"OCNMS CTD Data Exploration"

Eleanor (Ella) Crotty

Contents

Setup

```
# Warnings and startup messages suppressed
library(tidyverse)
library(patchwork) # Put plots together
library(scales) # Rescale datetime axes
library(ggrepel)
library(readxl)
library(here) # Project/filepath management
library(maps)
library(RColorBrewer) # Color palettes
library(colorRamps) # Color palettes
```

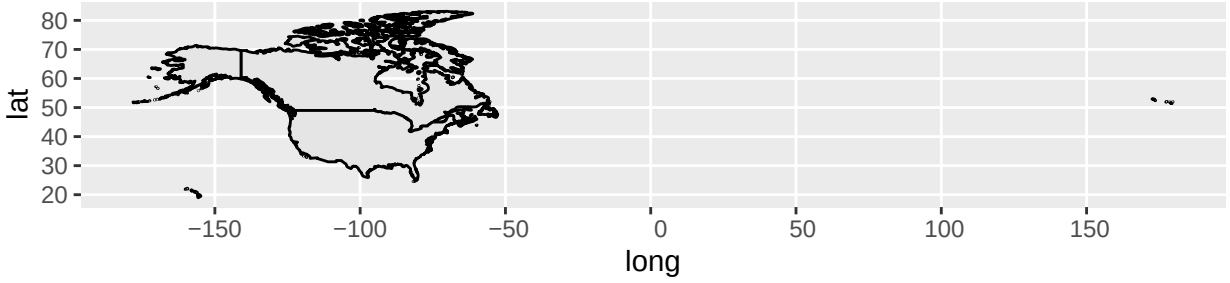
```
wd <- "OCNMS_Hypoxia/CTD_Data"
OME_CTD <- read.csv(here(wd, "OCNMS_OME_ctd_output_copy.csv"))
OCNMS_OME_CTD <- read.csv(here(wd, "OCNMS_OMEsites_ctd_output_copy.csv"))
OCNMS_All_CTD <- read.csv(here(wd, "OCNMS_Allsites_ctd_output_copy.csv"))

# Problem: all the longitudes are positive, they need to be negative
fixlong <- function(df) {
  df$longitude <- df$longitude*-1
  df
}

head(fixlong(OME_CTD))
```

```
##              startTime              filename         cast_info station_id
## 1 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
## 2 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
## 3 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
## 4 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
## 5 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
## 6 2021-10-18T18:21:17Z CAST_CE042_211018.cnv CAST_CE042_211018      CE042
```

```
##         date    timeJ pressure  depth temperature conductivity salinity
## 1 2021-10-18 291.7648    -4.39 -4.353     12.2923     0.000788   0.0095
## 2 2021-10-18 291.7648    -4.39 -4.353     12.3163     0.000788   0.0095
## 3 2021-10-18 291.7648    -4.12 -4.086     12.3431     0.000788   0.0095
## 4 2021-10-18 291.7648    -4.12 -4.086     12.3713     0.000788   0.0095
## 5 2021-10-18 291.7648    -4.12 -4.086     12.3996     0.000788   0.0095
## 6 2021-10-18 291.7648    -4.12 -4.086     12.4278     0.000788   0.0095
##   soundSpeed   oxygen fluorescence beamAttenuation beamTransmission descentRate
## 1    1456.07 -2.85347            0          55.262          -1.2928  1.1369e-15
## 2    1456.16 -2.85191            0          55.262          -1.2928 -1.7760e-16
## 3    1456.26 -2.85026            0          55.262          -1.2928  1.0800e-01
## 4    1456.37 -2.84843            0          55.262          -1.2928  1.6200e-01
## 5    1456.48 -2.84659            0          55.262          -1.2928  1.6200e-01
## 6    1456.58 -2.84477            0          55.262          -1.2928  1.0800e-01
##   bpos scan flag   station_name latitude longitude bottom_depth_m
## 1    0    2    0 Cape Elizabeth  47.3531 -124.4887             42
## 2    0    3    0 Cape Elizabeth  47.3531 -124.4887             42
## 3    0    4    0 Cape Elizabeth  47.3531 -124.4887             42
## 4    0    5    0 Cape Elizabeth  47.3531 -124.4887             42
## 5    0    6    0 Cape Elizabeth  47.3531 -124.4887             42
## 6    0    7    0 Cape Elizabeth  47.3531 -124.4887             42
```

```
OME_CTD <- fixlong(OME_CTD)
OCNMS_OME_CTD <- fixlong(OCNMS_OME_CTD)
OCNMS_All_CTD <- fixlong(OCNMS_All_CTD)
# All better!
```

```
mapUC <- map_data("world", region = c("usa", "canada"))

ggplot(mapUC, aes(x = long, y = lat, group = group)) +
  geom_polygon(fill = "gray90", color = "black") +
  coord_sf() # coord_quickmap is an approximation to preserve straight lines, which works
  ↪  best for small areas close to the equator. projection can be defined (see
  ↪  mapproj::mapproject() for list) and R now recommends using coord_sf(). coord_sf()
  ↪  takes xlim, ylim, crs
```



