Exercise 4: Housekeeping

Begin by creating a new do-file which you should then save as *Stata_Exercise4.do*. Add some appropriate comments at the beginning of the file. Remember to keep saving the do-file as you go along. Run through these exercises referring to chapter 5 in the module notes.

Add commands to change the working directory to the *Exercise 4* folder and to load the dataset *bl_combined.dta*.

Exercise 4.1 Renaming Variables

- Use the `rename` command to do the following:
  - Rename *wt* as *wt_kg, ht* as *ht_cm* and *smkstat* as *smoke5*
  - Add the prefix bl_ to the variables *creat, hb, pot, sodium* and *totbil*.
  - Change all the variable names to lower case

Exercise 4.2 Labelling Variables

- Use the `label variable` command to:
  - Add the following variable labels: *birthdt "Date of birth"*, *age "Age-years", agegroup "Age group-years", sbp "Systolic Blood Pressure-mmHg" , dbp "Diastolic Blood Pressure-mmHg"*, *hrate "Heart Rate-bpm", diab "Diabetes", miprev "Previous MI" and hfhosp "Previous Hospitalisation for heart failure"*

Exercise 4.3 Labelling Values

- Defining and Attaching Value Labels:
  - Use the `label define` command to create a value label
    - *hfdiag_lab* with values and labels as follows: 1 = "Ischaemic", 2 = "Non-ischemic"
  - Use the `tabulate` command to obtain a table of the variable *hfdiag* – why no labels?
  - Use the `label values` command to attach the value label *hfdiag_lab* to the variable *hfdiag.*
  - Now repeat the `tabulate` command.
  - Define and attach a value label to the variable *smoke5*. (Note: refer to the data dictionary on page 10 in your module notes)
  - Create a value label called *noyes_lab* with values and labels 0 = "No" and 1 = "Yes".
  - Attach this value label to the variables *angina*, *strisch*, *strhem*, *stremb*, *stroth*, *hyptn*, *diab*, *copd*, *cabg* and *pci*.
  - Try using the `foreach` command to loop through the list of variables and so save time.
  - Modify the value label that is attached to *agegroup* so that the labels are "<65 years", "65-69 years", "70-74 years", "75+ years"

Exercise 4.4 Viewing Labels

- Using the describe command see which variables have value labels attached and to find the names of the value labels.
- Use the `label list` command to obtain a list of all the value labels in the dataset in memory.
- Use the `label list` command to obtain a listing of the *smoke5* value label.

Exercise 4.5 Adding Notes

- Add a note to the dataset "Combined baseline data from HF study"
- Add a note to the variable *wc* "Some WC measured in metres"
- Add a note to the variable *bl_creat* "Missing values were recorded as 8888 and 9999"
- View the notes

Add a command to save the modified dataset as bl_combined1.dta.

Save the do-file.