

AbsenteeismAtWork

Ahmet Selimhan ŞAHİN

04/02/2022

```
library(lubridate)
library(dplyr)
library(tidyr)
library(knitr)
library(ggplot2)
library(rmarkdown)
library(dygraphs)
library(forcats)
library(readr)
```

```
Absenteeism_at_work <- read_delim("C:/Users/Selimhan/Desktop/files/IE-421 Data Science/Absent
eeism_at_work.csv" , delim = ";", escape_double = FALSE, locale = locale(decimal_mark = ",",
grouping_mark = "."), trim_ws = TRUE)
```

```
employeeNames <- read_delim("C:/Users/Selimhan/Desktop/files/IE-421 Data Science/employeeName
s.csv", delim = ";", escape_double = FALSE, trim_ws = TRUE)
```

```
d <- Absenteeism_at_work
```

```
d <- d %>%
  left_join(employeeNames , by = "ID")
```

```
d2 <- d %>%
  group_by(name) %>%
  summarize(
    count = n() ,
    hours = sum(`Absenteeism time in hours` , na.rm = TRUE) ,
    av.hours = round(hours/count , 1) ,
    distance = mean(`Distance from Residence to Work` , na.rm = TRUE) ,
    age = max(Age , na.rm = TRUE) ,
    dfailure = max(`Disciplinary failure` , na.rm = TRUE)

  )

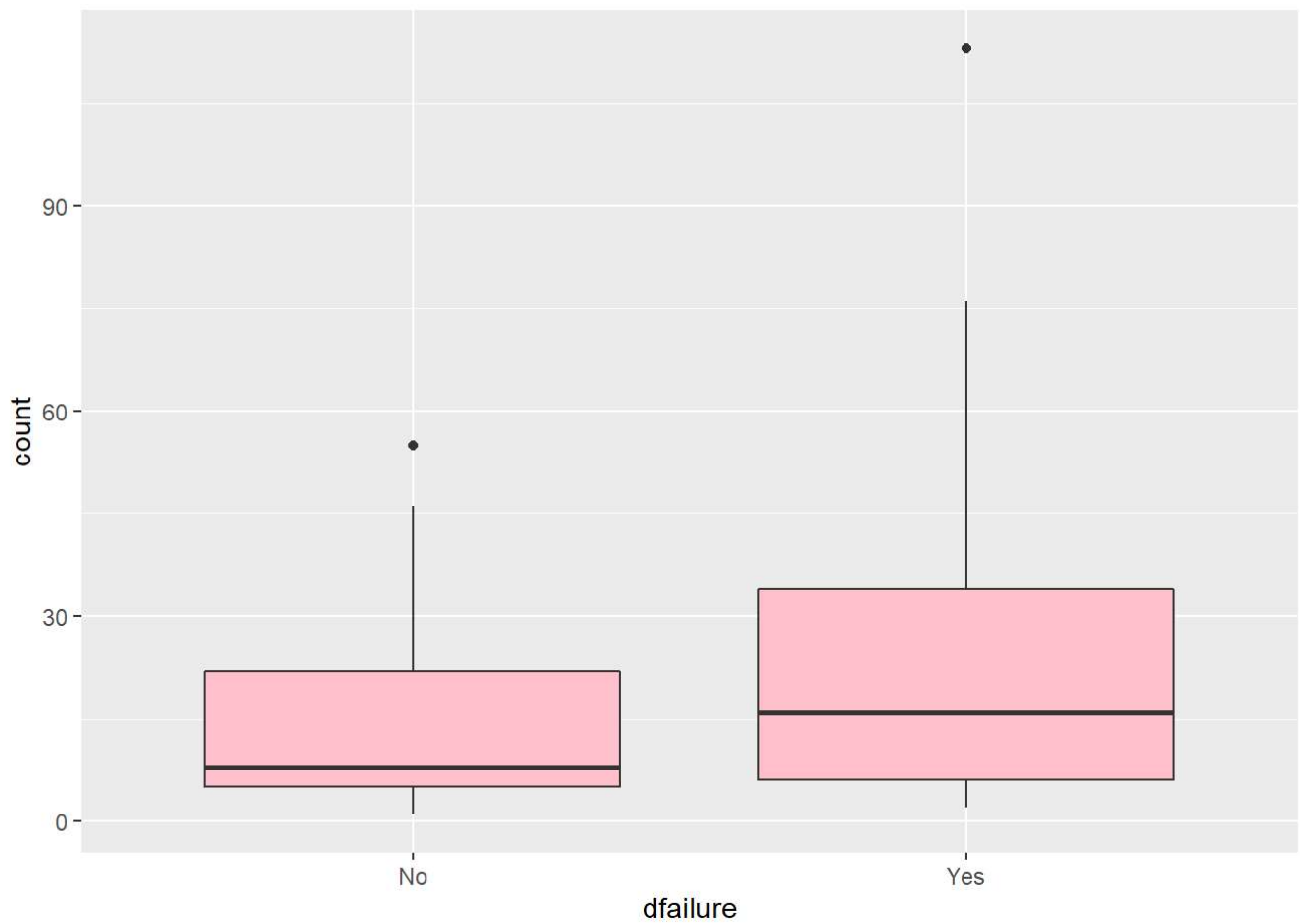
kable(d2)
```

name	count	hours	av.hours	distance	age	dfailure
Arya	6	30	5.0	5.0	39	1
Bran	76	347	4.6	26.0	28	1
Brienne	15	183	12.2	17.0	31	1
Bronn	42	306	7.3	50.0	36	1
Cersei	1	0	0.0	14.0	40	0
Daario	29	476	16.4	12.0	34	0
Daenerys	23	121	5.3	11.0	37	1
Davos	34	311	9.1	13.0	50	1
Drogo	113	482	4.3	51.0	38	1
Eddard	5	83	16.6	26.0	43	0
Gendry	3	16	5.3	11.0	33	1