Some excerpts from the R for Data Science tutorial

```
nz <- map_data("nz")

ggplot(nz, aes(x = long, y = lat, group = group)) +
  geom_polygon(fill = "white", color = "black") +
  coord_quickmap() # This will fix the weird stretch usually

# Making a ggplot with label changes
```

```r
histogram <- function(df, var, binwidth) {
  label <- rlang::englue("A histogram of {{var}} with binwidth {binwidth}")

  df |>
    ggplot(aes(x = {{ var }})) +
    geom_histogram(binwidth = binwidth) +
    labs(title = label)
}

diamonds |> histogram(carat, 0.1)
diamonds |> histogram(price, 1000)
```

```r
df <- tribble(
  ~id, ~measurement, ~value,
  "A",        "bp1",    100,
  "B",        "bp1",    140,
  "B",        "bp2",    115,
  "A",        "bp2",    120,
  "A",        "bp3",    105
)
```

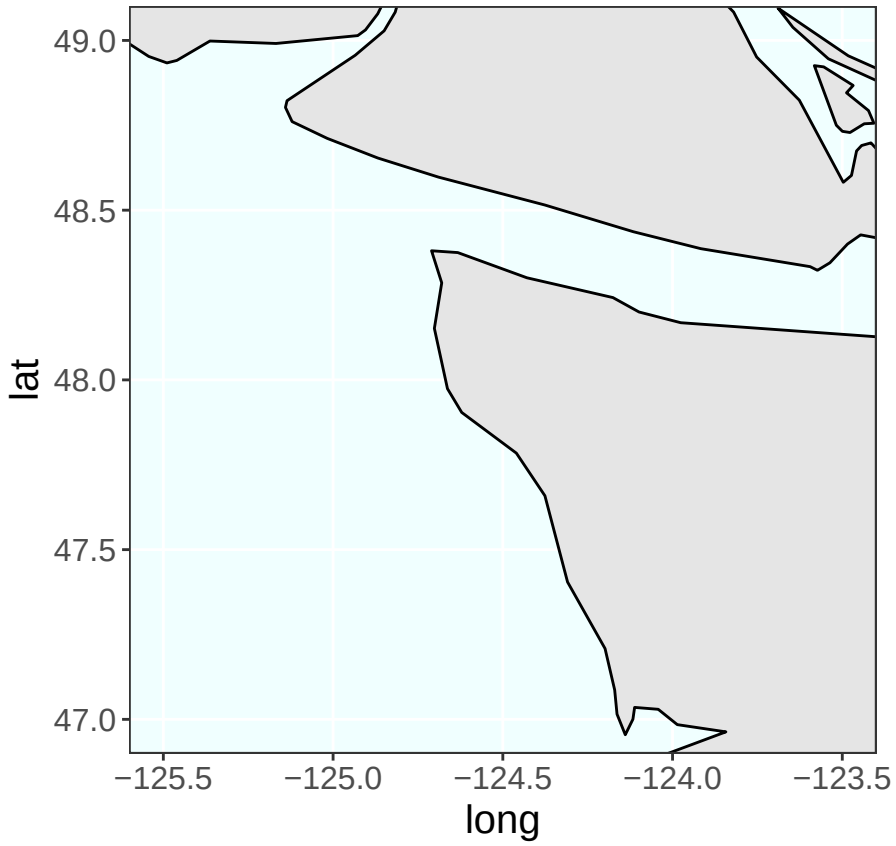
Functions

```r
# Variable to make a base map of OCNMS, which I can then add data points to

rangeOC <- tribble(
  ~MinLong, ~MaxLong, ~MinLat, ~MaxLat,
  min(OME_CTD$longitude), max(OME_CTD$longitude), min(OME_CTD$latitude),
  ↪  max(OME_CTD$latitude)
)

OCNMS_x <- c(-123.5, -125.5)
OCNMS_y <- c(47,49)

mapOCNMS <- ggplot(mapUC, aes(x = long, y = lat, group = group)) +
  geom_polygon(fill = "gray90", color = "black") +
  coord_sf(xlim = OCNMS_x,
           ylim = OCNMS_y) +
  theme_bw() +
  theme(text = element_text(size=15),
        panel.background = element_rect(fill = "azure1",
                                        colour = "azure1"),
        legend.key = element_rect(fill = "white",
                                  color = "white"),
        # It added in blue behind the dots in the key and I don't want that
        panel.grid.major = element_line(size = 0.5,
                                        linetype = 'solid',
                                        colour = "white"),
        panel.grid.minor = element_line(size = 0.25,
                                        linetype = 'solid',
                                        colour = "white")) # +
```

