NC APS Meeting 2012 - Meta-Analysis Workshop

Using R - the *metafor* package

L. V. Madden and P. A. Paul – Ohio State University

This document provides an overview of how to perform ***univariate fixed***- and ***random-effects*** meta-analysis in ***R***, with and without moderator variables, for evidence synthesis in Plant Pathology. It serves to complement the ***SAS*** exercises conducted during the workshop, providing ***R*** users with options for performing data explorations and analyses similar to those demonstrated in ***SAS***.

Using a data set (called SweetCornRust.txt or SweetCornRust.xlsx in your "***With R"*** folder) provided by Shah and Dillard (Plant Disease 90: 1413-1418 [2006]), the ***metafor*** package, developed by Viechtbauer (Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, **36**(3), 1–48 [2010]. http://www.jstatsoft.org/v36/i03/), will be used to fit meta-analytical models. We will demonstrate:

1. How to download, install, and load the ***metafor*** package.
2. Basic data importation and manipulation in ***R.***
3. How to graphically explore raw data with forest plot, using the ***forest*** function from the ***metafor*** package.
4. How to fit meta-analytical models using the ***rma*** function.
    - ***a.*** How to fit a fixed-effect model.
    - ***b.*** How to fit random-effect models using the restricted maximum likelihood (REML) and maximum likelihood (ML) parameter estimation methods.
    - ***c.*** How to graphically explore model results with forest, funnel and radial plots, using the ***forest***, ***funnelplot***, and ***radial*** functions from the ***metafor***, ***rmeta***, and ***meta*** package, respectively.
    - ***d.*** How to fit mixed-effect models, i.e, random-effect models with continuous and categorical moderator variables.

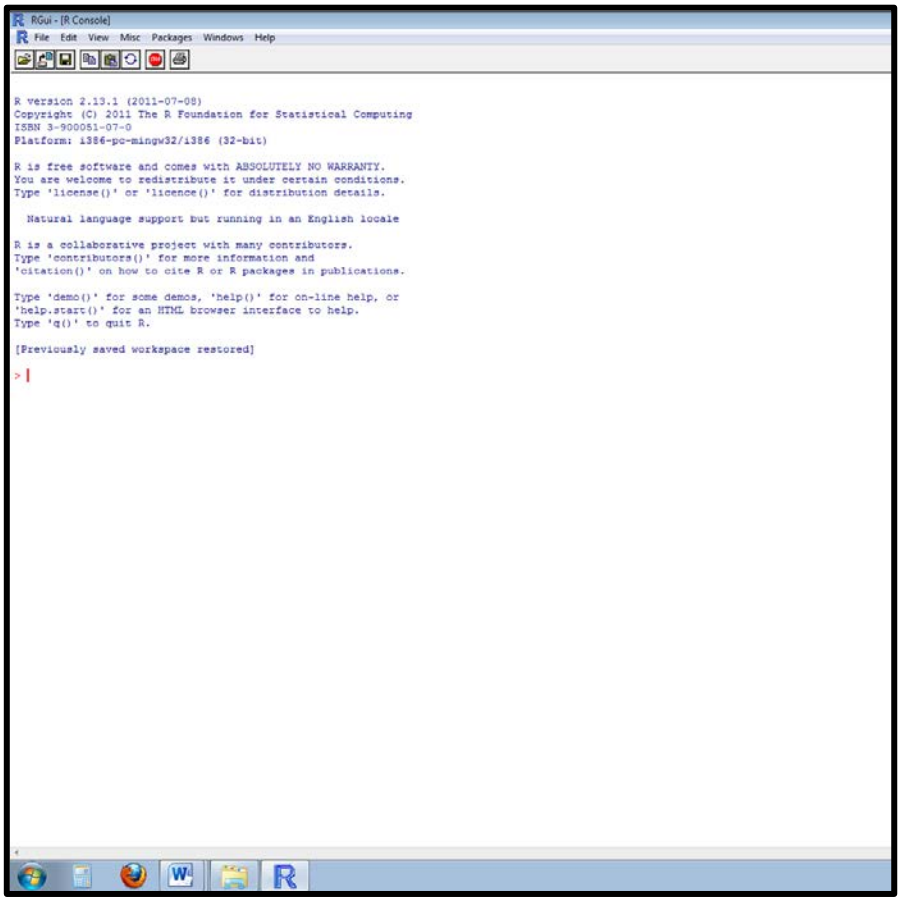
Included in the ***With R*** folder of your workshop package are:

1. MS Excel and text versions of the raw data - *SweetCornRust.xlsx* and *SweetCornRust.txt*
2. A zip file of the ***metafor*** package - *metafor_1.6-0.zip*
3. Metafor documents - *Viechtbauer2010.pdf* and *metafor.pdf*
4. Help files for the ***rma***, ***forest***, ***funnelplot,*** **and** ***radial*** functions - help(metafor).docx, help(forest).docx, help(funnelplot).docx, and help(radial).docx. (*Note: the names of these files, without ".docx" are the commands used in **R** to call the help files from the different libraries.*)

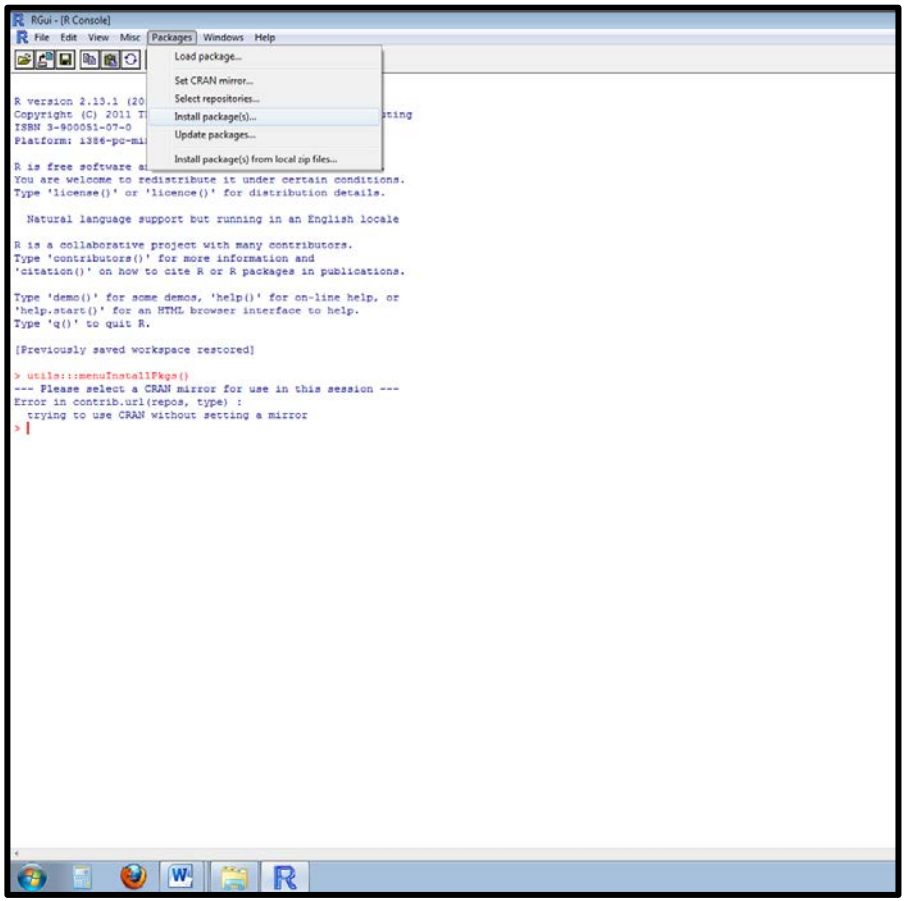
Note:

- *Details of the calculations used here are explained in Madden & Paul (Phytopathology 102:16-30 [2011]), and references given in that publication.*

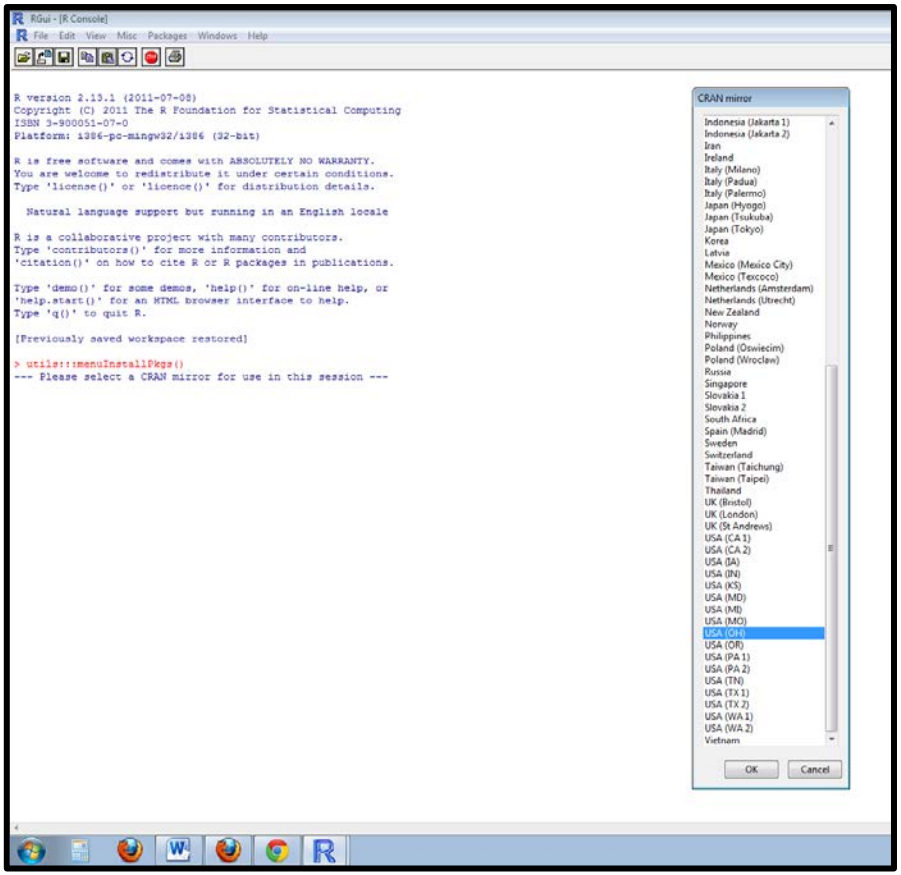
1. Download, install, and load the *metafor* package



