```
                                                                  ___  ____  ____  ____  ____(R)
                                                                 /__    /   ____/   /   ____/
                                                                ___/   /   /___/   /   /___/
                                                                  Statistics/Data Analysis

      name:  <unnamed>
       log:  \\med-fs1.med.ualberta.ca\Pediatrics\Symbiota\SyMBIOTA\Datasets\Charlene-Vienna nature
>  19 2022.smcl
  log type:  smcl
 opened on:  19 Jan 2022, 13:57:39

1 . do "C:\Users\SARAHC~1\AppData\Local\Temp\STD1ac0_000000.tmp"

2 . /*Do-file for testing a single mediator model involving binary dependent variable
> in Stata
>
> version 1.0
> do file was written using Stata Version 17
>
> by Mike Crowson, Ph.D., The University of Oklahoma
>
> This work is licensed under the Creative Commons Attribution 4.0 International
> License. To view a copy of this license, visit
> http://creativecommons.org/licenses/by/4.0/.
>
> General information:
>
> This do-file was created to allow one to test the indirect effect of a continuous
> IV/X variable on a binary DV/Y through a single continuous mediator (Med). It was
> designed to produce output consistent with Model 4 (involving a single mediator)
> from Andrew Hayes (see https://www.processmacro.org/index.html) Process macro.
> Professor Hayes was not involved in the creation of this do-file.
>
> How to use the do-file:
>
> The next section allows you to specify your model and request the number of
> bootstrap replications you desire for testing the indirect effect. DO NOT make
> any modifications to any of the code in any of the remaining portions in this
> file, or you risk output containing errors or error messages.
>
> A dummy coded (binary) IV/X is permitted, but DO NOT use the i.prefix
> you may be used to when performing regression with factor variables.
> When you run the code the the indirect effect will be calculated, and 95%
> bootstrap confidence intervals will be produced in the output. [Included are
> normal theory; percentile; and bias-corrected confidence intervals.]
> The output will also contain the individual regression models comprising the
> full mediation model. This code permits you to include a binary DV coded as 0
> (e.g. 'failure' or 'no'') and 1 (e.g., 'success' or 'yes''). Your binary DV can
> also be coded as 1 and 2 (instead of 0 and 1); but no other codings are permissible.
>
> The regression output will refer to IV (for your independent variable, or X), Med
> (for your mediating variable, or M), and DV (for your dependent variable, or Y).
> The additional covariates you include in your model specification will be listed
> as named in your dataset. The initial output will be broken down into two
> regression 'sub-models'. Model 1 is an OLS regression where the mediator is
> regressed onto your IV and any covariates. Model 2 is a binary logistic
> regression, where the DV is regressed onto your IV, mediator, and any covariates.
> Cases with missing data on any of the variables included in your analysis will be
> discarded (i.e., listwise deleted). However, those cases will not be eliminated
> from your dataset.
>
> Standardized regression coefficients from the OLS regression (Model 1)
> and odds ratios (associated with the logitistic regression (Model 2) can be found
> in the SUPPLEMENTAL portion of the output. Descriptive statistics and
> correlations (based on the total effective sample size associated with
> your mediation analysis) are also provided in this section.
>
> Assumptions: The single mediator model here assumes correct model specification:
> (a) You have correctly specified the order of the temporal relations among the X, M,
> and Y; (b) Your specification of the model does not omit important variables that
> might alter the relations among the X, M, and Y variables; (c) The specification
```

```
   > reflects the true causal relations among the variables (MacKinnon, 2008).
   >
   > In addition to the assumptions above, aach regression model (i.e., Model 1 OLS
   > regression and Model 2 logistic regression) comprising the full mediation model
   > has have their own assumptions that must be met (MacKinnon, 2008). Failure to
   > meet the assumptions in those models can result in incorrect judgements concerning
   > population parameters associated with those models, and could impact the estimated
   > indirect effect and any inference you draw from it.
   >
   > Included in the options in the next section is the option to obtain various
   > regression diagnostics for models 1 and 2. The available diagnostics are not
   > exhaustive, but should provide you with some means for evaluating the individual
   > regression models. Included in the output that is generated are general
   > descriptions for how you might use the information generated in diagnosing
   > potential problems.
   >
   > */
 3 .
 4 . cls

 5 . capture drop IV DV Med

 6 . capture macro drop covs reps diagnostics

 7 . capture graph drop graph1 graph2

 8 . preserve

 9 . ******************************************************************************
10 . ******************************************************************************
11 . /*ENTER YOUR MODEL SPECIFICATIONS AND OPTIONS IN THIS SECTION. DO NOT CHANGE
   > OTHER SECTIONS OF THE DO-FILE*/
12 .
13 . /*Specify your IV, Med, and DV (binary) here. Use the exact variable names from
   > your data file.*/
14 .
15 . generate IV = u_any_nat_500

16 . generate Med = simpson_actino
   (314 missing values generated)

17 . generate DV = inhalant3y2mm_2
   (211 missing values generated)

18 .
19 . *Indicate any covariates you wish to include in your model here.
20 .
21 . global covs = ""

22 .
23 . /*Indicate the number of bootstrap replications you wish here. (The default
   > when you download this Do-file is set at 50. You will want to raise this when
   > carrying out your own analysis; e.g., 1000 or 2000 or >)
   >
   > NOTE: It may take some time for the bootstrap replications to complete, especially
   > if you request a large number of replications. You will need to be patient! */
```

```
24 .
25 . global reps = 5000

26 .
27 . /*Provide regression diagnostics? Type 0 if no; 1 if yes.*/
28 .
29 . global diagnostics=1

30 . ***************************************************************************
31 . ***************************************************************************
32 .
33 . keep if !missing(IV)
   (0 observations deleted)

34 . keep if !missing(DV)
   (211 observations deleted)

35 . keep if !missing(Med)
   (171 observations deleted)

36 .
37 . foreach x in $covs{
  2.                 drop if `x'==.
  3. }

38 .
39 .
40 . quietly sum DV

41 . if r(min)==1{
42 . recode DV (1=0) (else=1)
43 . }

44 .
45 .
46 . *Equation 1: Multiple regression with mediator regressed onto IV and covs.
47 .
48 . reg Med IV $covs
```

