R Code – 2022 NASA International Space Apps Challenge

> nasa <- read.csv("Dropbox/data\_science/projects/nasa\_space\_apps\_2022/data/WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2\_df.csv", head = TRUE)

> nasa\_data <- nasa %>% select(-OA\_prec\_PM1\_AMS, -OA\_DL\_PM1\_AMS, -Sulfate\_prec\_PM1\_AMS, -Sulfate\_DL\_PM1\_AMS, -Nitrate\_prec\_PM1\_AMS, -Nitrate\_DL\_PM1\_AMS, -Ammonium\_prec\_PM1\_AMS, -Ammonium\_DL\_PM1\_AMS, -NR\_Chloride\_prec\_PM1\_AMS, -NR\_Chloride\_DL\_PM1\_AMS, -Potassium\_prec\_PM1\_AMS, -Potassium\_DL\_PM1\_AMS, -MSA\_prec\_PM1\_AMS, -MSA\_DL\_PM1\_AMS, -ClO4\_prec\_PM1\_AMS, -ClO4\_DL\_PM1\_AMS, -Iodine\_prec\_PM1\_AMS, -Iodine\_DL\_PM1\_AMS, -Bromine\_prec\_PM1\_AMS, -Bromine\_DL\_PM1\_AMS, -Seasalt\_prec\_PM1\_AMS, -Seasalt\_DL\_PM1\_AMS, -SDDataFlag\_AMS, -CloudFlag\_AMS, -Flowrate\_AMS, -StdtoVol\_AMS, -InletRH\_AMS)

> PM1\_AMS <- c("OA\_PM1\_AMS", "Sulfate\_PM1\_AMS", "Nitrate\_PM1\_AMS", "Ammonium\_PM1\_AMS", "NR\_Chloride\_PM1\_AMS", "Potassium\_PM1\_AMS", "MSA\_PM1\_AMS", "ClO4\_PM1\_AMS", "Iodine\_PM1\_AMS", "Bromine\_PM1\_AMS", "Seasalt\_PM1\_AMS", "AmmBalance\_PM1\_AMS", "OtoC\_Ratio\_PM1\_AMS", "HtoC\_Ratio\_PM1\_AMS", "OAtoOC\_PM1\_AMS", "OSc\_PM1\_AMS", "OrgNitrFraction\_PM1\_AMS", "f43\_PM1\_AMS", "f44\_PM1\_AMS", "f57\_PM1\_AMS", "f60\_PM1\_AMS", "f82\_PM1\_AMS", "f91\_PM1\_AMS", "fC2H3O\_PM1\_AMS", "fCO2\_PM1\_AMS", "fC4H9\_PM1\_AMS", "fC2H4O2\_PM1\_AMS", "fC5H6O\_PM1\_AMS", "fC7H7\_PM1\_AMS")

//> nasa\_data <- nasa\_data %>% filter(ALT\_AMS != -9999 & Density\_PM1\_AMS != -8888 & OADensity\_PM1\_AMS != -8888)

# maybe use this instead of the above command to filter across all columns

# This may have worked, but I forgot the negative signs and ended up trying something else

//> nasa\_data <- nasa\_data %>% filter(across(everything(), ~ . != 9999)) %>% filter(across(everything(), ~ . != 8888))

# Ended up trying this and it worked also:

> nasa\_data %>% summarise(across(where(is.numeric), ~ min(.x)))

Time\_Start Time\_Stop Time\_Mid LAT\_AMS LON\_AMS ALT\_AMS OA\_PM1\_AMS Sulfate\_PM1\_AMS Nitrate\_PM1\_AMS Ammonium\_PM1\_AMS NR\_Chloride\_PM1\_AMS Potassium\_PM1\_AMS

1 66009.87 66010.87 66010.37 -9999 -9999 279.2 0.9745 -0.088 -0.0789 -0.0093 -0.093 -12.6691

MSA\_PM1\_AMS ClO4\_PM1\_AMS Iodine\_PM1\_AMS Bromine\_PM1\_AMS Seasalt\_PM1\_AMS AmmBalance\_PM1\_AMS Density\_PM1\_AMS OADensity\_PM1\_AMS OtoC\_Ratio\_PM1\_AMS

1 -8888 -8888 -8888 -8888 -8888 -8888 1.114 1.019 -8888

HtoC\_Ratio\_PM1\_AMS OAtoOC\_PM1\_AMS OSc\_PM1\_AMS OrgNitrFraction\_PM1\_AMS f43\_PM1\_AMS f44\_PM1\_AMS f57\_PM1\_AMS f60\_PM1\_AMS f82\_PM1\_AMS f91\_PM1\_AMS

1 -8888 -8888 -8888 -9999 -8888 -8888 -8888 -8888 -8888 -8888

fC2H3O\_PM1\_AMS fCO2\_PM1\_AMS fC4H9\_PM1\_AMS fC2H4O2\_PM1\_AMS fC5H6O\_PM1\_AMS fC7H7\_PM1\_AMS

1 -8888 -8888 -8888 -8888 -8888 -8888

> nasa\_data <- nasa\_data %>% filter(across(where(is.numeric), ~ .x != -8888))

Warning message:

Using `across()` in `filter()` is deprecated, use `if\_any()` or `if\_all()`.

> nasa\_data %>% summarise(across(where(is.numeric), ~ min(.x)))

Time\_Start Time\_Stop Time\_Mid LAT\_AMS LON\_AMS ALT\_AMS OA\_PM1\_AMS Sulfate\_PM1\_AMS Nitrate\_PM1\_AMS Ammonium\_PM1\_AMS NR\_Chloride\_PM1\_AMS Potassium\_PM1\_AMS

1 66059.42 66060.42 66059.92 -9999 -9999 279.2 1.4755 0.1343 -0.0789 0.0589 -0.093 -12.6691

MSA\_PM1\_AMS ClO4\_PM1\_AMS Iodine\_PM1\_AMS Bromine\_PM1\_AMS Seasalt\_PM1\_AMS AmmBalance\_PM1\_AMS Density\_PM1\_AMS OADensity\_PM1\_AMS OtoC\_Ratio\_PM1\_AMS

1 -0.2794 -0.1554 -0.0789 -0.0908 -1.1968 0.148 1.217 1.078 0.329

HtoC\_Ratio\_PM1\_AMS OAtoOC\_PM1\_AMS OSc\_PM1\_AMS OrgNitrFraction\_PM1\_AMS f43\_PM1\_AMS f44\_PM1\_AMS f57\_PM1\_AMS f60\_PM1\_AMS f82\_PM1\_AMS f91\_PM1\_AMS

1 1.134 1.641 -1.221 -9999 0.001 0.0189 -0.0287 -0.0176 -0.0094 -0.0162

fC2H3O\_PM1\_AMS fCO2\_PM1\_AMS fC4H9\_PM1\_AMS fC2H4O2\_PM1\_AMS fC5H6O\_PM1\_AMS fC7H7\_PM1\_AMS

1 -0.0014 0.0127 -0.0373 -0.0057 -0.0054 -0.0094

> nasa\_data <- nasa\_data %>% filter(across(where(is.numeric), ~ .x != -9999))

