lab-07-simpsons.Rmd

rawan saad

17 March 2021

Packages

```
library(tidyverse)
library(mosaicData)
```

Exercises

1.

```
view(Whickham)
```

Your answer: Observational , because we just watching the people and write the notes also we didn't invoiving in any sitwation or controlling any variables.

2.

```
nrow(Whickham)
```

```
## [1] 1314
```

Your answer; 1314 , represent recorded participants' age , smoking status at baseline.

3.

```
ncol(Whickham)
```

```
## [1] 3
```

Your answer:
3 ,

```
class(Whickham$age)
```

```
## [1] "integer"
```

```
class(Whickham$smoker)
```

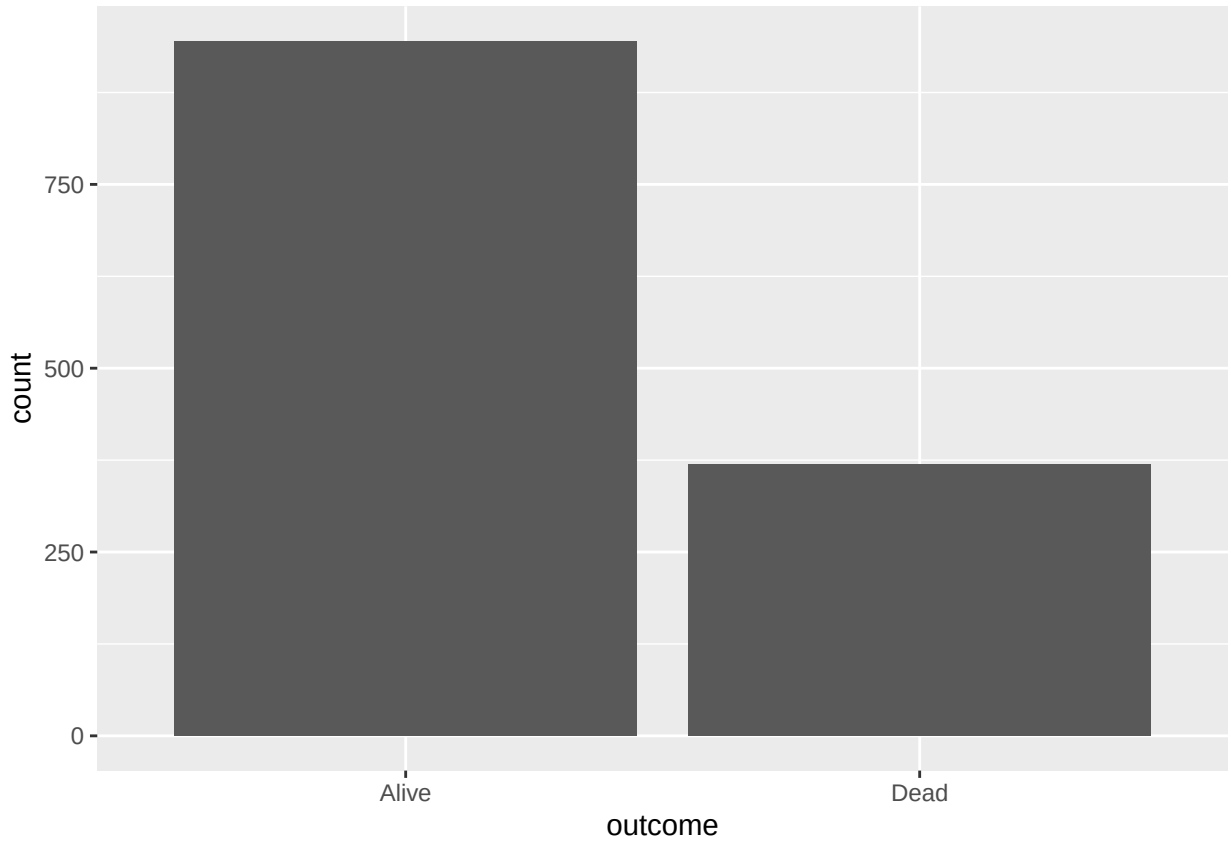
```
## [1] "factor"
```

```
class(Whickham$outcome)
```

