

ica6\_Shuangyu\_Zhao

shuangyu\_zhao

2023-01-26

2.

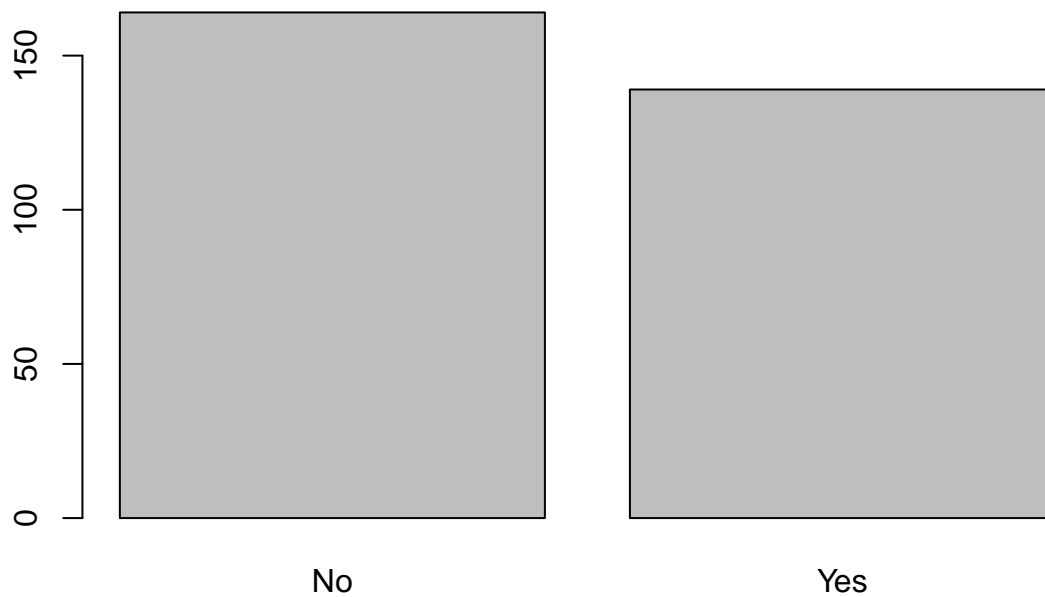
```
heart <- read.csv("/Users/apple/Desktop/STT811 appl_stat_model/data/Heart.csv")
head(heart)
```

```
##   X Age Sex   ChestPain RestBP Chol Fbs RestECG MaxHR ExAng Oldpeak Slope Ca
## 1 1  63  1     typical   145  233  1         2   150    0     2.3    3  0
## 2 2  67  1 asymptomatic   160  286  0         2   108    1     1.5    2  3
## 3 3  67  1 asymptomatic   120  229  0         2   129    1     2.6    2  2
## 4 4  37  1   nonanginal   130  250  0         0   187    0     3.5    3  0
## 5 5  41  0   nontypical   130  204  0         2   172    0     1.4    1  0
## 6 6  56  1   nontypical   120  236  0         0   178    0     0.8    1  0
##           Thal AHD
## 1         fixed  No
## 2         normal Yes
## 3 reversible Yes
## 4         normal  No
## 5         normal  No
## 6         normal  No
```

```
table(heart$AHD)
```

```
##
##   No Yes
## 164 139
```

```
barplot(table(heart$AHD))
```



balanced.

3.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
## intersect, setdiff, setequal, union
```

```
glimpse(heart)
```

```
## Rows: 303
```

```
## Columns: 15
```

```
## $ X      <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 1~
```

```
## $ Age     <int> 63, 67, 67, 37, 41, 56, 62, 57, 63, 53, 57, 56, 56, 44, 52, ~
```

```
## $ Sex     <int> 1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, ~
```

