HW 3

2019150432 임효진

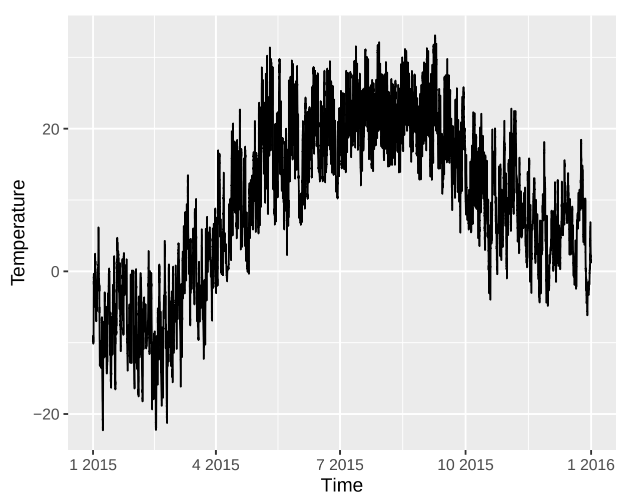
October 31, 2020

1

```r
library(macleish)
library(ggplot2)
library(dplyr)
head(whately_2015)
```

```
## # A tibble: 6 x 8
##   when                temperature wind_speed wind_dir rel_humidity pressure
##   <dttm>                    <dbl>      <dbl>    <dbl>        <dbl>    <int>
## 1 2015-01-01 00:00:00       -9.32       1.40     225.         54.6      985
## 2 2015-01-01 00:10:00       -9.46       1.51     248.         55.4      985
## 3 2015-01-01 00:20:00       -9.44       1.62     258.         56.2      985
## 4 2015-01-01 00:30:00       -9.3        1.14     244.         56.4      985
## 5 2015-01-01 00:40:00       -9.32       1.22     238.         56.9      984
## 6 2015-01-01 00:50:00       -9.34       1.09     242.         57.2      984
## # ... with 2 more variables: solar_radiation <dbl>, rainfall <int>
```

```r
whately_2015%>%ggplot(aes(x=when, y=temperature))+
    geom_line()+
    xlab("Time")+
    ylab("Temperature")
```