Warning message:

Using `across()` in `filter()` is deprecated, use `if\_any()` or `if\_all()`.

s> nasa\_data %>% summarise(across(where(is.numeric), ~ min(.x)))

Time\_Start Time\_Stop Time\_Mid LAT\_AMS LON\_AMS ALT\_AMS OA\_PM1\_AMS Sulfate\_PM1\_AMS Nitrate\_PM1\_AMS Ammonium\_PM1\_AMS NR\_Chloride\_PM1\_AMS Potassium\_PM1\_AMS

1 66229.25 66230.25 66229.75 33.5997 -121.5409 279.2 1.4755 0.1343 -0.0789 0.0658 -0.093 -12.6691

MSA\_PM1\_AMS ClO4\_PM1\_AMS Iodine\_PM1\_AMS Bromine\_PM1\_AMS Seasalt\_PM1\_AMS AmmBalance\_PM1\_AMS Density\_PM1\_AMS OADensity\_PM1\_AMS OtoC\_Ratio\_PM1\_AMS

1 -0.2794 -0.1554 -0.0485 -0.0908 -1.1968 0.148 1.217 1.078 0.329

HtoC\_Ratio\_PM1\_AMS OAtoOC\_PM1\_AMS OSc\_PM1\_AMS OrgNitrFraction\_PM1\_AMS f43\_PM1\_AMS f44\_PM1\_AMS f57\_PM1\_AMS f60\_PM1\_AMS f82\_PM1\_AMS f91\_PM1\_AMS

1 1.134 1.641 -1.221 0 0.001 0.0189 -0.0287 -0.0176 -0.0094 -0.0162

fC2H3O\_PM1\_AMS fCO2\_PM1\_AMS fC4H9\_PM1\_AMS fC2H4O2\_PM1\_AMS fC5H6O\_PM1\_AMS fC7H7\_PM1\_AMS

1 0.0077 0.0127 -0.0373 -0.0057 -0.0054 -0.0039

> nasa\_data\_gather <- nasa\_data %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS)) %>% filter(measurement != -8888 & measurement != -9999)

> nasa\_data\_gather %>% ggplot(aes(x = ALT\_AMS, y = measurement, color = OADensity\_PM1\_AMS)) + facet\_wrap(~ component, scales = "free") + geom\_point(size = 0.1) + theme\_bw() + theme(strip.text.x = element\_text(size = 5), axis.text = element\_text(size = 5), panel.grid = element\_blank()) + labs(title = "FIREX-AQ: Aersols by Altitude (data:WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2)", y = "", x = "Altitude (meters)", color = "Organic Aerosol\nDensity (g/m3)")

> nasa\_data\_per\_OAden <- nasa\_data %>% mutate(across(all\_of(PM1\_AMS), ~ . / OADensity\_PM1\_AMS, .names = "{col}\_per\_OAden"))

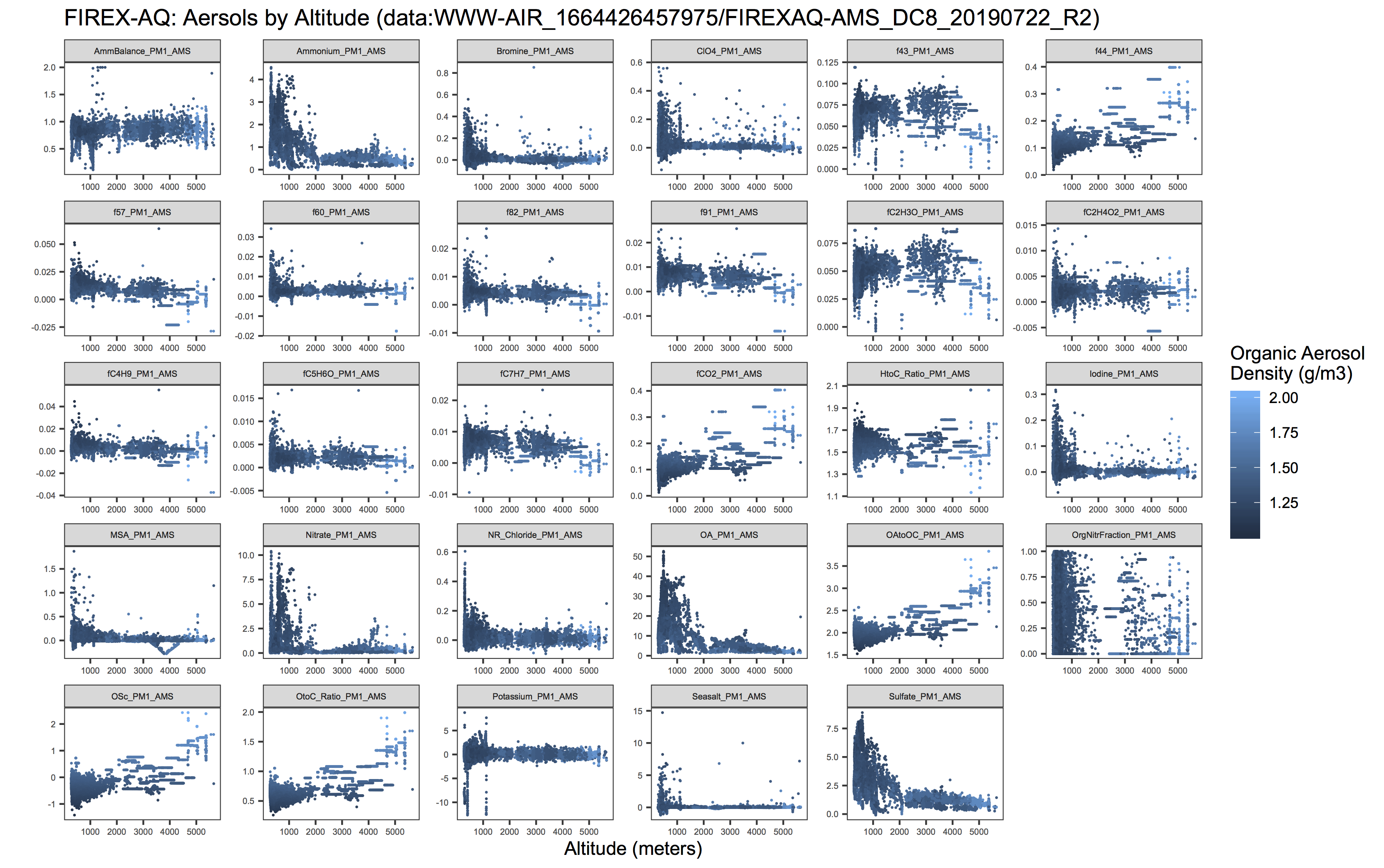
> nasa\_data\_per\_OAden\_gather <- nasa\_data\_per\_OAden %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS)) %>% filter(measurement != -8888 & measurement != -9999)

> nasa\_data <- nasa\_data %>% filter(ALT\_AMS != -9999 & Density\_PM1\_AMS != -8888 & OADensity\_PM1\_AMS != -8888)

