

```
/* Load and combine the 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;
```

```
/* Convert start_stop_technology to numeric */
```

```
data carsdb;  
  set carsdb;  
  if start_stop_technology = 'Y' then start_stop_numeric = 1;  
  else if start_stop_technology = 'N' then start_stop_numeric = 0;  
  
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;  
  
run;
```

```
/* Combine the datasets */
```

```
data combined;  
  set carsdb hybriddb realworlddb;  
  
run;
```

```
/* Clean the combined dataset: Handle missing values */
```

```
proc sql;  
  delete from combined  
  where start_stop_numeric is missing  
  or city_mpg is missing;  
  
quit;
```

```
/* Perform t-test for Hypothesis 2: Start-Stop Technology impact on City MPG */
```

```
proc ttest data=combined;  
  class start_stop_numeric;  
  var city_mpg;  
  
run;
```

```
/* Perform Correlation Analysis for Hypothesis 2 */
```

```
proc corr data=combined;  
  var start_stop_numeric city_mpg;  
  
run;
```