

# fish\_feed\_PFAS\_analysis

2023-03-02

```
library(lme4)
```

```
## Warning: package 'lme4' was built under R version 4.2.3
```

```
## Loading required package: Matrix
```

```
library(ggplot2)
```

```
# Read in raw data
```

```
feeds_disagg <- read.csv('feeds_restructured.csv')
```

```
### # Because manufacturers B and E only occur in one or 2 feeds, I will omit them from my  
# analyses
```

```
# Likewise, because protein sources 'ap' and 'pp' each only occur in one feed, I will omit them  
# from my analyses.
```

```
excluded_ps <- c('ap', 'pp')
```

```
excluded_mfr <- c('B', 'E')
```

```
subset_feeds <- subset(feeds_disagg, !(protein_source %in% excluded_ps) & !(mfr %in% excluded_mfr))
```

```
# Factor mfr and protein source
```

```
subset_feeds$mfr <- factor(subset_feeds$mfr)
```

```
subset_feeds$protein_source <- factor(subset_feeds$protein_source)
```

```
# PFOS Mixed Effects Model
```

```
pfos_mmod <- lmer(conc ~ protein_per + mfr + protein_source + (1 | feed_num),  
  subset = comp == 'PFOS', data = subset_feeds)
```

```
summary(pfos_mmod, correlation = FALSE)
```

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula: conc ~ protein_per + mfr + protein_source + (1 | feed_num)
```