> nasa\_data\_gather <- nasa\_data %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS)) %>% filter(measurement != -8888 & measurement != -9999)

//> nasa\_data\_gather <- nasa\_data %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS)) %>% filter(measurement != -8888 & measurement != -9999 & ALT\_AMS != -9999 & Density\_PM1\_AMS != -8888 & OADensity\_PM1\_AMS != -8888)

> nasa\_data\_gather %>% ggplot(aes(x = ALT\_AMS, y = measurement, color = OADensity\_PM1\_AMS)) + facet\_wrap(~ component, scales = "free") + geom\_point(size = 0.1) + theme\_bw() + theme(strip.text.x = element\_text(size = 5), axis.text = element\_text(size = 5), panel.grid = element\_blank()) + labs(title = "FIREX-AQ: Aersols by Altitude (data:WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2)", y = "", x = "Altitude (meters)", color = "Organic Aerosol\nDensity (g/m3)")



> PM1\_AMS\_df <- data.frame(unique(nasa\_data\_gather$component))

> colnames(PM1\_AMS\_df) <- c("components")

> PM1\_AMS\_df

components

1 OA\_PM1\_AMS

2 Sulfate\_PM1\_AMS

3 Nitrate\_PM1\_AMS

4 Ammonium\_PM1\_AMS

5 NR\_Chloride\_PM1\_AMS

6 Potassium\_PM1\_AMS

7 MSA\_PM1\_AMS

8 ClO4\_PM1\_AMS

9 Iodine\_PM1\_AMS

10 Bromine\_PM1\_AMS

11 Seasalt\_PM1\_AMS

12 AmmBalance\_PM1\_AMS

13 OtoC\_Ratio\_PM1\_AMS

14 HtoC\_Ratio\_PM1\_AMS

15 OAtoOC\_PM1\_AMS

16 OSc\_PM1\_AMS

17 OrgNitrFraction\_PM1\_AMS

18 f43\_PM1\_AMS

19 f44\_PM1\_AMS

20 f57\_PM1\_AMS

21 f60\_PM1\_AMS

22 f82\_PM1\_AMS

23 f91\_PM1\_AMS

24 fC2H3O\_PM1\_AMS

25 fCO2\_PM1\_AMS

26 fC4H9\_PM1\_AMS

27 fC2H4O2\_PM1\_AMS

28 fC5H6O\_PM1\_AMS

29 fC7H7\_PM1\_AMS

> nasa\_data\_per\_OADen <- nasa\_data %>% mutate(across(PM1\_AMS\_df, .names = "{col}\_per\_OADen"/OADensity\_PM1\_AMS))

# Repeat the below code for several missions each that targeted wildfires & agri fires

> nasa <- read.csv("Dropbox/data\_science/projects/nasa\_space\_apps\_2022/data/WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2\_df.csv", head = TRUE)

> nasa\_data <- nasa %>% select(-OA\_prec\_PM1\_AMS, -OA\_DL\_PM1\_AMS, -Sulfate\_prec\_PM1\_AMS, -Sulfate\_DL\_PM1\_AMS, -Nitrate\_prec\_sPM1\_AMS, -Nitrate\_DL\_PM1\_AMS, -Ammonium\_prec\_PM1\_AMS, -Ammonium\_DL\_PM1\_AMS, -NR\_Chloride\_prec\_PM1\_AMS, -NR\_Chloride\_DL\_PM1\_AMS, -Potassium\_prec\_PM1\_AMS, -Potassium\_DL\_PM1\_AMS, -MSA\_prec\_PM1\_AMS, -MSA\_DL\_PM1\_AMS, -ClO4\_prec\_PM1\_AMS, -ClO4\_DL\_PM1\_AMS, -Iodine\_prec\_PM1\_AMS, -Iodine\_DL\_PM1\_AMS, -Bromine\_prec\_PM1\_AMS, -Bromine\_DL\_PM1\_AMS, -Seasalt\_prec\_PM1\_AMS, -Seasalt\_DL\_PM1\_AMS, -SDDataFlag\_AMS, -CloudFlag\_AMS, -Flowrate\_AMS, -StdtoVol\_AMS, -InletRH\_AMS)

> PM1\_AMS <- c("OA\_PM1\_AMS", "Sulfate\_PM1\_AMS", "Nitrate\_PM1\_AMS", "Ammonium\_PM1\_AMS", "NR\_Chloride\_PM1\_AMS", "Potassium\_PM1\_AMS", "MSA\_PM1\_AMS", "ClO4\_PM1\_AMS", "Iodine\_PM1\_AMS", "Bromine\_PM1\_AMS", "Seasalt\_PM1\_AMS", "AmmBalance\_PM1\_AMS", "OtoC\_Ratio\_PM1\_AMS", "HtoC\_Ratio\_PM1\_AMS", "OAtoOC\_PM1\_AMS", "OSc\_PM1\_AMS", "OrgNitrFraction\_PM1\_AMS", "f43\_PM1\_AMS", "f44\_PM1\_AMS", "f57\_PM1\_AMS", "f60\_PM1\_AMS", "f82\_PM1\_AMS", "f91\_PM1\_AMS", "fC2H3O\_PM1\_AMS", "fCO2\_PM1\_AMS", "fC4H9\_PM1\_AMS", "fC2H4O2\_PM1\_AMS", "fC5H6O\_PM1\_AMS", "fC7H7\_PM1\_AMS")

> nasa\_data <- nasa\_data %>% filter(across(where(is.numeric), ~ .x != -9999)) %>% filter(across(where(is.numeric), ~ .x != -8888))

> nasa\_data\_gather <- nasa\_data %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS))

> nasa\_data\_per\_OAden <- nasa\_data %>% mutate(across(all\_of(PM1\_AMS), ~ . / OADensity\_PM1\_AMS, .names = "{col}\_per\_OAden"))

> nasa\_data\_per\_OAden\_gather <- nasa\_data\_per\_OAden %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS))

> nasa\_data\_per\_OAden\_gather %>% filter(str\_detect(component, 'per\_OAden')) %>% ggplot(aes(x = ALT\_AMS, y = measurement, color = OADensity\_PM1\_AMS)) + facet\_wrap(~ component, scales = "free") + geom\_point(size = 0.1) + theme\_bw() + theme(strip.text.x = element\_text(size = 5), axis.text = element\_text(size = 5), panel.grid = element\_blank()) + labs(title = "FIREX-AQ: Aersols by Altitude (data:WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2)", y = "Units per Organic Aerosol Density (relative measure)", x = "Altitude (meters)", color = "Organic Aerosol\nDensity (g/m3)")

# Wrote an R script that takes care of almost all the above commands (except first and last)

> nasa <- read.csv("Dropbox/data\_science/projects/nasa\_space\_apps\_2022/data/WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2\_df.csv", head = TRUE)

> source(file="/Users/gaellemuller-greven/Dropbox/data\_science/projects/nasa\_space\_apps\_2022/nasa.R")