Source	SS	df	MS			
Model	.188421076	1	.188421076	Number of obs	=	287
Residual	7.94308163	285	.027870462	F(1, 285)	=	6.76
				Prob > F	=	0.0098
				R-squared	=	0.0232
				Adj R-squared	=	0.0197
Total	8.1315027	286	.028431828	Root MSE	=	.16694

Med	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
IV	-.0514408	.019784	-2.60	0.010	-.0903822	-.0124994
_cons	.4903036	.014586	33.61	0.000	.4615936	.5190136

```
49 .
50 . *Path a coefficient
```

```
51 . di e(b)[1,1]
   -.0514408

52 .
53 . /*Equation 2: Logistic regression results. Unstandardized regression
   > coefficients, confidence intervals (for b) and tests are provided. */
54 .
55 . logistic DV IV Med $covs, coef

   Logistic regression                                    Number of obs =       287
                                                          LR chi2(    2)    =   8.53
                                                          Prob > chi2   =     0.0141
   Log likelihood =  -35.755018                            Pseudo R2     =    0.1065

   ------------------------------------------------------------------------------
             DV | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
   -------------+----------------------------------------------------------------
             IV |  -1.208572   .8270381    -1.46   0.144    -2.829537    .4123925
            Med |   4.405875   2.195819     2.01   0.045     .1021485    8.709601
          _cons |  -5.299851   1.386639    -3.82   0.000    -8.017614   -2.582089
   ------------------------------------------------------------------------------

56 .
57 . *Path b coefficient
58 . di e(b)[1,2]
   4.4058749

59 .
60 . *Path c-prime coefficient
61 . di e(b)[1,1]
   -1.2085724

62 .
63 . capture program drop medDO

64 . program medDO, rclass
     1.
65 . reg Med IV $covs
     2. scalar pa = _b[IV]
     3.
66 . logistic DV IV Med $covs, coef
     4. scalar pb = _b[Med]
     5.
67 . return scalar IE = pa*pb
     6.
68 . end

69 .
70 . bootstrap Ind_Eff=r(IE), reps($reps): medDO
   (running medDO on estimation sample)

   Bootstrap replications ( 5,000)
   ----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
   ..xxx..x..x.x...x.......x.x..xxx.........x..x.....    50
   .......x.x...x.....x................x..........x..   100
   ....x...x.......x...xx....x.....................xx   150
   ...............x...xx....x..x.x.....x..xx.........   200
   ......................x.x..........x...xx.x....x..   250
   ..................x.xx........x..x.x..............   300
   ...x.......x...............x....................x.   350
   ......x........x.x..xx...x...x.........x..........   400
   ......xx.....x.............x..x.x.................   450
   .xx...............xx............x....xxx..........   500
   x...........................xx..x..........xx..x..   550
   .......x.......x.x......x.......x....x.x......xx..   600
   ..............xx...............x................x.   650
   ..x.x..xx.............x........................x..   700
```

```
x.....x...........x....x.......xx..x.....x........   750
........x.....................x....x..............   800
...x.x.......xxx..x.........x.....x.........x.....   850
..x...xx........x.x.............xx......x.....x...   900
.......x.x.............x..x.............x.........   950
...x....x..........x............................. 1,000
.....x.................................x.......xxx 1,050
...x....x...........................x..x......... 1,100
.x..x......x...xxx..x...x.xx.........x.x.........x 1,150
......x..x.x........x.................x.x.......x. 1,200
......x....x...x......x.............x.....x...x... 1,250
.....x..x...........................x...x........ 1,300
.xx...............x.....xx..........x.......xx.... 1,350
.......xxx.................x...............x...... 1,400
...................x...................x......... 1,450
....x...........x..............x.....x........... 1,500
.....x.............................x..x....xxx.... 1,550