```
## Data: subset_feeds
```

```
## Subset: comp == "PFOS"
```

```
##
```

```
## REML criterion at convergence: -26.9
```

```
##
```

```
## Scaled residuals:
```

```
##      Min      1Q   Median      3Q      Max
```

```
## -2.43372 -0.14235  0.00383  0.40326  1.55618
```

```
##
```

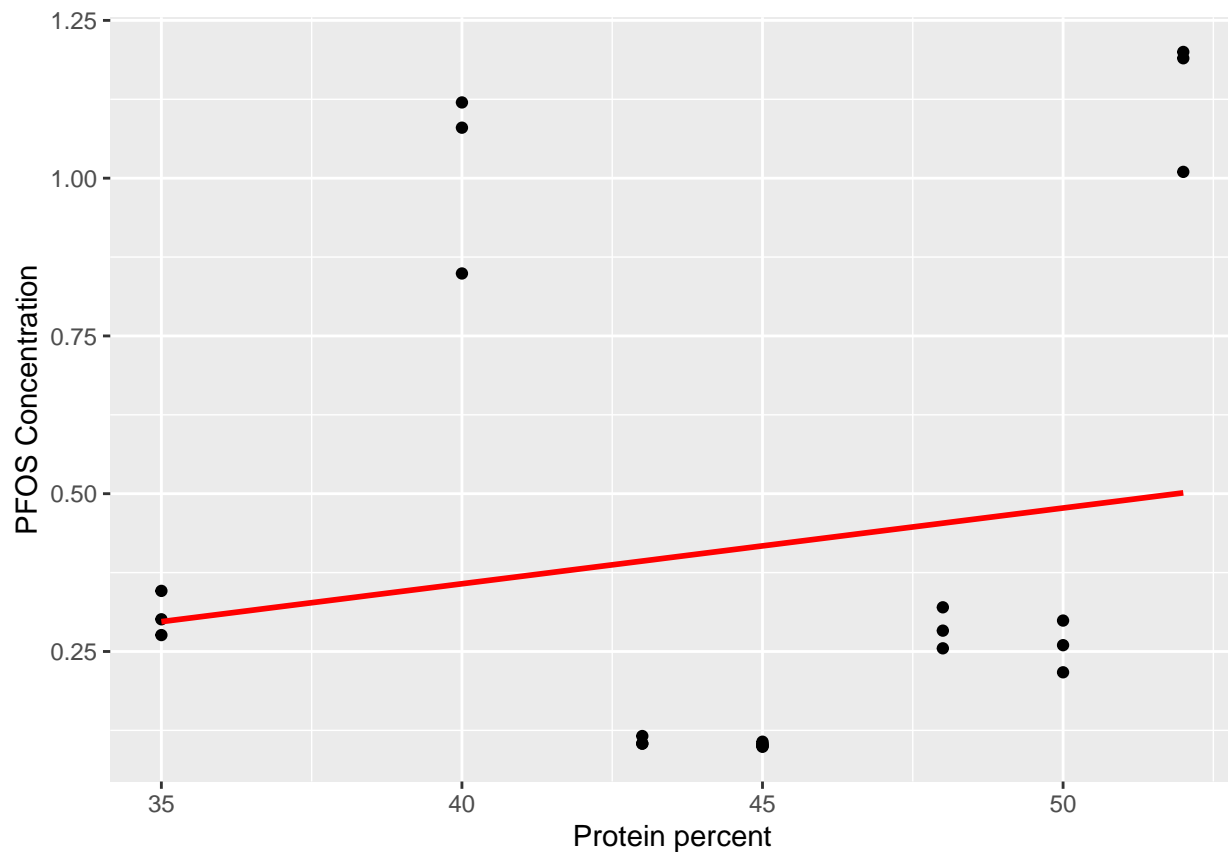
```
## Random effects:
```

```
## Groups   Name      Variance Std.Dev.
```

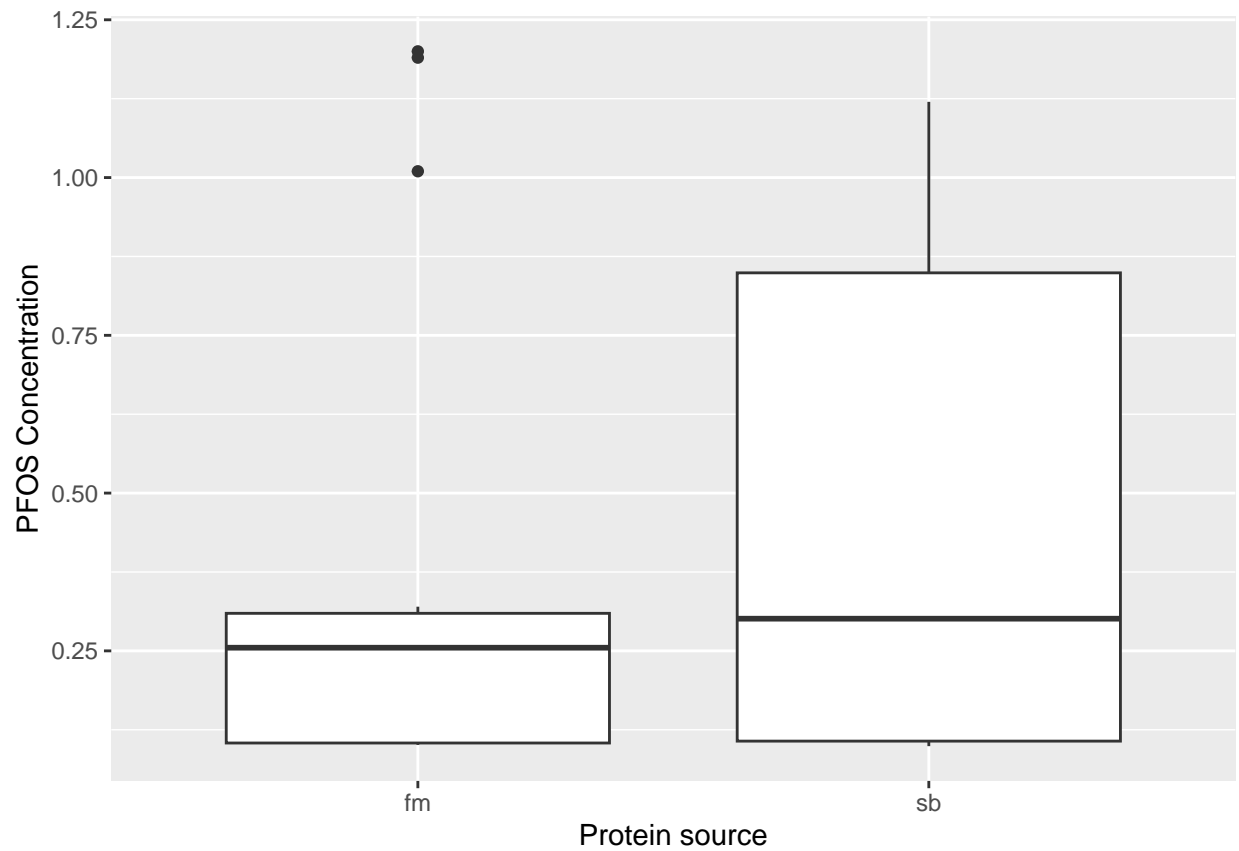
```
## feed_num (Intercept) 0.037399 0.19339
## Residual              0.004613 0.06792
## Number of obs: 24, groups: feed_num, 8
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   -4.3305     1.1400  -3.799
## protein_per    0.1212     0.0258   4.696
## mfrC          -0.9565     0.3472  -2.755
## mfrD          -1.4687     0.3096  -4.744
## protein_sourcesb 0.4491     0.3074   1.461
```

```
# Figures
ggplot(subset(subset_feeds, comp == "PFOS"), aes(x = protein_per, y = conc)) +
  geom_point() +
  xlab("Protein percent") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  ylab("PFOS Concentration")
```

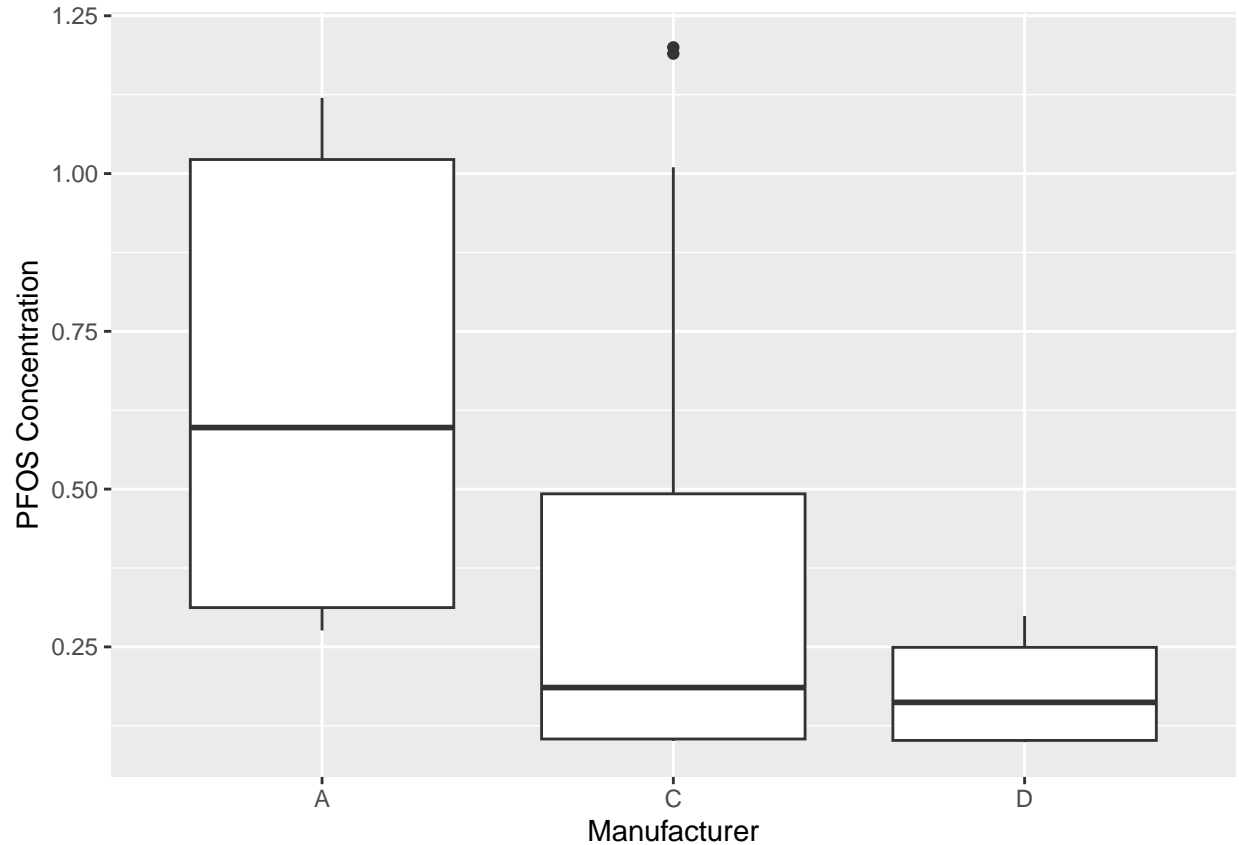
```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
ggplot(subset(subset_feeds, comp == "PFOS"), aes(x = protein_source, y = conc)) +
  geom_boxplot() +
  xlab("Protein source") +
  ylab("PFOS Concentration")
```