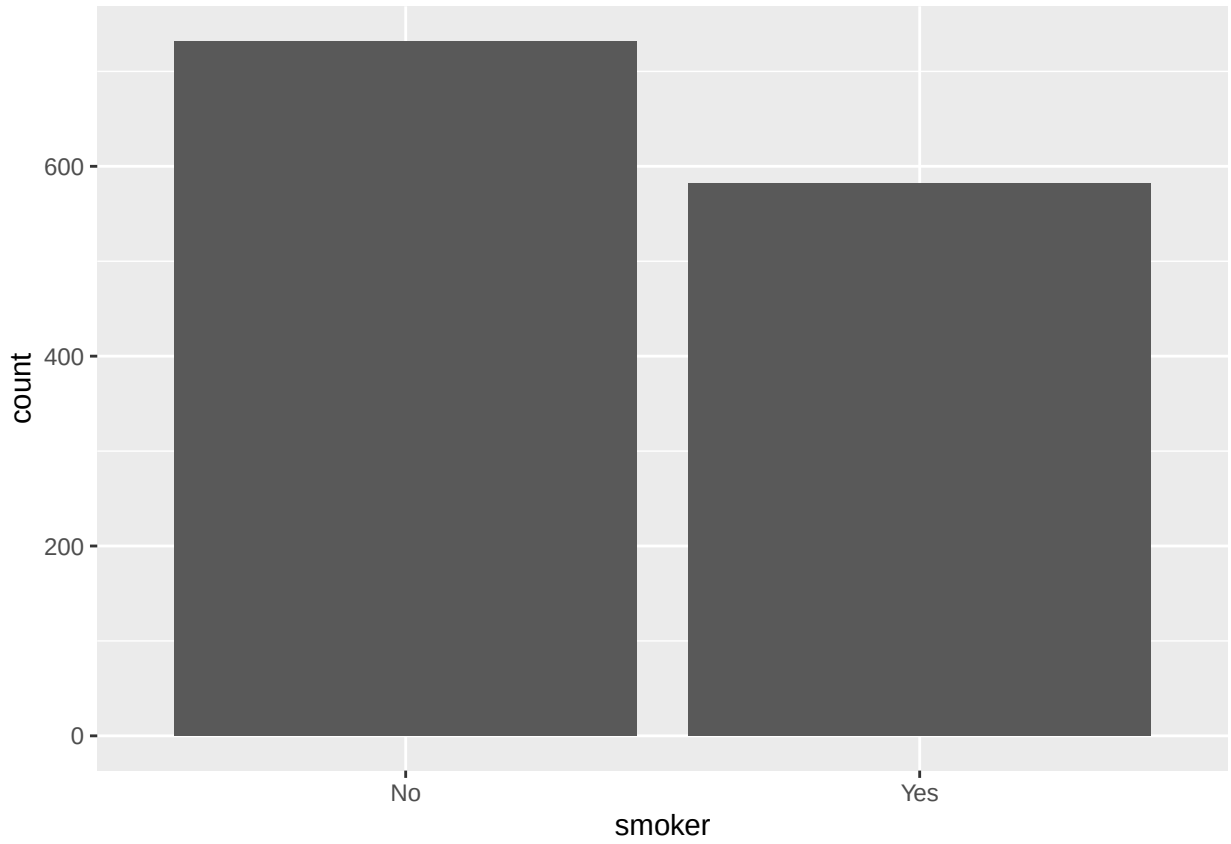
```
## [1] "factor"
```

Your answer: age (Numerical) , smoker and outcome are categorical age (integer) , smoker and outcome are (factor)