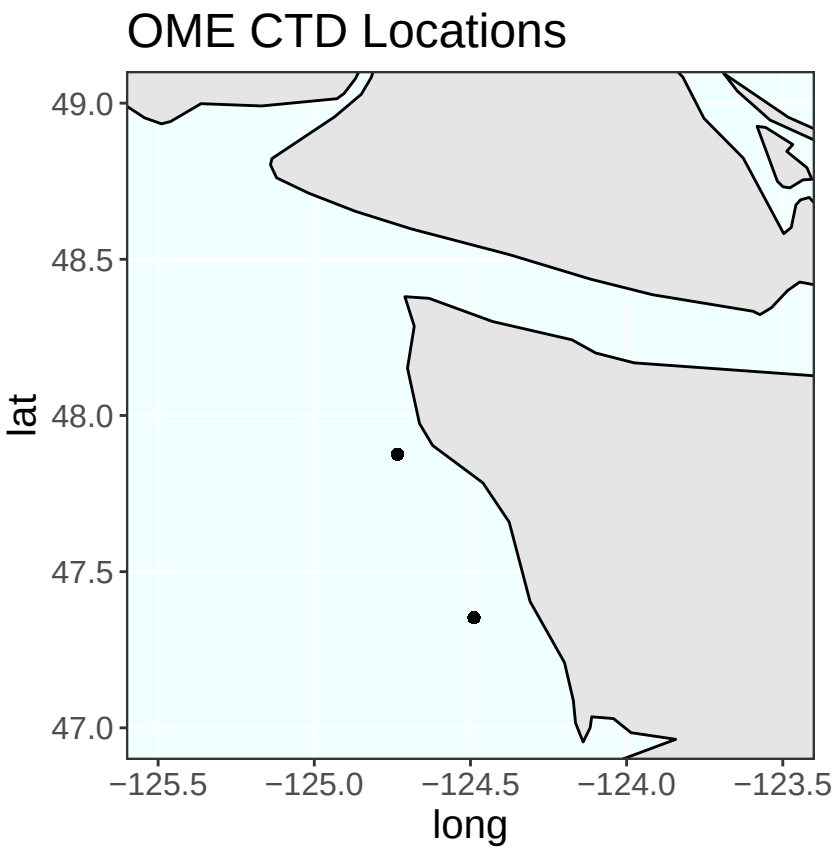
```
## Warning: The `size` argument of `element_line()` is deprecated as of ggplot2 3.4.0.
```

```
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

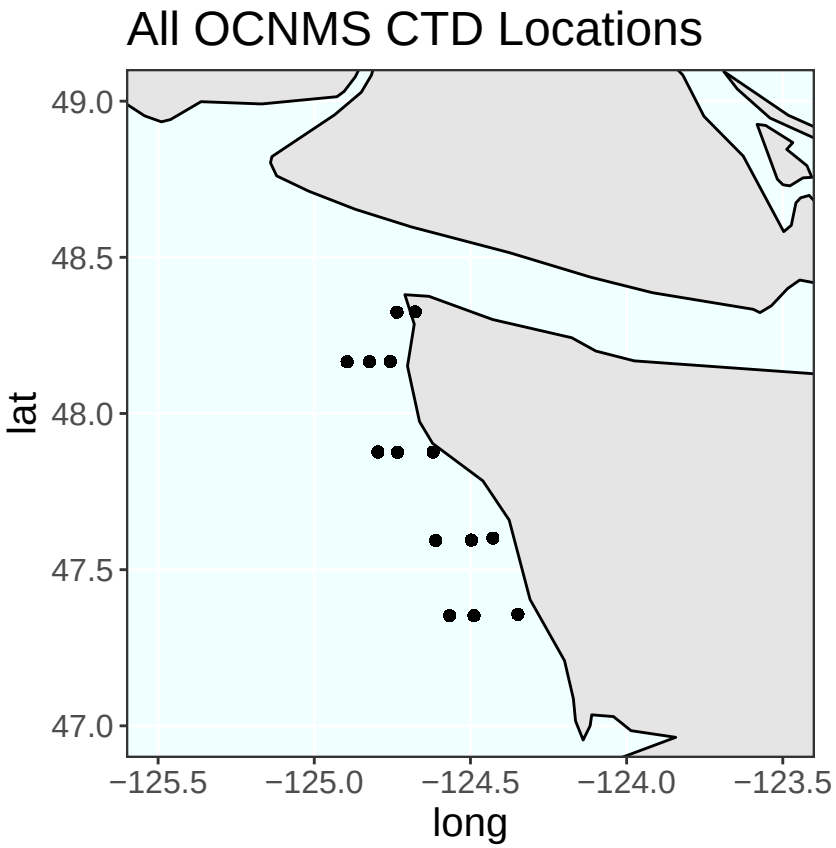
```
  # coord_quickmap() # This will fix the weird stretch usually, can also do
  ↪  coord_fixed(ratio = 1.3)