> nasa\_data\_per\_OAden\_gather %>% filter(str\_detect(component, 'per\_OAden')) %>% ggplot(aes(x = ALT\_AMS, y = measurement, color = OADensity\_PM1\_AMS)) + facet\_wrap(~ component, scales = "free") + geom\_point(size = 0.1) + theme\_bw() + theme(strip.text.x = element\_text(size = 5), axis.text = element\_text(size = 5), panel.grid = element\_blank()) + labs(title = "FIREX-AQ: Aersols by Altitude (data:WWW-AIR\_1664426457975/FIREXAQ-AMS\_DC8\_20190722\_R2)", y = "Units per Organic Aerosol Density (relative measure)", x = "Altitude (meters)", color = "Organic Aerosol\nDensity (g/m3)")

# R Script (located at Dropbox/data\_science/projects/nasa\_space\_apps\_2022/nasa.R)

filenames = c("FIREXAQ-AMS\_DC8\_20190722\_R2", "FIREXAQ-AMS\_DC8\_20190724\_R2", "FIREXAQ-AMS\_DC8\_20190725\_R2", "FIREXAQ-AMS\_DC8\_20190729\_R3", "FIREXAQ-AMS\_DC8\_20190730\_R2", "FIREXAQ-AMS\_DC8\_20190802\_R3", "FIREXAQ-AMS\_DC8\_20190803\_R2")

filename = ""

plot\_airborne\_data <- function(filename){

filename = filename

filename\_csv = paste(filename, "\_df.csv", sep = "")

filename\_title = paste(filename, ")", sep = "")

nasa <- read.csv(paste("/Users/gaellemuller-greven/Dropbox/data\_science/projects/nasa\_space\_apps\_2022/data/WWW-AIR\_1664426457975/", filename\_csv, sep = ""), head = TRUE)

nasa\_data <- nasa %>% select(-OA\_prec\_PM1\_AMS, -OA\_DL\_PM1\_AMS, -Sulfate\_prec\_PM1\_AMS, -Sulfate\_DL\_PM1\_AMS, -Nitrate\_prec\_PM1\_AMS, -Nitrate\_DL\_PM1\_AMS, -Ammonium\_prec\_PM1\_AMS, -Ammonium\_DL\_PM1\_AMS, -NR\_Chloride\_prec\_PM1\_AMS, -NR\_Chloride\_DL\_PM1\_AMS, -Potassium\_prec\_PM1\_AMS, -Potassium\_DL\_PM1\_AMS, -MSA\_prec\_PM1\_AMS, -MSA\_DL\_PM1\_AMS, -ClO4\_prec\_PM1\_AMS, -ClO4\_DL\_PM1\_AMS, -Iodine\_prec\_PM1\_AMS, -Iodine\_DL\_PM1\_AMS, -Bromine\_prec\_PM1\_AMS, -Bromine\_DL\_PM1\_AMS, -Seasalt\_prec\_PM1\_AMS, -Seasalt\_DL\_PM1\_AMS, -SDDataFlag\_AMS, -CloudFlag\_AMS, -Flowrate\_AMS, -StdtoVol\_AMS, -InletRH\_AMS)

PM1\_AMS <- c("OA\_PM1\_AMS", "Sulfate\_PM1\_AMS", "Nitrate\_PM1\_AMS", "Ammonium\_PM1\_AMS", "NR\_Chloride\_PM1\_AMS", "Potassium\_PM1\_AMS", "MSA\_PM1\_AMS", "ClO4\_PM1\_AMS", "Iodine\_PM1\_AMS", "Bromine\_PM1\_AMS", "Seasalt\_PM1\_AMS", "AmmBalance\_PM1\_AMS", "OtoC\_Ratio\_PM1\_AMS", "HtoC\_Ratio\_PM1\_AMS", "OAtoOC\_PM1\_AMS", "OSc\_PM1\_AMS", "OrgNitrFraction\_PM1\_AMS", "f43\_PM1\_AMS", "f44\_PM1\_AMS", "f57\_PM1\_AMS", "f60\_PM1\_AMS", "f82\_PM1\_AMS", "f91\_PM1\_AMS", "fC2H3O\_PM1\_AMS", "fCO2\_PM1\_AMS", "fC4H9\_PM1\_AMS", "fC2H4O2\_PM1\_AMS", "fC5H6O\_PM1\_AMS", "fC7H7\_PM1\_AMS")

nasa\_data <- nasa\_data %>% filter(across(where(is.numeric), ~ .x != -9999)) %>% filter(across(where(is.numeric), ~ .x != -8888))

nasa\_data\_gather <- nasa\_data %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS))

nasa\_data\_per\_OAden <- nasa\_data %>% mutate(across(all\_of(PM1\_AMS), ~ . / OADensity\_PM1\_AMS, .names = "{col}\_per\_OAden"))

nasa\_data\_per\_OAden\_gather <- nasa\_data\_per\_OAden %>% keep(is.numeric) %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS))

nasa\_data\_per\_OAden\_gather %>% filter(str\_detect(component, 'per\_OAden')) %>% ggplot(aes(x = ALT\_AMS, y = measurement, color = OADensity\_PM1\_AMS)) + facet\_wrap(~ component, scales = "free") + geom\_point(size = 0.1) + theme\_bw() + theme(strip.text.x = element\_text(size = 5), axis.text = element\_text(size = 5), panel.grid = element\_blank()) + labs(title = paste("FIREX-AQ: Aersols by Altitude (data:WWW-AIR\_1664426457975/",filename\_title), y = "Units per Organic Aerosol Density (relative measure)", x = "Altitude (meters)", color = "Organic Aerosol\nDensity (g/m3)")

ggsave(paste(filename,".png", sep = ""), device = "png", dpi = "retina", width = 3000, height = 2000, units = "px", path = "/Users/gaellemuller-greven/Dropbox/data\_science/projects/nasa\_space\_apps\_2022/plots/WWW-AIR\_1664426457975")

}

for(filename in filenames){

plot\_airborne\_data(filename)

}

- - - - - -

# called from the console like so:

> source(file="/Users/gaellemuller-greven/Dropbox/data\_science/projects/nasa\_space\_apps\_2022/nasa.R")

# WORKS BEAUTIFULLY…. ABOUT TO CRY HAPPY TEARS

# Plotting Flight Tracks

> df %>% filter(LAT\_AMS != -9999 & LON\_AMS != -9999 & LAT\_AMS > 30 & LON\_AMS < -80) %>% ggplot(aes(x = LON\_AMS, y = LAT\_AMS, color = mission)) + geom\_point(size = 0.1, alpha = 0.5) + scale\_color\_discrete(guide = "none") + theme\_classic() + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "")

> df %>% filter(LAT\_AMS != -9999 & LON\_AMS != -9999 & LAT\_AMS > 30 & LON\_AMS < -80 & Sulfate\_PM1\_AMS != -8888) %>% ggplot(aes(x = LON\_AMS, y = LAT\_AMS, color = Sulfate\_PM1\_AMS)) + geom\_point(size = 0.1, alpha = 0.5) + scale\_color\_viridis\_c(option = "plasma", trans = "log") + theme\_classic() + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Sulfate PM1 (log)")