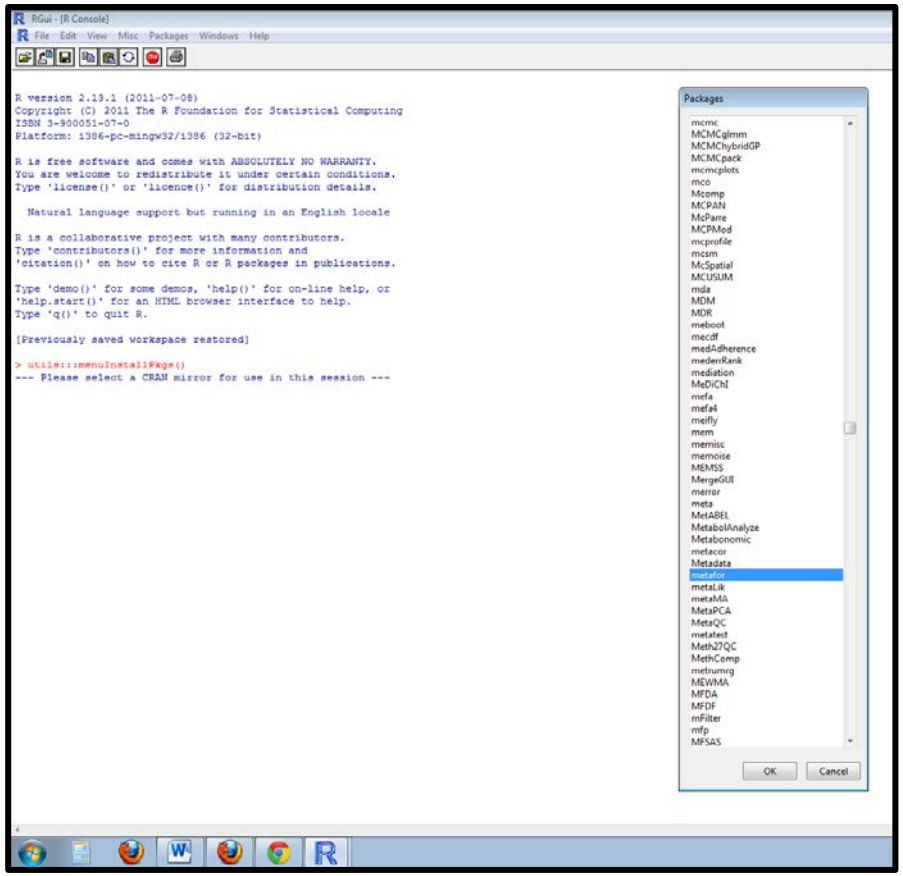
1. **The Main R Window (Console)**



2. **Select *Install package(s)…* from the *Packages* dropdown menu**



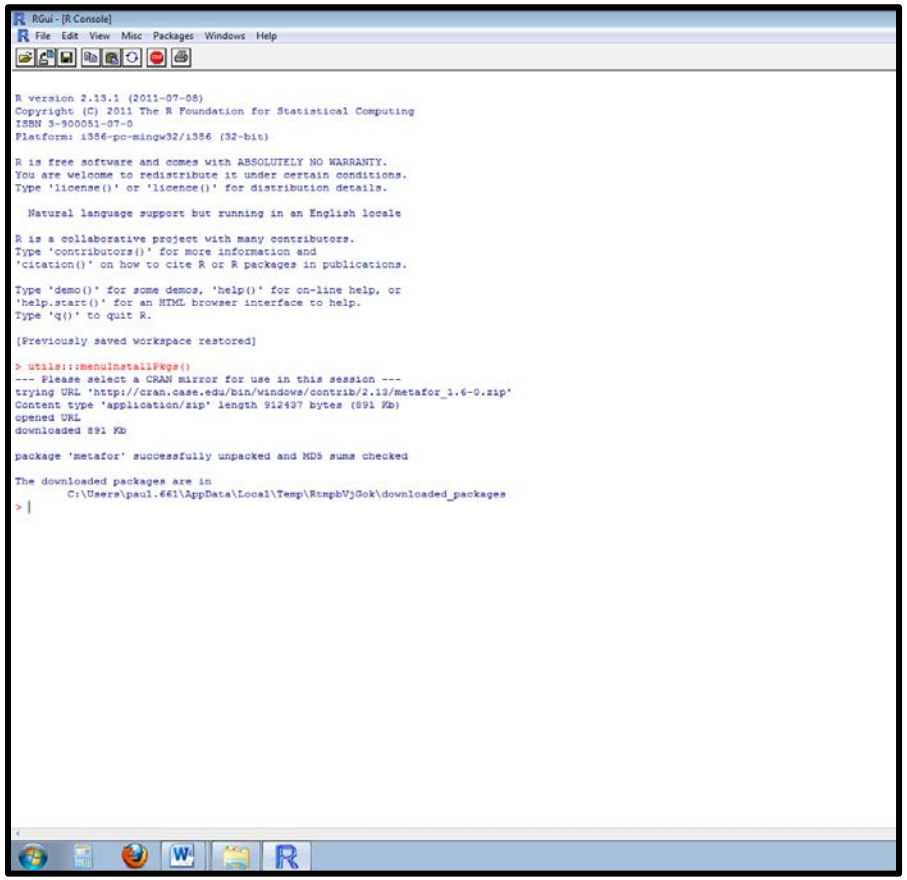
3. Select a download location from *CRAN mirror* menu and hit OK



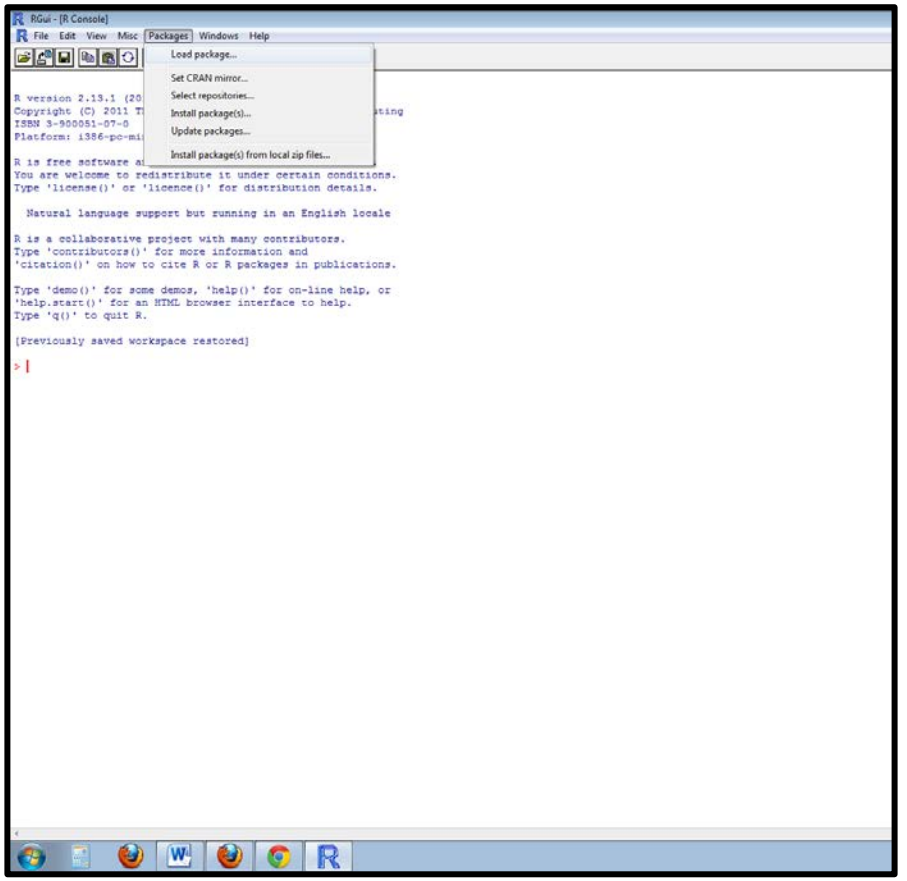
4. Select *metafor* from the *Packages* menu and hit OK

Note:

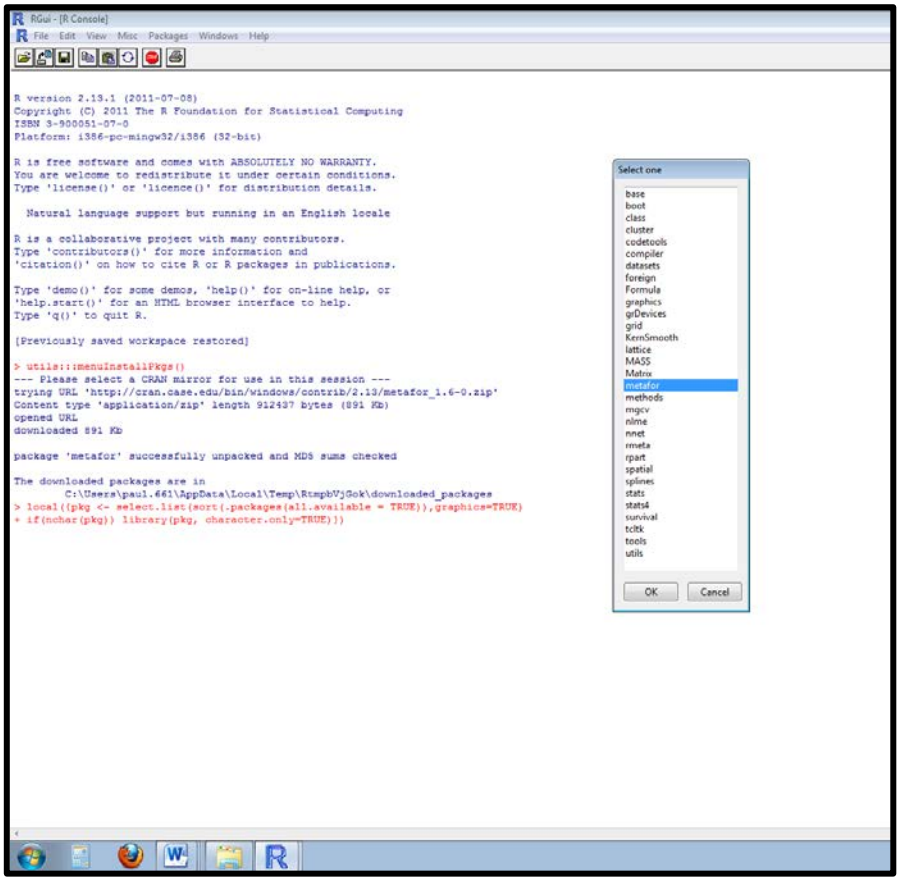
- *If you do not have access to the internet, you can select "Install package(s) from local zip files…" from the **Packages** dropdown menu to install the **metafor package** from your "**With R**" folder.*



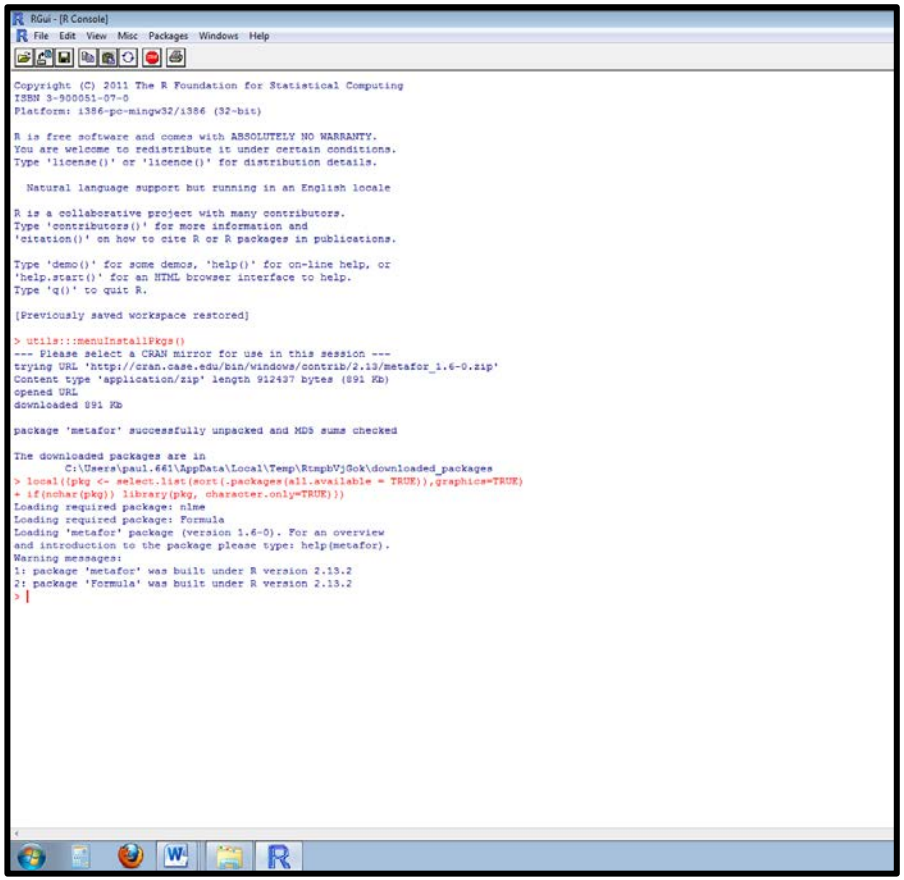
5. Path showing where the file (metafor_1.6-0.zip) is stored



6. Select *Load package(s)…* from the *Packages* dropdown menu



7. Select *metafor* from the *Select one* menu and hit OK



8. The package is now loaded and ready for use

Note:

- *You can select "Install package(s) from local zip files…" from the **Packages** dropdown menu to install the **metafor package** from a folder on your computer.*
- *Once the **metafor package** is installed, you will have to reload it (repeat steps 6 and 7) every time you restart **R***

2. Data importation and Manipulation

```
> RustData<-read.table("C:/Users/paul.661/Desktop/With R/SweetCornRust.txt", header=TRUE)
```

- Give your file a name (an R Object)
- *R* function (one of several) for importing data
- Data file path, specific for each computer
- Tells *R* that the first row of your data is the names of your variables (headers)

