```
/* Check the contents of each dataset */
proc contents data=WORK.IMPORT; run;
proc contents data=WORK.IMPORT1; run;
proc contents data=WORK.IMPORT2; run;
/* Rename variables in each dataset to match a standard format */
/* Adjust variable names in CarsDatabase */
data carsdb;
    set WORK.IMPORT;
    rename "City MPG (FT1)"N = city_mpg
           "Engine Displacement"N = engine_displacement
           "Annual Fuel Cost (FT1)"N = annual_fuel_cost
           "Tailpipe CO2 (FT1)"N = CO2_emissions;
run;

/* Adjust variable names in Hybrid Vehicle Fuel Efficiency */
data hybriddb;
    set WORK.IMPORT1;
    rename "Average Fuel Efficiency"N = avg_fuel_efficiency;
run;

/* Adjust variable names in Real World Fuel Efficiency */
data realworlddb;
    set WORK.IMPORT2;
    rename "ACTUAL FUEL ECONOMY Geotab"N = actual_fuel_economy;
run;

/* Combine the datasets */
data combined;
    set carsdb hybriddb realworlddb;
run;

/* Clean the combined dataset: Handle missing values */
proc sql;
    delete from combined
    where engine_displacement is missing
    or CO2_emissions is missing
    or city_mpg is missing
    or annual_fuel_cost is missing;
quit;

/* Check the cleaned data */
proc contents data=combined; run;
/* Perform Linear Regression for Hypothesis 1 */
proc reg data=combined;
    model CO2_emissions = engine_displacement / clb;
run;
quit;
/* Rename variables to standardize across datasets */
data carsdb;
    set WORK.IMPORT;
    rename "City MPG (FT1)"N = city_mpg
           "Engine Displacement"N = engine_displacement
           "Annual Fuel Cost (FT1)"N = annual_fuel_cost
```

```
            "Tailpipe CO2 (FT1)"N = CO2_emissions
            "Start Stop Technology"N = start_stop_technology;
run;

data hybriddb;
    set WORK.IMPORT1;
    rename "Average Fuel Efficiency"N = avg_fuel_efficiency;
run;

data realworlddb;
    set WORK.IMPORT2;
    rename "ACTUAL FUEL ECONOMY Geotab"N = actual_fuel_economy
           "Hybrid/Non-Hybrid"N = hybrid;
run;

/* Combine the datasets */
data combined;
    set carsdb hybriddb realworlddb;
run;

/* Clean the combined dataset: Handle missing values */
proc sql;
    delete from combined
    where engine_displacement is missing
    or CO2_emissions is missing
    or city_mpg is missing
    or annual_fuel_cost is missing;
quit;

/* Perform Linear Regression for Hypothesis 1 */
proc reg data=combined;
    model CO2_emissions = engine_displacement / clb;
run;
quit;

/* Perform t-test for Hypothesis 2: Start-Stop Technology impact on City MPG */
proc ttest data=combined;
    class start_stop_technology;
    var city_mpg;
run;

/* Perform t-test for Hypothesis 3: Annual Fuel Costs between Hybrid and Non-Hybrid */
proc ttest data=combined;
    class hybrid;
```