Creation of Sample ID Tables

The final versions of these tables are what was used to make the microplastic datasets that were input into Netica. Below I will explain how these tables were created:

Code is in R Project called ***3. Creation of MP Summary Tables.Rproj*** in the R Script called ***Summary Table Code.Rmd***

1. Particle data (**SFEI.ID.particles.byRiskRegion.shp**) was uploaded into the script from the Input file within the project. This datafile was originally created in **2. Add Location Data to Microparticle Datafile.Rproj** where the SFEI particle data was cut down to only include samples in our study site.
2. Then I created objects from the main particle dataset based on what particles were characterized as

* For Example, I create a TWP only object by filter only for Rubber or Potentially Rubber Particles

A screen shot of a computer

Description automatically generated

* I then subset by each sample matrix

1. Then I made a summary table based on Sample ID. See code below:

A picture containing text, screenshot, font

Description automatically generated

1. Once the sample ID tables were created in R, I filled in more info based on the SFEI report (<https://www.sfei.org/sites/default/files/biblio_files/Microplastic%20Levels%20in%20SF%20Bay%20-%20Final%20Report.pdf> ). See the below description to see more how each was created

**Stormwater\_SampleIDtable**

* Table was created using table 2.1 in chapter 2 of the SFEI Report and sample IDs found in SW data found [HERE](https://github.com/NSF-Microplastics-Project/ClipToProjectBoundary/blob/main/Data/SFEI_MicroparticleData_2020SEP08_sw.csv). Total particle count was taken from Table A-2.1 in the SFEI report
* These data do not have blanks included
* Stormwater data were collected during rainfall events that were expected to exceed 1.3 cm in 6 hours. The particles per liter is aggregated data of samples collected throughout a storm
* There was a field duplicate taken at Lin12M at Coliseum Way
  + - Sample IDs: *19-POC-198-125&355* and *19-POC-199-125&355*
    - In the SFEI report, this was dealt with by averaging the duplicates.
    - The # of liters sieved for this site was reported to be 68 liters (SFEI report; Chapter 2, Table 2.1). It is not clear if this is an average of the 2 samples.

**Sediment\_SampleIDtable**

* Did not include sample information from site LSB02 and CB15 because samples were not analyzed
* There were two field duplicates take from two different sites
  + Station Code: *SPB15*
    - Sample IDs: *17MMP-S-SPB15-MP-1* and *17MMP-S-SPB15-MP-2*
  + Station Code: *SOSL16*
    - Sample IDs: *17MMP-S-SOSL16-MP-1* and *17MMP-S-SOSL16-MP-2*
* Field duplicates were left as separate observations and not average, etc. with other observations
* Samples that did not have any particles analyzed for composition were not included
  + Station Code: *CB15* and *LSB02*
  + Sample ID: *15RMPMC-CB15-MP1* and *17MMP-S-LSB02-MP-1*
* Samples outside of the study area were removed
  + Station Code: *TB102* and *LSB02*
  + Sample ID: 17MMP-S-TB102-MP 17MMP-S-TB101-MP

**Effluent\_SampleIDtable**

* For total liters collected for each sample ID data was take from table 3.2 in the SFEI report for the column titled “Sampled Volume (L)”
  + Liters: total volume passed through the sampleing sieves to calculate the concentration of the microparticles per liter (microparticles/L)
  + Flow was measured using a flow meter except at Sunnyvale, where an ISCO sampler was used and flow was calculated using a container of known volume and a stopwatch.
* Field duplicate was left as separate observations and not average, etc. with other observations
* Field duplicate Sample ID: *PA\_A\_FDUP*
* Each observation was kept as a separate observation even if they came from the same facility
* 2 EBDA represents the aggregated effluent from six wastewater treatment plants that is discharged to the Bay (i.e., San Leandro, Oro Loma, Hayward, Union, Dublin San Ramon, and Livermore).
* Could not confirm discharge point for this facility (Station Code: *EBMUDWWTP*) because the discharge location was not clear in the information related to this facility. Discharge risk region was assumed based on the location of the facility address and NPDES permit information that mentioned EBMUD
* Could not confirm discharge point for this facility (Station Code: SFPUCWWTP-SE) because the NPDES permit information of this facility had expired. Discharge risk region was assumed based on the location of the facility address and NPDES permit information for the expired permit

**SurfaceWater\_SampleIDtable**

* Dry season sampling for the Bay and marine sanctuaries occurred between **August 21, 2017, and November 5, 2017**. Wet season sampling occurred between **November 16, 2017, and March 31, 2018**, following storm events.
  + A wet weather event was defined as 1.3 cm of rainfall within 24 hours, and such an event triggered field sampling. With one exception, all wet weather events met these criteria.
  + At sites SB10, SB11, SB12, and SB13, approximately 1.6 cm of rainfall occurred over multiple days March 13–17, 2018 (i.e., not within a 24-hour period). This event was sampled because the probability of another storm event occurring was low.
  + This whole bullet point was taken word for word from SFEI report (page 118)