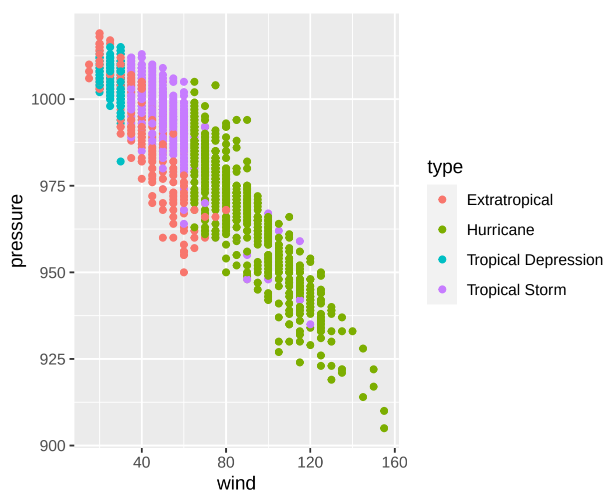
2

(a)

```
library(nasaweather)
head(storms)
```

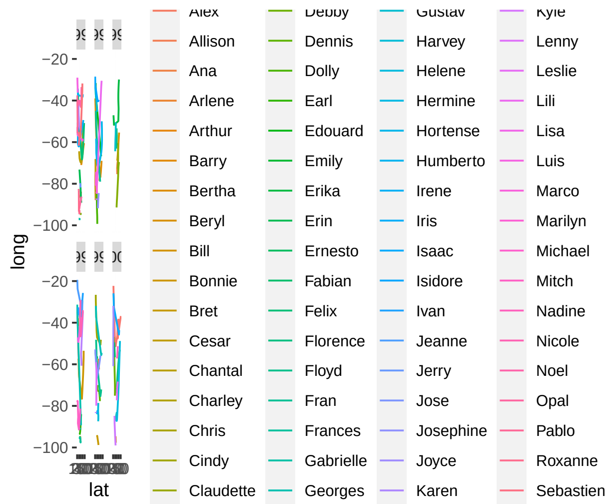
```
## # A tibble: 6 x 11
##    name    year month   day  hour   lat  long pressure  wind type         seasday
##    <chr>  <int> <int> <int> <int> <dbl> <dbl>    <int> <int> <chr>          <int>
## 1 Allis~  1995     6     3     0  17.4 -84.3     1005    30 Tropical De~       3
## 2 Allis~  1995     6     3     6  18.3 -84.9     1004    30 Tropical De~       3
## 3 Allis~  1995     6     3    12  19.3 -85.7     1003    35 Tropical St~       3
## 4 Allis~  1995     6     3    18  20.6 -85.8     1001    40 Tropical St~       3
## 5 Allis~  1995     6     4     0  22    -86       997    50 Tropical St~       4
## 6 Allis~  1995     6     4     6  23.3 -86.3      995    60 Tropical St~       4
```

```
storms%>%ggplot(aes(wind, pressure))+
    geom_point(aes(col=type))
```



(b)

```
storms%>%filter(type=="Tropical Storm")%>%
    ggplot(aes(lat, long))+
    geom_path(aes(col=name))+
    facet_wrap(~year)
```

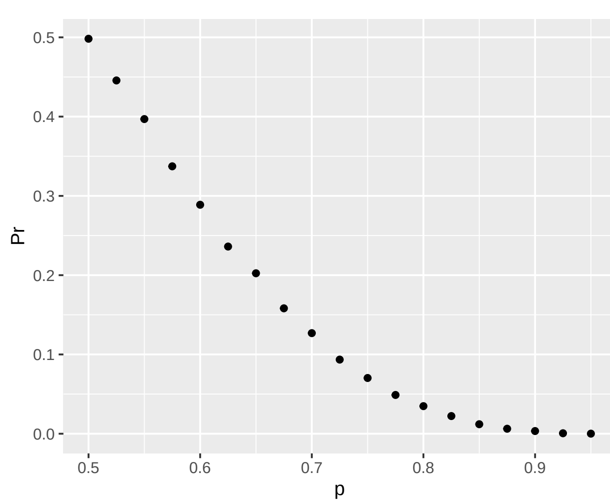


3

(a)

```
prob_win <- function(p){
B <- 10000
result <- replicate(B, {
b_win <- sample(c(1,0), 7, replace = TRUE, prob = c(1-p, p))
sum(b_win)>=4
})
mean(result)
}

Pr=sapply(seq(0.5, 0.95, 0.025), prob_win)
df=data.frame(p=seq(0.5, 0.95, 0.025), Pr)
df %>% ggplot(aes(p, Pr))+
  geom_point()
```



(b)

```
prob_win <- function(N, p=0.75){
B <- 10000
result <- replicate(B, {
b_win <- sample(c(1,0), N, replace = TRUE, prob = c(1-p, p))
sum(b_win)>=(N+1)/2
})
mean(result)
}

sapply(seq(1, 25, 2), prob_win)
```

```
##  [1] 0.2476 0.1585 0.1052 0.0735 0.0491 0.0348 0.0231 0.0175 0.0121 0.0086
## [11] 0.0055 0.0038 0.0038
```

4

(1)

```
library(dslabs)
data("polls_us_election_2016")
polls <- polls_us_election_2016 %>%
  filter(enddate >= "2016-10-31" & state == "U.S.")
N <- polls$samplesize[1]
x_hat=polls$rawpoll_clinton[1]/100
se_hat=sqrt(x_hat*(1-x_hat)/N)
x_hat+c(-1, 1)*qnorm(1-0.025)*se_hat
```

```
## [1] 0.4492385 0.4907615
```

