dummy-coded binary predictors for 3 dementia groups
gen demNF=.  // Create 2 new empty variables
gen demNC=.
// Replace for demgroup = none
replace demNF=0 if demgroup==1
replace demNC=0 if demgroup==1
// Replace for demgroup = future
replace demNF=1 if demgroup==2
replace demNC=0 if demgroup==2
// Replace for demgroup = current
replace demNF=0 if demgroup==3
replace demNC=1 if demgroup==3
// Label all variables
label variable age85     "age85: Age in Years (0=85)"
label variable grip9     "grip9: Grip Strength in Pounds (0=9)"
label variable sexMW     "sexMW: Sex (0=Men, 1=Women)"
label variable demNF     "demNF: Dementia Contrast for None=0 vs Future=1"
label variable demNC     "demNC: Dementia Contrast for None=0 vs Current=1"
label variable cognition "cognition: Cognition Outcome"
label variable demgroup  "demgroup: Dementia Group 1N 2F 3C"

// Select cases complete on all variables to be used
egen nmiss=rowmiss(cognition age grip sexMW demgroup)
drop if nmiss>0
```

R Syntax for Importing and Preparing Data for Analysis (after loading packages *readxl*, *psych*, *Hmisc*, *supernova*, *multcomp*, *lmhelpers*, and *TeachingDemos*):

```
# Set working directory (to import and export files to)
# Paste in the folder address where "Example4b_Data.xlsx" is saved in quotes
setwd("C:/Dropbox/24_PSQF6243/PSQF6243_Example4b")
```

```
# Import Example4b_Data.xlsx data from working directory -- path = file name
Example4b = read_excel(path="Example4b_Data.xlsx", sheet="Example4b")
# Convert to data frame to use for analysis
Example4b = as.data.frame(Example4b)

# Center quantitative predictors near their means
Example4b$age85=Example4b$age-85
Example4b$grip9=Example4b$grip-9

# Create 2 indicator-dummy-coded binary predictors for 3 dementia groups
Example4b$demNF=NA; Example4b$demNC=NA  # Create 2 new empty variables
Example4b$demNF[which(Example4b$demgroup==1)]=0  # Replace each for none group
Example4b$demNC[which(Example4b$demgroup==1)]=0
Example4b$demNF[which(Example4b$demgroup==2)]=1  # Replace each for future group
Example4b$demNC[which(Example4b$demgroup==2)]=0
Example4b$demNF[which(Example4b$demgroup==3)]=0  # Replace each for current group
Example4b$demNC[which(Example4b$demgroup==3)]=1
# demNF: None=0 vs Future=1
# demNC: None=0 vs Current=1

# Select cases complete on all variables to be used
Example4b = Example4b[complete.cases(Example4b[ ,
            c("cognition","age","grip","sexMW","demgroup")]),]
```

Note: I also wrote five custom functions to automate calculations of effect sizes from lm or glht output—please see code online for these (as used in this example).

Syntax and Output for Descriptive Statistics:

```
display "STATA Descriptive Statistics for Quantitative Variables"
summarize cognition age grip sexMW  // add detail after comma for more info

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
   cognition |        550    24.82182    10.98903          0         44
         age |        550    84.92679    3.430029   80.01649   96.96728
        grip |        550    9.112727    2.982954          0         19
       sexMW |        550    .5872727    .4927727          0          1
```

> **psych::describe** is used to make sure R is using the function from the psych package (instead of Hmisc, which has a function by the same name

```
print("R Descriptive Statistics for Quantitative Variables")
psych::describe(x=Example4b[ , c("cognition","age","grip","sexMW")])

          vars   n  mean    sd median trimmed   mad   min   max range  skew kurtosis   se
cognition    1 550 24.82 10.99  25.00   25.18 11.86  0.00 44.00 44.00 -0.26    -0.62 0.47
age          2 550 84.93  3.43  84.33   84.49  2.88 80.02 96.97 16.95  1.10     0.72 0.15
grip         3 550  9.11  2.98   9.00    9.11  2.97  0.00 19.00 19.00 -0.01    -0.17 0.13
sexMW        4 550  0.59  0.49   1.00    0.61  0.00  0.00  1.00  1.00 -0.35    -1.88 0.02
```

```
display "STATA Descriptive Statistics for Categorical Variables"
tabulate sexMW
tabulate demgroup

 sexMW: Sex |
   (0=Men,  |
  1=Women)  |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        227       41.27       41.27
          1 |        323       58.73      100.00
------------+-----------------------------------
      Total |        550      100.00
```

```
   demgroup: |
   1N 2F 3C |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        399       72.55       72.55
          2 |        109       19.82       92.36
          3 |         42        7.64      100.00
------------+-----------------------------------
      Total |        550      100.00
```

```
print("R Descriptive Statistics for Categorical Variables")
prop.table(table(x=Example4b$sexMW,useNA="ifany"))
```

```
     0      1
0.4127 0.5873
```

```
prop.table(table(x=Example4b$demgroup,useNA="ifany"))
```

```
      1       2       3
0.72545 0.19818 0.07636
```

Syntax and Partial Output for Creating Bivariate Pearson Correlations:

```
display "STATA Bivariate Correlations"
pwcorr cognition age grip sexMW, sig
```

