**MEL Manager – Analysis Exercise**

**This exercise should be completed in Stata. Please use a do file and capture all Syntax in the do file and submit the do file along with your responses.**

**Please respond to the following questions on this sheet. Make sure to capture all Stata Syntax in your do file.**

Step 1: Import “data set 1” into Stata.

* Question: What is the unit of observation in data set 1?

**Answer: hhid**

* Question: How many observations exist in this dataset?

**Answer: 420 observations, 7 variables**

Step 2: The data is currently wide by household. Reshape the data so it is long by household

* Question: After the reshape there are two observations with missing hh id. Why is that?

**Answer: Two hh observations missing hh id because:**

* + - 1. **The hh ids were NOT filled out in the CR forms during data collection.**
      2. **The hh ids were NOT keyed in during data entry.**
      3. **The households in the two groups were NOT assigned unique identifiers.**

Step 3: Drop the observations with missing hh id from the dataset.

**Answer: 2 observations dropped**

Step 4: Check for duplicate hh ids.

* Question: How many sets of duplicates exist in the dataset?

**Answer: 4**

Step 5: For each set of duplicates, keep the first observation and delete any subsequent duplicate observations

**Answer: Deleted**

Step 6: Merge with “data set 2”

* Question: How many matched observations are there?

Step 7: Keep the matched observations only

**Please submit this worksheet with the responses to the questions above answered. Please also submit your do file containing all of the syntax you used to complete this exercise. Good Luck!**