..............................xx.x.............x.. 1,600
...................x..............x.x.......x.x... 1,650
......x.........x....x..........x....xx.......x... 1,700
.x.x.......x....x......xx.....x.x...........x..... 1,750
...........x..................x..........x....... 1,800
x.............xx...................x..x..x.....x.. 1,850
.........x.........x...x..x...................... 1,900
x.........x...x................................. 1,950
.....xx....x.....x.xx...x....x...xx.x.....x....... 2,000
.x.x.x..............x........x...........x......x. 2,050
.......xxx................x..x.............x...x.. 2,100
.......x.......................x............x..... 2,150
.................x............x.x................ 2,200
.................x.x.......x...................x.. 2,250
....x....xx...............x...............x....... 2,300
x.......x..x.....x.xx........x.xx......x.x....... 2,350
.x...x.x..x.................x.xx...x....x...x..... 2,400
x....xx.....xxx......xx.x........x............... 2,450
.........x....................x..x.x............. 2,500
...xxx.....x.x......x.....x....xx.x..x.x.x..x..... 2,550
......x...x....x....x.....x............xx........ 2,600
.....x....x............x...........x..x.....xx.... 2,650
x...x................x.......xx.x...........xx..x. 2,700
x...x.x....x..x............x..x......xx..x...x.... 2,750
....x....x...........x...........x............... 2,800
......x.x...................................x.x... 2,850
x.xx.......xx......x...x....xx..........xx.x...... 2,900
.....x.x.............x..........x................ 2,950
.........x......x...x.x.....................x..... 3,000
............x.xx.x.....x.......x....x........x.x.. 3,050
...x..........x......x.........x..x.............. 3,100
..........x........x......x.....x.......x......x.. 3,150
.x..x..........x......x............xx.....x..x.... 3,200
x.....x.............x.x..x............xx......... 3,250
........x..........x.......xxx................x..x 3,300
.xx...............xxx........x....x......x....x... 3,350
.....xx..........x..........x........x.xx........ 3,400
xx..x.................x......xx........x......... 3,450
.......x..x.............x......x....x....xx...... 3,500
....................x....x....x..............x.... 3,550
.......................xx..x...x..x........x...... 3,600
......x.x.x......x.............................. 3,650
....x.x.........xx..x....x......x.x.....x....x.... 3,700
............x......xxx.xx..........x............. 3,750
.x........................x.....x.x..xx.......... 3,800
...............x.x........xx.....x.....xx...x.x... 3,850
..x..xx...............x........x............x...x. 3,900
x....x.x.......x.................x.........x...... 3,950
.x.............x..x..........x...........x.......x 4,000
.........x..x.......xx............x......x........ 4,050
```

```
.x.....................x..................x....... 4,100
................x..x...............x.x...x....x... 4,150
.....x.x...x........x......xx......x.........x.... 4,200
......xx..........x...x.........xxxx.............. 4,250
.x.x....x...........x..x........xx...x............ 4,300
.x.......x....x...............................x... 4,350
.....x.............x...x...x...........x.......... 4,400
.........................x..x..................... 4,450
..x.....x..x.........x.x...........x....x...xx.... 4,500
............x....xx......................x.....x.. 4,550
....x............x.......xxx.........x.........x.. 4,600
x...........x...x............x.................... 4,650
........................x......x..x.x...x..x..x... 4,700
....x.x...........xx...x....xx.......xx........... 4,750
..x.......x....xx..x....................x..x..x..x 4,800
..............x....x.............................. 4,850
............x..x.....x...x.x...................... 4,900
.......................................x..x....... 4,950
......x.......x.......x...xx.x...xx.x.........x... 5,000
```