- **Examine the data**

```
> RustData
   Study State   Variety Year    slope      SE lowerlimit upperlimit  MinD  MaxD MeanD
1      1    NY    Zenith 1999  2.13607 2.18587   -3.03268    7.30483  0.00  1.46  0.39
2      2    NY    Zenith 1998  0.47288 0.18572    0.03372    0.91204  0.01 19.04  3.05
3      8    NY    Squeen 2000  1.17716 0.14263    0.83989    1.51443  1.50 24.30  6.58
4     11    NY      Bold 2001  0.33829 0.05394    0.21626    0.46031 20.49 77.31 45.91
5    22a    NY    Zenith 1997  0.78559 0.24789   -0.00331    1.57450  1.40 19.40  8.10
6    22b    NY     Rival 1997 -0.50102 0.42602   -1.85679    0.85476  1.50  6.30  3.73
7     31    MI  HMX83865 1993  0.31888 0.01484    0.27163    0.36612  5.40 57.50 18.73
8     35    MI    YBelle 1992  0.14559 0.12224   -0.15352    0.44470  2.60 85.80 22.96
9     50    NY   Jubilee 1984 -0.79923 0.31152   -1.79063    0.19217  8.08 15.50 11.73
10    59    IL   FSSweet 1984  0.63267 0.02300    0.58577    0.67958  5.59 62.83 38.07
11    60    IL   FSSweet 1985  1.08843 0.06339    0.95246    1.22439 10.45 55.20 26.51
12    61    IL   FSSweet 1986  0.66299 0.25871    0.09357    1.23241  1.55 17.65  9.54
13    62    IL      Gcup 1984  0.85302 0.04194    0.76626    0.93979 15.45 59.91 39.79
14    63    IL      Gcup 1985  0.55555 0.03612    0.48142    0.62967  7.60 50.08 27.63
15    64    IL      Gcup 1986  0.15786 0.28335   -0.44282    0.75854  1.67 20.57  6.22
16    65    IL  Stylepak 1984  0.62416 0.01996    0.58344    0.66488 23.63 67.68 38.94
17    66    IL  Stylepak 1985  0.37280 0.02808    0.31402    0.43157  7.50 49.84 25.70
18    67    IL  Stylepak 1986  0.78146 0.14652    0.47233    1.09058  1.42 17.40  6.72
19    70    IL SnowWhite 2001  0.59867 0.01999    0.55697    0.64036 23.00 65.00 36.95
20    71    IL  Sterling 2001  0.40333 0.01810    0.36559    0.44108 18.00 59.00 35.81
```

- **Get summary statistics**

```
> summary(RustData)
     Study     State       Variety        Year          slope                SE
 1       : 1   IL:11   FSSweet :3   Min.   :1984   Min.   :-0.7992   Min.   :0.01484
 11      : 1   MI: 2   Gcup    :3   1st Qu.:1985   1st Qu.: 0.3334   1st Qu.:0.02681
 2       : 1   NY: 7   Stylepak:3   Median :1989   Median : 0.5771   Median :0.09281
 22a     : 1           Zenith  :3   Mean   :1991   Mean   : 0.5403   Mean   :0.23149
 22b     : 1           Bold    :1   3rd Qu.:1998   3rd Qu.: 0.7825   3rd Qu.:0.25060
 31      : 1           HMX83865:1   Max.   :2001   Max.   : 2.1361   Max.   :2.18587
 (Other):14           (Other) :6


   lowerlimit         upperlimit          MinD             MaxD          MeanD
 Min.   :-3.03268   Min.   :0.1922   Min.   : 0.000   Min.   : 1.46   Min.   : 0.390
 1st Qu.:-0.04086   1st Qu.:0.4564   1st Qu.: 1.500   1st Qu.:18.69   1st Qu.: 6.685
 Median : 0.29283   Median :0.7191   Median : 5.495   Median :49.96   Median :20.845
 Mean   :-0.03732   Mean   :1.1178   Mean   : 7.842   Mean   :41.59   Mean   :20.653
 3rd Qu.: 0.56359   3rd Qu.:1.1240   3rd Qu.:11.700   3rd Qu.:60.64   3rd Qu.:36.095
 Max.   : 0.95246   Max.   :7.3048   Max.   :23.630   Max.   :85.80   Max.   :45.910
```

3. *Graphical exploration of raw data*

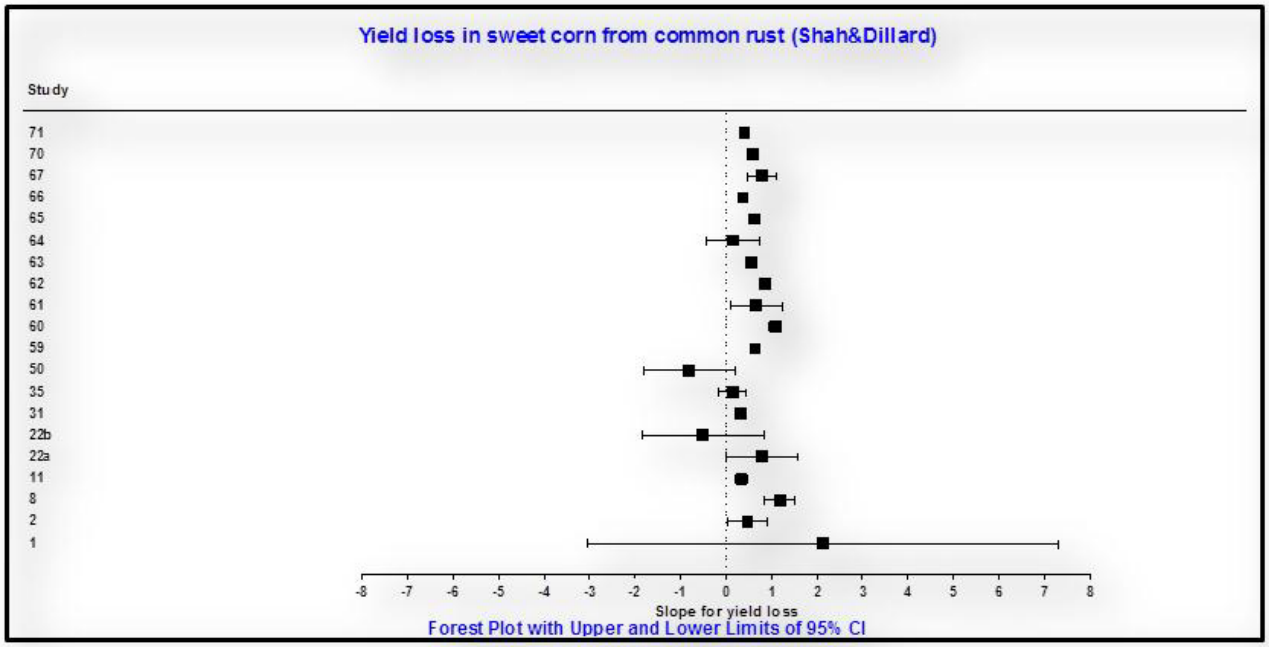
- **Forest Plots**

```
> forest(rev(RustData$slope), ci.lb=rev(RustData$lowerlimit),
  ci.ub=rev(RustData$upperlimit), annotate=FALSE, xlab="Slope for yield loss", font=2,
  slab=rev(RustData$Study), alim=c(-8, 8), cex.lab=1.5, pch=15, step=17, psize=2,
  cex=1.25, cex.axis=1.25)

> text(-15.5, 22, "Study", pos = 4, cex=1.5, font = 2)

> title("Yield loss in sweet corn from common rust (Shah&Dillard)", sub = "Forest
  Plot with Upper and Lower Limits of 95% CI", cex.main = 2, font.main= 2, col.main=
  "blue",cex.sub = 1.75, font.sub = 2, col.sub = "blue")
```

