lab-07-simpsons.Rmd

Manar mansour

3 March 2021

ID:2201002883

Packages

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

Exercises

1.

```
?Whickham
```

Your answer: observational , Because they notice a person's health status after a period of time 2.

```
nrow(Whickham)
```

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

Your answer; obs 1314
Each row represents whether the person is alive or dead, is he a smoker or not, and how old he is 3.

```
names(Whickham)
```

```
## [1] "outcome" "smoker"  "age"
```

Your answer: 3 var , "outcome" "smoker" "age"

```
unique(Whickham$outcome)
```

```
## [1] Alive Dead
## Levels: Alive Dead
```

```
unique(Whickham$smoker)
```

```
## [1] Yes No
## Levels: No Yes
```

```
unique(Whickham$age)
```

```
##  [1] 23 18 71 67 64 38 45 76 28 27 34 20 72 48 66 30 33 68 61 43 47 22 39 80 59
## [26] 56 62 51 32 60 37 36 50 55 73 52 25 53 31 54 69 79 75 21 29 24 26 49 84 40
## [51] 44 74 46 35 77 57 42 81 19 63 78 83 82 70 58 41 65
```
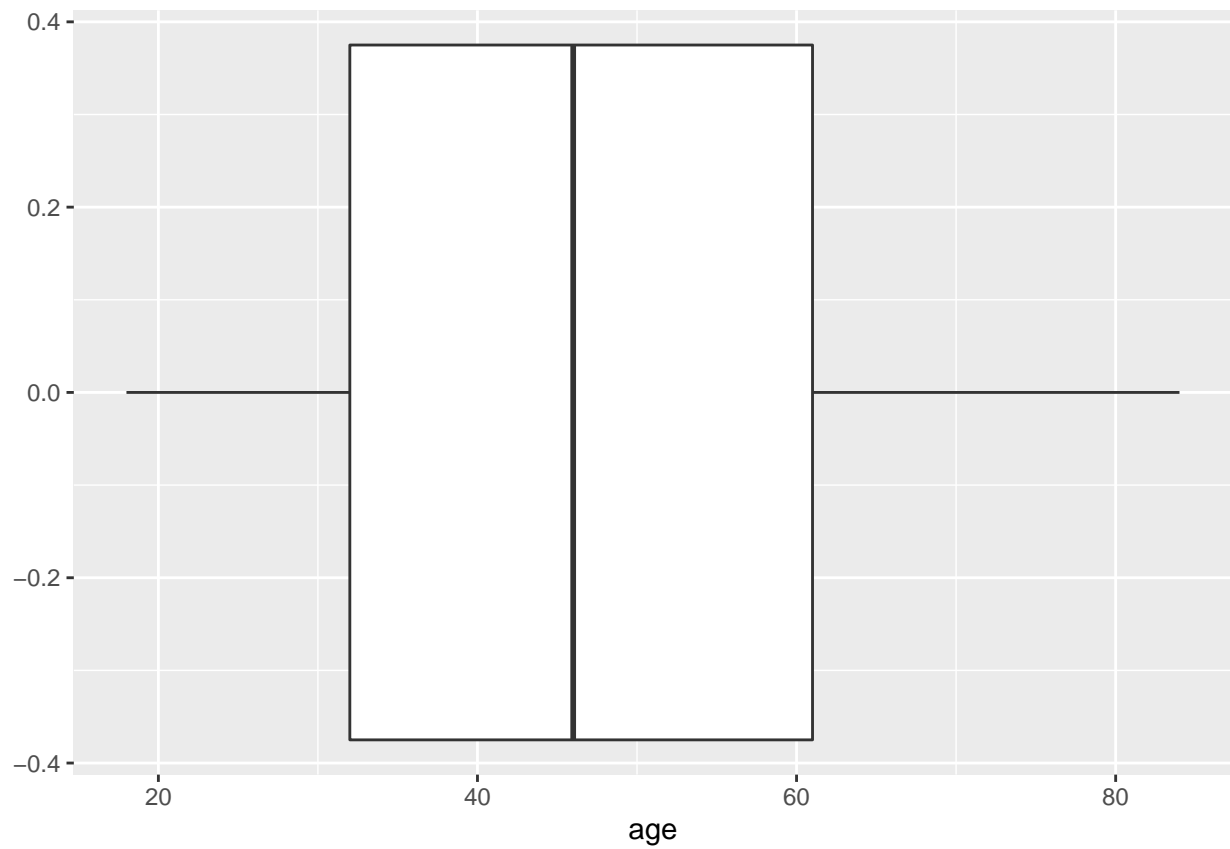
Your answer: alive dead,yes,no