Gilly	10	42	4.2	16.0	32	0
Gregor	6	25	4.2	29.0	48	1
Hodor	7	31	4.4	27.0	29	1
Jamie	3	6	2.0	50.0	32	1
Jeor	55	344	6.3	10.0	37	0
Joffrey	2	0	0.0	35.0	39	1
Jorah	5	16	3.2	48.0	49	0
Margaery	7	34	4.9	51.0	31	0
Melisandre	24	186	7.8	52.0	28	0
Missandei	24	73	3.0	25.0	47	1
Oberyn	16	118	7.4	16.0	28	1
Petyr	8	262	32.8	14.0	58	0
Ramsay	2	16	8.0	15.0	46	0
Robb	30	254	8.5	25.0	41	1
Samwell	5	21	4.2	17.2	41	1
Sansa	8	72	9.0	29.0	33	0
Shae	40	450	11.2	36.0	33	1
Stannis	20	126	6.3	22.0	40	0
Theon	37	253	6.8	31.0	40	1
Tommen	1	0	0.0	45.0	53	0
Tormund	3	16	5.3	15.0	50	1

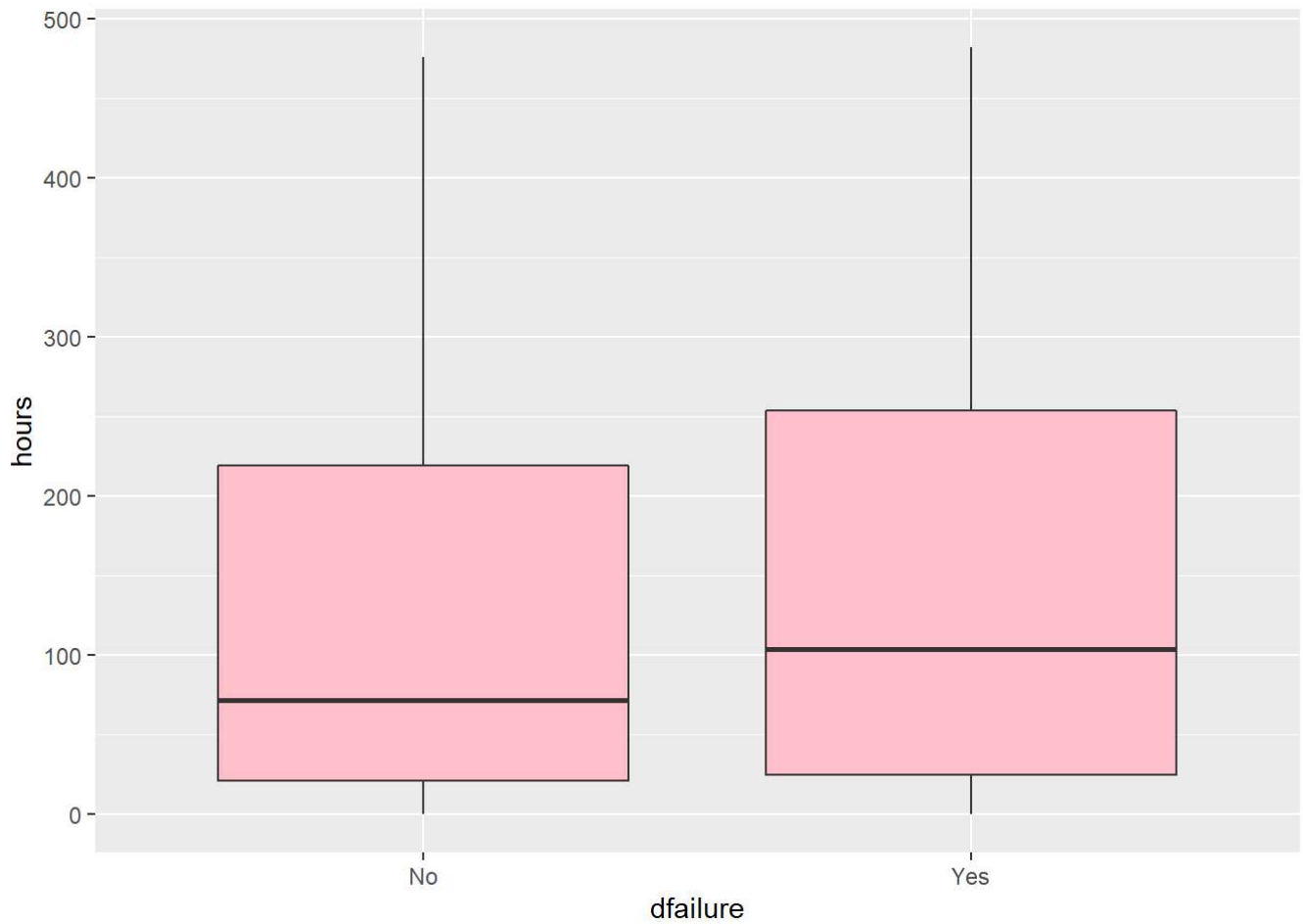
name	count	hours	av.hours	distance	age	dfailure
Tyrion	19	104	5.5	20.0	43	1
Tywin	7	27	3.9	42.0	27	0
Varys	8	40	5.0	49.0	36	1