> df %>% filter(LAT\_AMS != -9999 & LON\_AMS != -9999 & LAT\_AMS > 30 & LON\_AMS < -80 & Ammonium\_PM1\_AMS != -8888) %>% ggplot(aes(x = LON\_AMS, y = LAT\_AMS, color = Ammonium\_PM1\_AMS)) + geom\_point(size = 0.1, alpha = 0.5) + scale\_color\_viridis\_c(option = "plasma", trans = "log") + theme\_classic() + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Ammonium PM1 (log)")

> df %>% filter(LAT\_AMS != -9999 & LON\_AMS != -9999 & LAT\_AMS > 30 & LON\_AMS < -80 & Nitrate\_PM1\_AMS != -8888) %>% ggplot(aes(x = LON\_AMS, y = LAT\_AMS, color = Nitrate\_PM1\_AMS)) + geom\_point(size = 0.1, alpha = 0.5) + scale\_color\_viridis\_c(option = "plasma", trans = "log") + theme\_classic() + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Nitrate PM1 (log)")

> df %>% filter(LAT\_AMS != -9999 & LON\_AMS != -9999 & LAT\_AMS > 30 & LON\_AMS < -80 & Density\_PM1\_AMS != -8888) %>% ggplot(aes(x = LON\_AMS, y = LAT\_AMS, color = Density\_PM1\_AMS)) + geom\_point(size = 0.1, alpha = 0.5) + scale\_color\_viridis\_c(option = "plasma", trans = "log") + theme\_classic() + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Density PM1 (log)")

# plot on map with sf

# basic map with no datapoints with coord limits for lower 48

> ggplot(data = world) + geom\_sf(alpha = 0.2) + coord\_sf(xlim = c(-123, -69), ylim = c(26, 49))

# below code WORKS but need to filter out -8888 and -9999 NA flags

> ggplot(data = world) + geom\_sf(lwd = 0.2, color = rgb(156, 156, 156, maxColorValue = 255), alpha = 0.6) + coord\_sf(xlim = c(-123, -69), ylim = c(26, 49)) + geom\_point(data = df, aes(x = LON\_AMS, y = LAT\_AMS, color = Seasalt\_PM1\_AMS), size = 0.1, alpha = 0.5) + scale\_color\_viridis\_c(option = "plasma") + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Seasalt PM1 (log)")

missions <- c("FIREXAQ-AMS\_DC8\_20190722\_R2", "FIREXAQ-AMS\_DC8\_20190724\_R2", "FIREXAQ-AMS\_DC8\_20190725\_R2", "FIREXAQ-AMS\_DC8\_20190729\_R3", "FIREXAQ-AMS\_DC8\_20190730\_R2", "FIREXAQ-AMS\_DC8\_20190802\_R3", "FIREXAQ-AMS\_DC8\_20190803\_R2", "FIREXAQ-AMS\_DC8\_20190806\_R3", "FIREXAQ-AMS\_DC8\_20190807\_R3", "FIREXAQ-AMS\_DC8\_20190808\_R2", "FIREXAQ-AMS\_DC8\_20190812\_R2", "FIREXAQ-AMS\_DC8\_20190813\_R2", "FIREXAQ-AMS\_DC8\_20190815\_R2", "FIREXAQ-AMS\_DC8\_20190816\_R3", "FIREXAQ-AMS\_DC8\_20190819\_R2", "FIREXAQ-AMS\_DC8\_20190821\_R3", "FIREXAQ-AMS\_DC8\_20190823\_R3", "FIREXAQ-AMS\_DC8\_20190826\_R3", "FIREXAQ-AMS\_DC8\_20190829\_R3", "FIREXAQ-AMS\_DC8\_20190830\_R3", "FIREXAQ-AMS\_DC8\_20190831\_R3", "FIREXAQ-AMS\_DC8\_20190903\_R3", "FIREXAQ-AMS\_DC8\_20190905\_R2")

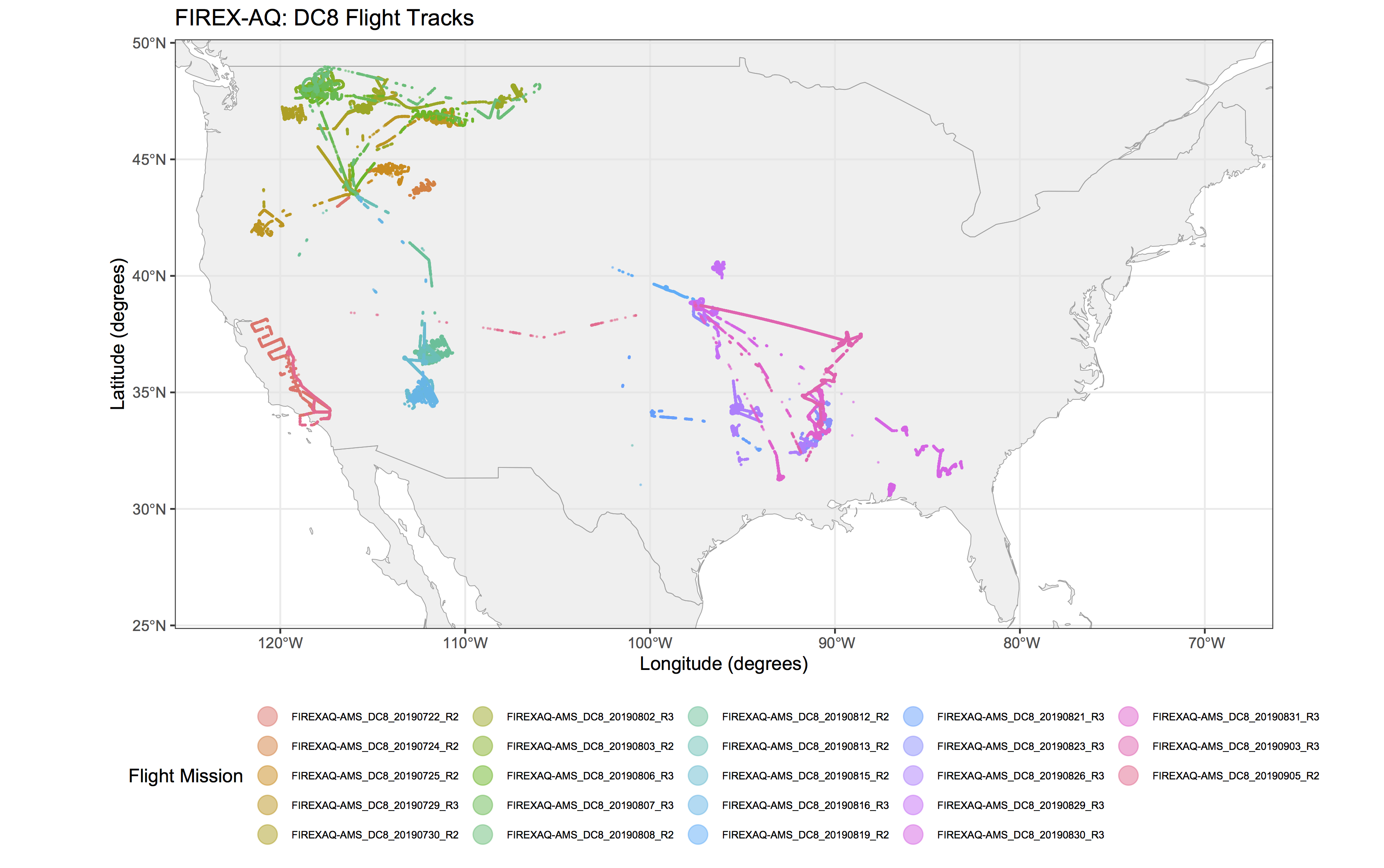
> missions\_wildfire <- c("FIREXAQ-AMS\_DC8\_20190722\_R2", "FIREXAQ-AMS\_DC8\_20190724\_R2", "FIREXAQ-AMS\_DC8\_20190725\_R2", "FIREXAQ-AMS\_DC8\_20190729\_R3", "FIREXAQ-AMS\_DC8\_20190730\_R2", "FIREXAQ-AMS\_DC8\_20190802\_R3", "FIREXAQ-AMS\_DC8\_20190803\_R2", "FIREXAQ-AMS\_DC8\_20190806\_R3", "FIREXAQ-AMS\_DC8\_20190807\_R3", "FIREXAQ-AMS\_DC8\_20190808\_R2", "FIREXAQ-AMS\_DC8\_20190812\_R2", "FIREXAQ-AMS\_DC8\_20190813\_R2", "FIREXAQ-AMS\_DC8\_20190815\_R2", "FIREXAQ-AMS\_DC8\_20190816\_R3")