```
             | cognit~n      age     grip    sexMW
-------------+------------------------------------
   cognition |   1.0000
             |
             |
         age |  -0.1705   1.0000
             |   0.0001
             |
        grip |   0.2418  -0.1841   1.0000
             |   0.0000   0.0000
             |
       sexMW |  -0.2363   0.0456  -0.4032   1.0000
             |   0.0000   0.2858   0.0000
```

Note that the binary predictors for dementia group are not included in this correlation matrix. This is because each of their meanings would differ if the predictor were included in separate correlations than when both predictors are together in the same model:

demgroup	demNF	demNC
1 = None	0	0
2 = Future	1	0
3 = Current	0	1

```
print("R Pearson Correlation Matrix with P-values using rcorr from Hmisc package")
print("Must convert data frame to matrix to use rcorr")
rcorr(x=as.matrix(Example4b[,c("cognition","age","grip","sexMW")]), type="pearson")
```

```
 Pearson correlation matrix
           cognition   age  grip sexMW
 cognition      1.00 -0.17  0.24 -0.24
 age           -0.17  1.00 -0.18  0.05
 grip           0.24 -0.18  1.00 -0.40
 sexMW         -0.24  0.05 -0.40  1.00

 P-values for Pearson correlations
           cognition age    grip   sexMW
 cognition           0.0000 0.0000 0.0000
 age       0.0000           0.0000 0.2858
 grip      0.0000    0.0000        0.0000
 sexMW     0.0000    0.2858 0.0000
```

What can we conclude about the linear relationships between each pair of variables?

Cognition and age:

Cognition and grip:

Cognition and sex:

Age and grip:

Age and sex:

Grip and sex:

Our first predictor is quantitative age—we know from the Pearson correlation matrix that there is a significant linear relationship between age and cognition, but is that linear trend sufficient to fully describe how age relates to cognition? We will use a GLM to check for a possible bivariate curvilinear relationship between cognition and age instead by including both age and age^2 (each centered so that 0 = 85 years, near the mean of age).

Syntax and Output for <u>Linear + Quadratic Age</u> (0=85 years) Predicting Cognition:

$$\boldsymbol{Cognition_i = \beta_0 + \beta_1(Age_i - 18) + \beta_2(Age_i - 18)^2 + e_i}$$

```
display "STATA Linear + Quadratic Age (0=85) Predicting Cognition"
regress cognition c.age85 c.age85#c.age85, level(95)
```

```
      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(2, 547)       =      8.34
       Model |  1961.68397         2  980.841985   Prob > F        =    0.0003
    Residual |  64334.8542       547  117.613993   R-squared       =    0.0296
-------------+----------------------------------   Adj R-squared   =    0.0260
       Total |  66296.5382       549  120.758722   Root MSE        =    10.845

--------------------------------------------------------------------------------
      cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
          age85 |  -.6094648   .1775233    -3.43   0.001    -.9581756    -.260754
c.age85#c.age85 |   .0175194   .0318874     0.55   0.583    -.0451174    .0801562
          _cons |   24.57136    .600584    40.91   0.000     23.39163    25.75109
--------------------------------------------------------------------------------
```

```
print("R Linear + Quadratic Age (0=85) Predicting Cognition")
ModelQAge = lm(data=Example4b, formula=cognition~1+age85+I(age85^2))
supernova(ModelQAge)           # supernova prints sums of squares and residual variance
SummaryCI(ModelQAge, level=.95) # custom function to add CIs to fixed effects table
```

Interpret $\boldsymbol{\beta_0}$ = Intercept:

Interpret $\boldsymbol{\beta_1}$ = slope of age85:

Interpret $\boldsymbol{\beta_2}$ = slope of age85^2:

Did $\boldsymbol{R^2}$ change significantly relative to linear age only (for which $\boldsymbol{r^2 = -0.17045^2 = .0291}$)?

Our second predictor is quantitative grip strength—we know from the Pearson correlation matrix that there is a significant linear relationship between grip strength and cognition, but is that linear trend sufficient to describe how grip strength relates to cognition? We will estimate a GLM to check for a possible bivariate curvilinear relationship instead by including both grip and grip^2 (each centered so that 0 = 9 pounds per square inch, near the mean of grip strength).

Syntax and Output for <u>Linear + Quadratic</u> Grip (0=9 pounds) Predicting Cognition:

$$\boldsymbol{Cognition_i = \beta_0 + \beta_1(Grip_i - 9) + \beta_2(Grip_i - 9)^2 + e_i}$$

```
display "STATA Linear + Quadratic Grip (0=9) Predicting Cognition"
regress cognition c.grip9 c.grip9#c.grip9, level(95)

print("R Linear + Quadratic Grip (0=9) Predicting Cognition")
ModelQGrip = lm(data=Example4b, formula=cognition~1+grip9+I(grip9^2))
```

```
supernova(ModelQGrip)            # supernova prints sums of squares and residual variance
                                        SS  df        MS      F   PRE     p
 ---------- --------------- | --------- --- -------- ------ ----- -----
      Model (error reduced) |  3897.774   2 1948.887 17.084 .0588 .0000
      grip9                 |  3838.870   1 3838.870 33.652 .0580 .0000
 I(grip9^2)                 |    20.543   1   20.543  0.180 .0003 .6715
      Error (from model)    | 62398.764 547  114.075
 ---------- --------------- | --------- --- -------- ------ ----- -----
      Total (empty model)   | 66296.538 549  120.759

