# Introduction to Scientific Computing, PSCB57, Fall 2017
# Geotab Project

You will work on the Geotab Project for majority of the fall term. It is a team-based project with about 5-6 students per team. There are three different deliverables. These deliverables will be marked and will contribute to your final grade.

Geotab (`https://www.geotab.com`) is a telematic company with headquarters in Oakville Ontario (psst, they are hiring!). Geotab offers a small device that plugs into the diagnostics port of a car or truck. It collects data from sensors in the car, from a built in GPS receiver, and an accelerometer. This data is then sent in real time to a Geotab database over the cellular network. The data can be used to manage drivers and vehicles in a company's fleet of cars and trucks.

You will use a Geotab device to collect your own data and then built your own data reduction pipeline. This will give you an idea of what working with real world data is like. We use the telematic data because it is relatively easy to understand and get started, especially when compared to more scientific datasets e.g. from astrophysics, genomics, or chemistry. Geotab has kindly provided us with 12 devices, so each group can collect their own data.

You have to come up with your own project idea. You decide what to do with the data that you collect. You can try to incorporate some of the techniques learned in this course, but you will probably need to go a bit further and use a method or package that we did not cover in the lectures. Keep in mind that you have to finish the project in just one term. It is more important to get something simple working, rather than getting stuck while implementing a more complicated idea. The most complicated idea will not necessarily get the best mark.

You are expected to use Jupyter notebooks and python for this project. If you think there is a good justification to work in another language, please confirm this with the TA or Professor Rein first.

The primary resource for you to get help with this project are the tutorial. The TAs will be available for your questions during the tutorials, via e-mail and during their office hours. If the TAs cannot help you, talk to Professor Rein, either after the class, during the tutorials or during his office hours. If you get stuck, or think your project is not going to succeed, please ask for help early.

## Deliverable 1 - Project Proposal

1. As a first step, form a group of 5-6 students. Try to make sure your group has a diverse skill set. For example, try to find students outside of your own program.

2. Once you have formed a group, get a Geotab device from Prof. Rein, and plug it into a car, either your own, or that of a friend or family member. If it is not your own car, you have to get the car owner's written consent.

3. Get an account and make sure you can download data. This is important. The interface provided by Geotab is very polished and has many features. However, for this project, you should primarily work with the raw data. Make sure the device works and collect data.

4. Once you downloaded the data, try to import it to a Jupyter notebook and play around with it a little. Then brain-stormed possible ideas in your group. Here is a list of examples.

   - For a car that is used in a daily commute, write a program that calculates on which day of the week the commute is shortest/longest. For that you need to read in trip data, find out when the commute starts and when it ends. You want to make sure to not confuse the commute with a shopping trip. You might also want to think about over how many weeks you need to measure the commute to get good answers.
   - Again, for a car that is used in a daily commute, compare different routes and find out which one is faster. You need to automatically detect when a trip starts and ends. And similar to above, you want to think about how much data you need to come to a conclusion.
   - Find all gas stops a car made. For that, you need to find a list of all gas stations in Toronto together with their coordinates. Then you would check if a car comes near to one of the gas stations. If it stops there for more than a 2 minutes, it is probably a gas stop.
   - Looking at the accelerometer data, try to find a way to measure if the road has pot holes. You need to find a way to disentangle normal accelerating and breaking from pot holes and think about how much data you need to give a confident pot hole location. This kind of data would be very useful for cities.
   - Try to find a correlation between outside air temperature and fuel efficiency. Think about what possible other factors could play into this and whether a correlation implies causation in this case.

5. Once you've all agreed on an idea, write it down in the form of a proposal.

   - The proposal should be about one page long, but no more than two pages.
   - It should include a name for your team and the names of your team members.
   - Make a case for why your idea is interesting, either from a business or scientific point of view.
   - List which parameters you need for your project (e.g. time, position, accelerometer data, etc).
   - List which other dataset you might need (e.g. a list of gas stations) and where you might get that data from (one idea: check out Open Data Toronto).
   - Estimate how much data you will be collecting and explain if this is enough to achieve your goal.

6. Please hand in the proposal in PDF format by e-mail to Professor Rein. The deadline is October 20th, 6pm.

## Deliverable 2 - Report

1. Write a report between 5 and 10 pages long that outlines your project, the implementation and the results. The page limit includes plots, tables, and maps.
2. List your team name and all team members. State clearly which team member worked on which part.
3. Structure your report into abstract, introduction, methods and results.
4. Say how your final project differs from the proposal. For example, did you have to use other data sets than you originally thought?
5. Was the amount of data you collected sufficient? Make sure to include a statistical justification to your answer.
6. Please send the report in PDF format by e-mail to Professor Rein. The deadline for the final report is December 4th, 6pm.

## Deliverable 3 - Presentation

1. The presentations will take place during the last lecture on December 4th.
2. Each team will have 10 minutes to present their idea and the results. Make sure to time your presentation. You cannot go over the time limit.
3. You may use slides. Please send the slides in PDF format by e-mail to Professor Rein, no later than December 3rd, 6pm. If you use powerpoint or keynote, please export the slides to PDF.
4. If you want to show a live demo of your work and you need to plug your laptop into the projector, you are responsible for making sure it works. There is no extra time for setting it up!
5. There will be a 2 minute question and answer session at the end.
6. It is ok if not all team members are presenting, however all team members have to be present at the presentation.