```
log using "D:\Jason\workshop\scales and indices\2022\stata code.log", replace


****************************************************************
* Read in the data in which no variables have missing values
****************************************************************

webuse bg2, clear
des
sum
gen id = _n


******************
* Sum scores
******************

gen  sum_score1 = bg2cost1 + bg2cost2 + bg2cost3 + bg2cost4 + bg2cost5 + bg2cost6
egen sum_score2 = rowtotal( bg2cost1 bg2cost2 bg2cost3 bg2cost4 bg2cost5 bg2cost6)

sum sum_score1 sum_score2

****************
* factor score
*****************

factor bg2cost1-bg2cost6
predict fact1 fact2

*****************************
* principal component scores
*****************************
pca bg2cost1- bg2cost6
predict pc1 pc2, score


***********************
* comparison
***********************
list id sum_score1 sum_score2 fact1 fact2 pc1 pc2 in 1/60
```

```
****************************************************************
* Read in the data in which some variables have missing values
****************************************************************

webuse bg2, clear
gen id = _n

**********************************************
* generate some missing values in the data
**********************************************
replace bg2cost1 =. in 1/30
replace bg2cost2 =. in 11/40
replace bg2cost3 =. in 21/50

sum
list bg2cost* in 1/60




******************************************************************************************
* generate an indicator showing how many items have missing values for the respondent
******************************************************************************************
egen missing = rowmiss(bg2cost1 bg2cost2 bg2cost3 bg2cost4 bg2cost5 bg2cost6)

list id bg2cost1 bg2cost2 bg2cost3 bg2cost4 bg2cost5 bg2cost6 missing in 1/60

*******************
* Sum scores
*******************

gen  sum_score1 = bg2cost1 + bg2cost2 + bg2cost3 + bg2cost4 + bg2cost5 + bg2cost6
egen sum_score2 = rowtotal( bg2cost1 bg2cost2 bg2cost3 bg2cost4 bg2cost5 bg2cost6)

sum sum_score1 sum_score2

****************
* factor score
*****************

factor bg2cost1-bg2cost6,
predict fact1 fact2


*****************************
* principal components core
*****************************
pca bg2cost1- bg2cost6
predict pc1 pc2, score


**********************
* comparison
**********************
list id missing sum_score1 sum_score2 fact1 fact2 pc1 pc2 in 1/60


log close
```