```
ggplot(subset(subset_feeds, comp == "PFOS"), aes(x = mfr, y = conc)) +  
  geom_boxplot() +  
  xlab("Manufacturer") +  
  ylab("PFOS Concentration")
```



```
# PFNA Mixed Effects Model
pfna_mmod <- lmer(conc ~ protein_per + mfr + protein_source + (1 | feed_num),
  subset = comp == 'PFNA', data = subset_feeds)

summary(pfna_mmod, correlation = FALSE)

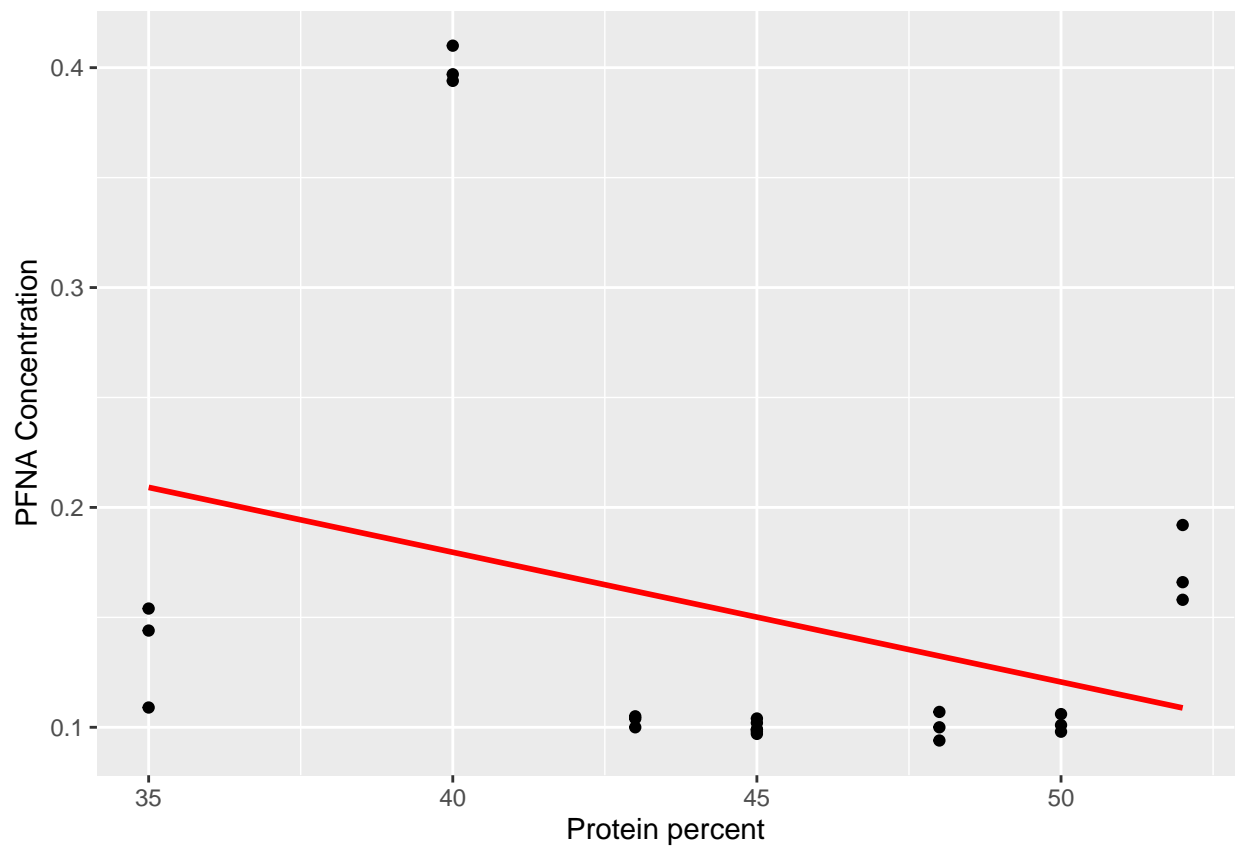
## Linear mixed model fit by REML ['lmerMod']
## Formula: conc ~ protein_per + mfr + protein_source + (1 | feed_num)
## Data: subset_feeds
## Subset: comp == "PFNA"
##
## REML criterion at convergence: -89.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.40015 -0.26595 -0.06954  0.31616  1.74703
##
## Random effects:
## Groups Name Variance Std.Dev.
## feed_num (Intercept) 0.0070673 0.08407
## Residual 0.0001285 0.01133
## Number of obs: 24, groups: feed_num, 8
##
## Fixed effects:
## Estimate Std. Error t value
## (Intercept) -0.46621 0.48714 -0.957
```