SummaryCI(ModelQGrip, level=.95) # custom function to add CIs to fixed effects table

            Estimate  StdErr t-value    p-value  LowerCI UpperCI
(Intercept) 24.57891 0.56607 43.4200 2.247e-179 23.46696 25.6908
grip9        0.88762 0.15301  5.8011  1.116e-08  0.58706  1.1882 btw, p < .0001
I(grip9^2)   0.01606 0.03785  0.4244  6.715e-01 -0.05828  0.0904 btw, p = .6715
```

Interpret β_0 = Intercept:

Interpret β_1 = slope of grip9:

Interpret β_2 = slope of grip9^2:

Did R^2 change significantly relative to linear grip only (for which $r^2 = 0.24183^2 = .0585$)?

Our third predictor is three-category dementia group—we do not yet know if there is a significant bivariate relationship between dementia group and cognition—a Pearson correlation would not have provided the desired interpretation of each binary predictor (i.e., it would have lumped together both groups coded 0 within each binary predictor). Instead, we can examine their bivariate relation by predicting cognition with dementia group in a GLM:

demgroup	demNF	demNC
1 = None	0	0
2 = Future	1	0
3 = Current	0	1

Syntax and Output with Three-Category Dementia Group Predicting Cognition:

$$Cognition = \beta_0 + \beta_1(DemNF_i) + \beta_2(DemNC_i) + e_i$$

None Mean: $\hat{y}_N = \beta_0 + \beta_1(0) + \beta_2(0) = \beta_0$ ← fixed effect #1

Future Mean: $\hat{y}_F = \beta_0 + \beta_1(1) + \beta_2(0) = \beta_0 + \beta_1$ ← linear combination

Current Mean: $\hat{y}_C = \beta_0 + \beta_1(0) + \beta_2(1) = \beta_0 + \beta_2$ ← linear combination

Diff of None vs. Future: $(\beta_0 + \beta_1) - (\beta_0) = \beta_1$ ← fixed effect #2

Diff of None vs. Current: $(\beta_0 + \beta_2) - (\beta_0) = \beta_2$ ← fixed effect #3

Diff of Future vs. Current: $(\beta_0 + \beta_2) - (\beta_0 + \beta_1) = \beta_2 - \beta_1$ ← linear combination

```
display "STATA Dementia Group (Ref=None) Predicting Cognition"
regress cognition c.demNF c.demNC, level(95)
estimates store DemOnly       // Save all model results for R2 in effect sizes below

      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(2, 547)       =     58.52
       Model |  11685.5109         2  5842.75547   Prob > F        =    0.0000
    Residual |  54611.0272       547  99.8373441   R-squared       =    0.1763
-------------+----------------------------------   Adj R-squared   =    0.1732
       Total |  66296.5382       549  120.758722   Root MSE        =    9.9919
```

```
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       demNF |  -5.675059   1.079888    -5.26   0.000    -7.796294   -3.553824
       demNC |  -16.38847   1.620894   -10.11   0.000    -19.57241   -13.20453
       _cons |   27.19799   .5002189    54.37   0.000     26.21541    28.18058
------------------------------------------------------------------------------
```

Interpret β_0 = Intercept:

Interpret β_1 = slope of None vs. Future:

Interpret β_2 = slope of None vs. Current:

```
lincom _cons*1 + c.demNF*0 + c.demNC*0 // Cognition Mean for None    = beta0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   27.19799   .5002189    54.37   0.000     26.21541    28.18058
------------------------------------------------------------------------------

lincom _cons*1 + c.demNF*1 + c.demNC*0 // Cognition Mean for Future  = beta0+beta1
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   21.52294    .957047    22.49   0.000       19.643    23.40287
------------------------------------------------------------------------------

lincom _cons*1 + c.demNF*0 + c.demNC*1 // Cognition Mean for Current = beta0+beta2
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   10.80952   1.541778     7.01   0.000     7.780993    13.83805
------------------------------------------------------------------------------

estimates restore DemOnly     // Restore model results for post-estimations below
lincom c.demNF*1  + c.demNC*0 // Mean Diff: None vs. Future    = beta1
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))
   display "Partial r = "        (r(estimate)/r(se))/sqrt((r(estimate)/r(se))^2+r(df))
   display "Semi-Partial r = "   (r(estimate)/r(se))*sqrt((1-e(r2))/r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -5.675059   1.079888    -5.26   0.000    -7.796294   -3.553824
------------------------------------------------------------------------------
Partial d = -.44939482
Partial r = -.21923118
Semi-Partial r = -.20393549

lincom c.demNF*0  + c.demNC*1 // Mean Diff: None vs. Current   = beta2
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))
   display "Partial r = "        (r(estimate)/r(se))/sqrt((r(estimate)/r(se))^2+r(df))
   display "Semi-Partial r = "   (r(estimate)/r(se))*sqrt((1-e(r2))/r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -16.38847   1.620894   -10.11   0.000    -19.57241   -13.20453
------------------------------------------------------------------------------
Partial d = -.86460962
Partial r = -.39681246
Semi-Partial r = -.39236008

lincom c.demNF*-1 + c.demNC*1 // Mean Diff: Future vs. Current = beta2-beta1
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))
```