mapOCNMS
```




```
mapOCNMS +
  geom_point(data = OME_CTD, aes(x = longitude, y = latitude, group = NA)) + # aes(group
  ↪  = NA) or geom_point(inherit = FALSE) keeps it from looking for a group since the
  ↪  base layer has groups
  ggtitle("OME CTD Locations")
```




```
mapOCNMS +
  geom_point(data = OCNMS_All_CTD, aes(x = longitude, y = latitude, group = NA)) + #
  ↪  aes(group = NA) or geom_point(inherit = FALSE) keeps it from looking for a group
  ↪  since the base layer has groups
  ggtitle("All OCNMS CTD Locations")
```

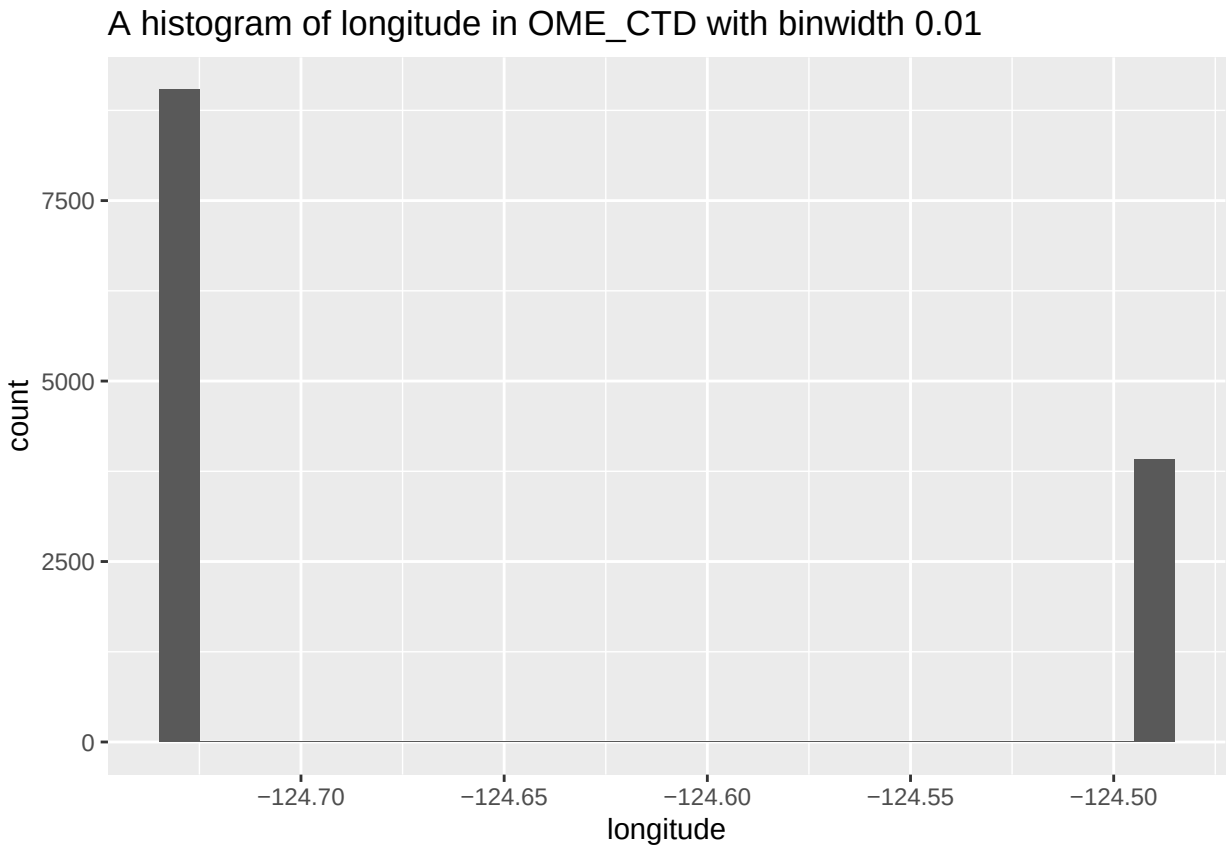