(2)

```
polls %>% mutate(x_hat=polls$rawpoll_clinton/100,
                 se_hat=sqrt(x_hat*(1-x_hat)/N),
                 lower=x_hat-qnorm(1-0.025)*se_hat,
                 upper=x_hat+qnorm(1-0.025)*se_hat)%>%
  select(pollster, enddate, x_hat, lower, upper)
```

```
##                                                    pollster    enddate  x_hat
## 1                                  ABC News/Washington Post 2016-11-06 0.4700
## 2                                   Google Consumer Surveys 2016-11-07 0.3803
## 3                                                     Ipsos 2016-11-06 0.4200
## 4                                                    YouGov 2016-11-07 0.4500
## 5                                          Gravis Marketing 2016-11-06 0.4700
## 6  Fox News/Anderson Robbins Research/Shaw & Company Research 2016-11-06 0.4800
## 7                                   CBS News/New York Times 2016-11-06 0.4500
```

```
## 8                               NBC News/Wall Street Journal 2016-11-05 0.4400
## 9                                                   IBD/TIPP 2016-11-07 0.4120
## 10                                          Selzer & Company 2016-11-06 0.4400
## 11                                         Angus Reid Global 2016-11-04 0.4800
## 12                                       Monmouth University 2016-11-06 0.5000
## 13                                           Marist College 2016-11-03 0.4400
## 14                                The Times-Picayune/Lucid 2016-11-07 0.4500
## 15                                   USC Dornsife/LA Times 2016-11-07 0.4361
## 16                    RKM Research and Communications, Inc. 2016-11-05 0.4760
## 17                                     CVOTER International 2016-11-06 0.4891
## 18                                          Morning Consult 2016-11-05 0.4500
## 19                                             SurveyMonkey 2016-11-06 0.4700
## 20                 Rasmussen Reports/Pulse Opinion Research 2016-11-06 0.4500
## 21                                            Insights West 2016-11-07 0.4900
## 22                             RAND (American Life Panel) 2016-11-01 0.4370
## 23 Fox News/Anderson Robbins Research/Shaw & Company Research 2016-11-03 0.4550
## 24                                 CBS News/New York Times 2016-11-01 0.4500
## 25                                 ABC News/Washington Post 2016-11-05 0.4700
## 26                                                    Ipsos 2016-11-04 0.4300
## 27                                 ABC News/Washington Post 2016-11-04 0.4800
## 28                                                   YouGov 2016-11-06 0.4290
## 29                                                 IBD/TIPP 2016-11-06 0.4070
## 30                                 ABC News/Washington Post 2016-11-03 0.4700
## 31                                                 IBD/TIPP 2016-11-03 0.4440
## 32                                                 IBD/TIPP 2016-11-05 0.4300
## 33                                 ABC News/Washington Post 2016-11-02 0.4700
## 34                                 ABC News/Washington Post 2016-11-01 0.4700
## 35                                 ABC News/Washington Post 2016-10-31 0.4600
## 36                                                    Ipsos 2016-11-03 0.4320
## 37                                                 IBD/TIPP 2016-11-04 0.4420
## 38                                                   YouGov 2016-11-01 0.4600
## 39                                                 IBD/TIPP 2016-10-31 0.4460
## 40                                                    Ipsos 2016-11-02 0.4550
## 41                 Rasmussen Reports/Pulse Opinion Research 2016-11-03 0.4400
## 42                                The Times-Picayune/Lucid 2016-11-06 0.4500
## 43                                                    Ipsos 2016-11-01 0.4470
## 44                                                 IBD/TIPP 2016-11-02 0.4400
## 45                                                 