Group member:

Yun-Tung, Jemma, Allie, Duncan

**Friday 29th September 2023, 2-4pm:**

*Everyone:*

We went over the group assignment outline, brainstormed subject ideas and looked at different data resources online. We seemed to settle on something related to environmental emissions, e.g., carbon dioxide (CO2), carbon monoxide (CO1), methane (CH4), particulates, etc.

Some potentially relevant datasets...

* Carbon monoxide (CO1) concentrations (mg/m3) across NZ at different locations throughout the day from 2004-2020:<https://data.mfe.govt.nz/table/106251-carbon-monoxide-concentrations-2004-2020/>
* Minimum, maximum, and average temperature (C) per-day across NZ from 1909-2019:<https://data.mfe.govt.nz/table/105056-daily-temperature-1909-2019/>

**Friday 6th October 2023:**

Duncan:

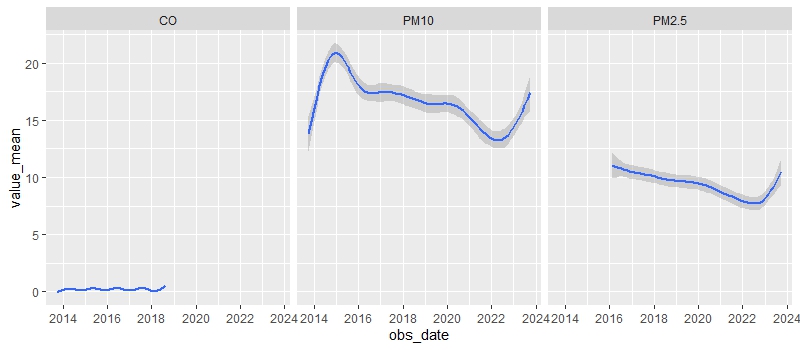
* Found the environment canterbury air quality data repository: <https://data.ecan.govt.nz/Catalogue/Search?Query=air&CollectionId=0>
* Created a Google Drive folder containing air pollution data on a 10 min basis: <https://data.ecan.govt.nz/Catalogue/Method?MethodId=29#tab-desc>

Allie:

* Looked over ECan data, added some location labels and noted null datasets
* “Pollution\_data” nulls: burnside, data4, data7, data10, data12, data14 over the past 10yrs

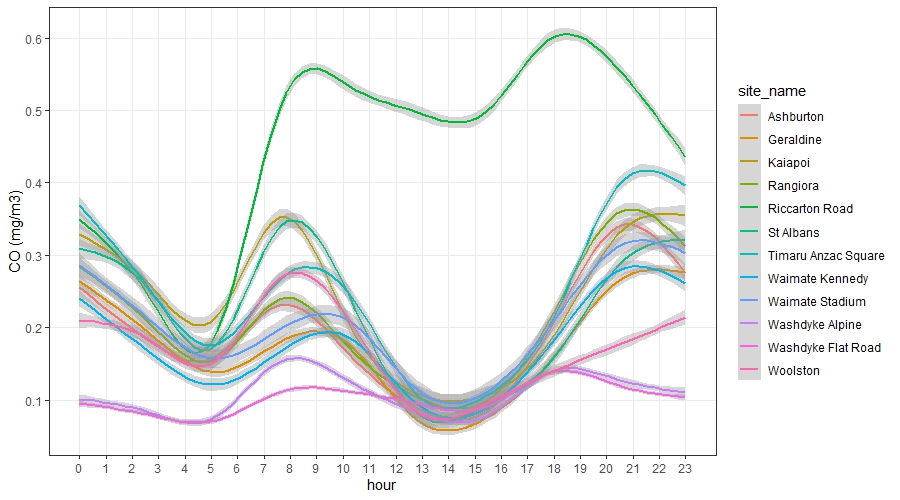
**Sunday 8th October 2023:**

Allie:

* I gathered data from eCan sites from the past 10 years (1/10/2013 to 1/10/2023) and labeled files accordingly. I found that far fewer sites had no entries, with the exception of Burnside, Lincoln, and Timaru Grey Rd. Based on <https://data.ecan.govt.nz/Catalogue/Method?MethodId=23#tab-data>, these sites stopped recording in 2010, 2010, and 2006 respectively. Not sure whether we want to investigate these any further (or why they stopped recording, for that matter)
* Began poking around in said data and started writing some R code. Did some wrangling (à la the course name) using the Geraldine site just as a test. It looks like different tests begin/end at separate times. For example, CO stops mid-2018 and PM2.5 starts in 2016. I just used a facet wrap for now; I’ll need to add some consideration for the different scales the different measures use. Probably easiest just to graph separately. Here’s the graph:
* There’s a visible drop in both mean particulate values after 2020, which depressingly begins to increase again around 2022.
* Will combine the datasets and graph on a contaminant-by-contaminant basis, with grouping by location.

