DHS downsampling work

Marlin

2022-07-11

```
library(DSIWastewater)
library(zoo)
library(lubridate)
library(dplyr)
library(tidyr)
library(ggplot2)
```

```
source("DownSamplingFuncs.R")
```

```
#Analyse day of week effect with new data
#see if mean changes on day of week
#push off to later
#a <- workset4_data%>%
#  group_by(date)%>%
#  summarise(m = mean(sars_cov2_adj_load_log10))
#1, 4
```

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

Mad_data <- DHSWaste_data%>%
  buildWorkSheet4()%>%
  filter(WWTP == "Madison MSD WWTF")
```

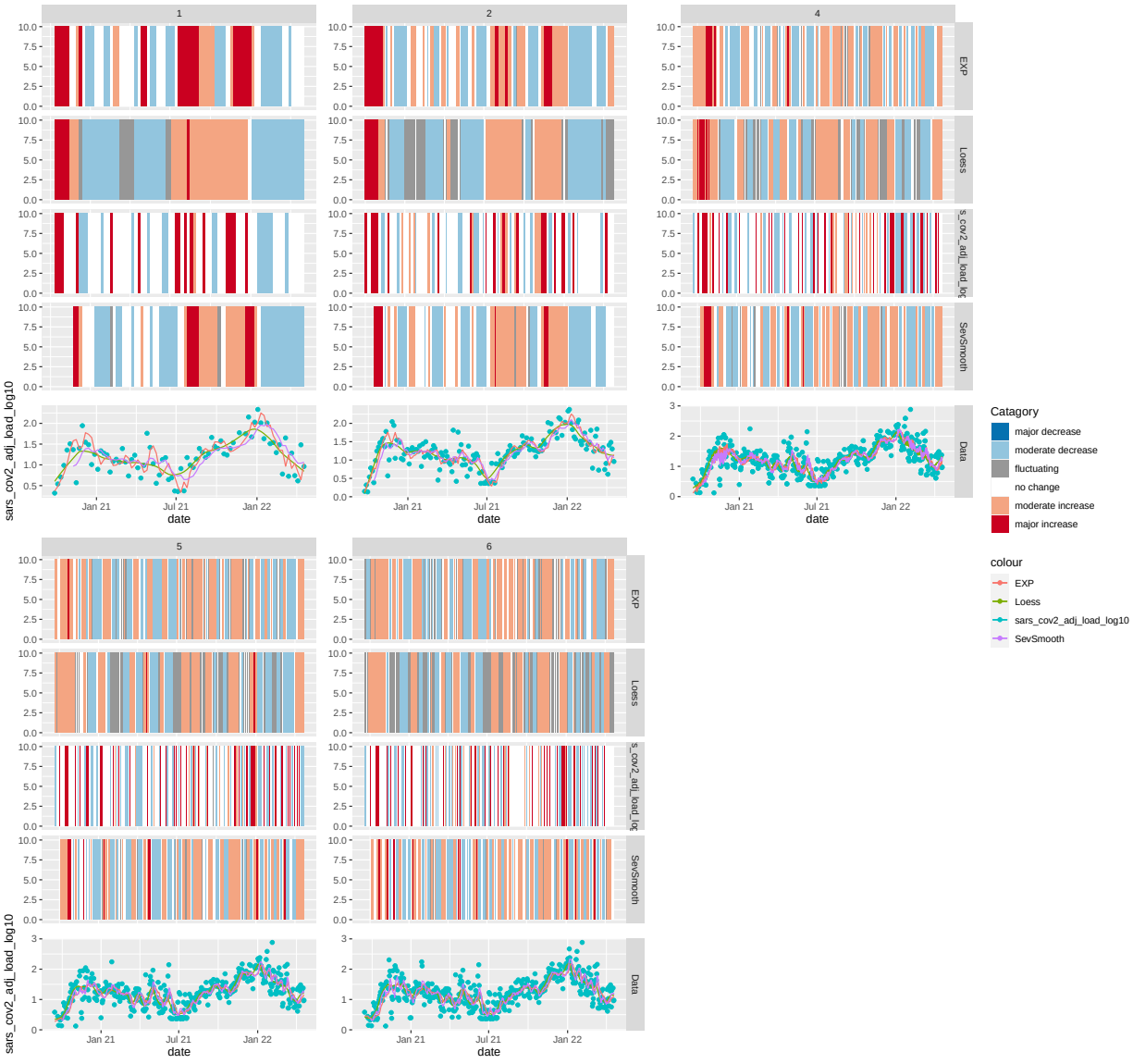
```
Full_Mad_data <- unlist(lapply(1:6,    # Get all combinations
              combinat::combn,
              x = 1:6,
              simplify = FALSE),
       recursive = FALSE)%>%
  lapply(FUN = PrepDataSmoothings,
         DF = Mad_data)%>%
  bind_rows()

Full_reg_data <- Full_Mad_data%>%
                          buildRegressionEstimateTable(
                               RunOn = c("sars_cov2_adj_load_log10",
                                "SevSmooth",
                                "EXP",
                                "Loess"),
                               SplitOn = "TrueName")%>%
  mutate(data = nchar(TrueName))
```

```
sliceVec <- c("2","25","1235","1256","12356","123456")
Singular_Mad_data <- Full_Mad_data%>%
  filter(TrueName %in% sliceVec)

Singular_reg_data <- Full_reg_data%>%
  filter(TrueName %in% sliceVec)
```

```
createDHSMethod_Plot(Singular_reg_data, Singular_Mad_data,
                     PointVal = c( "sars_cov2_adj_load_log10"),
                     LineVal = c("Loess", "EXP", "SevSmooth"),
                     FacGridFormula = Method ~ data)
```




```
Messure_reg_estimates_data <- Full_reg_data%>%
  prepDataForMessure(BreakOn = "data")
```

```
Messure_data <- Messure_reg_estimates_data%>%
  group_by(Method, data)%>%
  mutate(diff = abs(Catagory - Loess),
         BigDiff = diff >= 2,
         vol = abs(Catagory-lag(Catagory)),
         Bigvol = vol>= 2)%>%
  summarise(diff = mean(diff),
            PerBigDiff = mean(BigDiff),
            vol = mean(vol, na.rm = TRUE),
            PerBigvol = mean(Bigvol, na.rm = TRUE)
            )


Messure_data%>%
  pivot_longer(col = -c(Method, data))%>%
  ggplot(aes(x = data, y = value))+
  geom_line(aes(color = Method))+
  facet_wrap(~name, scales = "free")
```