

Example Data Vignette

#Example data This data comes from our wastewater and case data merged together by site and date. Then filtered with by 3 sites; Green Bay, Janesville, and Mauston. Then a new column is added; geo_mean which is the geometric mean of N1 and N2. This gives roughly 500 rows to work with for examples across different sized sites over a large time frame.

```
library(DSIWastewater)
library(ggplot2)
```

```
data(Example_data, package = "DSIWastewater")
```

```
## Warning in data(Example_data, package = "DSIWastewater"): data set
## 'Example_data' not found
```

```
head(Example_data)
```

```
##      date      site tests prob_case conf_case prob_death conf_death
## 1 2020-08-23 Green Bay   365         0         2         0         0
## 2 2020-08-24 Green Bay   142         0        18         0         0
## 3 2020-08-30 Green Bay   290         0         0         0         0
## 4 2020-08-31 Green Bay   194         2        37         0         0
## 5 2020-09-06 Green Bay   187         0        13         0         0
## 6 2020-09-07 Green Bay   318         0         0         0         0
##      regions  county   pop      N1      N2 PMMoV hf183   flow tss  ph
## 1 Northeastern Green Bay 189000 7449.6 4464.0   NA   NA 18.363 300 7.6
## 2 Northeastern Green Bay 189000 7891.2 4732.8   NA   NA 22.004 398 7.1
## 3 Northeastern Green Bay 189000 5023.2 4586.4   NA   NA 19.903 272 7.3
## 4 Northeastern Green Bay 189000 8743.2 4524.0   NA   NA 22.663 186 7.4
## 5 Northeastern Green Bay 189000 8409.6 2736.0   NA   NA 16.496 408 7.1
## 6 Northeastern Green Bay 189000 12072.0 6146.4   NA   NA 17.660 534 7.5
##      conductivity temperature      epaid sample_id      pcr_type
## 1             NA             NA WI0065251      260 droplet digital PCR
## 2             NA             NA WI0065251      261 droplet digital PCR
## 3             NA             NA WI0065251      276 droplet digital PCR
## 4             NA             NA WI0065251      277 droplet digital PCR
## 5             NA             NA WI0065251      284 droplet digital PCR
## 6             NA             NA WI0065251      285 droplet digital PCR
##      n1_sars_cov2_units n1_sars_cov2_lod n2_sars_cov2_units n2_sars_cov2_lod bod
## 1 copies/L wastewater      FALSE copies/L wastewater      FALSE 210
## 2 copies/L wastewater      FALSE copies/L wastewater      FALSE 310
## 3 copies/L wastewater      FALSE copies/L wastewater      FALSE 230
## 4 copies/L wastewater      FALSE copies/L wastewater      FALSE 210
## 5 copies/L wastewater      FALSE copies/L wastewater      FALSE 210
## 6 copies/L wastewater      FALSE copies/L wastewater      FALSE 240
##      do      created_on last_modified_on lab_submitter n1_lod n1_loq n2_lod n2_loq
```

```

## 1 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2236.8 7449.6 3724.8 6700.8
## 2 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2371.2 7891.2 3945.6 7104.0
## 3 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2337.6 7792.8 3895.2 7015.2
## 4 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2325.6 7764.0 3878.4 6986.4
## 5 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2068.8 6897.6 3448.8 6208.8
## 6 NA 12/4/2020 17:38 1/7/2022 13:55 UWM 2128.8 7089.6 3547.2 6384.0
## bcov_rec_rate bcov_spike_conc zipcode state capacity_mgd sample_type
## 1 NA NA NA <NA> NA <NA>
## 2 NA NA NA <NA> NA <NA>
## 3 NA NA NA <NA> NA <NA>
## 4 NA NA NA <NA> NA <NA>
## 5 NA NA NA <NA> NA <NA>
## 6 NA NA NA <NA> NA <NA>
## composite_freq sample_matrix sample_location sample_location_specify
## 1 <NA> <NA> <NA> <NA>
## 2 <NA> <NA> <NA> <NA>
## 3 <NA> <NA> <NA> <NA>
## 4 <NA> <NA> <NA> <NA>
## 5 <NA> <NA> <NA> <NA>
## 6 <NA> <NA> <NA> <NA>
## wwtp_comments concentration_method extraction_method lod_ref quant_stan_type
## 1 <NA> <NA> <NA> <NA> <NA>
## 2 <NA> <NA> <NA> <NA> <NA>
## 3 <NA> <NA> <NA> <NA> <NA>
## 4 <NA> <NA> <NA> <NA> <NA>
## 5 <NA> <NA> <NA> <NA> <NA>
## 6 <NA> <NA> <NA> <NA> <NA>
## quant_stan_ref inhibition_method n1_sars_cov2_error n1_ntc_amplify
## 1 <NA> <NA> NA <NA>
## 2 <NA> <NA> NA <NA>
## 3 <NA> <NA> NA <NA>
## 4 <NA> <NA> NA <NA>
## 5 <NA> <NA> NA <NA>
## 6 <NA> <NA> NA <NA>
## n1_num_ntc_amplify n1_num_no_target_control n2_sars_cov2_error n2_ntc_amplify
## 1 NA NA NA <NA>
## 2 NA NA NA <NA>
## 3 NA NA NA <NA>
## 4 NA NA NA <NA>
## 5 NA NA NA <NA>
## 6 NA NA NA <NA>
## n2_num_ntc_amplify n2_num_no_target_control avg_sars_cov2_conc
## 1 NA NA NA
## 2 NA NA NA
## 3 NA NA NA
## 4 NA NA NA
## 5 NA NA NA
## 6 NA NA NA
## avg_sars_cov2_below_lod inhibition_detect inhibition_adjust
## 1 NA <NA> <NA>
## 2 NA <NA> <NA>
## 3 NA <NA> <NA>
## 4 NA <NA> <NA>
## 5 NA <NA> <NA>

```

```
## 6      NA      <NA>      <NA>
##   analytical_comments sample_collect_time test_result_date equiv_sewage_amt
## 1      <NA>      NA secs      <NA>      NA
## 2      <NA>      NA secs      <NA>      NA
## 3      <NA>      NA secs      <NA>      NA
## 4      <NA>      NA secs      <NA>      NA
## 5      <NA>      NA secs      <NA>      NA
## 6      <NA>      NA secs      <NA>      NA
##   geo_mean
## 1 5766.716
## 2 6111.258
## 3 4799.834
## 4 6289.216
## 5 4796.735
## 6 8613.904
```

```
ggplot(Example_data, aes(x=date, y=geo_mean, color = site)) +
  geom_point()
```

```
## Warning: Removed 24 rows containing missing values ('geom_point()').
```