```
display "Partial r = "        (r(estimate)/r(se))/sqrt((r(estimate)/r(se))^2+r(df))
display "Semi-Partial r = "   (r(estimate)/r(se))*sqrt((1-e(r2))/r(df))
```

```
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -10.71341   1.814668    -5.90   0.000    -14.27798   -7.148842
------------------------------------------------------------------------------
Partial d = -.50485545
Partial r = -.24475042
Semi-Partial r = -.22910354
```

```
print("R 3-Category Dementia Group (Ref=None) Predicting Cognition")
ModelDem = lm(data=Example4b, formula=cognition~1+demNF+demNC)
supernova(ModelDem)          # supernova prints sums of squares and residual variance
```

```
                                SS  df        MS       F   PRE     p
 ----- --------------- | --------- --- --------- ------- ----- -----
 Model (error reduced) | 11685.511   2  5842.755  58.523 .1763 .0000
 demNF                 |  2757.252   1  2757.252  27.617 .0481 .0000
 demNC                 | 10206.116   1 10206.116 102.227 .1575 .0000
 Error (from model)    | 54611.027 547    99.837
 ----- --------------- | --------- --- --------- ------- ----- -----
 Total (empty model)   | 66296.538 549   120.759
```

```
SummaryCI(ModelDem, level=.95) # custom function to add CIs to fixed effects table
```

```
            Estimate StdErr t-value    p-value LowerCI UpperCI
(Intercept)   27.198 0.5002  54.372 9.846e-223  26.215  28.181
demNF         -5.675 1.0799  -5.255  2.122e-07  -7.796  -3.554
demNC        -16.388 1.6209 -10.111  3.793e-22 -19.572 -13.205
```

```
print("R Ask for predicted cognition per group and group differences")
print("In number lists below, values are multiplier for each fixed effect IN ORDER")
glhtModelDem = glht(model=ModelDem, linfct=rbind(
  "Cognition Mean: None"         = c(1, 0, 0),
  "Cognition Mean: Future"       = c(1, 1, 0),
  "Cognition Mean: Current"      = c(1, 0, 1),
  "Mean Diff: None vs Future"    = c(0, 1, 0),
  "Mean Diff: None vs Current"   = c(0, 0, 1),
  "Mean Diff: Future vs Current" = c(0,-1, 1)))

glhtSummaryCI(glhtModelDem, level=.95)  # custom function to add CIs to glht output table
```

```
                             Estimate StdErr   p-value LowerCI UpperCI
Cognition Mean: None           27.198 0.5002 0.000e+00  26.215  28.181
Cognition Mean: Future         21.523 0.9570 0.000e+00  19.643  23.403
Cognition Mean: Current        10.810 1.5418 6.998e-12   7.781  13.838
Mean Diff: None vs Future      -5.675 1.0799 2.122e-07  -7.796  -3.554
Mean Diff: None vs Current    -16.388 1.6209 0.000e+00 -19.572 -13.205
Mean Diff: Future vs Current  -10.713 1.8147 6.239e-09 -14.278  -7.149
```

```
glhtEffectSizes(glhtObject=glhtModelDem, modelObject=ModelDem,
                level=.95) # custom function to compute glht effect sizes
```

```
                         Estimate   p-value Partial-d Partial-r SemiPartial-r Partial-R2 SemiPartial-R2
Cog Mean: None             27.198 0.000e+00    4.6496    0.9186        2.1100    0.84386        4.45201
Cog Mean: Future           21.523 0.000e+00    1.9231    0.6931        0.8727    0.48041        0.76162
Cog Mean: Current          10.810 6.998e-12    0.5995    0.2871        0.2721    0.08245        0.07402
Diff: None vs Future       -5.675 2.122e-07   -0.4494   -0.2192       -0.2039    0.04806        0.04159
Diff: None vs Current     -16.388 0.000e+00   -0.8646   -0.3968       -0.3924    0.15746        0.15395
Diff: Future vs Current   -10.713 6.239e-09   -0.5049   -0.2448       -0.2291    0.05990        0.05249
```

The next models follow the order used in Hoffman (2015) chapter 2 to illustrate how one would conduct a sequential analysis involving multiple predictors. At each step (starting with age, then adding grip strength, then adding sex, then adding dementia group), we could answer any research questions about how the new predictor variable contributes uniquely after controlling for the previous predictors.

STATA Syntax and Output for Sequential Models Predicting Cognition:

Final Model: $Cognition_i = \beta_0 + \beta_1(Age_i - 85) + \beta_2(Grip_i - 9) + \beta_3(SexMW_i) + \beta_4(DemNF_i) + \beta_5(DemNC_i) + e_i$

```
display "STATA code for how to use nestreg to get F-test for R2 change"
// Format below is outcome (what to add to model at each step)
nestreg: regress cognition (age85) (grip9) (sexMW) (demNF demNC)
estimates store Combined    // Save all model results for R2 in effect sizes below