IBD/TIPP 2016-11-01 0.4400
## 46                 Rasmussen Reports/Pulse Opinion Research 2016-11-02 0.4200
## 47                                                    Ipsos 2016-10-31 0.4400
## 48                                The Times-Picayune/Lucid 2016-11-05 0.4500
## 49                 Rasmussen Reports/Pulse Opinion Research 2016-10-31 0.4400
## 50                                  Google Consumer Surveys 2016-10-31 0.3769
## 51                                     CVOTER International 2016-11-05 0.4925
## 52                 Rasmussen Reports/Pulse Opinion Research 2016-11-01 0.4400
## 53                                     CVOTER International 2016-11-04 0.4906
## 54                                The Times-Picayune/Lucid 2016-11-04 0.4500
## 55                                   USC Dornsife/LA Times 2016-11-06 0.4323
```

```
## 56                                        CVOTER International 2016-11-03 0.4853
## 57                                    The Times-Picayune/Lucid 2016-11-03 0.4400
## 58                                       USC Dornsife/LA Times 2016-11-05 0.4263
## 59                                        CVOTER International 2016-11-02 0.4878
## 60                                       USC Dornsife/LA Times 2016-11-04 0.4256
## 61                                        CVOTER International 2016-11-01 0.4881
## 62                                    The Times-Picayune/Lucid 2016-11-02 0.4400
## 63                                            Gravis Marketing 2016-10-31 0.4600
## 64                                       USC Dornsife/LA Times 2016-11-03 0.4338
## 65                                    The Times-Picayune/Lucid 2016-11-01 0.4300
## 66                                       USC Dornsife/LA Times 2016-11-02 0.4247
## 67                                            Gravis Marketing 2016-11-02 0.4700
## 68                                       USC Dornsife/LA Times 2016-11-01 0.4236
## 69                                    The Times-Picayune/Lucid 2016-10-31 0.4200
## 70                                       USC Dornsife/LA Times 2016-10-31 0.4328
##        lower     upper
## 1  0.4492385 0.4907615
## 2  0.3601059 0.4004941
## 3  0.3994690 0.4405310
## 4  0.4293053 0.4706947
## 5  0.4492385 0.4907615
## 6  0.4592177 0.5007823
## 7  0.4293053 0.4706947
## 8  0.4193513 0.4606487
## 9  0.3915257 0.4324743
## 10 0.4193513 0.4606487
## 11 0.4592177 0.5007823
## 12 0.4792010 0.5207990
## 13 0.4193513 0.4606487
## 14 0.4293053 0.4706947
## 15 0.4154716 0.4567284
## 16 0.4552250 0.4967750
## 17 0.4683060 0.5098940
## 18 0.4293053 0.4706947
## 19 0.4492385 0.4907615
## 20 0.4293053 0.4706947
## 21 0.4692052 0.5107948
## 22 0.4163668 0.4576332
## 23 0.4342854 0.4757146
## 24 0.4293053 0.4706947
## 25 0.4492385 0.4907615
## 26 0.4094059 0.4505941
## 27 0.4592177 0.5007823
## 28 0.4084118 0.4495882
## 29 0.3865640 0.4274360
## 30 0.4492385 0.4907615
## 31 0.4233319 0.4646681
## 32 0.4094059 0.4505941
```

```
## 33 0.4492385 0.4907615
## 34 0.4492385 0.4907615
## 35 0.4392677 0.4807323
## 36 0.4113943 0.4526057
## 37 0.4213414 0.4626586
## 38 0.4392677 0.4807323
## 39 0.4253227 0.4666773
## 40 0.4342854 0.4757146
## 41 0.4193513 0.4606487
## 42 0.4293053 0.4706947
## 43 0.4263182 0.4676818
## 44 0.4193513 0.4606487
## 45 0.4193513 0.4606487
## 46 0.3994690 0.4405310
## 47 0.4193513 0.4606487
## 48 0.4293053 0.4706947
## 49 0.4193513 0.4606487
## 50 0.3567413 0.3970587
## 51 0.4717034 0.5132966
## 52 0.4193513 0.4606487
## 53 0.4698047 0.5113953
## 54 0.4293053 0.4706947
## 55 0.4116926 0.4529074
## 56 0.4645100 0.5060900
## 57 0.4193513 0.4606487
## 58 0.4057282 0.4468718
## 59 0.4670072 0.5085928
## 60 0.4050326 0.4461674
## 61 0.4673069 0.5088931
## 62 0.4193513 0.4606487
## 63 0.4392677 0.4807323
## 64 0.4131841 0.4544159
## 65 0.4094059 0.4505941
## 66 0.4041383 0.4452617
## 67 0.4492385 0.4907615
## 68 0.4030453 0.4441547
## 69 0.3994690 0.4405310
## 70 0.4121897 0.4534103
```

