**Issues with the ETL**

The function.py file can be found [here.](https://bitbucket-students.deakin.edu.au/projects/D2IC-PG/repos/d2i---melbourne-city/browse/T2_2021/ETL/function.py) At the moment, it is recording parking sensor data split into daily files ([found here](https://s3.console.aws.amazon.com/s3/buckets/opendataplayground.deakin?region=ap-southeast-2&prefix=parkingsensor/daily/&showversions=false)). There is a [data dump file](https://s3.console.aws.amazon.com/s3/object/opendataplayground.deakin?region=ap-southeast-2&prefix=parkingsensor/parkingsensor.csv) that the website is currently using, however, it is not currently updating as the ETL would break once the file got to a certain size. To mitigate this daily dumping was encouraged but then there were issues with reading each file for 28 days’ worth of data and the website would time out. To solve this issue the daily dump file should be converted to a 28-day file (by deleting data that’s more than 28 days) whilst still dumping daily files. This has not been done and should be done in trimester 3 if the team were to continue with the parking sensor use case.

There was an idea to automate the [unique bay ID file](https://s3.console.aws.amazon.com/s3/object/opendataplayground.deakin?region=ap-southeast-2&prefix=static_datasets/UniqueBays.csv) ([also here](https://bitbucket-students.deakin.edu.au/projects/D2IC-PG/repos/d2i---melbourne-city/browse/T2_2021/ETL/UniqueBays.csv)) so as to update them as new parking sensors are uploaded to the system (an issue of vandals breaking transponders). But instead, this can be done manually so as not to slow down the ETL. This file is not being used by the website.