Use any programming language (e.g. **Python, Matlab, C/C#, R, Excel VBA**) to complete the tasks below. Do not use Excel. ***Task 1*** is worth **50 points**, while ***Tasks 2 and 3*** can get **25 points** each; completing ***Task 4*** can earn you **50 bonus points**.

Hope you have fun while coding!

The parent folder contains **…\RawDataTest.csv**; a brief description of what is contained in the file is included below. This file contains **second-by-second raw engine data** collected from a sample passenger car.

* The raw data contains a header that is essentially the names of 40 different variables for which data is included in the rows below.
  + **List of Variables**: TRIP\_ID; START\_DATE; START\_TIME; END\_DATE; END\_TIME; Seconds ; Time; Latitude; Longitude; Course; GPS Speed ; Altitude GPS; Barometer Pressure (kPA); Barometer Altitude (m); Speed (km/h); Engine Load; RPM; Battery State (SOC) (%); Manifold air pressure ; In-take air temperature (deg C); Air mass flow (g/s); Throttle position (%); Coolant temperature (degree C); Acceleration X (m/s2); Acceleration Y (m/s2); Acceleration Z (m/s2); Slope (m/m); Injection Cut (s/n); Instant Consumption (L/100km); Accumulated Consumption (L); Average Consumption (L/100km); Gear shifting; Idle (s/n); Cold Running (s/n); FUEL\_TYPE; DISPLACEMENT (cc); CAR\_DATE; I N\_AREA; AREA\_IDS; REFERENCE\_POINTS
* The data for each of the respective variables, including the header, is delimited by semi-colon (;).
* Note: Consecutive semi-colons represent missing or non-recorded values.
* The decimal values are delimited by comma (,) not by a dot (.) – e.g. 1.524 is written as 1,524.

***Task 1 – Data cleaning (50 points)***

Clean the raw data (in RawDataTest.csv) and **save the clean data set in a separate .csv file** in the parent folder. The clean data table must meet the following requirements:

1. Separate columns of the data for the following variables; data types are also specified:
   1. **TRIP\_ID** as character; **Seconds** as integer; **START\_DATE** as date/time; **START\_TIME** as date/time; **Speed (km/h)** as numeric; **RPM** as numeric; **Latitude** as numeric; **Longitude** as numeric
2. Empty values should be represented as NA, NaN, etc. and not as zeroes.
3. Correct the decimal values. The raw data for **Speed (km/h)**, **RPM**, **Latitude** and **Longitude** has decimal values represented by comma (,) and not as dot (.).

***Task 2 – Analysis (25 points)***

* Compute **total travel time**, **idle time** and **maximum speed** in an output table for unique all **Trip\_IDs**.
  + **Total travel time** can be computed from the maximum value of the variable **Seconds** for a given **Trip\_ID.**
  + **Idle time** is calculated based on total time for which **Speed (km/hr)** values are zero and corresponding **RPM** values are non-zero.

***Task 3 – Plotting (25 points)***

For **TRIP\_ID** with maximum **total travel time** (computed as in *Task 2* above, create a speed vs. time plot. The output must include legends with its **START\_TIME** and **total travel time**.

***Task 4– Spatial Plotting (50 bonus points)***

The raw data you just cleaned also contain GPS coordinates (latitude and latitude). For the same **TRIP\_ID** as computed in *Task 3*, plot the car’s route on a map (NOT via maps.google.com!!).