Most of the codes used with these functions are for formatting different components of the graph

Notes:

- ci.lb = specify what to use as the lower bound of the error bar - in this case the lower limit of the 95% confidence interval found in the *lowerlimit* column of the *RustData* dataset.
- ci.ub = specify what to use as the upper bound of the error bar - in this case the upper limit of the 95% confidence interval found in the *upperlimit* column of the *RustData* dataset
- rev = tells the program to plot the points in reverse order, i.e. with study 71 at the top and study 1 at the bottom of the graph.



4. *Fitting meta-analytical models using the rma function*

a. Fixed-effect model

```
> CornRust_Fixed<- rma(slope, SE^2, method="FE", data=RustData)
> CornRust_Fixed

Fixed-Effects Model (k = 20)

Test for Heterogeneity:
Q(df = 19) = 523.0911, p-val < .0001

Model Results:

estimate       se     zval     pval    ci.lb    ci.ub
  0.4928   0.0075  66.0119   <.0001   0.4782   0.5074      ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Specify model and parameter estimation method

Tell the program what to use as the effect size (*z*) and the sampling variance (S^2) :- for the slope, $S^2 = SE^2$

Note:

- Highlighted are the estimated mean effect size (estimate), the standard error of the effect size (se), the test statistic (*Z,* the standard normal in this case), the significance level, and the lower (ci.ub) and upper (ci.ub) limits of the 95% confidence interval around the effect size.

b.i. Random-effect model, with Restricted Maximum Likelihood (REML) as the Parameter Estimation Method

```
> CornRust_Random_reml<- rma(slope, SE^2, method="REML", data=RustData)
> CornRust_Random_reml

Random-Effects Model (k = 20; tau^2 estimator: REML)
tau^2 (estimate of total amount of heterogeneity): 0.1181 (SE = 0.0456)
tau (sqrt of the estimate of total heterogeneity): 0.3436
I^2 (% of total variability due to heterogeneity): 98.95%
H^2 (total variability / sampling variability):    95.63

Test for Heterogeneity:
Q(df = 19) = 523.0911, p-val < .0001

Model Results:
estimate       se      zval      pval     ci.lb     ci.ub
  0.5164   0.0856    6.0327    <.0001    0.3486    0.6842      ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Restricted maximum likelihood as the estimation method

Between-study variance (σ^2)

b.2. Random-effect model, with Maximum Likelihood (ML) as the Parameter Estimation Method

```
> CornRust_Random_ml<- rma(slope, SE*SE, method="ML", data=RustData)
> CornRust_Random_ml

Random-Effects Model (k = 20; tau^2 estimator: ML)
tau^2 (estimate of total amount of heterogeneity): 0.1075 (SE = 0.0406)
tau (sqrt of the estimate of total heterogeneity): 0.3279
I^2 (% of total variability due to heterogeneity): 98.85%
H^2 (total variability / sampling variability):    87.13

Test for Heterogeneity:
Q(df = 19) = 523.0911, p-val < .0001

Model Results:
estimate       se      zval      pval     ci.lb     ci.ub
  0.5190   0.0821    6.3193    <.0001    0.3580    0.6799      ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

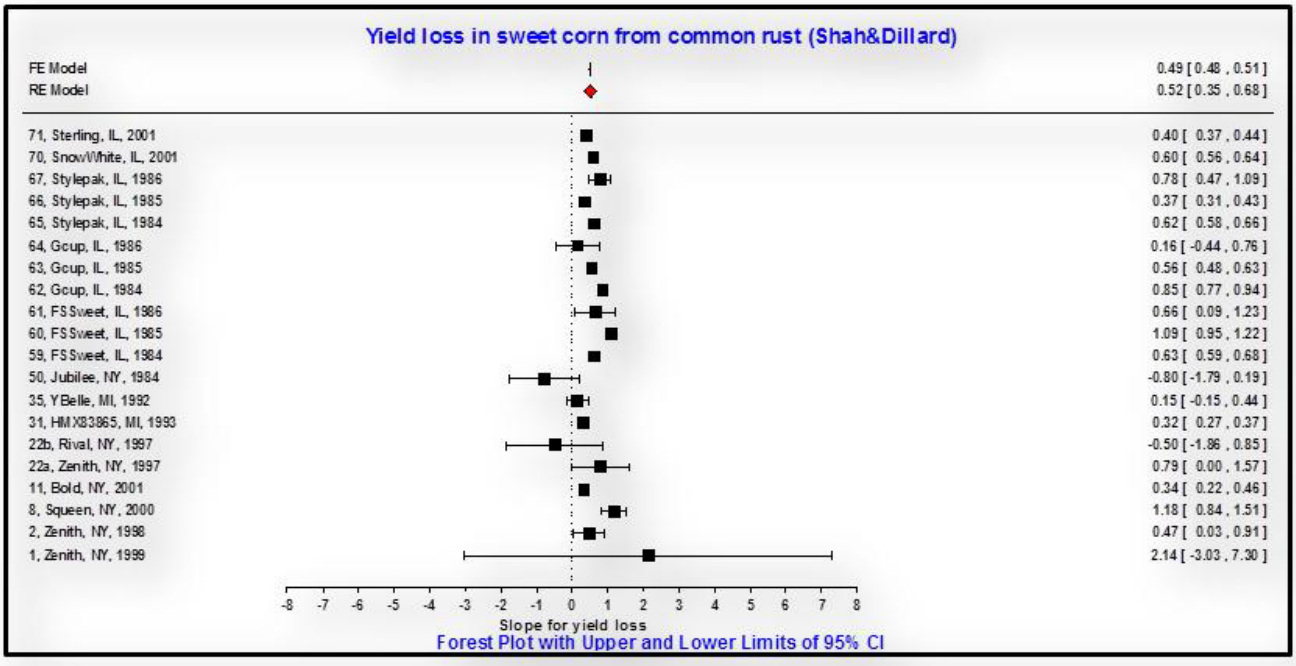
Maximum likelihood as the estimation method

c. Graphical exploration of the results

- **Forest plot**

