Untitled

Bas Machielsen

28-2-2020

Import data

```
data <- read_excel("Copy of Cotton-mills-wage-data-detailed.xlsx",
                   sheet = 2)
```

Here I reshape the data so that ggplot will treat it

```
data <- data %>%
    pivot_longer(cols = 2:24,
                 names_to = "year",
                 values_to = "rate")
```

Very important equation:

$$E = mc^{2-i} \tag{1}$$

Here I filter the data to exclude the categories mentioned below

```
data2 <- data %>%
    filter(Classes != "Average female wages",
           Classes != "Average male wages",
           Classes != "gender wage gap (f/m)")
```

Here I specify the data which I want to use to generate the names

```
test <- data %>%
    group_by(year) %>%
    mutate(cutoffpos = mean(rate, na.rm=TRUE) + 1.5*IQR(rate, na.rm = TRUE),
           minimum = min(rate, na.rm = TRUE)) %>%
    filter(rate >= cutoffpos | rate == minimum )
```

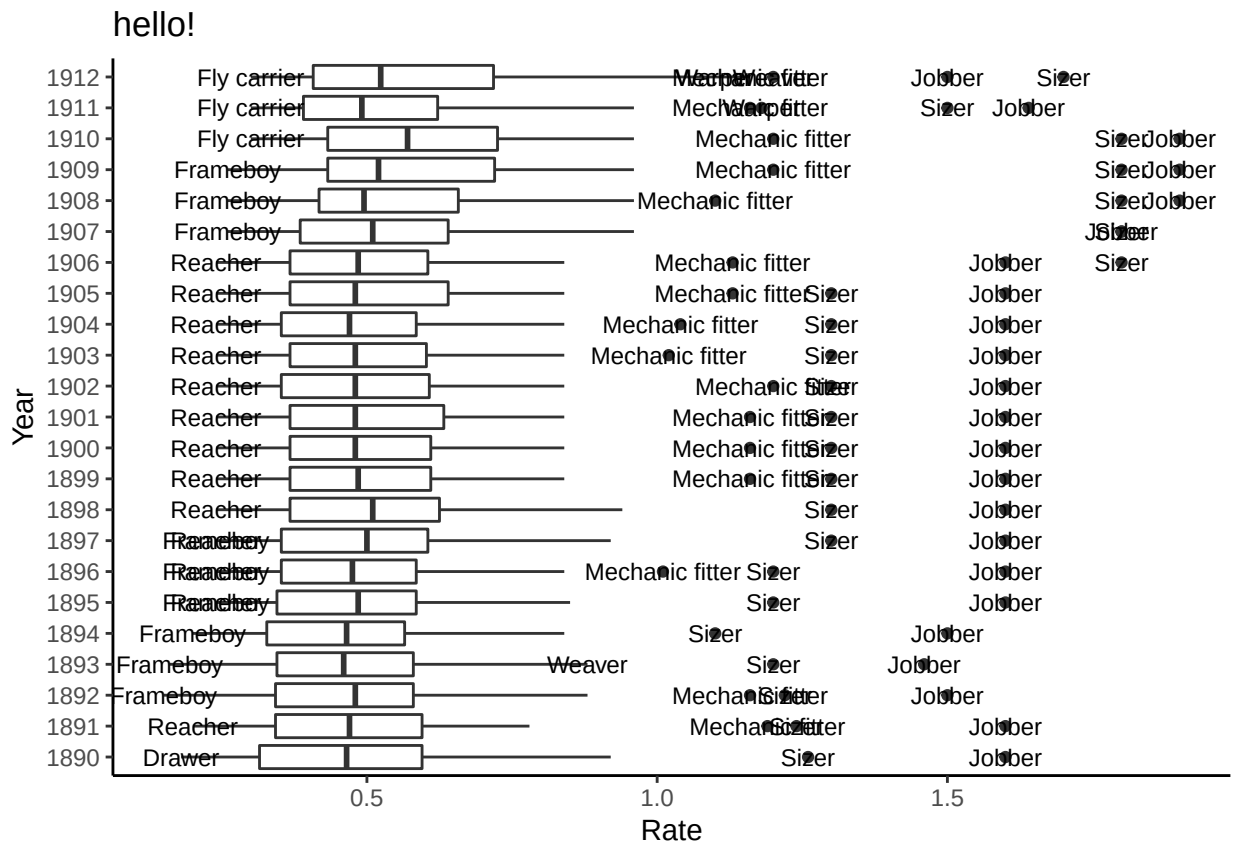
Here I make figure 1

```
figure1 <- ggplot(data2, aes(x = year,
                             y = rate))  +
    geom_boxplot() +
    geom_text(test,
              mapping = aes(
                  x = year,
                  y = rate,
                  label = Classes),
              size = 3) +
    coord_flip() +
    theme_classic() +
```

```
    ggtitle("hello!") + # In this line, you change the title
    xlab("Year") +
    ylab("Rate")
ggsave("figure1.png", figure1)
```

```
## Saving 6.5 x 4.5 in image
```

```
figure1
```



Here I make figure 2

```
figure2 <- ggplot(data2, aes(x = Classes,
                             y = rate,
                             fill = Classes)) +
    geom_boxplot() +
    theme(axis.text.x = element_text(angle = 90, hjust = 1))
ggsave("figure2.png", figure2)
```

```
## Saving 6.5 x 4.5 in image
```

```
figure2
```