(3)

```
polls %>% mutate(x_hat=polls$rawpoll_clinton/100,
                 se_hat=sqrt(x_hat*(1-x_hat)/N),
                 lower=x_hat-qnorm(1-0.025)*se_hat,
                 upper=x_hat+qnorm(1-0.025)*se_hat,
                 hit=0.482>=lower&0.482<=upper)%>%
  select(pollster, enddate, x_hat, lower, upper, hit)
```

```
##                                                       pollster    enddate  x_hat
## 1                                    ABC News/Washington Post 2016-11-06 0.4700
## 2                                     Google Consumer Surveys 2016-11-07 0.3803
## 3                                                       Ipsos 2016-11-06 0.4200
## 4                                                      YouGov 2016-11-07 0.4500
## 5                                           Gravis Marketing 2016-11-06 0.4700
## 6  Fox News/Anderson Robbins Research/Shaw & Company Research 2016-11-06 0.4800
## 7                                     CBS News/New York Times 2016-11-06 0.4500
## 8                                NBC News/Wall Street Journal 2016-11-05 0.4400
## 9                                                    IBD/TIPP 2016-11-07 0.4120
## 10                                          Selzer & Company 2016-11-06 0.4400
## 11                                         Angus Reid Global 2016-11-04 0.4800
## 12                                       Monmouth University 2016-11-06 0.5000
## 13                                            Marist College 2016-11-03 0.4400
## 14                                 The Times-Picayune/Lucid 2016-11-07 0.4500
## 15                                      USC Dornsife/LA Times 2016-11-07 0.4361
## 16                      RKM Research and Communications, Inc. 2016-11-05 0.4760
## 17                                      CVOTER International 2016-11-06 0.4891
## 18                                           Morning Consult 2016-11-05 0.4500
## 19                                              SurveyMonkey 2016-11-06 0.4700
## 20                 Rasmussen Reports/Pulse Opinion Research 2016-11-06 0.4500
## 21                                             Insights West 2016-11-07 0.4900
## 22                                 RAND (American Life Panel) 2016-11-01 0.4370
## 23 Fox News/Anderson Robbins Research/Shaw & Company Research 2016-11-03 0.4550
## 24                                    CBS News/New York Times 2016-11-01 0.4500
## 25                                   ABC News/Washington Post 2016-11-05 0.4700
## 26                                                      Ipsos 2016-11-04 0.4300
## 27                                   ABC News/Washington Post 2016-11-04 0.4800
## 28                                                     YouGov 2016-11-06 0.4290
## 29                                                   IBD/TIPP 2016-11-06 0.4070
## 30                                   ABC News/Washington Post 2016-11-03 0.4700
## 31                                                   IBD/TIPP 2016-11-03 0.4440
## 32                                                   IBD/TIPP 2016-11-05 0.4300
## 33                                   ABC News/Washington Post 2016-11-02 0.4700
## 34                                   ABC News/Washington Post 2016-11-01 0.4700
## 35                                   ABC News/Washington Post 2016-10-31 0.4600
## 36                                                      Ipsos 2016-11-03 0.4320
## 37                                                   IBD/TIPP 2016-11-04 0.4420
## 38                                                     YouGov 2016-11-01 0.4600
## 39                                                   IBD/TIPP 2016-10-31 0.4460
## 40                                                      Ipsos 2016-11-02 0.4550
## 41                 Rasmussen Reports/Pulse Opinion Research 2016-11-03 0.4400
## 42                                 The Times-Picayune/Lucid 2016-11-06 0.4500
## 43                                                      Ipsos 2016-11-01 0.4470
## 44                                                   IBD/TIPP 2016-11-02 0.4400
## 45                                                   IBD/TIPP 2016-11-01 0.4400
## 46                 Rasmussen Reports/Pulse Opinion Research 2016-11-02 0.4200
## 47                                                      Ipsos 2016-10-31 0.4400
```