```
> forest(rev(RustData$slope), ci.lb=rev(RustData$lowerlimit),
  ci.ub=rev(RustData$upperlimit), xlab="Slope for yield loss",
  slab=paste((rev(RustData$Study)), (rev(RustData$Variety)), (rev(RustData$State)),
  (rev(RustData$Year)), sep = ", "), alim=c(-8, 8), cex = 1.25, cex.axis = 1.25,
  cex.lab=1.5, pch=15, step=17, psize=2, font = 2)

> addpoly.rma(CornRust_Fixed, row=23, cex=1.25, font = 2, col="blue")

> addpoly.rma(CornRust_Random_reml, row=22, cex=1.25, font = 2, col="blue")

> title("Yield loss in sweet corn from common rust (Shah&Dillard)", sub = "Forest
  Plot with Upper and Lower Limits of 95% CI", cex.main = 2,   font.main= 2,
  col.main= "blue",cex.sub = 1.75, font.sub = 2, col.sub = "blue")
```

Adds a polygon representing the estimated effect size to the forest plot



- **Funnel plot - using the *funnelplot* function from the *rmeta* library**

Load the *rmeta* library

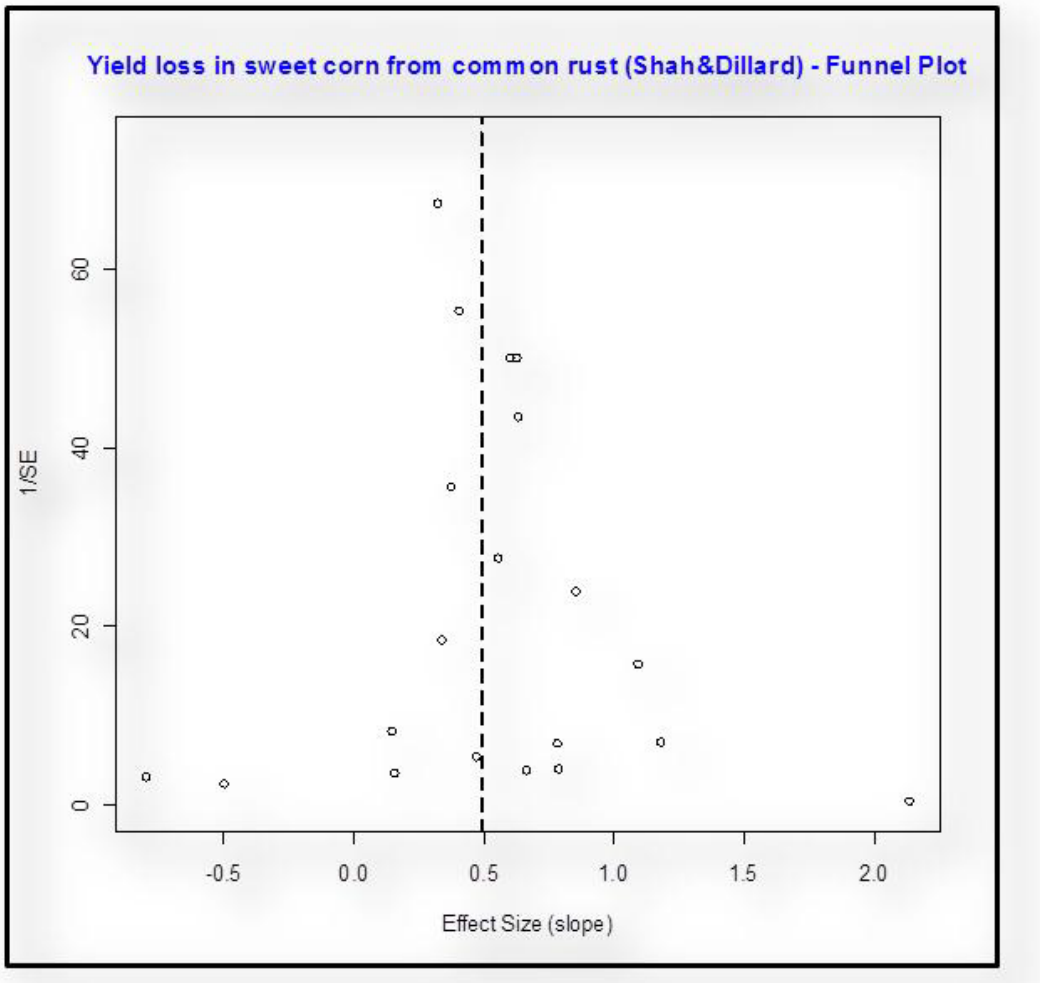
```
> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
```

Fit fixed-effect model using the meta.summaries function