Bootstrap results

Number of obs = **287**
Replications = **4,345**

Command: **medDO**
Ind_Eff: **r(IE)**

	Observed coefficient	Bootstrap std. err.	z	P>|z|	Normal-based [95% conf. interval]	
Ind_Eff	**-.2266417**	**.1603321**	**-1.41**	**0.157**	**-.5408869**	**.0876035**

Note: One or more parameters could not be estimated in 655 bootstrap replicates; standard-error estimates include only complete replications.

71 . estat bootstrap, percentile bc

Bootstrap results

Number of obs = **287**
Replications = **4345**

Command: **medDO**
Ind_Eff: **r(IE)**

	Observed coefficient	Bias	Bootstrap std. err.	[95% conf. interval]		
Ind_Eff	**-.22664174**	**-.0103938**	**.16033213**	**-.6229164**	**-.0036841**	(P)
				-.6881957	**-.0195255**	(BC)

Key: P: Percentile
BC: Bias-corrected
Note: One or more parameters could not be estimated in 655 bootstrap replicates; standard-error estimates include only complete replications.

```
72 .
73 . ***************************************************************************
74 . /*SUPPLEMENTAL:STANDARDIZED REGRESSION WEIGHTS PRODUCED FOR MULTIPLE REGRESSION
   > OUTPUT & ODDS RATIOS PRODUCED FOR LOGISTIC REGRESSION OUTPUT*/
75 .
76 . reg Med IV $covs, beta

      Source |       SS           df       MS      Number of obs   =       287
-------------+----------------------------------   F(1, 285)       =      6.76
       Model |  .188421076         1  .188421076   Prob > F        =    0.0098
    Residual |  7.94308163       285  .027870462   R-squared       =    0.0232
-------------+----------------------------------   Adj R-squared   =    0.0197
       Total |   8.1315027       286  .028431828   Root MSE        =    .16694

------------------------------------------------------------------------------
         Med | Coefficient  Std. err.      t    P>|t|                     Beta
-------------+----------------------------------------------------------------
          IV |  -.0514408    .019784    -2.60   0.010                -.1522227
       _cons |   .4903036    .014586    33.61   0.000                        .
------------------------------------------------------------------------------

77 .
78 . * Logistic regression results: Odds ratios (OR's) and confidence intervals
79 . logistic DV IV Med $covs

Logistic regression                                     Number of obs =    287
                                                        LR chi2(    2)  =   8.53
                                                        Prob > chi2   =   0.0141
Log likelihood =  -35.755018                              Pseudo R2     =   0.1065

------------------------------------------------------------------------------
          DV | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
          IV |   .2986233   .2469728    -1.46   0.144      .0590402    1.510427
         Med |   81.93079   179.9052     2.01   0.045      1.107548    6060.825
       _cons |   .0049923   .0069226    -3.82   0.000      .0003296    .0756159
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

80 .
81 . ***************************************************************************
82 . /*SUPPLEMENTAL: DESCRIPTIVE STATISTICS AND CORRELATIONS AMONG VARIABLES BASED
   > ON TOTAL EFFECTIVE SAMPLE SIZE ASSOCIATED WITH THE MEDIATION ANALYSES*/
83 .
84 . summarize IV DV Med $covs

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
          IV |        287     .543554    .4989695          0          1
          DV |        287    .0313589    .1745901          0          1
         Med |        287    .4623427    .1686174          0       .835

85 .
86 . pwcorr IV DV Med $covs, obs sig

             |       IV       DV      Med
-------------+---------------------------
          IV |   1.0000 
             |
             |      287
             |
          DV |  -0.1161   1.0000 
             |   0.0495
             |      287      287
             |
         Med |  -0.1522   0.1443   1.0000 
             |   0.0098   0.0144
```

```
                │       287       287       287
                │

 87 . ***************************************************************************
 88 . if $diagnostics==1{
 89 .
 90 . *REGRESSION DIAGNOSTICS FOR OLS REGRESSION/MODEL 1
 91 .
 92 . quietly reg Med IV $covs
 93 .
 94 . predict r, residuals
 95 . predict fit, xb
 96 . predict studentized, rstudent
 97 . predict cook, cooksd
 98 . /*Breusch-Pagan test for linear heteroskedasticity/ Koenker (1981) version.
   > Statistical significance is considered an indicator of a violation of the
   > assumption of constant variance (i.e., homoskedasticity).*/
 99 .
100 . estat hettest, iid

    Breuschâ€“Pagan/Cookâ€“Weisberg test for heteroskedasticity
    Assumption: i.i.d. error terms
    Variable: Fitted values of  Med

    H0: Constant variance

        chi2(1) =   4.11
    Prob > chi2 =  0.0425
101 .
102 . /*White's test for heteroskedasticity of more general functional form.
   > Statistical significance is considered an indicator of a violation of the
   > assumption of constant variance (i.e., homoskedasticity).*/
103 .
104 . imtest, white

    White's test
    H0: Homoskedasticity
    Ha: Unrestricted heteroskedasticity

        chi2(1) =   4.11
    Prob > chi2 =  0.0425

    Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	**4.11**	**1**	**0.0425**
Skewness	**0.98**	**1**	**0.3231**
Kurtosis	**3.75**	**1**	**0.0529**
Total	**8.84**	**3**	**0.0315**

