

# CMPT 826 - Data and Process Modeling and Analytics

[Home](#) / [My courses](#) / [CMPT 826 \(18/19WT2\)](#) / [Assignments](#) / [Assignment 2](#)

## Assignment 2

In this assignment you will examine the processing of operationalizing a behaviour, and the impact of different types of operationalizations.

### STEP 1: Preparation (20 Marks)

Using code from or similar to the MAUP assignment, filter out participants with less than 50% of possible battery data, and GPS points outside of Saskatoon, using the same bounding box as the MAUP assignment (52.058367, -106.7649138128), (52.214608, -106.52225318), and a reported accuracy of 100m or better. Aggregate location to the duty cycle level. Grid space at 100 m, and express all locations as positions on this grid.

### STEP 2: Trip Definition (10 Marks)

Operationalize three different trips, using the N-times definition with N equal to 1, 3 and 5 duty cycles. For each participant determine the number of trips, the length (in grid cells) of the trips, and the time (in duty cycles) of each trip. Provide commented code as part of the hand-in.

### STEP 3: Presentation (20 Marks)

Create the distributions for trip number (over participants), trip length (over participant-trips), and trip duration (over participant trips) and plot these distributions for each N as multiple curves on the same plot (3 plots with 3 curves). Change the axis to show what you consider the best representation of these distributions (e.g. linear for Gaussian, log-linear for exponential, log-log for power law). Explain how you concluded which representation best fit the data. Plot a heat map over all participants which includes only trips, and another which includes only non-trips for each N, for a total of 6 maps.

### STEP 4: Interpretation (30 Marks)

In a series of paragraphs, address the following questions.

4.1 What is the impact of changing N on the distribution of trip number, length and duration? What trips are being captured and which are being ignored? Given an example of a research question where the differences would be important, and an example of a question where they would be unimportant.

4.2 What distinguishing features did you see in the heatmaps? Where there points included in either map (trip, not trip) at any N that seemed out of place? How would you change the operationalization to eliminate these points?

4.3 Using the DADeP as headings, describe the entire process of this assignment as a series of algorithms. Assuming you have access to a function `probability_of_location(location)` describe the process using the dynamic location formalism described in class.

Quality of writing and presentation (10 Marks)

Quality of code and code comments (10 Marks)

## Submission status

|                   |                       |
|-------------------|-----------------------|
| Submission status | Submitted for grading |
|-------------------|-----------------------|

|                  |  |
|------------------|--|
| Grading status   | Graded   |
| Due date         | Monday, 4 March 2019, 12:00 AM                 |
| Time remaining   | Assignment was submitted 3 hours 41 mins early |
| Last modified    | Sunday, 3 March 2019, 8:18 PM                  |
| File submissions |  |

Submission comments

+

[Comments \(0\)](#)

Edit submission

You can still make changes to your submission

## Feedback

|                   |   |
|-------------------|---|
| Grade             | 95.00 / 100.00  |
| Graded on         | Friday, 29 March 2019, 2:21 PM  |
| Feedback comments | <div><div><div></div><div>+</div></div><div>Step 1: Preparation (18 of 20 Marks)<ul style="list-style-type: none"><li>Filter participants with less than 50% of possible battery data (4)</li><li>Remove all GPS traces ...</li></ul></div></div> |

◀ Assignment 1

Jump to...

◀▶

Assignment 3 ▶

