

## Loading a dataset (Mauritania DHS)

```
use "NPIR.dta"
```

## Drawing a random sample (20%)

```
sample 20
```

## keeping target variables (subset)

```
keep v012 v013 v024 v025 v106 v190 v201 v445 v463a v463aa hw5_1 hw8_1  
hw11_1
```

## Rename variables

```
ren (v012 v013 v024 v025 v106 v190 v201 v445 v463a v463aa hw5_1 hw8_1  
hw11_1) (Age_cont age region residence education wealth parity BMI  
smoking cigar standing wasting underweight)
```

## Getting the list of variables

```
. ds
cage      residence  parity    cigar     underweight
age       education BMI       standing
region    wealth    smoking   wasting
```

## Change variable names to uppercase

```
. rename *, upper

. ds
AGE_CONT      RESIDENCE  PARITY    CIGAR     UNDERWEIGHT
AGE           EDUCATION  BMI       STUNTING
REGION        WEALTH    SMOKING   WASTING
```

## Change variable names to lowercase

```
. rename *, lower

. ds
age_cont      residence  parity    cigar     underweight
age           education  bmi       stunting
region        wealth    smoking   wasting
```

## Sorting variables

```
order age bmi stunting wasting
```

## Scanning the dataset

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
cage	15,714	28.44604	9.622998	15	49
age	15,714	3.324233	1.916793	1	7
region	15,714	6.785032	3.968804	1	14
residence	15,714	1.503436	.5000041	1	2
education	15,714	.9805269	.8198688	0	3
-----+-----					
wealth	15,714	3.123075	1.35172	1	5
parity	15,714	2.532328	2.802365	0	17
bmi	7,346	2633.295	683.8877	1216	9998
smoking	15,714	.0188367	.1359524	0	1
cigar	15,714	.0255186	.1955345	0	2
-----+-----					
stunting	6,660	47.55676	1198.36	-598	9998
wasting	6,660	40.61577	1197.623	-475	9998
underweight	6,662	89.80261	1189.883	-398	9998

## Recode the variables with missing values (9998) to missing

```
recode bmi stunting wasting underweight (9998=.)
```

## Summarize

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
cage	15,714	28.44604	9.622998	15	49
age	15,714	3.324233	1.916793	1	7
region	15,714	6.785032	3.968804	1	14
residence	15,714	1.503436	.5000041	1	2
education	15,714	.9805269	.8198688	0	3
-----+-----					
wealth	15,714	3.123075	1.35172	1	5
parity	15,714	2.532328	2.802365	0	17
bmi	7,336	2623.256	627.9282	1216	5771
smoking	15,714	.0188367	.1359524	0	1
cigar	15,714	.0255186	.1955345	0	2
-----+-----					
stunting	6,566	-94.89552	136.582	-598	580
wasting	6,566	-101.9359	121.8822	-475	510
underweight	6,568	-52.00167	103.6249	-398	478

## Recode stunting wasting underweight per WHO guidelines

```
recode stunting (min/-301=3 "Severely_stunted") (-300/-201=2 "Moderately
stunted") (-200/max=1 "Not stunted") if stunting!=9998, g(stunted)
```

```
recode underweight (min/-301=3 "Severely_underweight") (-300/-201=2
"Moderately_underweight") (-200/max=1 "Not_underweight") if
underweight!=9998, g(under_weight)
```

```
recode wasting (min/-301=3 "Severely_wasted") (-300/-201=2
"Moderately_wasted") (-200/max=1 "Not_wasted ") if wasting!=9998,
g(wasted)
```

## Summarize the newly created variables

```
. sum stunted wasted under_weight
```

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
stunted	6,566	1.258757	.5578999	1	3
wasted	6,566	1.250838	.5118356	1	3
under_weight	6,568	1.069123	.2722141	1	3

