Assignment 3

Assignment 3

When we merged the company data (AS_merged) with the exploration data expl_oper, we lost approximately 90 observations. Create a data frame of these 90 observations. Why were they dropped? What do they have in common? Are the loss of these observations problematic for our further analysis? How so?

TASK 1

```
expl_AS_full_join = expl_oper %>% left_join(AS_merged, by=c("NPD_id", "year"))

diff_set = setdiff(expl_AS_full_join, expl_AS)
```

Questions: Why were they dropped? What do they have in common? Are the loss of these observations problematic for our further analysis? How so?

Why were they dropped?

They were dropped because the NPD_id and year values in exploration (expl_oper) data set did not correspond with any similar values in the financial data set (AS_merged).For a left join / merge between two rows to be made both the NPD_id and year much find an equal pairing in the other data set.

What do they have in common?

### TASK 2

```
library(tidyverse)
setwd("C:/Users/Jon/OneDrive/NHH/Master/V2021/ENE434/LABS/Datafiles/")
brent_crude_set <- read.csv("Europe_Brent_Spot_Price_FOB.csv") %>%
  tail(.,-4) %>%
  mutate(year = as.numeric(rownames(.)),
         Europe.Brent.Spot.Price.FOB = as.numeric(Europe.Brent.Spot.Price.FOB)) %>%
  rename(spot_price = Europe.Brent.Spot.Price.FOB)
```