**Monday 9th October 2023**

Allie

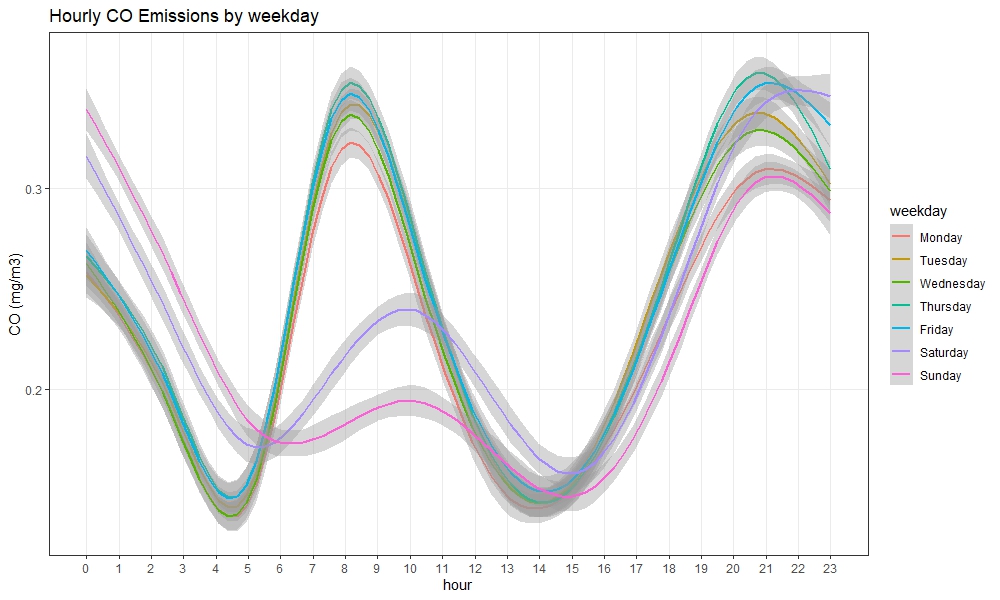
* Did some more stuff in R. Compiled all the data into one frame and started to generate plots. Plots regarding change over time need to be investigated further; they’re pretty messy when comparing sites.
* Comparing across the day looks more typical, however. 
* It looks pretty consistent with other reports on this topic, even: <https://www.stats.govt.nz/indicators/carbon-monoxide-concentrations/>
* Riccarton Rd is almost always congested, which makes sense. Also, the collection unit is like 2m away from the road. I’m not an expert on how pollutants permeate in the air, but I have to imagine that it would have some impact.

**Tuesday 10th October 2023**

Everyone:

* Discussed the direction of the project, went over some of the initial findings/points of interest. For example, Riccarton road is abnormally high in CO. The unit is very close to the road, which may contribute to this. Similarly, the Washdyke Alpine site is high in PM10, and in the center of an industrial area.
* Brought up some tasks that need doing, including:
  + Get some descriptions of the different measurement units, according to coordinates provided in the data
  + Create map of locations measured
  + Generate more plots, such as trends over time
  + Examine plots for outliers, see if location has any explanatory power
  + Gather some info about the contaminants (e.g., “what produces PM10?”). Doesn’t have to be super complicated, just something to discuss.
  + Create a Julia component(?)