```
# Histogram function
histogram <- function(df, var, binwidth) {
  label <- rlang::englue("A histogram of {{var}} in {{df}} with binwidth {binwidth}")

  df |>
    ggplot(aes(x = {{ var }})) +
    geom_histogram(binwidth = binwidth) +
    labs(title = label)
}
```

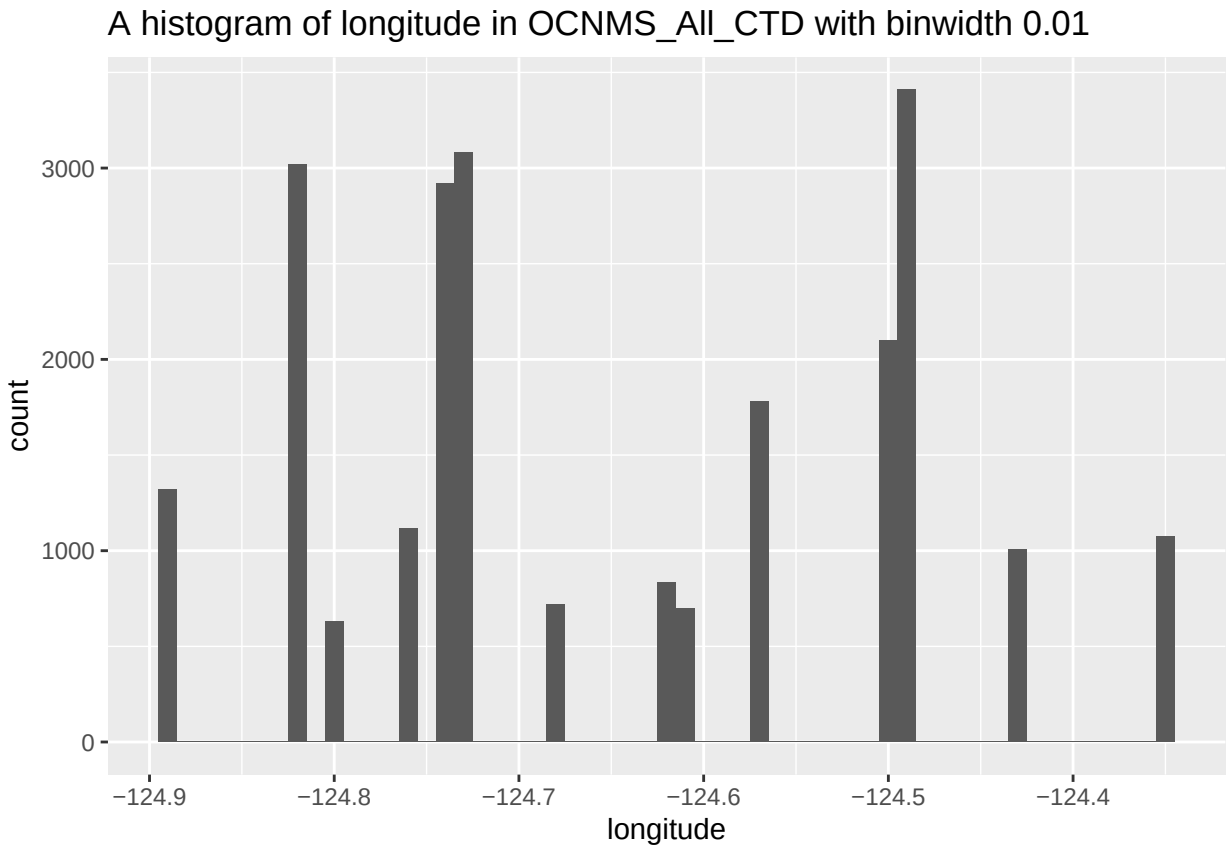
Oxygen Data Exploration