> missions\_agriburns <- c("FIREXAQ-AMS\_DC8\_20190819\_R2", "FIREXAQ-AMS\_DC8\_20190821\_R3", "FIREXAQ-AMS\_DC8\_20190823\_R3", "FIREXAQ-AMS\_DC8\_20190826\_R3", "FIREXAQ-AMS\_DC8\_20190829\_R3", "FIREXAQ-AMS\_DC8\_20190830\_R3", "FIREXAQ-AMS\_DC8\_20190831\_R3", "FIREXAQ-AMS\_DC8\_20190903\_R3", "FIREXAQ-AMS\_DC8\_20190905\_R2")

> nasa\_data <- nasa\_data %>% mutate(mission\_type = as.factor(ifelse(mission %in% missions\_wildfire, "wildfire", ifelse(mission %in% missions\_agriburns, "agriburn", ""))))

# map of flight missions with legend

> ggplot(data = world) + geom\_sf(lwd = 0.2, color = rgb(156, 156, 156, maxColorValue = 255), alpha = 0.6) + coord\_sf(xlim = c(-123, -69), ylim = c(26, 49)) + geom\_point(data = nasa\_data\_per\_OAden, aes(x = LON\_AMS, y = LAT\_AMS, color = mission), size = 0.1, alpha = 0.5) + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Flight Mission") + theme(legend.position="bottom", legend.text=element\_text(size=6)) + guides(color = guide\_legend(override.aes = list(size = 5)))



> nasa\_data\_mission\_summary <- nasa\_data %>% group\_by(mission) %>% summarize(mean\_LAT = mean(LAT\_AMS), mean\_LON = mean(LON\_AMS))

> as.data.frame(nasa\_data\_mission\_summary)

mission mean\_LAT mean\_LON

1 FIREXAQ-AMS\_DC8\_20190722\_R2 36.28655 -119.37623

2 FIREXAQ-AMS\_DC8\_20190724\_R2 43.79656 -112.60827

3 FIREXAQ-AMS\_DC8\_20190725\_R2 44.46708 -114.01934

4 FIREXAQ-AMS\_DC8\_20190729\_R3 44.20016 -116.73073

5 FIREXAQ-AMS\_DC8\_20190730\_R2 46.44514 -117.92307

6 FIREXAQ-AMS\_DC8\_20190802\_R3 47.36740 -114.04139

7 FIREXAQ-AMS\_DC8\_20190803\_R2 47.97129 -117.97825

8 FIREXAQ-AMS\_DC8\_20190806\_R3 46.81102 -114.46495

9 FIREXAQ-AMS\_DC8\_20190807\_R3 47.25013 -116.80051

10 FIREXAQ-AMS\_DC8\_20190808\_R2 47.40813 -115.14226

11 FIREXAQ-AMS\_DC8\_20190812\_R2 37.06341 -111.96700

12 FIREXAQ-AMS\_DC8\_20190813\_R2 37.31930 -112.43061

13 FIREXAQ-AMS\_DC8\_20190815\_R2 35.93673 -112.59932

14 FIREXAQ-AMS\_DC8\_20190816\_R3 35.21878 -112.33157

15 FIREXAQ-AMS\_DC8\_20190819\_R2 39.31130 -98.75152

16 FIREXAQ-AMS\_DC8\_20190821\_R3 33.24950 -94.32014

17 FIREXAQ-AMS\_DC8\_20190823\_R3 33.52473 -91.05090

18 FIREXAQ-AMS\_DC8\_20190826\_R3 33.66570 -93.82184

19 FIREXAQ-AMS\_DC8\_20190829\_R3 39.12000 -96.54619

20 FIREXAQ-AMS\_DC8\_20190830\_R3 31.91907 -85.76911

21 FIREXAQ-AMS\_DC8\_20190831\_R3 33.69196 -91.22169

22 FIREXAQ-AMS\_DC8\_20190903\_R3 35.81522 -90.87053

23 FIREXAQ-AMS\_DC8\_20190905\_R2 34.98109 -117.82508

> missions\_wildfire <- c("FIREXAQ-AMS\_DC8\_20190722\_R2", "FIREXAQ-AMS\_DC8\_20190724\_R2", "FIREXAQ-AMS\_DC8\_20190725\_R2", "FIREXAQ-AMS\_DC8\_20190729\_R3", "FIREXAQ-AMS\_DC8\_20190730\_R2", "FIREXAQ-AMS\_DC8\_20190802\_R3", "FIREXAQ-AMS\_DC8\_20190803\_R2", "FIREXAQ-AMS\_DC8\_20190806\_R3", "FIREXAQ-AMS\_DC8\_20190807\_R3", "FIREXAQ-AMS\_DC8\_20190808\_R2", "FIREXAQ-AMS\_DC8\_20190812\_R2", "FIREXAQ-AMS\_DC8\_20190813\_R2", "FIREXAQ-AMS\_DC8\_20190815\_R2", "FIREXAQ-AMS\_DC8\_20190816\_R3", "FIREXAQ-AMS\_DC8\_20190905\_R2")

> missions\_agriburns <- c("FIREXAQ-AMS\_DC8\_20190819\_R2", "FIREXAQ-AMS\_DC8\_20190821\_R3", "FIREXAQ-AMS\_DC8\_20190823\_R3", "FIREXAQ-AMS\_DC8\_20190826\_R3", "FIREXAQ-AMS\_DC8\_20190829\_R3", "FIREXAQ-AMS\_DC8\_20190830\_R3", "FIREXAQ-AMS\_DC8\_20190831\_R3", "FIREXAQ-AMS\_DC8\_20190903\_R3")