```
## $ ChestPain <chr> "typical", "asymptomatic", "asymptomatic", "nonanginal", "no~
## $ RestBP    <int> 145, 160, 120, 130, 130, 120, 140, 120, 130, 140, 140, 140, ~
## $ Chol      <int> 233, 286, 229, 250, 204, 236, 268, 354, 254, 203, 192, 294, ~
## $ Fbs       <int> 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, ~
## $ RestECG   <int> 2, 2, 2, 0, 2, 0, 2, 0, 2, 2, 0, 2, 2, 0, 0, 0, 0, 0, 0, 0, ~
## $ MaxHR     <int> 150, 108, 129, 187, 172, 178, 160, 163, 147, 155, 148, 153, ~
## $ ExAng     <int> 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, ~
## $ Oldpeak    <dbl> 2.3, 1.5, 2.6, 3.5, 1.4, 0.8, 3.6, 0.6, 1.4, 3.1, 0.4, 1.3, ~
## $ Slope     <int> 3, 2, 2, 3, 1, 1, 3, 1, 2, 3, 2, 2, 2, 1, 1, 1, 3, 1, 1, 1, ~
## $ Ca        <int> 0, 3, 2, 0, 0, 0, 2, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, ~
## $ Thal      <chr> "fixed", "normal", "reversable", "normal", "normal", "normal~
## $ AHD       <chr> "No", "Yes", "Yes", "No", "No", "No", "Yes", "No", "Yes", "Y~
```

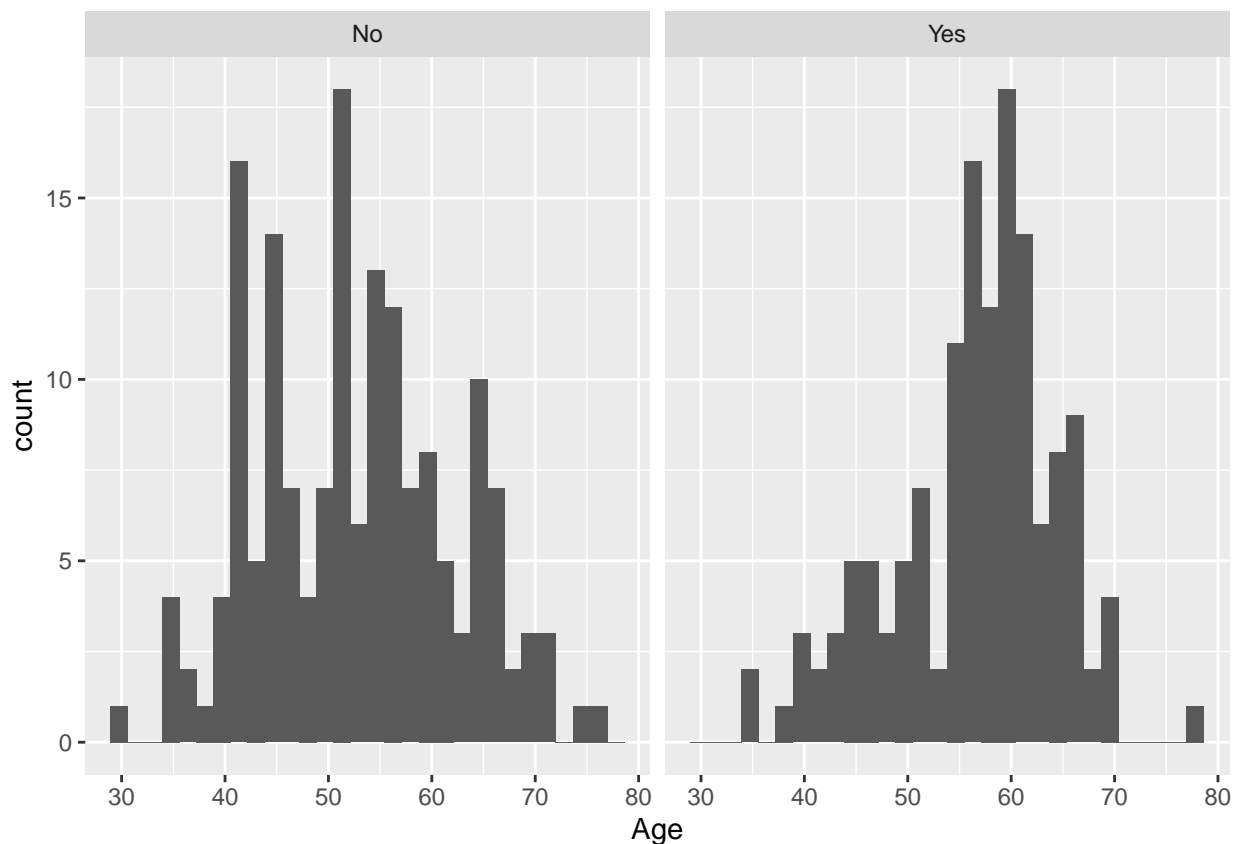
truly numeric: Age, RestBP, Chol, MaxHR, Oldpeak

categorical: Sex, ChestPain, Fbs, RestECG, ExAng, Slope, Ca, Thal

#### 4. age

```
library(ggplot2)
ggplot(data = heart, aes(x = Age)) + geom_histogram() + facet_grid(.~AHD)
```

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.