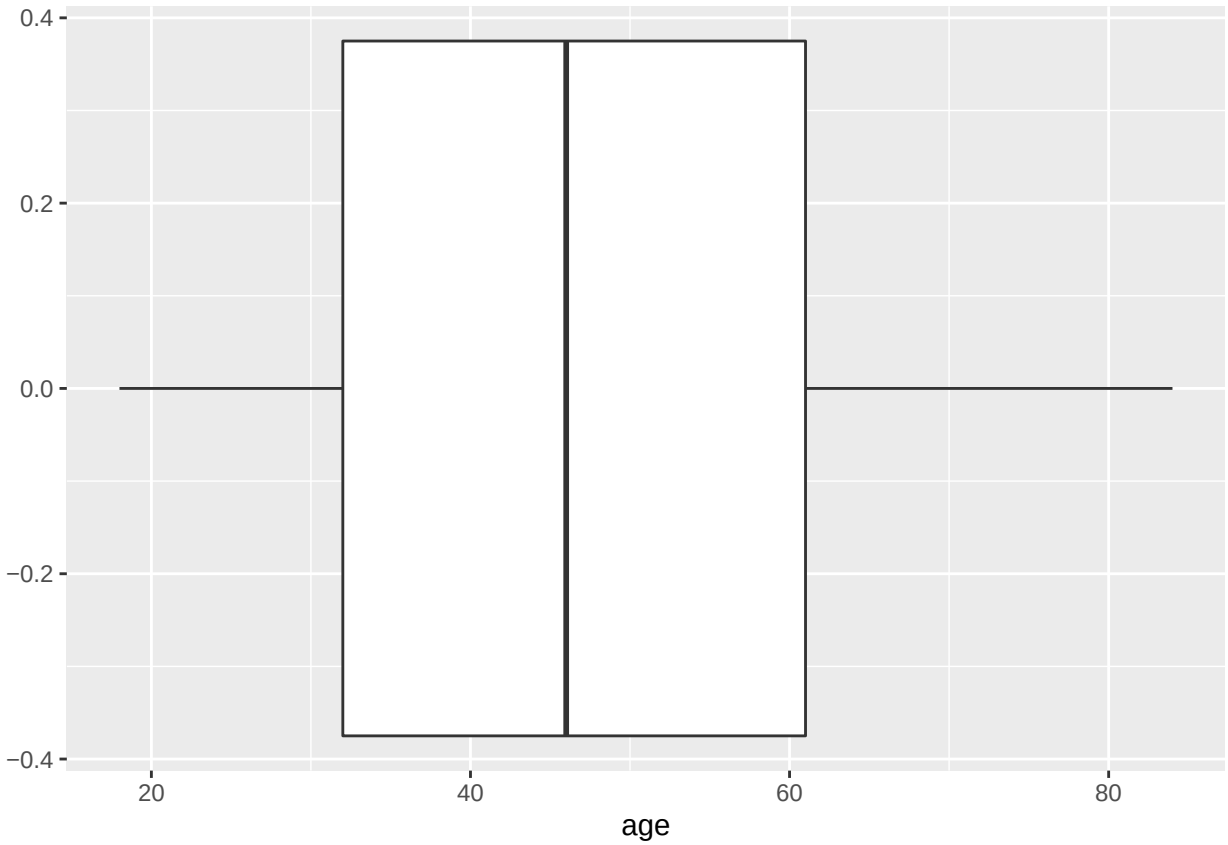
```
ggplot(Whickham, aes(x = outcome)) +
  geom_bar()
```




```
ggplot(Whickham, aes(x = smoker)) +
  geom_bar()
```