Histograms

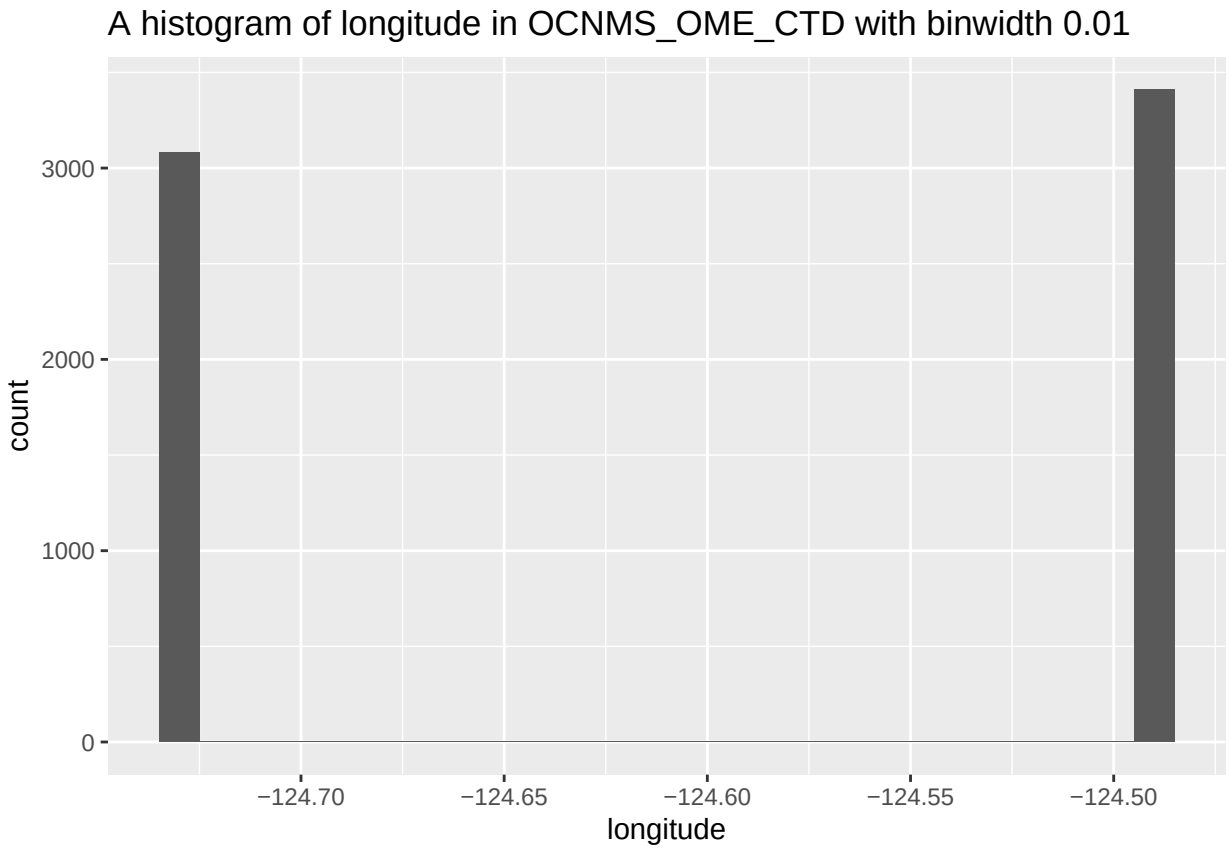
```
histogram(OME_CTD, longitude, 0.01)
```




```
histogram(OCNMS_All_CTD, longitude, 0.01)
```




```
histogram(OCNMS_OME_CTD, longitude, 0.01)
```


A histogram of longitude in OCNMS_OME_CTD with binwidth 0.01
count
3000
2000
1000
0
−124.70
−124.65
−124.60
−124.55
−124.50
longitude