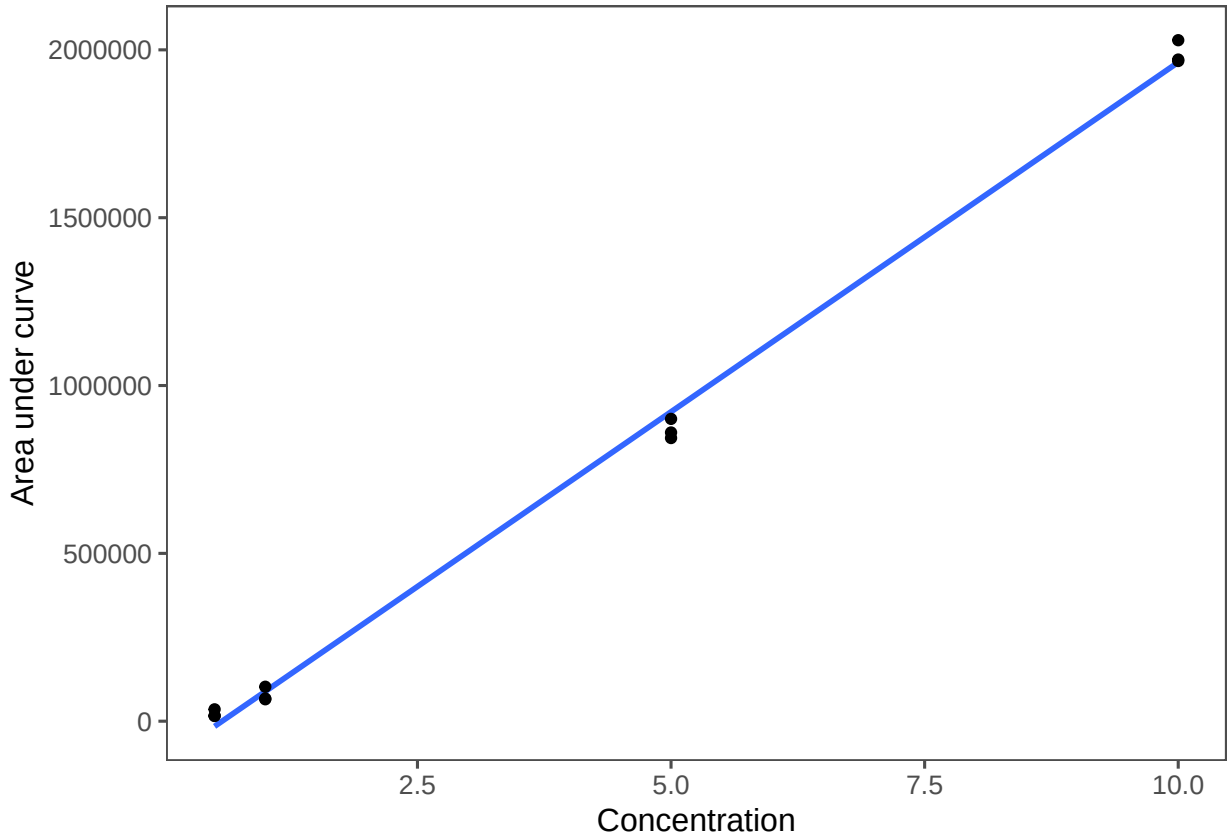
Week 11

Gillian McGinnis

11/20/2020

```
ggplot(standard, aes(x = conc, y = area))+
  stat_smooth(method = "lm", se = FALSE)+
  geom_point()+
  theme_few()+
  labs(x = "Concentration", y = "Area under curve")
```

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

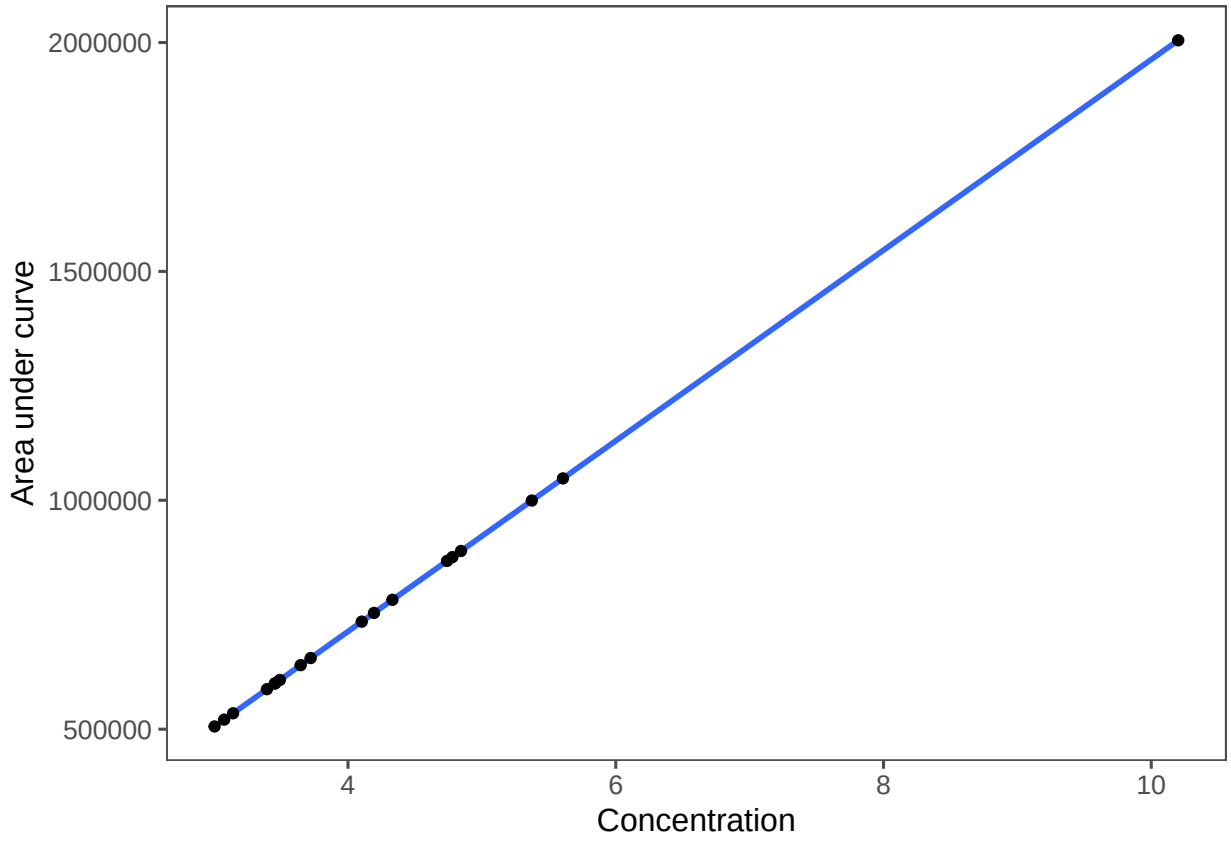



```
curve_results <- summary(lm(area ~ conc, data = standard))

model <- lm(area ~ conc, data = standard)

slope <- model$coefficients[2]
intercept <- model$coefficients[1]
slope_std <- summary(model)$coefficients[2,2]
intercept_std <- summary(model)$coefficients[1,2]
```

```
equation <- tibble(slope, slope_std, intercept, intercept_std)
```

```
conc_results <- lcms %>%
  mutate(conc = (area-intercept)/slope)

ggplot(conc_results, aes(x = conc, y = area))+
  stat_smooth(method = "lm", se = FALSE)+
  geom_point()+
  theme_few()+
  labs(x = "Concentration", y = "Area under curve")
```

