The data provided here is a subset of data collected as part of the Framingham Heart Study, and includes information on 4,434 adults aged between 32 and 70 years. Random sampling was used to recruit people to the study. Observations in this sample are independent from each other, except where repeated measures are recorded (e.g. blood pressure measurements at baseline and follow- up).

**Table 1**. Variables included in the dataset

|  |  |  |
| --- | --- | --- |
| **Variables** | **Meaning** | **Scale/coding** |
|  |  |  |
| **Sex** *(sex)* | Biological sex at first exam | 1=male  0= female |
| **Smoking status at baseline**  *(cursmoke1)* | *Smokers*: those who reported that they currently smoke  in their first clinical examination/interview? | 1 = smokers  0 = non-smokers |
| **Diabetes diagnosis at**  **baseline** *(diabetes1)* | Diabetes: person diagnosed with diabetes at first exam | 1 = yes  0 = no |
| **Systolic blood pressure at**  **baseline** *(sysbp1)* | Systolic blood pressure (mmHg) at first exam | in mm/Hg |
| **Diastolic blood pressure at**  **baseline** *(diabp1)* | Diastolic blood pressure (mmHg) at first exam | in mm/Hg |
| **Systolic blood pressure at**  **follow-up** *(sysbp2)* | Systolic blood pressure (mmHg) at second exam | in mm/Hg |
| **Diastolic blood pressure at**  **follow-up** *(diabp2)* | Diastolic blood pressure (mmHg) at second exam | in mm/Hg |

Using this subset of the Framingham dataset (attached with this assignment) you are asked to conduct an appropriate statistical test to answer each of the following questions:

1. Is baseline systolic blood pressure consistent with the recommended value of 120 mm/Hg?
2. Does baseline systolic blood pressure vary depending on whether a person has a diagnosis of diabetes?
3. Does baseline systolic blood pressure vary by sex?
4. Does diastolic blood pressure change over time (i.e. did it change between the first exam at baseline (diabp1) and the second exam at follow-up (diabp2))?
5. Does smoking status vary by sex?

For **each question above**:

* Write a statement to represent the null and alternative hypotheses
* State the inferential statistical test you have chosen to answer this question and why that test is appropriate
* Copy the results from Stata into your word file (copy the table of results as a picture and paste into word file –as we did in practical class)
* Interpret the results

Please submit your log file, as a record of the analyses you carried out in Stata, with your assignment onto Moodle. For this exercise you will not need to create a ‘do file’ making the dataset a working file, as all the variables are already created and the data is clean. You will need to create a ‘do file’ for the analyses to answer each of the questions. To run the analyses from your ‘do file’, make sure you have the dataset open (i.e. Framingham\_dataset\_CA2.dta), then open your ‘do file’ with all your commands and select the ‘execute (do)’ icon highlighted in red below. If you select execute (do) it will run all the commands in your ‘do file’. If you only want to run a select command, highlight that command and then select execute (do). All the results will then appear in the Stata Results window. And if you have opened and saved a log file, they will be recorded in the log file. Once complete, make sure you close the log file.

To create your log file:

* File > Log >Begin
* Save log file
* Run your analyses (i.e. statistical tests to answer questions 1-5) – as explained above using your do file.
* File > Log > Close
* To open log file File> Log >view