Block  1: age85
      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(1, 548)       =     16.40
       Model |  1926.18134         1  1926.18134   Prob > F        =    0.0001
    Residual |  64370.3568       548  117.464155   R-squared       =    0.0291
-------------+----------------------------------   Adj R-squared   =    0.0273
       Total |  66296.5382       549  120.758722   Root MSE        =    10.838
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |  -.5460905   .1348555    -4.05   0.000    -.8109876   -.2811935
       _cons |   24.78184   .4622431    53.61   0.000     23.87385    25.68982
------------------------------------------------------------------------------

Block  2: grip9
      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(2, 547)       =     22.14
       Model |  4965.35362         2  2482.67681   Prob > F        =    0.0000
    Residual |  61331.1846       547  112.122824   R-squared       =    0.0749
-------------+----------------------------------   Adj R-squared   =    0.0715
       Total |  66296.5382       549  120.758722   Root MSE        =    10.589
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |  -.4175853   .1340458    -3.12   0.002     -.680893   -.1542777
       grip9 |   .8024817   .1541361     5.21   0.000     .4997105    1.105253
       _cons |   24.70078   .4518795    54.66   0.000     23.81315    25.58842
------------------------------------------------------------------------------

Block  3: sexMW
      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(3, 546)       =     20.04
       Model |  6574.67775         3  2191.55925   Prob > F        =    0.0000
    Residual |  59721.8604       546  109.380697   R-squared       =    0.0992
-------------+----------------------------------   Adj R-squared   =    0.0942
       Total |  66296.5382       549  120.758722   Root MSE        =    10.459

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |  -.4337719   .1324638    -3.27   0.001    -.6939729   -.1735709
       grip9 |   .5460019   .1662766     3.28   0.001     .2193818    .8726221
       sexMW |   -3.79878   .9903591    -3.84   0.000    -5.744161   -1.853399
       _cons |   26.95943   .7388729    36.49   0.000     25.50805    28.41081
------------------------------------------------------------------------------
```

```
Block  4: demNF demNC
      Source |       SS           df       MS      Number of obs   =       550
-------------+----------------------------------   F(5, 544)       =     41.75
       Model |  18385.9793         5  3677.19586   Prob > F        =    0.0000
    Residual |  47910.5589       544  88.0708803   R-squared       =    0.2773
-------------+----------------------------------   Adj R-squared   =    0.2707
       Total |  66296.5382       549  120.758722   Root MSE        =    9.3846
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |   -.405734   .1188972    -3.41   0.001    -.6392878   -.1721802
       grip9 |   .6042256   .1497757     4.03   0.000      .310016    .8984351
       sexMW |  -3.657374   .8914326    -4.10   0.000    -5.408446   -1.906303
       demNF |  -5.721971   1.019078    -5.61   0.000    -7.723782    -3.72016
       demNC |  -16.47981   1.522754   -10.82   0.000    -19.47101   -13.48862
       _cons |   29.26433   .6985079    41.90   0.000     27.89222    30.63643
------------------------------------------------------------------------------
```

The table below provides an F-test for the change in R^2 at each step:

```
  +-------------------------------------------------------------+
  |       |          Block  Residual                     Change |
  | Block |       F     df        df   Pr > F       R2    in R2 |
  |-------+-----------------------------------------------------|
  |     1 |   16.40      1       548   0.0001   0.0291          |
  |     2 |   27.11      1       547   0.0000   0.0749   0.0458 |
  |     3 |   14.71      1       546   0.0001   0.0992   0.0243 |
  |     4 |   67.06      2       544   0.0000   0.2773   0.1782 |
```

The code below shows an alternative way to test the change in R^2 for multiple predictors using TEST instead of NESTREG, such as for the combination of age and sex here:

```
ereturn list                 // See what has been stored automatically
global SSresidual = e(rss)   // Save last model SS residual for effect sizes below
global SSfull = e(mss)       // Save last model SS model for effect sizes below

display "F-Test DFnum=2 for Age and Sex"
test (c.age85=0)(c.sexMW=0)  // Demographic variables

       F(  2,   544) =   13.79
            Prob > F =    0.0000

display "STATA Reduced Model to Get Model SS from Omitting Age and Sex"
quietly regress cognition c.grip9 c.demNF c.demNC, level(95)
global  SSeffect = $SSfull - e(mss)
display "Partial R2 = "      $SSeffect/($SSeffect+$SSresidual)
display "Semi-Partial R2 = " $SSeffect/($SSfull+$SSresidual)
Partial R2 = .04825628
Semi-Partial R2 = .03664158
```

Lastly, we can request eta-family effect sizes for the fixed slopes of age, grip, and sex (the R^2 versions should not be used for the binary dementia group predictors given their common reference group):

```
// Get partial correlations for age, grip, and sex (sr2 for dem slopes are not valid)
pcorr cognition c.age85 c.grip9 c.sexMW c.demNF c.demNC

                 Partial   Semipartial      Partial   Semipartial   Significance
   Variable |    Corr.         Corr.      Corr.^2       Corr.^2          Value
------------+-----------------------------------------------------------------
      age85 |   -0.1448       -0.1244       0.0210        0.0155         0.0007
      grip9 |    0.1704        0.1470       0.0290        0.0216         0.0001
      sexMW |   -0.1732       -0.1495       0.0300        0.0224         0.0000
      demNF |   -0.2340       -0.2046       0.0548        0.0419         0.0000
      demNC |   -0.4209       -0.3945       0.1772        0.1556         0.0000
```

And we can request Cohen's d effect sizes for group difference by sex and by dementia group:

```
estimates restore Combined    // Restore model results for post-estimations below
lincom c.sexMW*1              // Mean Diff: Men vs. Women = beta3
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -3.657374   .8914326    -4.10   0.000    -5.408446   -1.906303
------------------------------------------------------------------------------
Partial d = -.35181263

lincom c.demNF*1  + c.demNC*0 // Mean Diff: None vs. Future = beta4
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -5.721971   1.019078    -5.61   0.000    -7.723782    -3.72016
------------------------------------------------------------------------------
Partial d = -.48146926

lincom c.demNF*0  + c.demNC*1 // Mean Diff: None vs. Current = beta5
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -16.47981   1.522754   -10.82   0.000    -19.47101   -13.48862
------------------------------------------------------------------------------
Partial d = -.92801118

lincom c.demNF*-1 + c.demNC*1 // Mean Diff: Future vs. Current = beta5-beta4
   display "Partial d = "        (2*(r(estimate)/r(se)))/sqrt(r(df))
   display "Partial r = "        (r(estimate)/r(se))/sqrt((r(estimate)/r(se))^2+r(df))
   display "Semi-Partial r = "   (r(estimate)/r(se))*sqrt((1-e(r2))/r(df))

------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -10.75784   1.707957    -6.30   0.000    -14.11284   -7.402844
------------------------------------------------------------------------------
Partial d = -.54010571
Partial r = -.26071341
Semi-Partial r = -.22957202
```

R Syntax and Output for Sequential Models Predicting Cognition:

Final Model: $$Cognition_i = \beta_0 + \beta_1(Age_i - 85) + \beta_2(Grip_i - 9) + \beta_3(SexMW_i) + \beta_4(DemNF_i) + \beta_5(DemNC_i) + e_i$$

```
print("R Linear Age (0=85) Predicting Cognition")
ModelAge = lm(data=Example4b, formula=cognition~1+age85)
supernova(ModelAge)           # supernova prints sums of squares and residual variance
SummaryCI(ModelAge, level=.95) # custom function to add CIs to fixed effects table

                                SS  df       MS      F   PRE     p
 ----- --------------- | --------- --- -------- ------ ----- -----
 Model (error reduced) |  1926.181   1 1926.181 16.398 .0291 .0001
 Error (from model)    | 64370.357 548  117.464
 ----- --------------- | --------- --- -------- ------ ----- -----
 Total (empty model)   | 66296.538 549  120.759
```

```
            Estimate StdErr t-value    p-value LowerCI UpperCI
(Intercept)  24.7818 0.4622  53.612 3.928e-220  23.874 25.6898
age85        -0.5461 0.1349  -4.049  5.874e-05  -0.811 -0.2812
```

```
print("R Add Linear Grip Strength (0=9) Predicting Cognition")
ModelAgeGrip = lm(data=Example4b, formula=cognition~1+age85+grip9)
supernova(ModelAgeGrip)            # supernova prints sums of squares and residual variance
SummaryCI(ModelAgeGrip, level=.95) # custom function to add CIs to fixed effects table
```

```
                                  SS  df       MS      F   PRE     p
 ----- --------------- | --------- --- -------- ------ ----- -----
 Model (error reduced) |  4965.354   2 2482.677 22.142 .0749 .0000
 age85                 |  1088.123   1 1088.123  9.705 .0174 .0019
 grip9                 |  3039.172   1 3039.172 27.106 .0472 .0000
 Error (from model)    | 61331.185 547  112.123
 ----- --------------- | --------- --- -------- ------ ----- -----
 Total (empty model)   | 66296.538 549  120.759
```

```
            Estimate StdErr t-value    p-value LowerCI UpperCI
(Intercept)  24.7008 0.4519  54.662 8.417e-224 23.8132 25.5884
age85        -0.4176 0.1340  -3.115  1.934e-03 -0.6809 -0.1543
grip9         0.8025 0.1541   5.206  2.731e-07  0.4997  1.1053
```

```
print("R Add Binary Sex (0=M, 1=W) Predicting Cognition")
ModelAgeGripSex = lm(data=Example4b, formula=cognition~1+age85+grip9+sexMW)
supernova(ModelAgeGripSex)            # supernova prints sums of squares and residual variance
SummaryCI(ModelAgeGripSex, level=.95) # custom function to add CIs to fixed effects table
```

```
                                  SS  df       MS      F   PRE     p
 ----- --------------- | --------- --- -------- ------ ----- -----
 Model (error reduced) |  6574.678   3 2191.559 20.036 .0992 .0000
 age85                 |  1172.922   1 1172.922 10.723 .0193 .0011
 grip9                 |  1179.416   1 1179.416 10.783 .0194 .0011
 sexMW                 |  1609.324   1 1609.324 14.713 .0262 .0001
 Error (from model)    | 59721.860 546  109.381
 ----- --------------- | --------- --- -------- ------ ----- -----
 Total (empty model)   | 66296.538 549  120.759
```