```
## protein_per      0.01730      0.01102      1.569
## mfrC             -0.22826      0.14837     -1.538
## mfrD             -0.29708      0.13229     -2.246
## protein_sourcesb 0.08550      0.13137      0.651
```

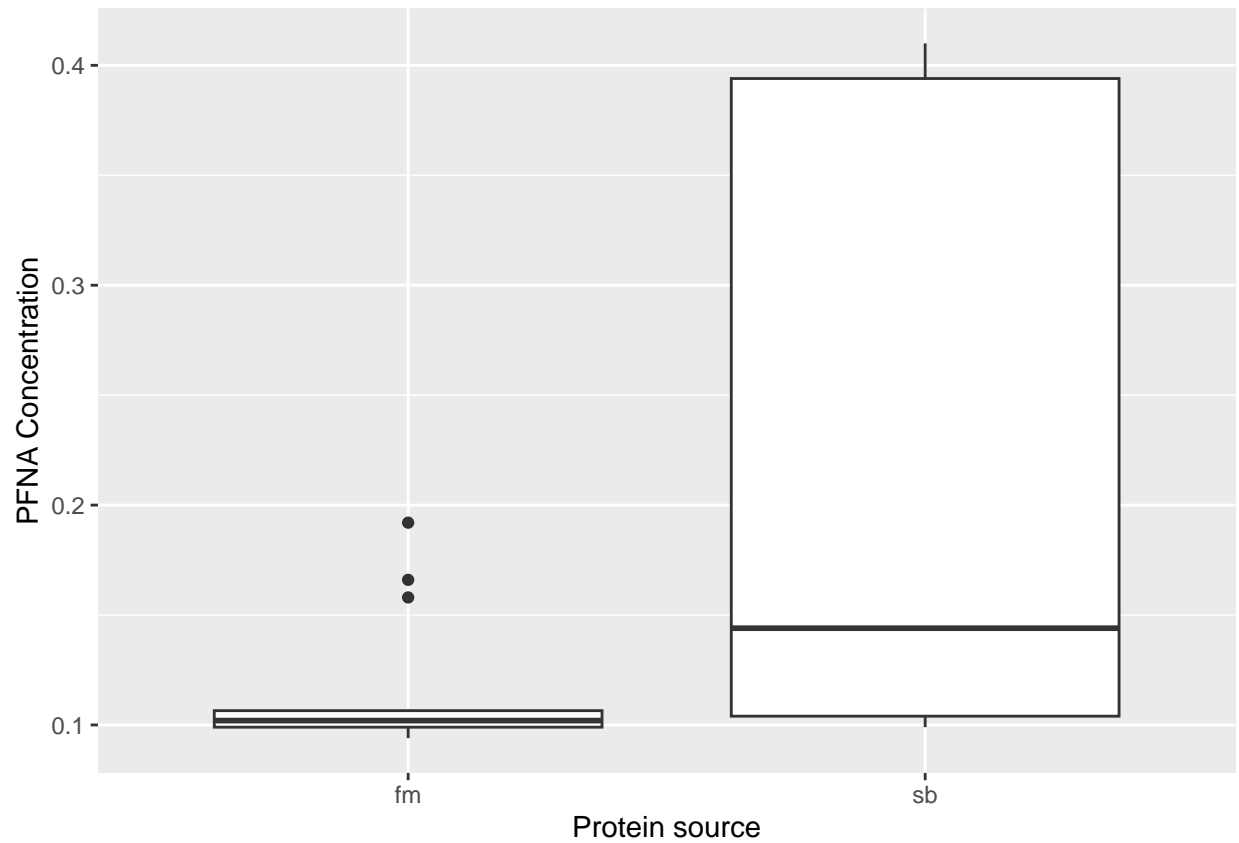
*# Figures*

```
ggplot(subset(subset_feeds, comp == "PFNA"), aes(x = protein_per, y = conc)) +
  geom_point() +
  xlab("Protein percent") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  ylab("PFNA Concentration")
```

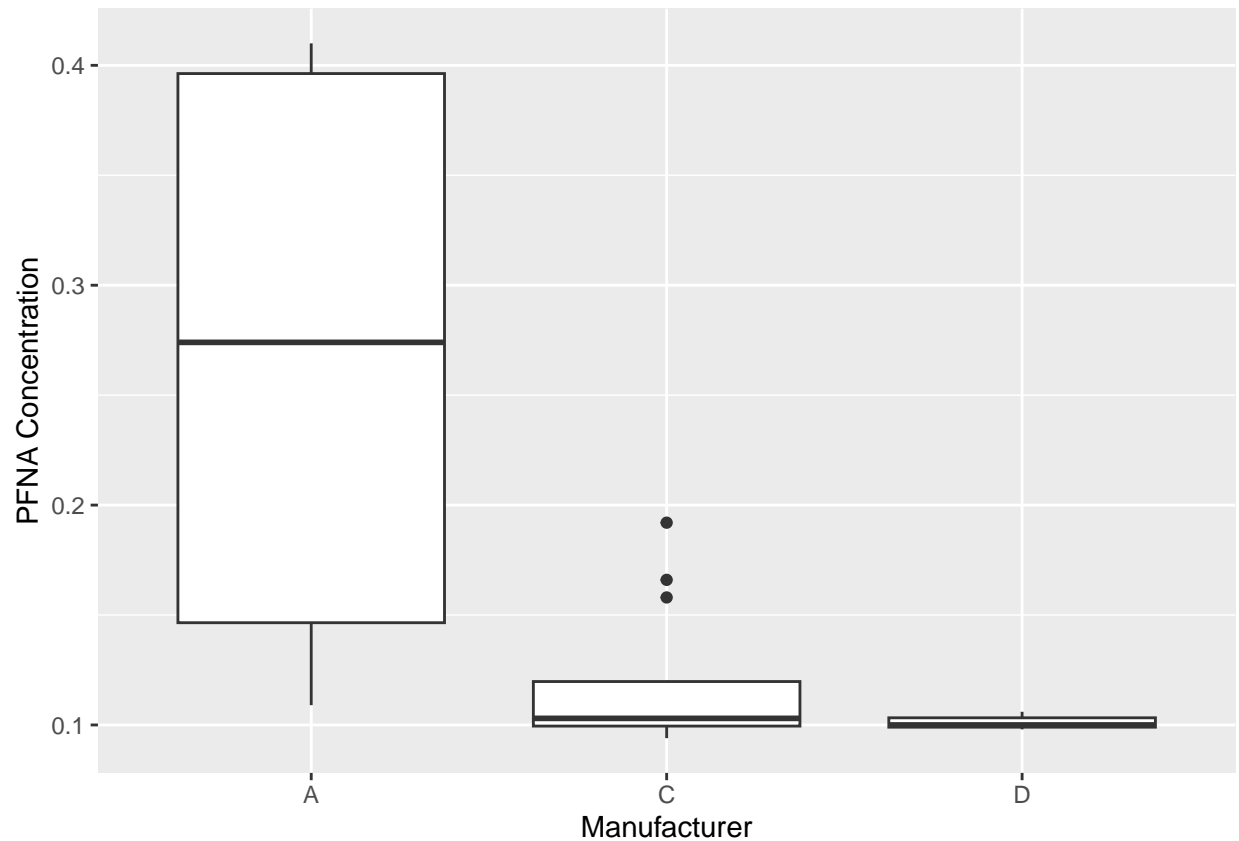
## 'geom\_smooth()' using formula = 'y ~ x'