> nasa\_data <- nasa\_data %>% mutate(mission\_type = as.factor(ifelse(mission %in% missions\_wildfire, "wildfire", ifelse(mission %in% missions\_agriburns, "agriburn", ""))))

> nasa\_data\_per\_OAden <- nasa\_data %>% mutate(across(all\_of(PM1\_AMS), ~ . / OADensity\_PM1\_AMS, .names = "{col}\_per\_OAden"))

> nasa\_data\_per\_OAden\_gather <- nasa\_data\_per\_OAden %>% gather(key = "component", value = "measurement", -c(Time\_Start, Time\_Stop, Time\_Mid, ALT\_AMS, LAT\_AMS, LON\_AMS, mission, mission\_type, Density\_PM1\_AMS, OADensity\_PM1\_AMS))

> str(nasa\_data\_per\_OAden\_gather)

'data.frame': 20534300 obs. of 12 variables:

$ Time\_Start : num 66229 66232 66233 66234 66235 ...

$ Time\_Stop : num 66230 66233 66234 66235 66236 ...

$ Time\_Mid : num 66230 66233 66234 66235 66236 ...

$ LAT\_AMS : num 34.6 34.6 34.6 34.6 34.6 ...

$ LON\_AMS : num -118 -118 -118 -118 -118 ...

$ ALT\_AMS : num 2999 3035 3046 3060 3071 ...

$ Density\_PM1\_AMS : num 1.66 1.65 1.65 1.65 1.63 ...

$ OADensity\_PM1\_AMS: num 1.57 1.57 1.57 1.57 1.57 ...

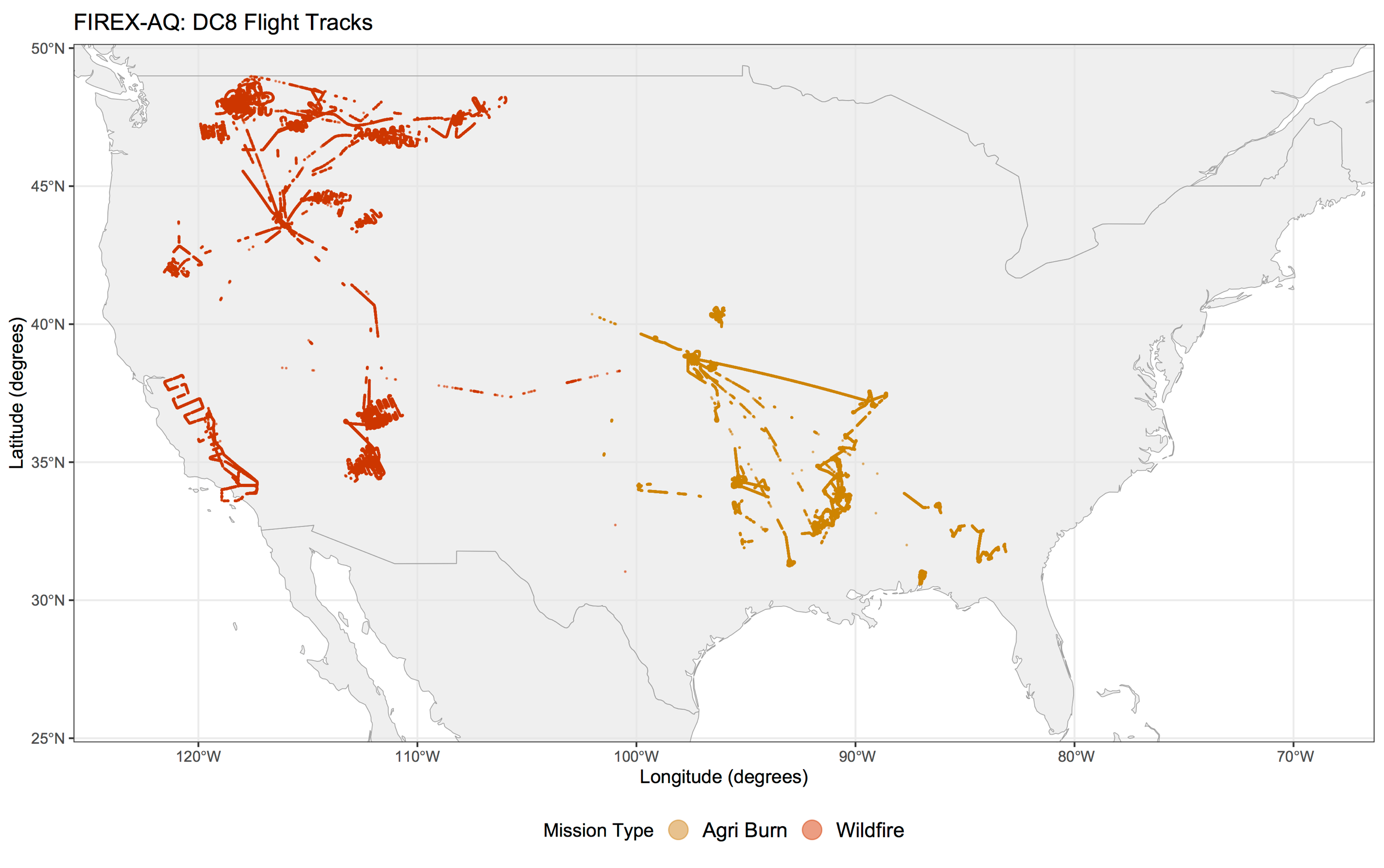
$ mission : Factor w/ 23 levels "FIREXAQ-AMS\_DC8\_20190722\_R2",..: 1 1 1 1 1 1 1 1 1 1 ...

$ mission\_type : Factor w/ 2 levels "agriburn","wildfire": 2 2 2 2 2 2 2 2 2 2 ...

$ component : chr "OA\_PM1\_AMS" "OA\_PM1\_AMS" "OA\_PM1\_AMS" "OA\_PM1\_AMS" ...

$ measurement : num 2.09 2.65 2.68 2.36 3.33 ...