```
## 48                                    The Times-Picayune/Lucid 2016-11-05 0.4500
## 49                  Rasmussen Reports/Pulse Opinion Research 2016-10-31 0.4400
## 50                                    Google Consumer Surveys 2016-10-31 0.3769
## 51                                       CVOTER International 2016-11-05 0.4925
## 52                  Rasmussen Reports/Pulse Opinion Research 2016-11-01 0.4400
## 53                                       CVOTER International 2016-11-04 0.4906
## 54                                    The Times-Picayune/Lucid 2016-11-04 0.4500
## 55                                      USC Dornsife/LA Times 2016-11-06 0.4323
## 56                                       CVOTER International 2016-11-03 0.4853
## 57                                    The Times-Picayune/Lucid 2016-11-03 0.4400
## 58                                      USC Dornsife/LA Times 2016-11-05 0.4263
## 59                                       CVOTER International 2016-11-02 0.4878
## 60                                      USC Dornsife/LA Times 2016-11-04 0.4256
## 61                                       CVOTER International 2016-11-01 0.4881
## 62                                    The Times-Picayune/Lucid 2016-11-02 0.4400
## 63                                           Gravis Marketing 2016-10-31 0.4600
## 64                                      USC Dornsife/LA Times 2016-11-03 0.4338
## 65                                    The Times-Picayune/Lucid 2016-11-01 0.4300
## 66                                      USC Dornsife/LA Times 2016-11-02 0.4247
## 67                                           Gravis Marketing 2016-11-02 0.4700
## 68                                      USC Dornsife/LA Times 2016-11-01 0.4236
## 69                                    The Times-Picayune/Lucid 2016-10-31 0.4200
## 70                                      USC Dornsife/LA Times 2016-10-31 0.4328
##        lower     upper   hit
## 1  0.4492385 0.4907615  TRUE
## 2  0.3601059 0.4004941 FALSE
## 3  0.3994690 0.4405310 FALSE
## 4  0.4293053 0.4706947 FALSE
## 5  0.4492385 0.4907615  TRUE
## 6  0.4592177 0.5007823  TRUE
## 7  0.4293053 0.4706947 FALSE
## 8  0.4193513 0.4606487 FALSE
## 9  0.3915257 0.4324743 FALSE
## 10 0.4193513 0.4606487 FALSE
## 11 0.4592177 0.5007823  TRUE
## 12 0.4792010 0.5207990  TRUE
## 13 0.4193513 0.4606487 FALSE
## 14 0.4293053 0.4706947 FALSE
## 15 0.4154716 0.4567284 FALSE
## 16 0.4552250 0.4967750  TRUE
## 17 0.4683060 0.5098940  TRUE
## 18 0.4293053 0.4706947 FALSE
## 19 0.4492385 0.4907615  TRUE
## 20 0.4293053 0.4706947 FALSE
## 21 0.4692052 0.5107948  TRUE
## 22 0.4163668 0.4576332 FALSE
## 23 0.4342854 0.4757146 FALSE
## 24 0.4293053 0.4706947 FALSE
```

```
## 25 0.4492385 0.4907615  TRUE
## 26 0.4094059 0.4505941 FALSE
## 27 0.4592177 0.5007823  TRUE
## 28 0.4084118 0.4495882 FALSE
## 29 0.3865640 0.4274360 FALSE
## 30 0.4492385 0.4907615  TRUE
## 31 0.4233319 0.4646681 FALSE
## 32 0.4094059 0.4505941 FALSE
## 33 0.4492385 0.4907615  TRUE
## 34 0.4492385 0.4907615  TRUE
## 35 0.4392677 0.4807323 FALSE
## 36 0.4113943 0.4526057 FALSE
## 37 0.4213414 0.4626586 FALSE
## 38 0.4392677 0.4807323 FALSE
## 39 0.4253227 0.4666773 FALSE
## 40 0.4342854 0.4757146 FALSE
## 41 0.4193513 0.4606487 FALSE
## 42 0.4293053 0.4706947 FALSE
## 43 0.4263182 0.4676818 FALSE
## 44 0.4193513 0.4606487 FALSE
## 45 0.4193513 0.4606487 FALSE
## 46 0.3994690 0.4405310 FALSE
## 47 0.4193513 0.4606487 FALSE
## 48 0.4293053 0.4706947 FALSE
## 49 0.4193513 0.4606487 FALSE
## 50 0.3567413 0.3970587 FALSE
## 51 0.4717034 0.5132966  TRUE
## 52 0.4193513 0.4606487 FALSE
## 53 0.4698047 0.5113953  TRUE
## 54 0.4293053 0.4706947 FALSE
## 55 0.4116926 0.4529074 FALSE
## 56 0.4645100 0.5060900  TRUE
## 57 0.4193513 0.4606487 FALSE
## 58 0.4057282 0.4468718 FALSE
## 59 0.4670072 0.5085928  TRUE
## 60 0.4050326 0.4461674 FALSE
## 61 0.4673069 0.5088931  TRUE
## 62 0.4193513 0.4606487 FALSE
## 63 0.4392677 0.4807323 FALSE
## 64 0.4131841 0.4544159 FALSE
## 65 0.4094059 0.4505941 FALSE
## 66 0.4041383 0.4452617 FALSE
## 67 0.4492385 0.4907615  TRUE
## 68 0.4030453 0.4441547 FALSE
## 69 0.3994690 0.4405310 FALSE
## 70 0.4121897 0.4534103 FALSE
```