```
ggplot(subset(subset_feeds, comp == "PFNA"), aes(x = protein_source, y = conc)) +
  geom_boxplot() +
  xlab("Protein source") +
  ylab("PFNA Concentration")
```



```
ggplot(subset(subset_feeds, comp == "PFNA"), aes(x = mfr, y = conc)) +  
  geom_boxplot() +  
  xlab("Manufacturer") +  
  ylab("PFNA Concentration")
```



#### # PFDA Mixed Effects Model

```
pfda_mmod <- lmer(conc ~ protein_per + mfr + protein_source + (1 | feed_num),
  subset = comp == 'PFDA', data = subset_feeds)
```

```
summary(pfda_mmod, correlation = FALSE)
```

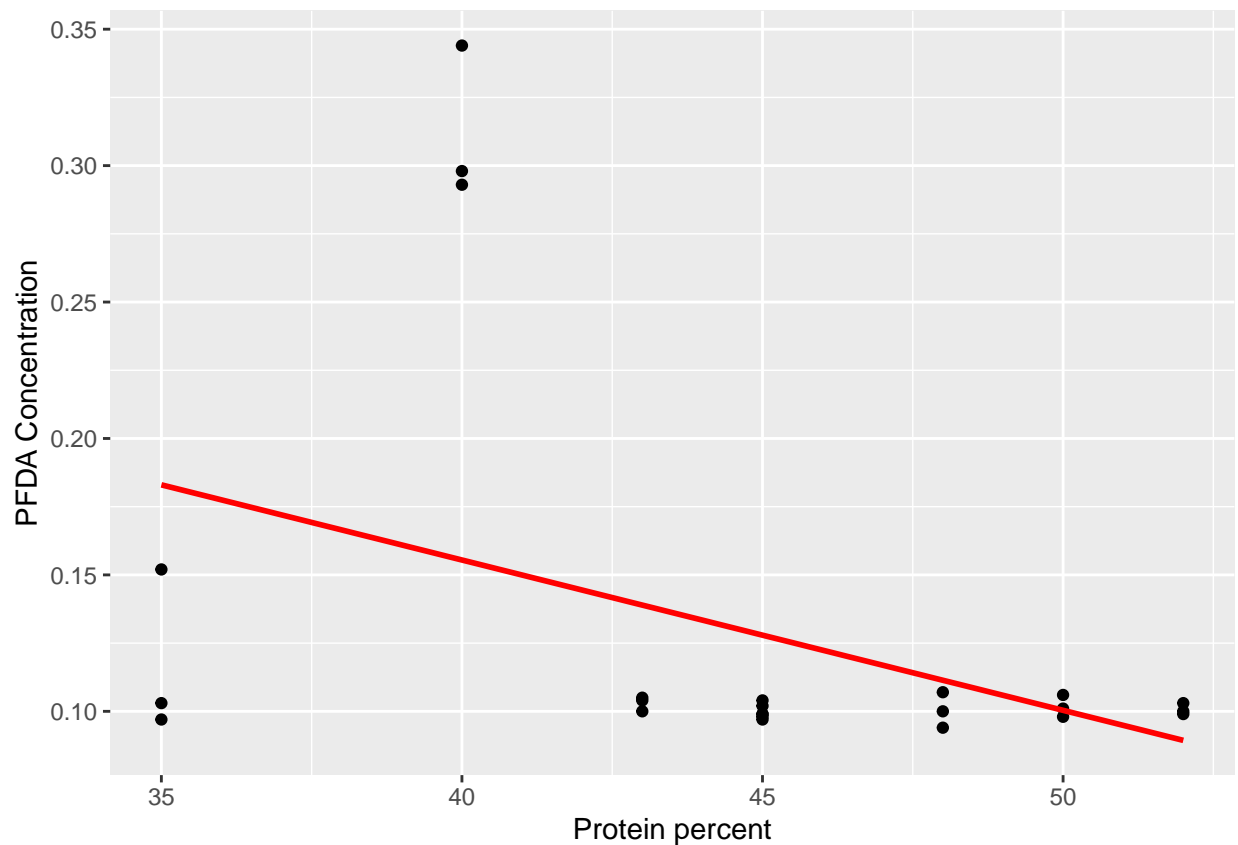
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: conc ~ protein_per + mfr + protein_source + (1 | feed_num)
## Data: subset_feeds
## Subset: comp == "PFDA"
##
## REML criterion at convergence: -81.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.43743 -0.18569 -0.06897  0.18049  2.24357
##
## Random effects:
## Groups Name Variance Std.Dev.
## feed_num (Intercept) 0.0049125 0.07009
## Residual 0.0002232 0.01494
## Number of obs: 24, groups: feed_num, 8
##
## Fixed effects:
## Estimate Std. Error t value
```

```
## (Intercept)      -0.132855    0.407975   -0.326
## protein_per       0.008197    0.009233    0.888
## mfrC              -0.151635    0.124259   -1.220
## mfrD              -0.175308    0.110795   -1.582
## protein_sourcesb  0.039983    0.110023    0.363
```