## Checking the co-occurrence of the three indicators

```

gen malnutrition = ""
foreach v of varlist stunted wasted under_weight{
  replace malnutrition = malnutrition + "`v' " if `v' == 1
}

replace malnutrition = trim(malnutrition)
encode malnutrition, g(Malnourished)

```

## Label the new variable

```

lab var Malnourished "co-occurrence of stunting, wasting or underweight"

```

## Tabulate the new variable

```

tab Malnourished

  co-occurrence of stunting, |
  wasting or underweight |      Freq.    Percent    Cum.
-----+-----
           stunted |          219     3.39     3.39
  stunted under_weight |         315     4.88     8.27
           stunted wasted |           90     1.39     9.67
stunted wasted under_weight |       4,635    71.80    81.47
           under_weight |          759    11.76    93.23
           wasted under_weight |          437     6.77   100.00
-----+-----
                    Total |       6,455   100.00

```

## Tabulate without labels

```
ta Malnourished, nol

co-occurrenc
e of
stunting,
wasting or
underweight |      Freq.      Percent      Cum.
-----+-----
      1 |      219      3.39      3.39
      2 |      315      4.88      8.27
      3 |       90      1.39      9.67
      4 |    4,635     71.80     81.47
      5 |       759     11.76     93.23
      6 |       437      6.77    100.00
-----+-----
    Total |    6,455    100.00
```

## Revise the labels

```
la def Malnourished 1 "stunted" 2 "stunted & underweight" 3 "stunted &
wasted" 4 "stunted, wasted & underweight" 5 "underweight" 6 "wasted &
underweight", replace
```

```
la val Malnourished Malnourished
```

## Labelate to check the new labels

ta Malnourished

co-occurrence of stunting,   wasting or underweight	Freq.	Percent	Cum.
-----+-----			
stunted	219	3.39	3.39
stunted & underweight	315	4.88	8.27
stunted & wasted	90	1.39	9.67
stunted, wasted & underweight	4,635	71.80	81.47
underweight	759	11.76	93.23
wasted & underweight	437	6.77	100.00
-----+-----			
Total	6,455	100.00	

## Drop selected variables

```
drop stunted wasted under_weight
```

## Drop variables that start with a specific letter(s)

```
drop s*
```

```
drop s* w*
```



Drop variables that end with a specific letter(s)

```
drop *t
```

```
drop *t *e
```

## Make a quantile variable of a continuous one

```
xtile stunted_quantile = stunting, n(5)
```

```
ta stunted_quantile
```

```
5 quantiles |
```

```
of stunting |      Freq.      Percent      Cum.
```

	Freq.	Percent	Cum.
1	1,316	20.04	20.04
2	1,320	20.10	40.15
3	1,325	20.18	60.33
4	1,304	19.86	80.19
5	1,301	19.81	100.00
Total	6,566	100.00	

## MAKE A DECILE VARIABLE OF A CONTINUOUS ONE

```
xtile stunted_decile = stunting, n(5)
```

```
tab stunted_decile
```

```
      10 |  
quantiles |  
of stunting |      Freq.      Percent      Cum.  
-----+-----  
      1 |      657      10.01      10.01  
      2 |      659      10.04      20.04  
      3 |      666      10.14      30.19  
      4 |      654       9.96      40.15  
      5 |      665      10.13      50.27  
      6 |      660      10.05      60.33  
      7 |      656       9.99      70.32  
      8 |      648       9.87      80.19  
      9 |      647       9.85      90.04  
     10 |      654       9.96     100.00  
-----+-----  
    Total |     6,566     100.00
```

## Listing variables containing the label 'wasting'

```
lookfor2 wasting
```

variable name	storage type	display format	value label	variable label
wasting deviation	int	%8.0g	wasting	weight/age standard deviation
wasted (weight/age Malnourished	int	%17.0g	wasted	RECODE of <code>wasting</code> standard deviation)
wasting or	long	%29.0g	Malnourished	co-occurrence of stunting, underweight

## Saving the dataset in Excel format

```
export excel using "mydata", firstrow(variables)
```