```
quantile(filter(heart, AHD == 'Yes')$Age, seq(0,1, by=0.1))
```

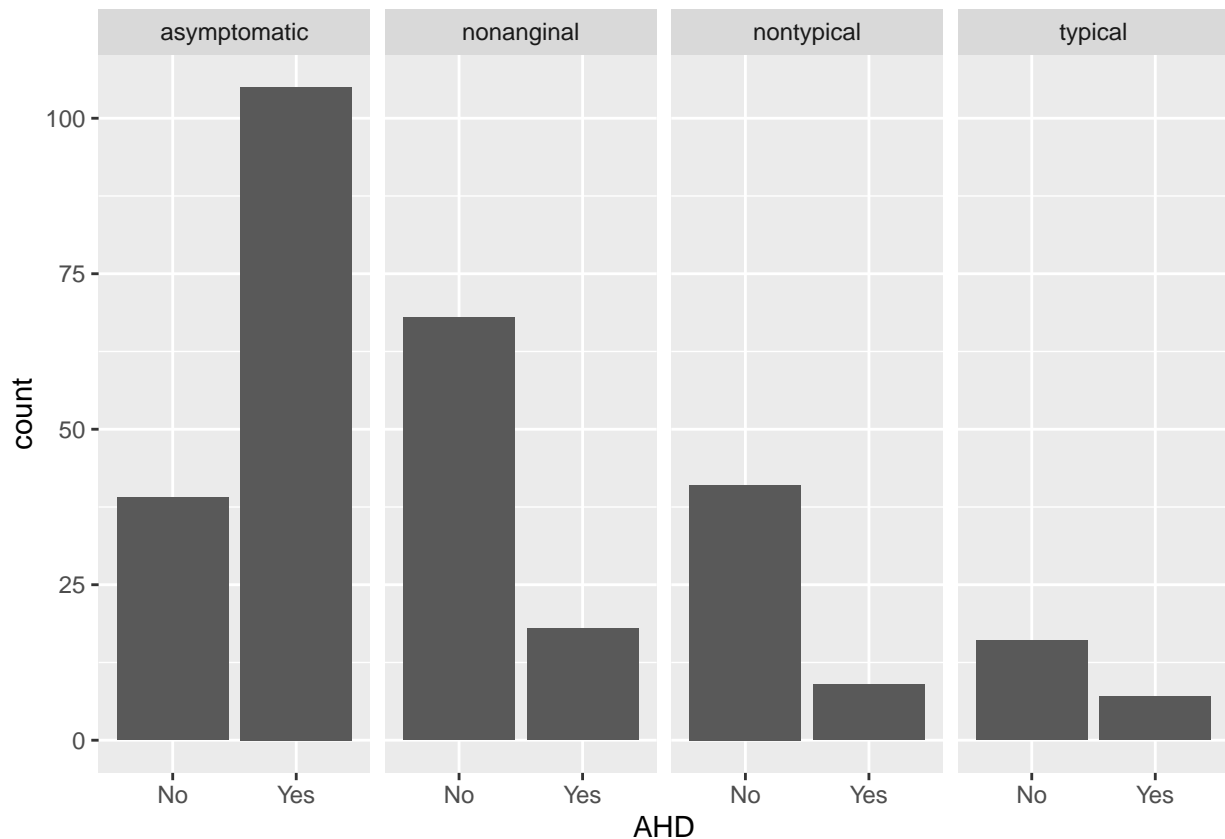
```
## 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
## 35.0 44.8 50.0 54.0 57.0 58.0 59.0 61.0 63.0 66.0 77.0
```

