

**Aims:**

The aim of this laboratory is for you to familiarize yourself with using Apache Derby and to time queries.

**Task 1: Unfinished work**

Make sure you have completed the work from the previous lab sheets (all of which contain essential building blocks for the first assignment). This includes having both MongoDB and Derby installed in an AWS instance (and have consolidated billing).

**Task 2: Become familiar with the data for the assignment**

The following tasks will help you get started on the assignment (although the assignment is an individual assignment it is fine to work on the following with other students to develop your understanding of the data):

- (1) To become familiar with the data to be used in the assignment, go to <https://data.melbourne.vic.gov.au/Transport/Pedestrian-Counting-System-Monthly-counts-per-hour/2aktrb> and read the page. This should help you understand what the data is about.
- (2) Download the compressed version of the data from the file Pedestrian\_Counting\_System - Monthly\_counts\_per\_hour.csv for the assignment (see assignment for details of where to download it from), and get the data on to your the AWS linux instance (of course you may choose to first download and manipulate the data on another computing environment instead).
- (3) Use the "View data" link on the website to find out for which of the following the questions you can use this interface to find answers (in the assignment you will need to write SQL queries to get these same answers):
  - a) Where is the sensor with ID number 40 located?
  - b) For how many days in 2019 has the City of Melbourne published the sensor data of pedestrian counts?
  - c) In 2019, what location had the most pedestrian (was the total number of pedestrian at that location across the whole year)?
  - d) What was the busiest day in the city in 2019 (give the date)?
  - e) What day of the week (i.e. Monday to Sunday) is the busiest?
- (4) Manually convert a couple of rows of data into insert statements for both MongoDB and Derby (this may not be for your final database design).
- (5) Consider what tools could be used to automatically convert all the rows into insert statements.