(4)

```
polls1=polls %>% mutate(x_hat=polls$rawpoll_clinton/100,
                 se_hat=sqrt(x_hat*(1-x_hat)/N),
                 lower=x_hat-qnorm(1-0.025)*se_hat,
                 upper=x_hat+qnorm(1-0.025)*se_hat,
                 hit=0.482>=lower&0.482<=upper)%>%
  select(pollster, enddate, x_hat, lower, upper, hit)

polls1%>%summarise(mean(hit))
```

```
##   mean(hit)
## 1 0.2857143
```

(5)

```
# 0.95
```

(6)

```
polls <- polls_us_election_2016 %>%
  filter(enddate >= "2016-10-31" & state == "U.S.")  %>%
  mutate(d_hat = rawpoll_clinton / 100 - rawpoll_trump / 100)

N <- polls$samplesize[1]
d_hat <- polls$d_hat[1]
x_hat <- (d_hat+1)/2
se_hat <- 2*sqrt(x_hat*(1-x_hat)/N)
d_hat + c(-1,1)*qnorm(0.975)*se_hat
```

```
## [1] -0.001564627  0.081564627
```

5

(1)

```
library(HistData)
library(tidyr)
library(broom)
data("GaltonFamilies")
set.seed(1)
galton_heights <- GaltonFamilies %>%
  group_by(family, gender) %>%
  sample_n(1) %>%
  ungroup()
galton_heights%>%
```

```
  gather(parent, parentHeight, 'father':'mother') %>%
  mutate(child = ifelse(gender == "female", "daughter", "son"))%>%
  group_by(parent, child)%>%
  do(tidy(lm(childHeight~parentHeight, data=.)))
```

```
## # A tibble: 8 x 7
## # Groups:   parent, child [4]
##   parent child    term         estimate std.error statistic  p.value
##   <chr>  <chr>    <chr>           <dbl>     <dbl>     <dbl>    <dbl>
## 1 father daughter (Intercept)    40.2      4.34        9.26 7.76e-17
## 2 father daughter parentHeight    0.345    0.0625      5.52 1.21e- 7
## 3 father son      (Intercept)    39.8      4.47        8.91 5.96e-16
## 4 father son      parentHeight    0.426    0.0646      6.59 4.74e-10
## 5 mother daughter (Intercept)    37.6      4.78        7.87 3.61e-13
## 6 mother daughter parentHeight    0.413    0.0745      5.55 1.07e- 7
## 7 mother son      (Intercept)    49.3      4.73       10.4  4.36e-20
## 8 mother son      parentHeight    0.312    0.0739      4.22 3.84e- 5
```

(2)