Ygritte	46	253	5.5	26.0	30	0

```
d2$dfailure <- factor(d2$dfailure , labels = c("No","Yes"))
```

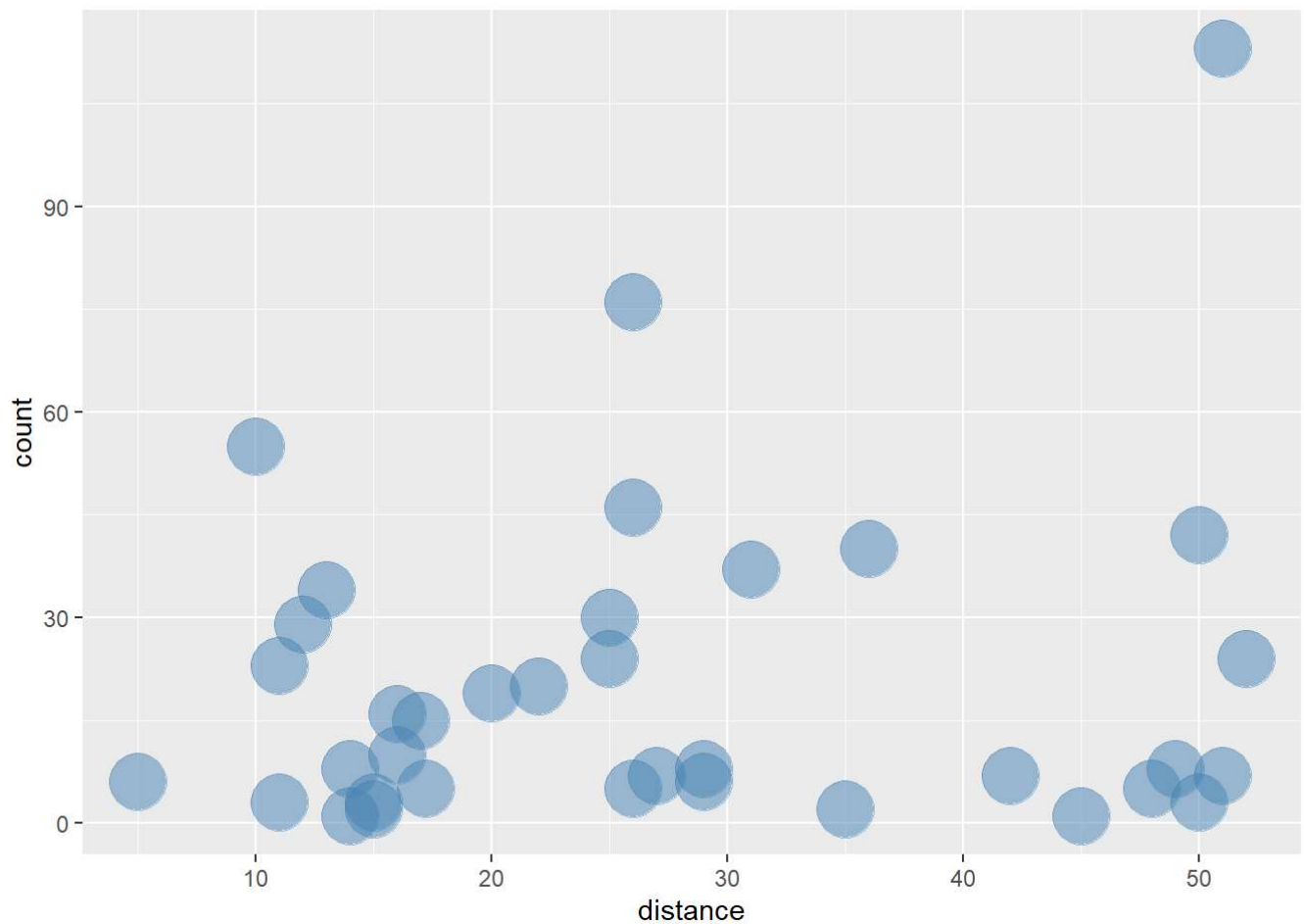
```
ggplot(d2) + geom_boxplot(aes(x = dfailure , y = count ) , fill = "pink" )
```



```
ggplot(d2) + geom_boxplot(aes(x = dfailure , y = hours ) , fill = "pink" )
```



```
ggplot(d2) + geom_point(aes(x = distance , y = count) , size = 10 , color = "steelblue" , alpha = 0.5)
```



```
ggplot(d2 , aes(x = distance , y = count , label = name , size = av.hours , color = dfailure)) + geom_point(alpha = 0.3) + geom_text(aes(label = name) , size = 3) + scale_size_continuous(name = "Average Hours" , range = c(0,20)) + scale_colour_discrete(name = "Disciplinary Failure")
```