#### #Figures

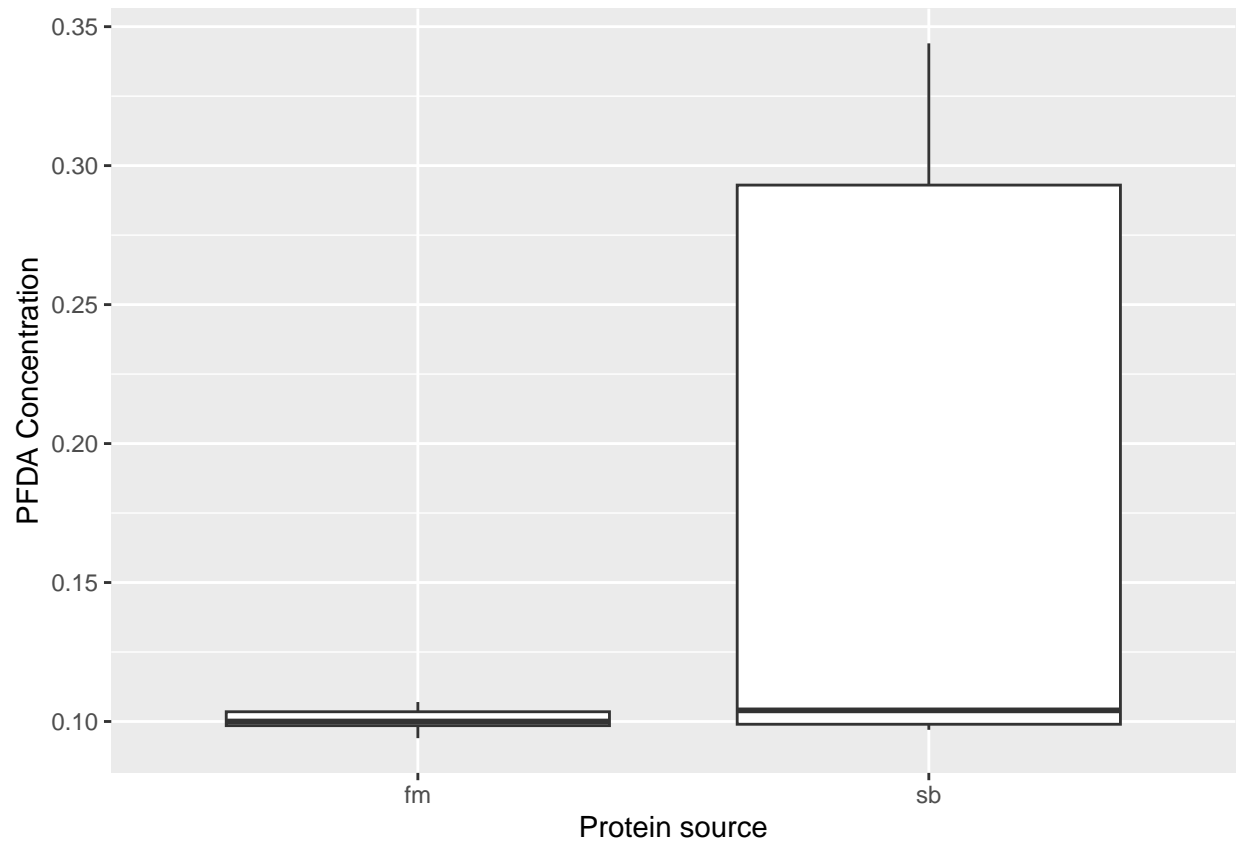
```
ggplot(subset(subset_feeds, comp == "PFDA"), aes(x = protein_per, y = conc)) +
  geom_point() +
  xlab("Protein percent") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  ylab("PFDA Concentration")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

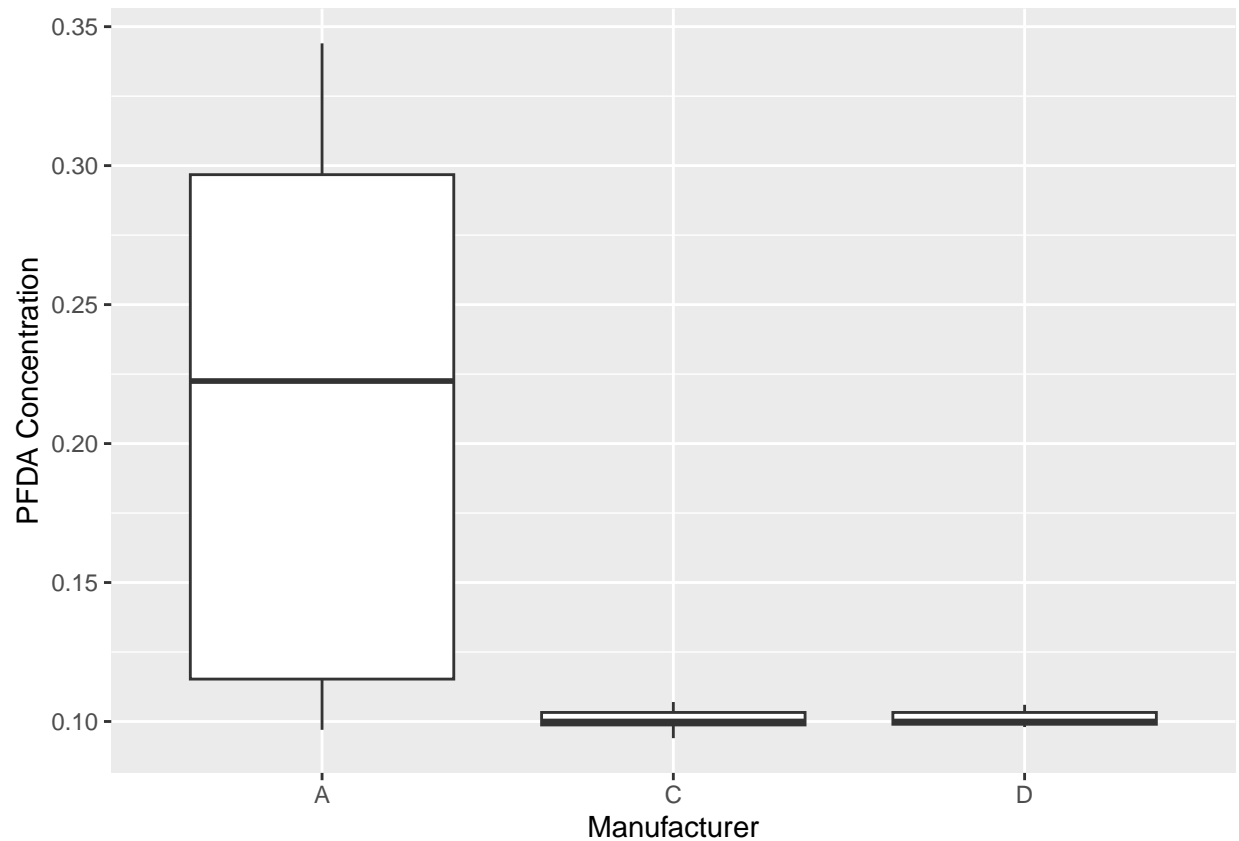