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

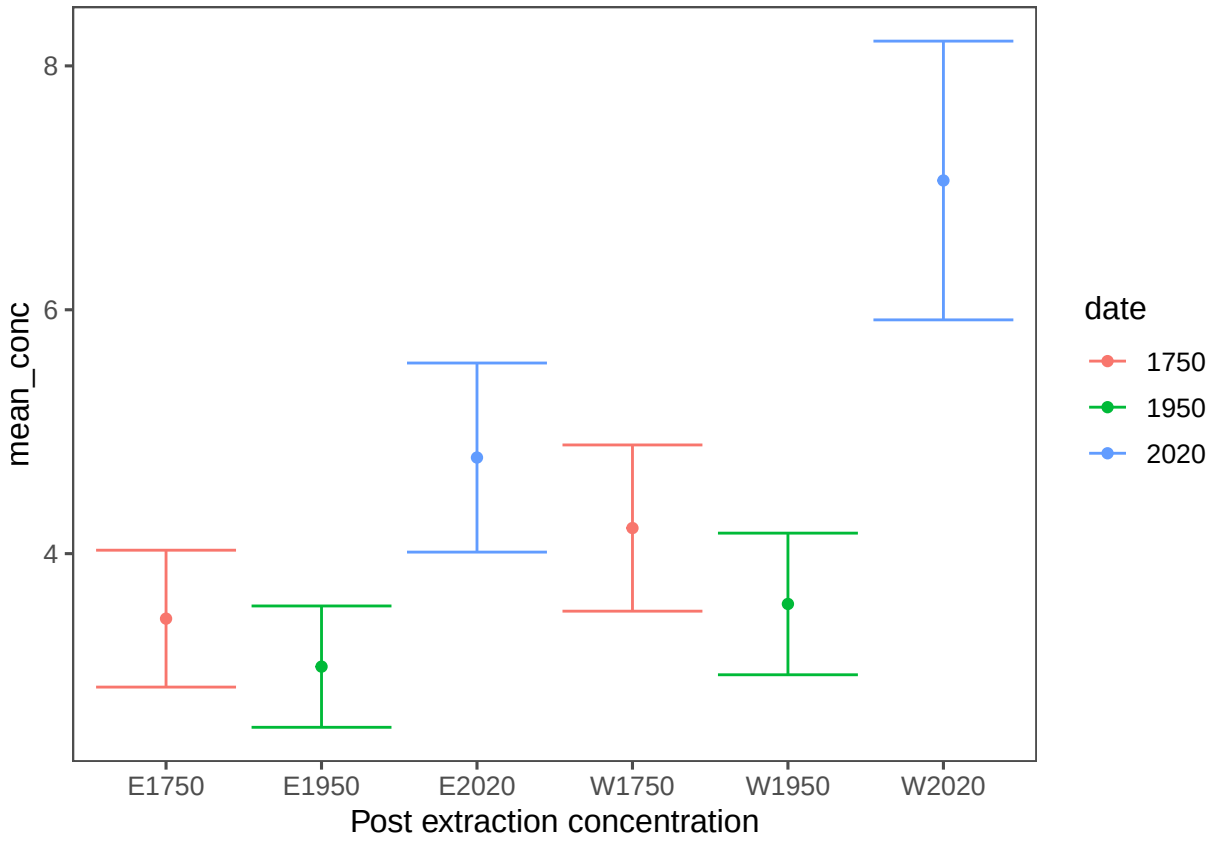



```
conc_summary <- conc_results %>%
  group_by(loc_date) %>%
  summarize(mean_conc = mean(conc),
            sd_conc = sd(conc),
            n = n()) %>%
  #mutate(rsd = sd_conc*mean_conc)
  mutate(#rsd = (100*sd_conc)/mean_conc,
         #e_yb = sqrt(rsd)^2 + intercept_std,
         e_yb = intercept_std, #pretending that there is no error in y b/c minions took very few sample
         yb = mean_conc-intercept,
         e_x = mean_conc*sqrt((e_yb/yb)^2 + (slope_std/slope)^2)) %>%
  select(loc_date, mean_conc, sd_conc, e_x, n) %>%
  rename(conc_error = e_x) %>%
  mutate(upper = mean_conc+conc_error,
```

```
        lower = mean_conc-conc_error) %>%
  mutate(loc = case_when(
    str_detect(loc_date, "E") ~ "E",
    str_detect(loc_date, "W") ~ "W"
  )) %>%
  mutate(date = case_when(
    str_detect(loc_date, "1750") ~ "1750",
    str_detect(loc_date, "1950") ~ "1950",