```
105 .
106 . /*Collinearity diagnostics / Tolerance and Variance inflation factors. VIF's
   > greater than 10 typically signal more severe problems with multicollinearity.
   > Cohen et al. (2003) indicate some authors view values around 6 or 7 as reflecting
   > more substantial problems.*/
```

```
107 .
108 . estat vif

        Variable |       VIF       1/VIF
    -------------+----------------------
              IV |      1.00    1.000000
    -------------+----------------------
        Mean VIF |      1.00
109 .
110 . /*Skewness and kurtosis of residuals. A skewness and/or kurtosis value greater
    > than 2 in absolute value is often regarded as an indication of more substantial
    > departure from normality. See Lomax & Hahs-Vaughn (2012).*/
111 .
112 . tabstat r, statistics( skewness kurtosis )

        Variable |  Skewness  Kurtosis
    -------------+--------------------
               r |  .0598029  2.580515
    ----------------------------------
113 .
114 . /*Normality tests for residuals. This output includes tests of normality based on
    > skew and then one on kurtosis and then a joint test based on both skew and kurtosis.
    > Null hypothesis for these tests in the current context is that the residuals are
    > normally distributed.*/
115 .
116 . sktest r

    Skewness and kurtosis tests for normality
                                                           ------ Joint test ------
        Variable |       Obs   Pr(skewness)   Pr(kurtosis)   Adj chi2(2)  Prob>chi2
    -------------+-----------------------------------------------------------------
               r |       287         0.6720         0.0948          2.99     0.2244
117 .
118 . /*List of cases with studentized residuals > 3 in absolute value. These are
    > potential residual outliers.*/
119 .
120 . list studentized if abs(studentized) >=3
121 .
122 . /*List of cases with Cook's d values >= 4/n
    > (see https://www.stata.com/manuals/rregresspostestimation.pdf).
    > Useful for identifying cases that are of high influence on the regression model.*/
123 .
124 . list cook if cook >= 4/e(N)

          +----------+
          |     cook |
          |----------|
     32.  |  .022436 |
     43.  | .0150248 |
     68.  | .0206549 |
     73.  | .0152909 |
     98.  | .0169219 |
          |----------|
    112.  | .0334303 |
    195.  | .0334303 |
    214.  | .0165228 |
    275.  | .0180425 |
    281.  | .0153337 |
          +----------+
```

```
125 .
126 . /*Listing cases with Cook's d values > 1 (a fairly common threshold for
    > identifying high influence cases; see e.g., Lomax & Hahs-Vaughn, 2012)*/
127 .
128 . list cook if cook >= 1
129 .
130 . ***************************************************************************
131 . *REGRESSION DIAGNOSTICS FOR LOGISTIC REGRESSION/MODEL 2
132 .
133 . quietly reg DV IV Med $covs
134 . *Collinearity diagnostics / Tolerance and Variance inflation factors
135 . estat vif
```

Variable	VIF	1/VIF
IV	**1.02**	**0.976828**
Med	**1.02**	**0.976828**
Mean VIF	**1.02**	