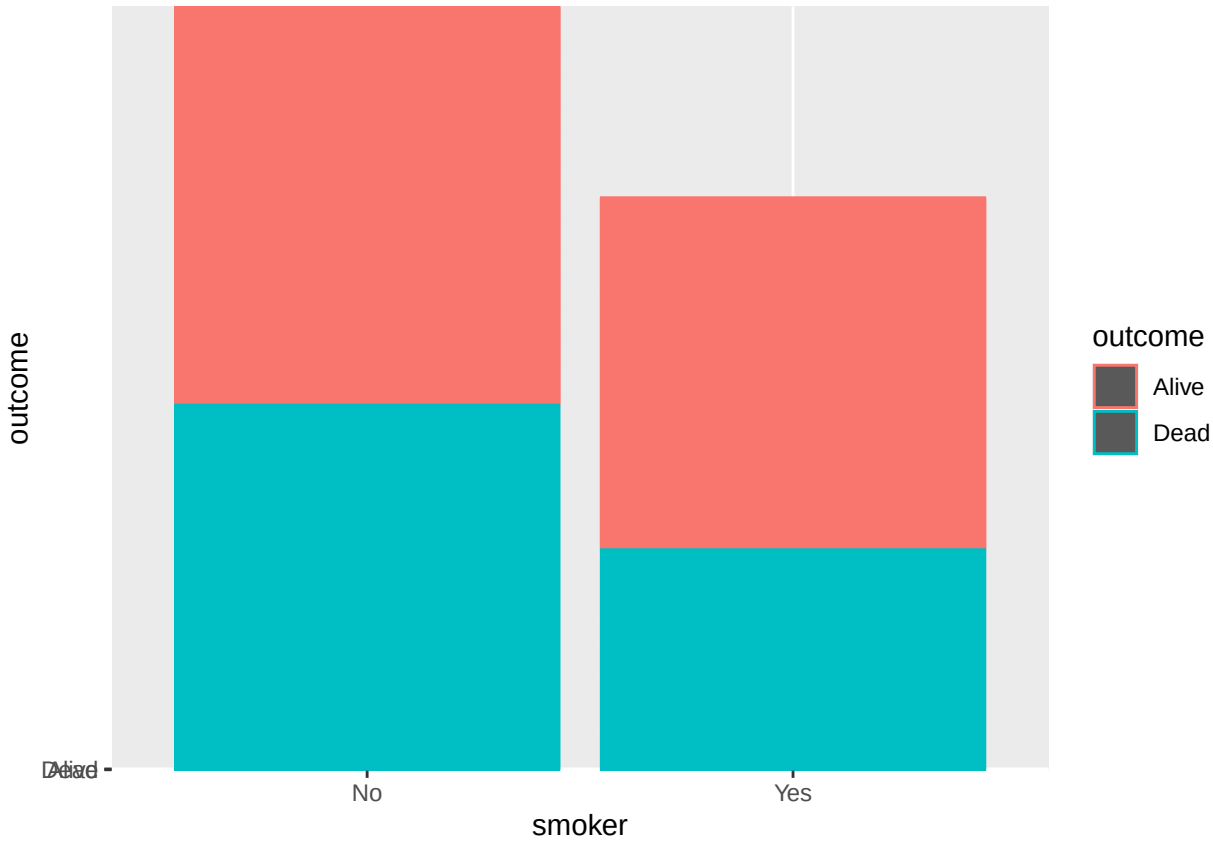



```
ggplot(Whickham, aes(x = age)) +
  geom_boxplot()
```



4. I expect the health will be worser and may be the person will be died after while , if he keeping smoke

```
ggplot(data=Whickham, aes(x=smoker, y=outcome, color=outcome)) + geom_bar(stat="identity")
```



Knit, commit, and push to github.

