\*For this example I have to create a fresh directory because this is a permanent file

obtained by my former SAS instructor at AU;

\*Starting with a LIBNAME statement to store the permanent file

and in order to read in the permanent file i will use the SET statement;

\*SaS example is the folder i created;

\*My sas permanent folder is example;

LIBNAME example "E:\SAS programming\SaS example";

**DATA** Biometry;

\*Read in step;

SET example.diabetes;

\*Using the DROP step to drop the variables I want to drop

I should have been able to use KEEP to keep the variables i want and the rest

should automatically drop. I got an error for that because the rest of the values

were not referenced. I wanted to get this example in to you as soon as possible after

i received my feedback;

\*I can continue to drop however many variables, I only chose 2;

DROP diagnosis\_1 diagnosis\_2;

\*data=Biometry is important because if i create and run another dataset,

the previous dataset won't be saved and will always run the latest dataset created;

**PROC** **PRINT** data=Biometry;

**RUN**;

\*Once this file runs, if you're able to run it on your machine, the log will say

the file contains 6222 observations. Not quite 1,000,000. This is because biostatisticians

do not typically work with big data in the medical industry, although they can.

\*From here i can do a PROC MEANS to get summary data

I can target a few variables and create a Table out of them using Table key

and VAR key with PROC FREQ

\*Overall, this is the first step i would make in creating a sub dataset out of

the larger dataset to focus on doing data analysis on