```
ggplot(subset(subset_feeds, comp == "PFDA"), aes(x = protein_source, y = conc)) +
  geom_boxplot() +
  xlab("Protein source") +
  ylab("PFDA Concentration")
```





```
ggplot(subset(subset_feeds, comp == "PFDA"), aes(x = mfr, y = conc)) +  
  geom_boxplot() +  
  xlab("Manufacturer") +  
  ylab("PFDA Concentration")
```



```
# PFUnA Mixed Effects Model
```

```
pfuna_mmod <- lmer(conc ~ protein_per + mfr + protein_source + (1 | feed_num),
  subset = comp == 'PFUnA', data = subset_feeds)
```

```
summary(pfuna_mmod, correlation = FALSE)
```

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula: conc ~ protein_per + mfr + protein_source + (1 | feed_num)
```

```
## Data: subset_feeds
```

```
## Subset: comp == "PFUnA"
```

```
##
```

```
## REML criterion at convergence: -81.6
```

```
##
```

```
## Scaled residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -2.29438 -0.21067 -0.05572  0.23134  1.58057
```

```
##
```

```
## Random effects:
```

```
## Groups Name Variance Std.Dev.
```

```
## feed_num (Intercept) 0.0032233 0.05677
```

```
## Residual 0.0002399 0.01549
```

```
## Number of obs: 24, groups: feed_num, 8
```

```
##
```

```
## Fixed effects:
```

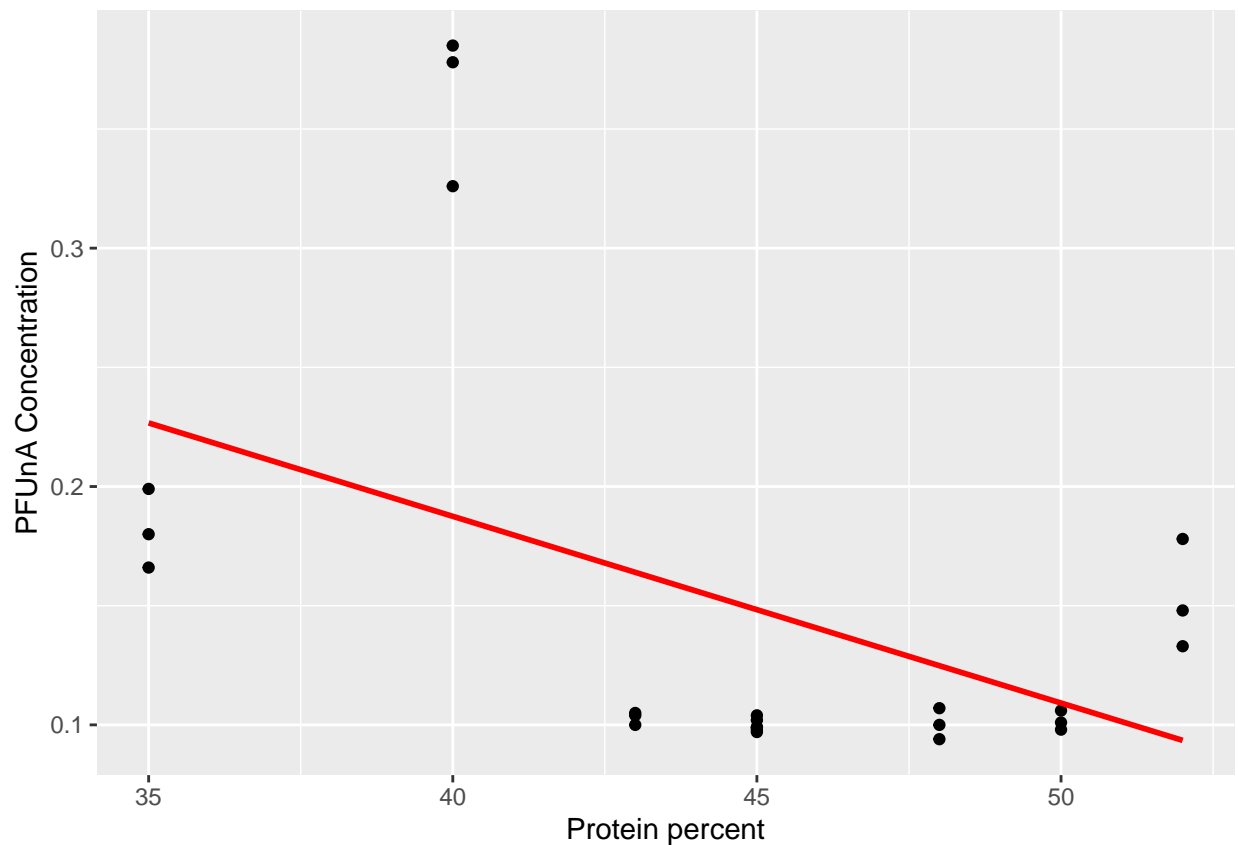
```
## Estimate Std. Error t value
```

```
## (Intercept)      -0.241510    0.332039   -0.727
## protein_per       0.012114    0.007514    1.612
## mfrC              -0.214013    0.101131   -2.116
## mfrD              -0.262521    0.090173   -2.911
## protein_sourcesb  0.059570    0.089544    0.665
```

