**/Jimma University**

**Institute of Health**

**Faculty of Public Health**

**Epidemiology department**

**Stata software Individual assignment**

**May 2023**

**Evaluated out of 20%**

# Problem I.

1. For this ssignment use the dataset **hwk2.dta**, available on the course web site. The data are a subset of the data from HERS, the trial of hormone therapy for secondary prevention of coronary heart disease. The dataset includes the following variables for the 2,032 diabetes-free women with glucose \_125 mg=dl:

\_ ht (treatment assignment: 1 = hormone therapy, 0 = placebo)

\_ glucose0 (fasting glucose at baseline, in mg=dl)

\_ glucchange (change in glucose from baseline to yr 1 in mg=dl)

\_ exercise (at least 3 times per week: 1 = yes, 0 = no)

\_ age (in years at baseline)

\_ raceth (race/ethnicity: 1 = white, 2 = African American, 3 = other)

\_ smoker (current smoker at baseline: 1 = yes, 0 = no)

\_ bmi (body mass index at baseline, in kg/m2)

This part of the exercise is motivated by the Diabetes Prevention Project, which showed a protective effect of exercise and weight loss in preventing progression to diabetes among women with fasting glucose levels between 100 and 125 mg/dl.

* 1. Categorize age into four using IQR and rename it as “agecat”
  2. Rename category as IQR 1 as lowest, IQR as medium, IQR as Normal and IQR 4 as Higher
  3. Show agecat in bar graph
  4. For the BMI for BMI0 and BMI1 analyse summary and compare using mean.
  5. Show BMI categorized by bar graph for BMI0 and BMI1
  6. Analyse the descriptive analysis(mean, median, mode, IQR, etc) for variables
  7. Check the normality distribution for fasting glucose levels

# Problem II.

For this Assigment you will need **hwk3.dta**, available on the course web site. The dataset includes observations for a 10% random sub-sample of the HERS cohort.Variables were age, race/ethnicity, alcohol use, exercise, WHR, and BMI, and obtain the residuals.

* + 1. Show age with glucose using scatter plot and write what you understand from the distribution.
    2. Show categorize BMI by histogram
    3. Show the glucose level for diabetes using two way bar graph method.
    4. Show the triglycerides (mg/dl) in box-plot graph and describe message from graph

**N.B**

**For Checking assumptions (if any)**

1. The exercise concerns the association between central obesity, as measured by waist-hip ratio (WHR), and triglyceride (TGL) levels. Both are components of the so-called metabolic syndrome and precursors of diabetes and cardiovascular disease.
2. Prepare a Word file with answers to each of the following questions. Paste relevant STATA output and graphs into the file to illustrate your points.

Submit only outputs **in pdf file** and also **commands** with separate file

1. Your file should have file name as:-

*”your name\_your Fathername\_stata\_assignment\_department\_year” .*

**Good Luck!**