

## Clase 4 STATA

Se trabajará con la base de datos que se llama “Duplicados”.

1. Reemplazar valores de una variable con el comando “replace”.

```
replace cintura= . if cintura>300 | cintura==0
```

2. Generar una nueva variable con el comando “gen”.

```
gen imc = peso/(talla/100)^2
```

3. Recodificar una variable con el comando “recode”

```
recode imc (min / 18.4999999=1 bajo_peso) (18.5 / 24.9999999=2 normopeso) (25 / 29.9999999=3 sobrepeso) (30 / max= 4 obeso), gen(EN)
```

```
tab EN
```

```
bysort EN: sum imc
```

4. Crear etiqueta con el comando “label define”.

```
label define Educación 1 "Bajo" 2 "Medio" 3 "Alto"
```

5. Adjuntar etiqueta que se llama “Educación” a la variable que se llama “nedu”

```
label values nedu Educación
```

6. Como chequear valores duplicados en las bases de datos.

```
duplicates tag id, generate(dup)
```

```
tab dup
```

```
sort id
```

```
list id edad nedu sexo peso talla coles if dup==1
```

Para dejar la base de datos lista debemos realizar:

```
replace id=2977 if id==2976 & peso==88.5
```

```
replace id=4167 if id==4166 & talla==150
```

```
list id edad nedu sexo peso talla coles if dup==1
```

```
duplicates drop id, forcé
```

```
drop dup
```