> ggplot(data = world) + geom\_sf(lwd = 0.2, color = rgb(156, 156, 156, maxColorValue = 255), alpha = 0.6) + coord\_sf(xlim = c(-123, -69), ylim = c(26, 49)) + geom\_point(data = nasa\_data\_per\_OAden, aes(x = LON\_AMS, y = LAT\_AMS, color = mission\_type), size = 0.1, alpha = 0.5) + scale\_color\_manual(values = c("agriburn" = "orange3", "wildfire" = "orangered3"), labels = c("Agri Burn", "Wildfire")) + labs(title = "FIREX-AQ: DC8 Flight Tracks", y = "Latitude (degrees)", x = "Longitude (degrees)", color = "Mission Type") + theme(legend.position="bottom", legend.text=element\_text(size=12)) + guides(color = guide\_legend(override.aes = list(size = 5)))



# STATS

> PM1\_AMS\_per\_OAden <- c("OA\_PM1\_AMS\_per\_OAden", "Sulfate\_PM1\_AMS\_per\_OAden", "Nitrate\_PM1\_AMS\_per\_OAden", "Ammonium\_PM1\_AMS\_per\_OAden", "NR\_Chloride\_PM1\_AMS\_per\_OAden", "Potassium\_PM1\_AMS\_per\_OAden", "MSA\_PM1\_AMS\_per\_OAden", "ClO4\_PM1\_AMS\_per\_OAden", "Iodine\_PM1\_AMS\_per\_OAden", "Bromine\_PM1\_AMS\_per\_OAden", "Seasalt\_PM1\_AMS\_per\_OAden", "AmmBalance\_PM1\_AMS\_per\_OAden", "OtoC\_Ratio\_PM1\_AMS\_per\_OAden", "HtoC\_Ratio\_PM1\_AMS\_per\_OAden", "OAtoOC\_PM1\_AMS\_per\_OAden", "OSc\_PM1\_AMS\_per\_OAden", "OrgNitrFraction\_PM1\_AMS\_per\_OAden", "f43\_PM1\_AMS\_per\_OAden", "f44\_PM1\_AMS\_per\_OAden", "f57\_PM1\_AMS\_per\_OAden", "f60\_PM1\_AMS\_per\_OAden", "f82\_PM1\_AMS\_per\_OAden", "f91\_PM1\_AMS\_per\_OAden", "fC2H3O\_PM1\_AMS\_per\_OAden", "fCO2\_PM1\_AMS\_per\_OAden", "fC4H9\_PM1\_AMS\_per\_OAden", "fC2H4O2\_PM1\_AMS\_per\_OAden", "fC5H6O\_PM1\_AMS\_per\_OAden", "fC7H7\_PM1\_AMS\_per\_OAden")

> nasa\_data\_per\_OAden\_lm <- nasa\_data\_per\_OAden %>% select(c(mission\_type, ALT\_AMS, Density\_PM1\_AMS, OADensity\_PM1\_AMS), all\_of(PM1\_AMS\_per\_OAden))

> logit <- glm(mission\_type ~ ., data=nasa\_data\_per\_OAden\_lm, family="binomial")

> summary(logit)

Call:

glm(formula = mission\_type ~ ., family = "binomial", data = nasa\_data\_per\_OAden\_lm)

Deviance Residuals:

Min 1Q Median 3Q Max

-6.1271 -0.2270 -0.0677 0.2057 8.4904

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 5.766e+01 1.476e+00 39.056 < 2e-16 \*\*\*

ALT\_AMS 1.593e-03 9.226e-06 172.692 < 2e-16 \*\*\*

Density\_PM1\_AMS 8.171e+00 5.016e-01 16.288 < 2e-16 \*\*\*

OADensity\_PM1\_AMS -3.809e+01 9.896e-01 -38.486 < 2e-16 \*\*\*

OA\_PM1\_AMS\_per\_OAden 1.220e-01 4.358e-03 28.007 < 2e-16 \*\*\*

Sulfate\_PM1\_AMS\_per\_OAden -4.829e-01 5.101e-02 -9.468 < 2e-16 \*\*\*

Nitrate\_PM1\_AMS\_per\_OAden 9.787e-01 5.896e-02 16.600 < 2e-16 \*\*\*

Ammonium\_PM1\_AMS\_per\_OAden -1.091e+00 1.454e-01 -7.503 6.21e-14 \*\*\*

NR\_Chloride\_PM1\_AMS\_per\_OAden -6.190e+00 1.925e-01 -32.150 < 2e-16 \*\*\*

Potassium\_PM1\_AMS\_per\_OAden -2.188e-01 1.500e-02 -14.585 < 2e-16 \*\*\*

MSA\_PM1\_AMS\_per\_OAden 2.800e-01 1.235e-01 2.268 0.023337 \*

ClO4\_PM1\_AMS\_per\_OAden -4.514e+00 2.632e-01 -17.149 < 2e-16 \*\*\*

Iodine\_PM1\_AMS\_per\_OAden 5.446e+00 5.705e-01 9.545 < 2e-16 \*\*\*

Bromine\_PM1\_AMS\_per\_OAden 9.748e-01 3.530e-01 2.761 0.005757 \*\*

Seasalt\_PM1\_AMS\_per\_OAden -3.027e-02 6.958e-02 -0.435 0.663566

AmmBalance\_PM1\_AMS\_per\_OAden 7.419e-02 7.406e-02 1.002 0.316453

OtoC\_Ratio\_PM1\_AMS\_per\_OAden -1.580e+02 1.050e+01 -15.056 < 2e-16 \*\*\*

HtoC\_Ratio\_PM1\_AMS\_per\_OAden 5.750e+01 5.267e+00 10.918 < 2e-16 \*\*\*

OAtoOC\_PM1\_AMS\_per\_OAden 2.868e+00 8.097e-01 3.543 0.000396 \*\*\*

OSc\_PM1\_AMS\_per\_OAden 7.027e+01 5.248e+00 13.390 < 2e-16 \*\*\*

OrgNitrFraction\_PM1\_AMS\_per\_OAden 2.815e-01 5.116e-02 5.503 3.73e-08 \*\*\*

f43\_PM1\_AMS\_per\_OAden -8.195e+00 1.735e+00 -4.722 2.33e-06 \*\*\*

f44\_PM1\_AMS\_per\_OAden 4.564e+00 3.160e+00 1.444 0.148649

f57\_PM1\_AMS\_per\_OAden -3.349e+01 4.125e+00 -8.119 4.71e-16 \*\*\*

f60\_PM1\_AMS\_per\_OAden -1.874e+02 7.320e+00 -25.596 < 2e-16 \*\*\*

f82\_PM1\_AMS\_per\_OAden 1.081e+02 8.695e+00 12.428 < 2e-16 \*\*\*

f91\_PM1\_AMS\_per\_OAden 1.043e+02 8.312e+00 12.547 < 2e-16 \*\*\*

fC2H3O\_PM1\_AMS\_per\_OAden -6.160e+01 1.896e+00 -32.493 < 2e-16 \*\*\*

fCO2\_PM1\_AMS\_per\_OAden 3.987e+01 3.175e+00 12.558 < 2e-16 \*\*\*

fC4H9\_PM1\_AMS\_per\_OAden 1.014e+02 4.968e+00 20.403 < 2e-16 \*\*\*

fC2H4O2\_PM1\_AMS\_per\_OAden 7.495e+01 7.641e+00 9.809 < 2e-16 \*\*\*

fC5H6O\_PM1\_AMS\_per\_OAden -1.916e+02 1.162e+01 -16.489 < 2e-16 \*\*\*

fC7H7\_PM1\_AMS\_per\_OAden 1.471e+02 7.415e+00 19.837 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 324901 on 241579 degrees of freedom

Residual deviance: 90981 on 241547 degrees of freedom

AIC: 91047

Number of Fisher Scoring iterations: 9