```
            Estimate StdErr t-value    p-value LowerCI UpperCI
(Intercept)  26.9594 0.7389  36.487 1.532e-148 25.5080 28.4108
age85        -0.4338 0.1325  -3.275  1.125e-03 -0.6940 -0.1736
grip9         0.5460 0.1663   3.284  1.090e-03  0.2194  0.8726
sexMW        -3.7988 0.9904  -3.836  1.398e-04 -5.7442 -1.8534
```

```
print("R Combined Model Predicting Cognition")
ModelCombined = lm(data=Example4b, formula=cognition~1+age85+grip9+sexMW+demNF+demNC)
supernova(ModelCombined)            # supernova prints sums of squares and residual variance
SummaryCI(ModelCombined, level=.95) # custom function to add CIs to fixed effects table
```

```
                                 SS  df        MS       F   PRE     p
 ----- --------------- | --------- --- --------- ------- ----- -----
 Model (error reduced) | 18385.979   5  3677.196  41.753 .2773 .0000
 age85                 |  1025.586   1  1025.586  11.645 .0210 .0007
 grip9                 |  1433.336   1  1433.336  16.275 .0290 .0001
 sexMW                 |  1482.498   1  1482.498  16.833 .0300 .0000
 demNF                 |  2776.568   1  2776.568  31.527 .0548 .0000
 demNC                 | 10315.200   1 10315.200 117.124 .1772 .0000
 Error (from model)    | 47910.559 544    88.071
 ----- --------------- | --------- --- --------- ------- ----- -----
```

```
              Estimate   StdErr  t-value     p-value   LowerCI   UpperCI
(Intercept)  29.264325 0.698508 41.89548 2.11219e-172 27.892222 30.636428
age85        -0.405734 0.118897 -3.41248  6.91650e-04 -0.639288 -0.172180
grip9         0.604226 0.149776  4.03420  6.26304e-05  0.310016  0.898435
sexMW        -3.657374 0.891433 -4.10281  4.70693e-05 -5.408446 -1.906303
```

```
demNF        -5.721971 1.019078  -5.61485  3.14037e-08  -7.723782  -3.720160
demNC       -16.479813 1.522754 -10.82238  7.44927e-25 -19.471010 -13.488616

# Get F-test and effect sizes for fixed slopes of interest using custom function
R2changeF(ReducedModel=ModelAge,        FullModel=ModelAgeGrip,    PredName="Grip")
R2changeF(ReducedModel=ModelAgeGrip,    FullModel=ModelAgeGripSex, PredName="Sex")
R2changeF(ReducedModel=ModelAgeGripSex, FullModel=ModelCombined,   PredName="Dementia")

F-Test and R2 Change for Grip Slopes
   R2-total R2-change DF-num DF-den F-value        p-value Partial-R2 SemiPartial-R2
2 0.0748961 0.0458421      1    547 27.1057 0.000000273055  0.0472138      0.0458421

F-Test and R2 Change for Sex Slopes
   R2-total R2-change DF-num DF-den F-value     p-value Partial-R2 SemiPartial-R2
2 0.0991708 0.0242746      1    546 14.7131 0.000139784  0.0262399      0.0242746

F-Test and R2 Change for Dementia Slopes
  R2-total R2-change DF-num DF-den F-value     p-value Partial-R2 SemiPartial-R2
2 0.277329  0.178159      2    544 67.0557 9.31175e-27   0.197772       0.178159

# Fit model without fixed slopes of interest (age85 and sexMW for demographics here)
ModelNoAgeSex = lm(data=Example4b, formula=cognition~1+grip9+demNF+demNC)
# Get F-test and effect sizes for fixed slopes of interest using custom function
R2changeF(ReducedModel=ModelNoAgeSex, FullModel=ModelCombined, PredName="Demographics")

F-Test and R2 Change for Demographics Slopes
  R2-total R2-change DF-num DF-den F-value       p-value Partial-R2 SemiPartial-R2
2 0.277329 0.0366416      2    544 13.7912 0.00000143696  0.0482563      0.0366416

print("Effect Sizes for Model Fixed Effects")
FixedEffectSizes(ModelCombined)  # custom function to add effect sizes for fixed slopes

             Estimate      p-value Partial-d Partial-r SemiPartial-r Partial-R2 SemiPartial-R2
(Intercept)  29.264325 2.11219e-172  3.592508  0.873727      1.526996  0.7633992      2.3317171
age85        -0.405734  6.91650e-04 -0.292618 -0.144768     -0.124377  0.0209576      0.0154697
grip9         0.604226  6.26304e-05  0.345930  0.170434      0.147038  0.0290479      0.0216201
sexMW        -3.657374  4.70693e-05 -0.351813 -0.173246     -0.149538  0.0300143      0.0223616
demNF        -5.721971  3.14037e-08 -0.481469 -0.234048     -0.204649  0.0547786      0.0418810
demNC       -16.479813  7.44927e-25 -0.928011 -0.420902     -0.394451  0.1771587      0.1555918

print("R Ask for missing model-implied group difference")
glhtModelCombined = glht(model=ModelCombined,
                         linfct=rbind("Future vs Current" = c(0,0,0, 0,-1,1)))
glhtSummaryCI(glhtModelCombined, level=.95) # custom function to add CIs to glht output