Allie:

* Generated some more plots, including plots by day of the week, e.g., 

Duncan:

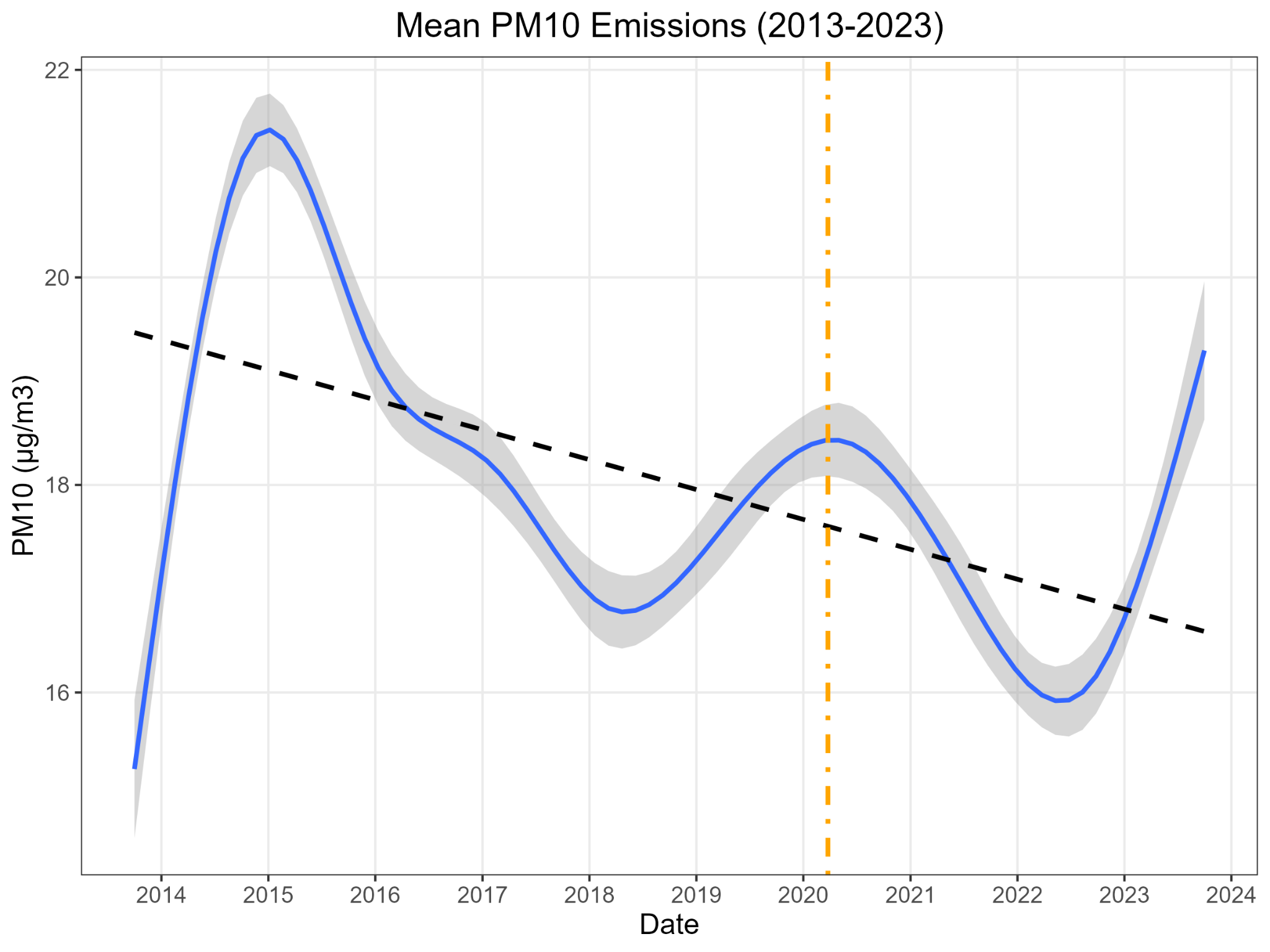
* Started making a map of site locations

Jemma

* Found data correlating to nz environmental standards for 5 major types of emissions: CO, NO2, O3, PM10, SO2 : <https://www.legislation.govt.nz/regulation/public/2004/0309/latest/DLM287036.html>

**Thursday 12th October 2023:**

Allie

* Created nicer plots across time. Since the different time series across locations made it a bit incomprehensible, I looked at the means over time. These included vertical lines for the start of lockdown (25th March 2020). For example:
* After covid (as given by the orange line), it looks like PM10 drops, NO2 spikes.
* Added a new document to the drive for a general report outline. Added the plots by section.
* Added the ECan hourly data combination code to a jupyterlab file and added to the GitHub.