    str_detect(loc_date, "2020") ~ "2020"
  ))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
ggplot(conc_summary, aes(x = loc_date, y = mean_conc, color = date))+
  geom_point()+
  geom_errorbar(ymin = conc_summary$lower, ymax = conc_summary$upper)+
  labs(x = "Post extraction concentration")+
  theme_few()+
  expand_limits(ymin = conc_summary$lower, ymax = conc_summary$upper)
```

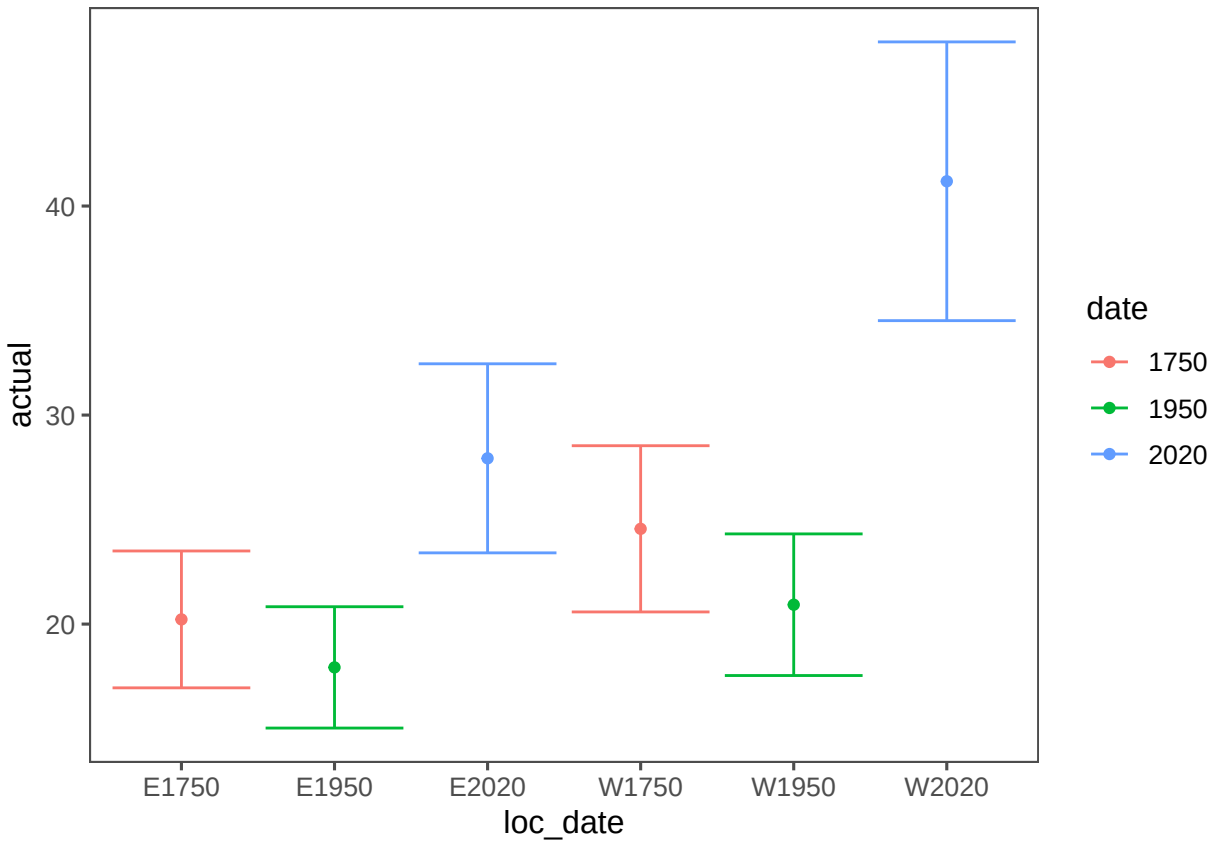



```
  #facet_wrap(~loc)
```

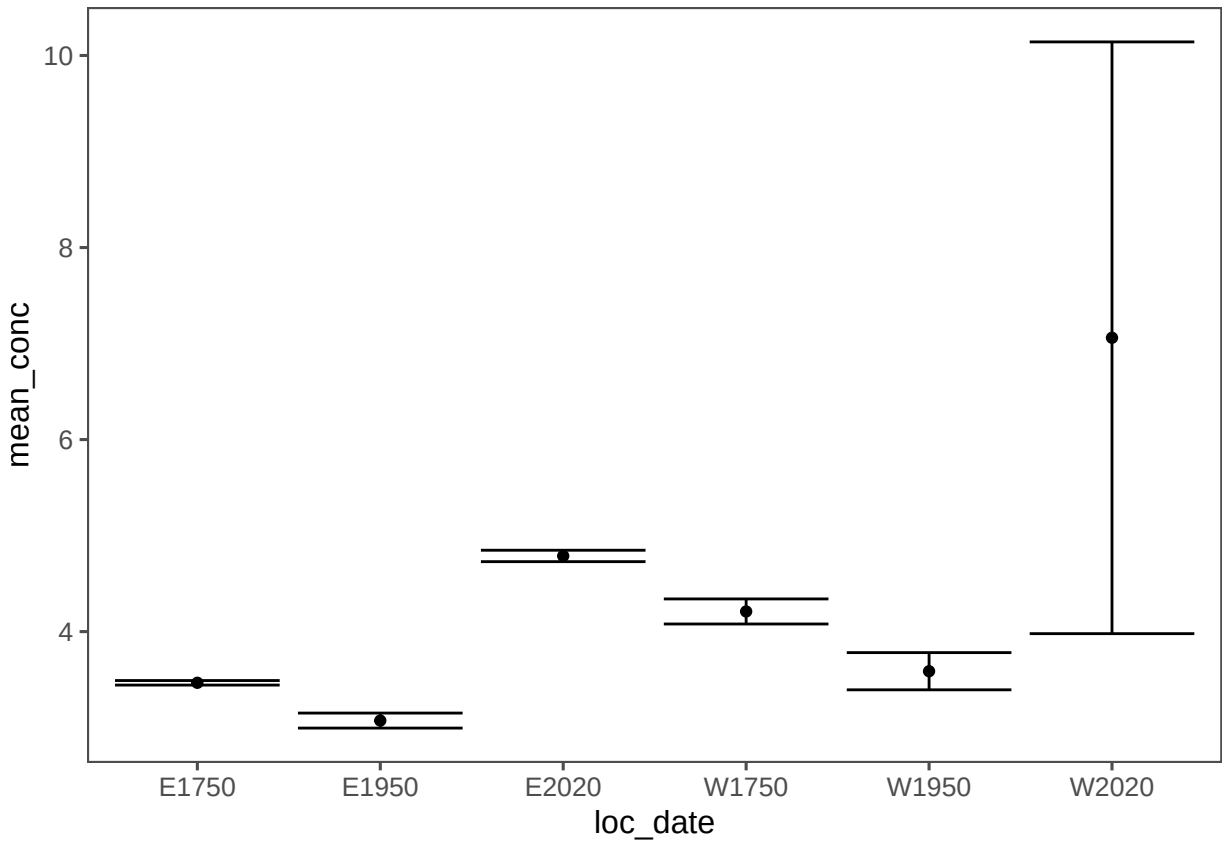
```
## boo
m <- equation$slope
b <- equation$intercept
y <- conc_summary$mean_conc

b_e <- equation$intercept_std
```

```r
m_e <- equation$slope_std

x <- (y-b)/m

RSD <- ((conc_summary$sd_conc)*conc_summary$mean_conc)
#CPS <- sample_data$cps
CONC <- conc_summary$mean_conc

e_yb <- sqrt((RSD)^2 + (b_e)^2)
    #yb <- CPS-b
yb <- CONC-b
e_x <- x*sqrt((e_yb/yb)^2 + (m_e/m)^2)


conc_summary_error <- conc_summary %>%
  mutate()
```