                  Estimate  StdErr    p-value  LowerCI  UpperCI
Future vs Current -10.7578 1.70796 6.1976e-10 -14.1128 -7.40284

glhtEffectSizes(glhtObject=glhtModelCombined, modelObject=ModelCombined,
                level=.95) # custom function to compute glht effect sizes

                 Estimate    p-value Partial-d Partial-r SemiPartial-r Partial-R2 SemiPartial-R2
Future vs Current -10.7578 6.1976e-10 -0.540106 -0.260713     -0.229572  0.0679715      0.0527033
```

Example Results Section [notes about what also to include]:

All analyses were conducted in [software program, version] using [function name]. Table 1 provides descriptive statistics and Pearson bivariate correlations among the cognition outcome and predictors of age, grip strength, and sex (0 = men, 1= women). As shown in Table 1, cognition was predicted to be significantly greater in persons who were younger, who were stronger, and in men relative to women. To provide a meaningful intercept in the models that follow, we centered age such that 0 = 85 years and grip strength such that 0 = 9 pounds per square inch (i.e., near their sample means).

Table 1: Descriptive Statistics and Bivariate Correlations (bold values indicate *p* < .0001)

Variable	Cognition	Age	Grip	Sex (W=1)
Cognition	1.000			
Age	**−0.170**	1.000		
Grip Strength	**0.242**	**−0.184**	1.000	
Sex (W=1)	**−0.236**	0.046	**−0.403**	1.000
Mean	24.822	84.927	9.113	0.587
SD	10.989	3.430	2.983	0.493
Min	0.000	80.016	0.000	0.000
Max	44.000	96.967	19.000	1.000

In separate linear regressions, we examined the potential for a quadratic effect of age and for a quadratic effect of grip strength. Neither quadratic slope was significant [could give *p*-values or effect sizes for completeness], indicating that linear slopes for age and grip strength were likely to be sufficient. In an analysis of variance, we examined the bivariate effect of dementia group (none = 72.55%, future = 19.82%, or current = 7.64%) in predicting cognition. We found significant mean differences in cognition across the three groups, $F(2, 547) = 58.52$, MSE = 99.84, $p < .0001$, $R^2 = .176$. Results (including d partial effect sizes for mean differences in standard deviation units) were as follows. Relative to the reference group of no dementia (M = 27.198, SE = 0.500), cognition was significantly lower by 5.675 (SE = 1.0799, $d = -0.449$) in the future group (M = 21.522, SE = 0.957) and significantly lower by 16.388 (SE = 1.621, $d = -0.865$) in the current group (M = 10.810, SE = 1.542). Cognition in the current group was also significantly lower than the future group by 10.713 (SE = 1.815, $d = -0.505$).

We then examined the incremental contribution of each predictor in sequential models, as summarized in Table 2. The model R^2 remained significant at each step. [More text would be added to describe how this sequence would address the research questions.]

Table 2: Results from Sequential Models Predicting Cognition

New Predictor	DF numerator	DF denominator	F	$p <$	R^2 change	Model R^2
Age	1	548	16.40	.000		.029
Grip Strength	1	547	27.11	.000	.046	.075
Sex	1	546	14.71	.000	.024	.099
Dementia Group	2	544	67.06	.000	.178	.277

The results from the final model with all four predictors (i.e., an analysis of covariance or multiple linear regression) are shown in Table 3. Semipartial squared correlation (sr^2) effect sizes were also obtained to describe the amount of variance captured by distinct sets of predictor slopes. We found that the unique contribution of each predictor remained significant in the same direction as their bivariate effects, such that cognition was predicted to be significantly higher in participants who were younger, who were stronger, in men, and in persons without dementia. The model accounted for a significant amount of variance in cognition, $F(5, 544) = 41.75$, MSE = 88.07, $p < .0001$, $R^2 = .277$, and the omnibus effect of dementia group remained significant, $F(2, 544) = 67.06$, $p < .0001$, $sr^2 = .178$. In addition, the demographic variables of age and sex had a smaller but significant unique joint contribution, $F(2, 544) = 13.79$, $p < .0001$, $sr^2 = .034$, indicating that controlling for dementia status and grip strength did not fully mitigate their effects.

[More text would be added to describe how the unique effects of each predictor would address the research questions.]

Table 3

Results from Combined Model Predicting Cognition

Fixed Effect	**Est**	**SE**	***p <***	***d***	***r***	***sr²***
Intercept	29.264	0.699				
Age (0=85 years)	−0.406	0.119	.001	−0.293	−0.145	0.016
Grip (0=9 pounds)	0.604	0.150	.001	0.346	0.170	0.022
Sex (0=Men)	−3.657	0.891	.001	−0.352	−0.173	0.022
None vs Future Dementia	−5.722	1.019	.001	−0.481	−0.234	
None vs Current Dementia	−16.480	1.523	.001	−0.928	−0.421	0.178
Future vs Current Dementia	−10.758	1.708	.001	−0.540	−0.261	

Note: *d* and *r* partial effect sizes were computed from the slope *t* test-statistics as follows: $d = \frac{2t}{\sqrt{DF_{den}}}$; $r = \frac{t}{\sqrt{t^2+DF_{den}}}$. Semi-partial squared correlations (sr^2) were computed using Type III sums of squares for each slope separately for age, grip strength, and sex, and for the joint combination of the two slopes for dementia group given their common reference group.