5.

```
Whickham %>%
  count(smoker, outcome)
```

```
##   smoker outcome   n
## 1     No   Alive 502
## 2     No    Dead 230
## 3    Yes   Alive 443
## 4    Yes    Dead 139
```

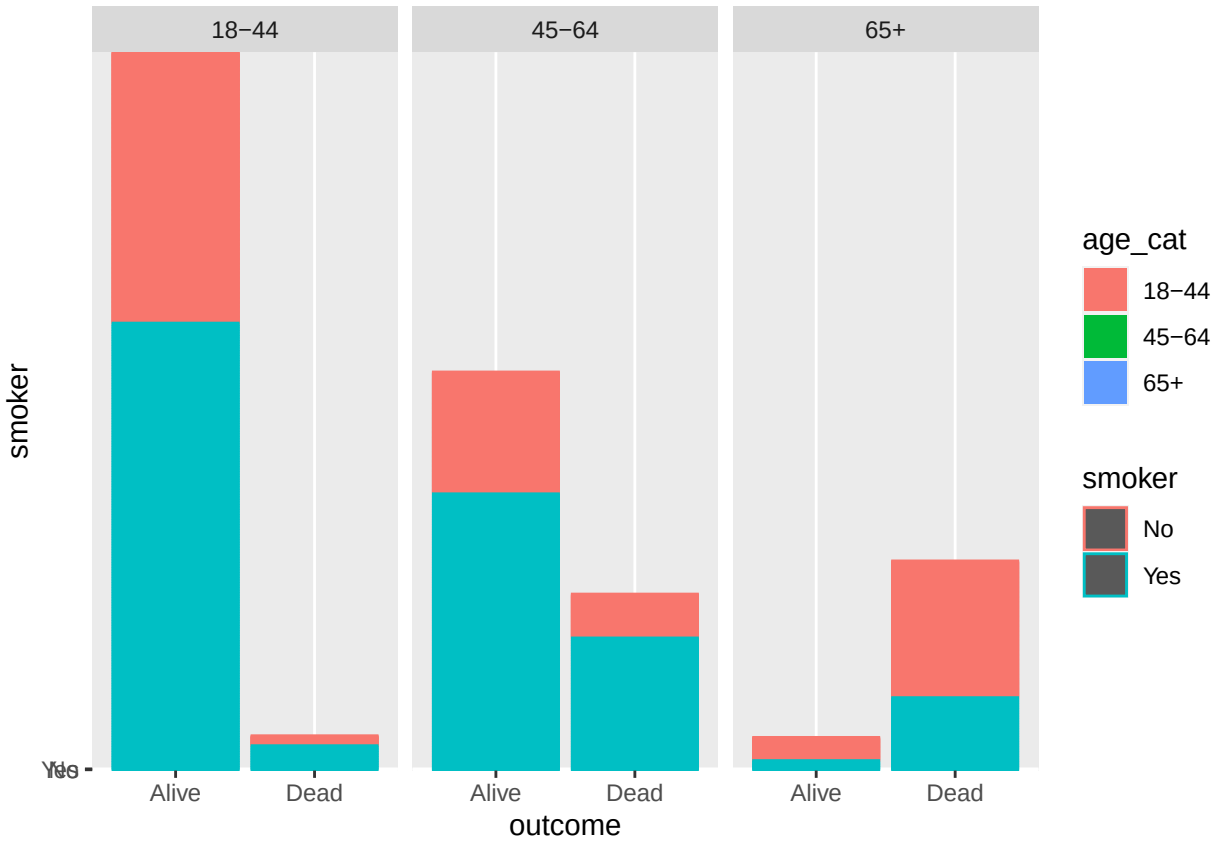
smoker (732) NO > 31,4 (Dead) »> (68.6) Alive smoker (582) YES > 23,8 (Dead) »> (76.2) Alive

I does expected this result because now the most died people not smoker. 6.

```
 Whickham <- Whickham %>% mutate(age_cat = case_when (age <= 44 ~ "18-44", age > 44. & age <= 64 ~ "45-
```

7.

```
ggplot(data =Whickham, aes(x=outcome, y= smoker,color=smoker, fill=age_cat)) + geom_bar(stat="identity"
```



what changes > the category of the age it's appear to us and we see the most of dead people not smoker in age (65+) .. but in age (45-64) and (18-44) the most dead people are smoker that is relationship between the smoking and health not clearly but can say that your health will be change to worst if you be smoker.

Knit, commit, and push to github.