Compare the average yearly (brent) oil price with expl_AS. Merging by year.

```
explAGG_spot = explAgg_AS %>%
  left_join(brent_crude_set, by = c("year"))
```

```
explAGG_spot %>% group_by(spot_price) %>%
  summarise(total_n_wells = sum(numbWells)) %>%
  ggplot(aes(x= spot_price, y= total_n_wells)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs()
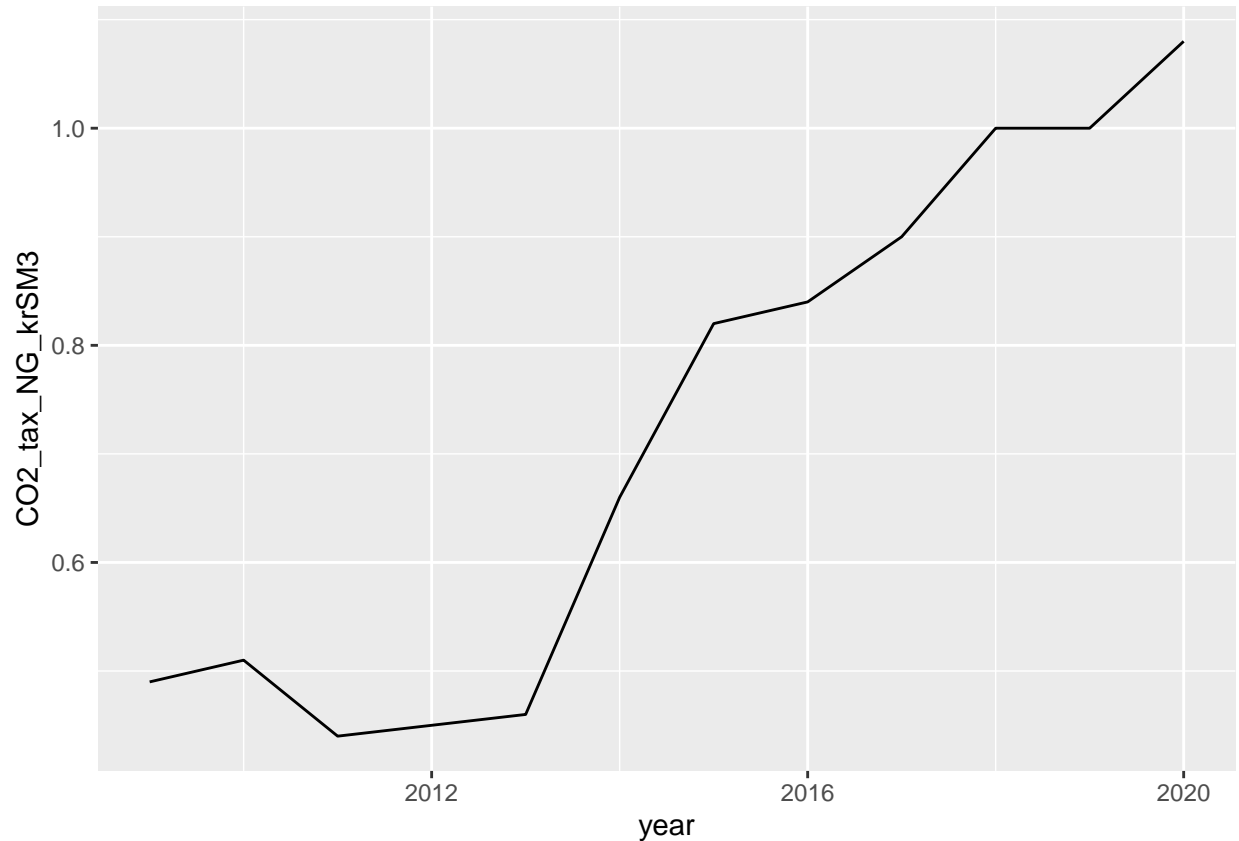```

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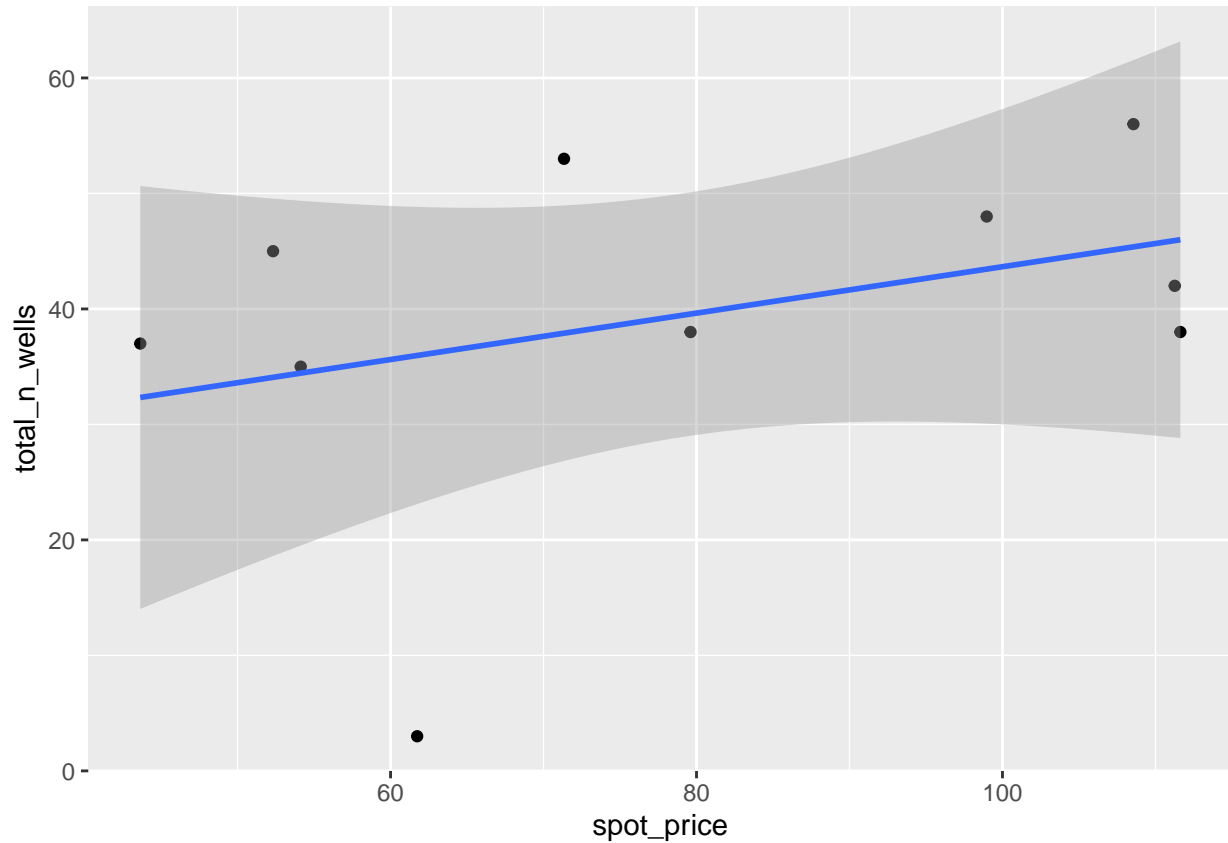


## TASK 3

Overview of the CO2 taxes in the period 2009-2020.

```
CO2_tax = tibble(
  year=2009:2020,
  CO2_tax_NG_krSM3=c(0.49, 0.51, 0.44, 0.45, 0.46, 0.66,.82, 0.84, 0.90, 1, 1, 1.08)
  )

CO2_tax %>% ggplot(aes(x=year, y=CO2_tax_NG_krSM3)) +
  geom_line()
```

Question: Do you think that this may effect exploratory drilling in the period studied? Do you think a correlation could be estimated with the given data (and with the addition of CO2-tax data)? Why or why not?

TASK 4

```r
reg1 = lm(numbWells ~ total_assets + profitability, data=explAgg_AS)
summary(reg1)
```

```
##
## Call:
## lm(formula = numbWells ~ total_assets + profitability, data = explAgg_AS)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.8817 -1.9410 -1.3936  0.4946 11.2441
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.428e+00  3.798e-01   6.391 5.13e-09 ***
## total_assets   2.673e-08  3.491e-09   7.656 1.18e-11 ***
## profitability -4.859e-03  1.015e-02  -0.479    0.633
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Residual standard error: 3.47 on 101 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.3789, Adjusted R-squared:  0.3667
## F-statistic: 30.81 on 2 and 101 DF,  p-value: 3.572e-11
```

Total_assets is more significant than profitability. The R-squared is quite low.