```r
airborne_conc <- conc_summary %>%
  mutate(actual = mean_conc*2*70/24,
         actual_e = conc_error*2*70/24,
         lower = actual-actual_e,
         upper = actual+actual_e) # 2mL dilution factor (1mL water + 1mL ethanol), 70 b/c cut 1in^2, 24

ggplot(airborne_conc, aes(x = loc_date, y = actual, color = date))+
  geom_point()+
  geom_errorbar(ymin = airborne_conc$lower, ymax = airborne_conc$upper)+
  expand_limits(ymin = airborne_conc$lower, ymax = airborne_conc$upper)+
  theme_few()
```




```
conc_ci <- conc_summary %>%
  mutate(se = qnorm(0.975)*sd_conc/sqrt(n),
         lower_ci = mean_conc - se,
         upper_ci = mean_conc + se)

ggplot(conc_ci, aes(x = loc_date, y = mean_conc))+
  geom_point()+
  geom_errorbar(ymin = conc_ci$lower_ci, ymax = conc_ci$upper_ci)+
  expand_limits(ymin = conc_ci$lower_ci, ymax = conc_ci$upper_ci)+
  theme_few()
```




```
grubbs_df <- conc_results %>%
  filter(loc_date == "W2020")
grubbs.test(grubbs_df$conc)
```

```
##
##  Grubbs test for one outlier
##
## data:  grubbs_df$conc
## G = 1.1536480, U = 0.0018223, p-value = 0.04078
## alternative hypothesis: highest value 10.2013934404882 is an outlier
```

```
#p val = 0.04078, that's an outlier, fellers
```

```
anova_df <- NULL
aov_test <- function(unique_site){
  filtered_df <- conc_results %>%
    mutate(actual = conc*2*70/24) %>%
    filter(loc == unique_site)

  anova <- aov(actual ~ date, data = filtered_df) %>%
    tidy()
  anova <- as.data.frame(anova) %>%
    mutate(site = unique_site)
  anova_df <<-rbind(anova_df, anova)
  #return(anova_df)
}

# I know I could write a function for this but it's being finicky
```

```r
aov_test("W")
aov_test("E")

anova_df <- anova_df %>%
    mutate(label = case_when(
    p.value < 0.05 ~ "p < 0.05", # Reject null hypothesiss; diff is significant
    p.value >= 0.05 ~ "Non-Sig" # Fail to reject null hyp; diff is not significant
  ))
```