```
quantile(filter(heart, AHD == 'No')$Age, seq(0,1, by=0.1))
```

```
## 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
## 29.0 41.0 43.6 46.0 51.0 52.0 54.0 58.0 62.0 65.7 76.0
```

ChestPain

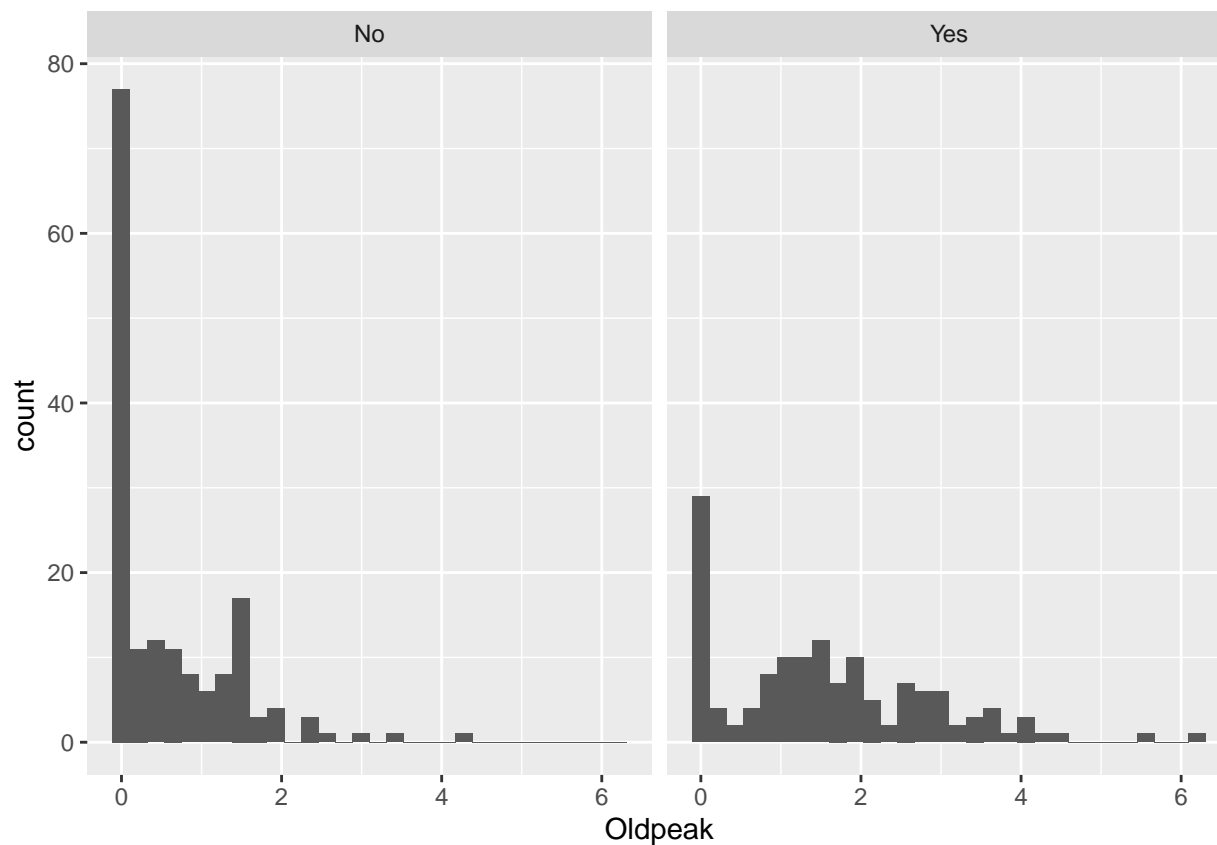
```
ggplot(data = heart, aes(x = AHD)) + geom_bar() + facet_grid(.~ChestPain)
```



OldPeak

```
ggplot(data = heart, aes(x = Oldpeak)) + geom_histogram() + facet_grid(.~AHD)
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
quantile(filter(heart, AHD == 'Yes')$Oldpeak, seq(0,1, by=0.1))
```

```
##    0%   10%   20%   30%   40%   50%   60%   70%   80%   90%  100%
## 0.00 0.00 0.10 0.80 1.00 1.40 1.80 2.16 2.80 3.40 6.20
```

```
quantile(filter(heart, AHD == 'No')$Oldpeak, seq(0,1, by=0.1))
```

```
##    0%   10%   20%   30%   40%   50%   60%   70%   80%   90%  100%
## 0.0  0.0  0.0  0.0  0.0  0.2  0.5  0.8  1.2  1.6  4.2
```

Thal

```
ggplot(data = heart, aes(x = AHD)) + geom_bar() + facet_grid(.~Thal)
```