```
136 .
137 . quietly logistic DV IV Med $covs, coef
138 . predict p
    (option pr assumed; Pr(DV))
139 . predict pearsonres, residuals
140 . generate logit=ln(p/(1-p))
141 . predict db, dbeta
142 .
143 . /*Below is a List of cases with Pearson residuals > 2.5. According to Pituch &
    > Stevens (2016), a case with a Pearson residual > |2.5| or |3.0| indicates it
    > is not adequately fit by the model (making it a potential outlier).*/
144 .
145 . list pearsonres if abs(pearsonres) >=2.5
```

	pearso~s
31.	**4.704206**
36.	**3.011036**
54.	**8.994146**
71.	**2.500748**
117.	**6.03278**
202.	**2.834261**
205.	**4.128695**
223.	**8.994146**
232.	**5.992768**
249.	**2.861295**

```
146 .
147 . /*Identifying potential cases having an undue influence on the regression
    > parameters in the logistic regression model.
    >
    > Here, we will use Pregibon (1981) dbeta values to identify cases that may be
    > exerting undue influence on the logistic regression model parameters. In the
    > Stata rlogisticpostestimation manual
    > [https://www.stata.com/manuals/rlogisticpostestimation.pdf], these values are
    > described as "the difference in the coefficient vector that is due to deletion
    > of the observation along with all others that share the same covariate pattern"
    > (p. 2).
    >
    > In the output below, you will find Pregibon (1981) values for cases exceeding the
    > 95th percentile on the index which can aid in the identification of those cases
    > that are exerting greater influence on the logistic model. My recommendation is
    > to use these in conjunction with the scatterplot and index plot that have been
    > generated to determine which, if any, cases are exerting undue influence on the
    > model parameters.*/
```

```
.
. quietly summarize db, detail
. list db if db >= r(p95)

     +-----------+
     |        db |
     |-----------|
 31. |   .194689 |
 36. |  .1689801 |
 54. |  .7465738 |
 71. |  .2519116 |
117. |  .5633729 |
     |-----------|
202. |  .1906338 |
205. |  .1504559 |
214. |  .0125348 |
220. |  .0230089 |
223. |  .7465738 |
     |-----------|
232. |  .3660031 |
235. |  .0230089 |
242. |  .0230089 |
249. |  .1868844 |
272. |  .0230089 |
     +-----------+

.
. twoway (scatter db p), name(graph1)
. gen case_id = _n
. twoway (scatter db case_id), name(graph2)
. graph combine graph1 graph2, holes(2)
. }

. macro drop covs reps diagnostics

. capture graph drop graph1 graph2

. ********************************************************************************
.
. /*
> References and suggested readings
>
> Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). Applied multiple
> regression/correlation analysis for the behavioral sciences (3rd Ed).
> Mahwah, NJ: Lawrence Erlbaum Associates.
>
> Hayes, A.F. (2018). Introduction to mediation, moderation, and conditional
> process analysis: A regression-based approach (2nd edition). New York: The
> Guilford Press.
>
> Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed).
> Los Angeles: Sage.
>
> Koenker, R. (1981). A note on studentizing a test for heteroskedasticity.
> Journal of Economics, 17, 107-112.
>
> Lomax, R.G., & Hahs-Vaughn, D.L. (2012). An introduction to statistical
> concepts (3rd ed). New York: Routledge.
>
> MacKinnon, D.P. (2008). Introduction to statistical mediation analysis. New York:
> Lawrence Erlbaum Associates.
>
> Osborne, J.W. (2015). Best practices in logistic regression. Los Angeles: Sage.
>
> Pituch, K.A., & Stevens, J.P. (2016). Applied multivariate statistics for the
> social sciences (6th ed). Thousand Oaks, CA: Sage.
>
> Pregibon, D. (1981). Logistic regression diagnostics. The Annals of Statistics,
```

```
> 9, 705-724.
>
> Tabachnick, B.G., & Fidell, L.S. (2013). Using multivariate statistics
> (6th ed.). New York: Pearson.
>
```


```
end of do-file

162 . save "\\med-fs1.med.ualberta.ca\Pediatrics\Symbiota\SyMBIOTA\Datasets\Charlene-Vienna nature project\Br
    > rged_SBJan182022.dta", replace
    file \\med-fs1.med.ualberta.ca\Pediatrics\Symbiota\SyMBIOTA\Datasets\Charlene-Vienna nature project
        analysis\210726_CLEANED_child_uplvi_npri_cimd_Nielsen_allergens_metadata_merged_SBJan182022.dta

163 . exit
```