```
> RMETA_FIXED<-meta.summaries (slope, SE, method=c("fixed"), names=Study,
conf.level=0.95, data=RustData)
```

Get Funnel Plot

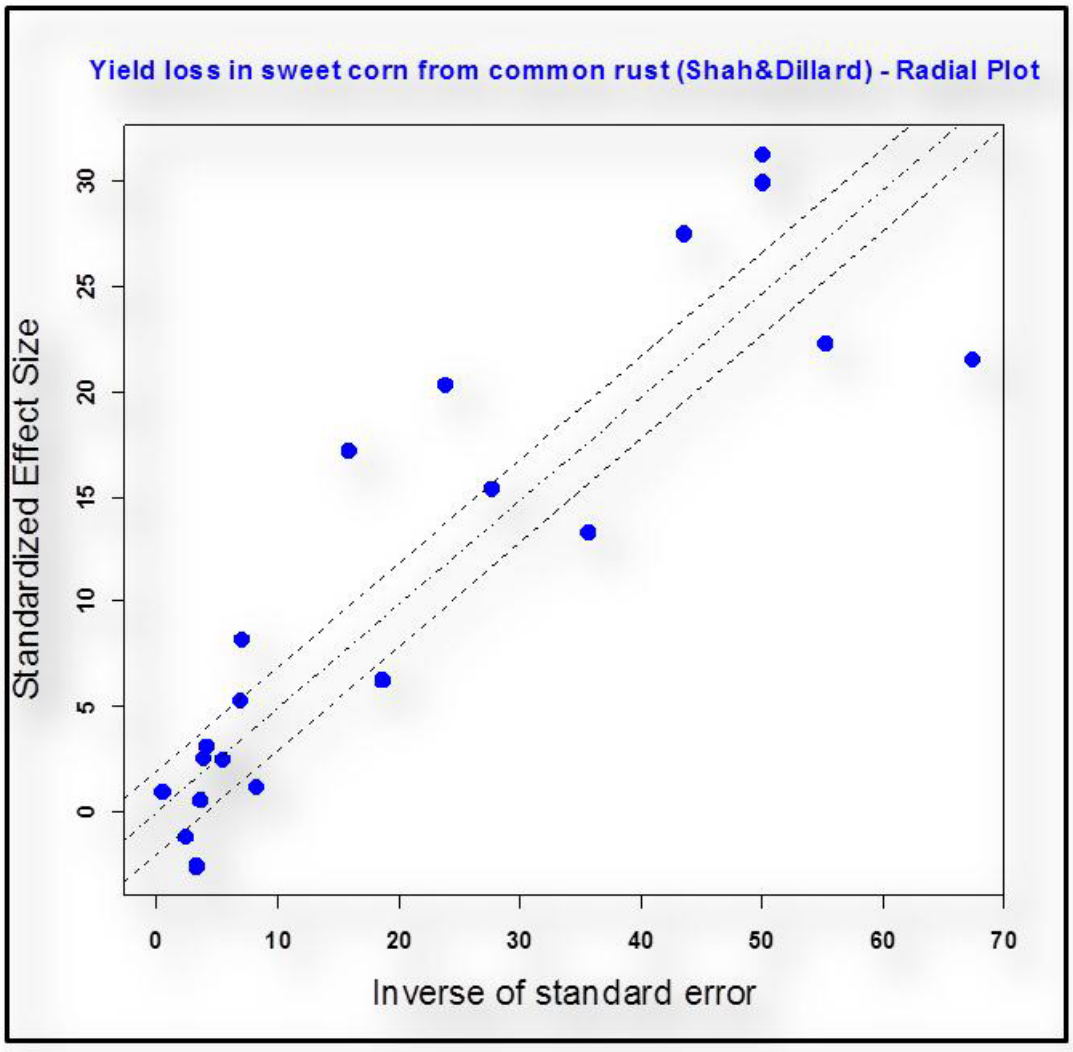
```
> funnelplot(RMETA_FIXED, xlab="Effect Size (slope)", ylab="1/SE")
Or
> funnelplot(RustData$slope, RustData$SE, size=1/RustData$SE, data=RustData,
xlab="Effect Size (Slope)", ylab="1/SE")
> title("Yield loss in sweet corn from common rust (Shah&Dillard) - Funnel Plot",
  cex.main = 1.2, font.main= 2, col.main= "blue")
```



***Note**: See **Section 1** for information on how to download, install and load the **rmeta** package*

- **Radial plot - using the *radial* function from the *meta* library**

```
Load the meta library
> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})

Fit meta-analytical models using the metagen function
> Meta_Both<-metagen(slope, SE, Study, data=RustData, level=0.95)

Get radial plot
> radial(Meta_Both, level=0.95, pch=16, ylab="Standardized Effect Size", cex=1.5,
  col="blue", font=2, cex.lab=1.5)
> title("Yield loss in sweet corn from common rust (Shah&Dillard) - Radial Plot",
  cex.main = 1.2, font.main= 2, col.main= "blue")
```



***Note**: See **Section 1** for information on how to download, install and load the **meta** package*

d. Mixed-effect model

- Random effect model, with mean disease intensity as a continuous moderator

```
> Random_MeanD<- rma(slope, SE^2, method="REML", mods=cbind(MeanD),
data=RustData)
> Random_MeanD

Mixed-Effects Model (k = 20; tau^2 estimator: REML)

tau^2 (estimate of residual amount of heterogeneity): 0.1297 (SE =
0.0511)
tau (sqrt of the estimate of residual heterogeneity): 0.3602

Test for Residual Heterogeneity:
QE(df = 18) = 363.9921, p-val < .0001

Test of Moderators (coefficient(s) 2):
QM(df = 1) = 0.2639, p-val = 0.6074

Model Results:

         estimate      se    zval    pval    ci.lb   ci.ub
intrcpt    0.4350  0.1778  2.4468  0.0144   0.0865  0.7834  *
MeanD      0.0033  0.0064  0.5137  0.6074  -0.0093  0.0159

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Specifies the moderator variable

- **Random effect model, with variety as a categorical moderator variable**

```
> Random_Variety<- rma(slope, SE^2, method="REML", mods=~factor(Variety)-1,
  data=RustData)
> Random_Variety

Mixed-Effects Model (k = 20; tau^2 estimator: REML)

tau^2 (estimate of residual amount of heterogeneity): 0.0507 (SE = 0.0349)
tau (sqrt of the estimate of residual heterogeneity): 0.2252

Test for Residual Heterogeneity:
QE(df = 8) = 135.4046, p-val < .0001

Test of Moderators (coefficient(s) 1,2,3,4,5,6,7,8,9,10,11,12):
QM(df = 12) = 113.2668, p-val < .0001

Model Results:
                          estimate      se     zval    pval    ci.lb    ci.ub
factor(Variety)Bold         0.3383  0.2315   1.4611  0.1440  -0.1155   0.7921
factor(Variety)FSSweet      0.8181  0.1470   5.5670  <.0001   0.5301   1.1062
factor(Variety)Gcup         0.6129  0.1476   4.1538  <.0001   0.3237   0.9021
factor(Variety)HMX83865     0.3189  0.2256   1.4132  0.1576  -0.1234   0.7611
factor(Variety)Jubilee     -0.7992  0.3844  -2.0793  0.0376  -1.5526  -0.0459
factor(Variety)Rival       -0.5010  0.4819  -1.0398  0.2985  -1.4454   0.4434
factor(Variety)SnowWhite    0.5987  0.2260   2.6485  0.0081   0.1556   1.0417
factor(Variety)Squeen       1.1772  0.2665   4.4166  <.0001   0.6548   1.6996
factor(Variety)Sterling     0.4033  0.2259   1.7856  0.0742  -0.0394   0.8461
factor(Variety)Stylepak     0.5730  0.1376   4.1659  <.0001   0.3034   0.8426
factor(Variety)YBelle       0.1456  0.2562   0.5683  0.5699  -0.3566   0.6477
factor(Variety)Zenith       0.6230  0.2189   2.8458  0.0044   0.1939   1.0521
```

Specifies the moderator as a categorical (factor) variable