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

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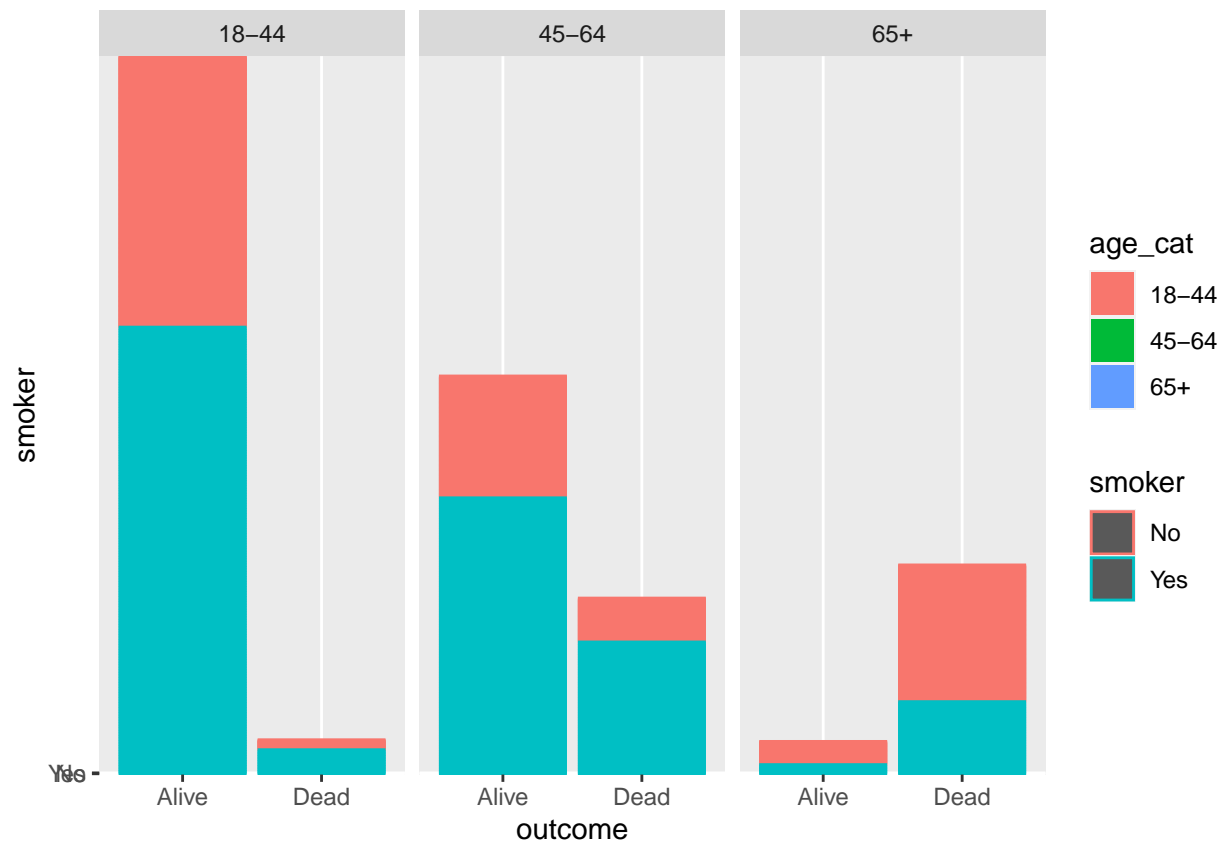
```
ggplot(Whickham, aes(x = age)) +
  geom_boxplot()
```

4.

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

ans: Their continued smoking will lead to their death

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
```

```
502+230
```

```
## [1] 732
```

```
230/732
```

```
## [1] 0.3142077
```

The number of deaths among non-smokers is greater than the number of smokers 6.

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

7.

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

Knit, commit, and push to github.