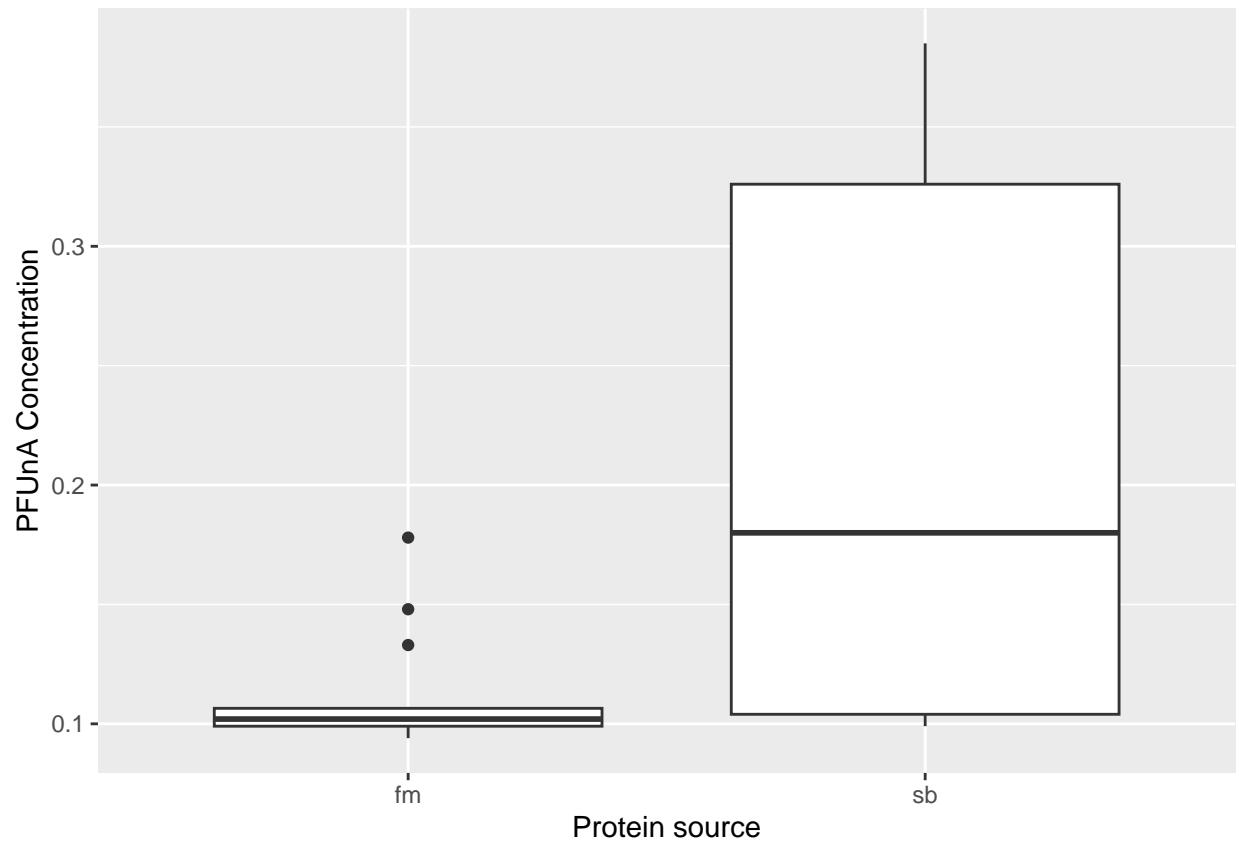
#### #Figures

```
ggplot(subset(subset_feeds, comp == "PFUnA"), aes(x = protein_per, y = conc)) +
  geom_point() +
  xlab("Protein percent") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  ylab("PFUnA Concentration")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
ggplot(subset(subset_feeds, comp == "PFUnA"), aes(x = protein_source, y = conc)) +
  geom_boxplot() +
  xlab("Protein source") +
  ylab("PFUnA Concentration")
```



```
ggplot(subset(subset_feeds, comp == "PFUnA"), aes(x = mfr, y = conc)) +  
  geom_boxplot() +  
  xlab("Manufacturer") +  
  ylab("PFUnA Concentration")
```

