**Samples\_Merged Clean Data Formatting**

**DOI:** digital object identifier of manuscript being referenced

**Sample\_ID:** unique key for sample being referenced

**Subsample\_ID:** unique key for subsample being referenced, if applicable

Combination of DOI, Sample\_ID, and Subsample\_ID must be unique for each row. If you have duplicate concentrations for one sample causing an error, assign unique Subsample\_ID to each concentration

**Location:** most precise data possible for location where sample was collected

**Source:** source of water in sample, either “tap water” or “bottled water”

**Date:** date sample was collected, if applicable

Date may be in the format of month/day/year, or month/year if exact date is not provided

**Concentration:** concentration of microplastics in sample

**Concentration\_Units:** units used in corresponding concentration

May be in mass/mass, count/mass, mass/volume, or count/volume

**Color\_ :** followed by color being referenced, column may include the proportion found in sample (numeric value between 0 and 1), marked “Present” if found in sample but manuscript did not specify proportion, or N/A if not found in sample

**Material\_ :** followed by material being referenced, column may include the proportion found in sample (numeric value between 0 and 1), marked “Present” if found in sample but manuscript did not specify proportion, or N/A if not found in sample

**Morphology\_ :** followed by morphology being referenced, column may include the proportion found in sample (numeric value between 0 and 1), marked “Present” if found in sample but manuscript did not specify proportion, or N/A if not found in sample

**Size\_ :** followed by size being referenced, column may include the proportion found in sample (numeric value between 0 and 1), marked “Present” if found in sample but manuscript did not specify proportion, or N/A if not found in sample