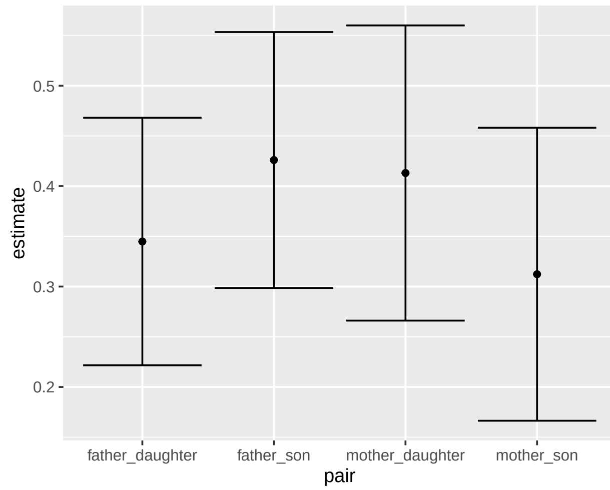
```
galton_heights%>%
  gather(parent, parentHeight, 'father':'mother') %>%
  mutate(child = ifelse(gender == "female", "daughter", "son"))%>%
  group_by(parent, child)%>%
  do(tidy(lm(childHeight~parentHeight, data=.), conf.int=T))
```

```
## # A tibble: 8 x 9
## # Groups:   parent, child [4]
##   parent child  term    estimate std.error statistic  p.value conf.low conf.high
##   <chr>  <chr>  <chr>      <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 father daugh~ (Inter~   40.2      4.34        9.26 7.76e-17   31.6      48.7
## 2 father daugh~ parent~    0.345    0.0625      5.52 1.21e- 7    0.222     0.468
## 3 father son    (Inter~   39.8      4.47        8.91 5.96e-16   31.0      48.6
## 4 father son    parent~    0.426    0.0646      6.59 4.74e-10    0.299     0.553
## 5 mother daugh~ (Inter~   37.6      4.78        7.87 3.61e-13   28.2      47.0
## 6 mother daugh~ parent~    0.413    0.0745      5.55 1.07e- 7    0.266     0.560
## 7 mother son    (Inter~   49.3      4.73       10.4  4.36e-20   39.9      58.6
## 8 mother son    parent~    0.312    0.0739      4.22 3.84e- 5    0.166     0.458
```

(3)

```
galton_heights%>%
  gather(parent, parentHeight, 'father':'mother') %>%
  mutate(child = ifelse(gender == "female", "daughter", "son"))%>%
  unite(pair, c("parent", "child")) %>%
  group_by(pair) %>%
  do(tidy(lm(childHeight~parentHeight, data=.), conf.int=T))%>%
```

```
  filter(term=="parentHeight")%>%
  ggplot(aes(x=pair, y=estimate, ymin=conf.low,
             ymax=conf.high))+
  geom_errorbar()+
  geom_point()
```



6

(1)

```
dat=research_funding_rates%>%
    rename(success_total=success_rates_total,
           success_men=success_rates_men,
           success_women=success_rates_women)%>%
    gather(key, value, -discipline)%>%
    separate(key, into = c("status", "gender"))%>%
    spread(status, value)%>%
    filter(gender!="total")
dat%>%group_by(gender) %>%
  summarise(total_award=sum(awards),
            not_award=sum(applications)-sum(total_award))
```

```
## # A tibble: 2 x 3
##   gender total_award not_award
##   <chr>        <dbl>     <dbl>
## 1 men            290      1345
## 2 women          177      1011
```

(2)

```r
dat2=dat%>%group_by(gender) %>%
  summarise(total_award=sum(awards),
            not_award=sum(applications)-sum(total_award))
dat2%>%group_by(gender)%>%
  summarise(diff=total_award/(total_award+not_award))
```

```
## # A tibble: 2 x 2
##   gender  diff
##   <chr>  <dbl>
## 1 men    0.177
## 2 women  0.149
```

(3)

```r
dat2%>%select(-gender) %>%
  do(tidy(chisq.test(.)))
```

```
## # A tibble: 1 x 4
##   statistic p.value parameter method
##       <dbl>   <dbl>     <int> <chr>
## 1      3.81  0.0509         1 Pearson's Chi-squared test with Yates' continuity~
```

(4)

```r
research_funding_rates%>%
  mutate(discipline=reorder(discipline, success_rates_total))%>%
  rename(success_total=success_rates_total,
         success_men=success_rates_men,
         success_women=success_rates_women)%>%
  gather(key, value, -discipline)%>%
  separate(key, into = c("status", "gender"))%>%
  spread(status, value)%>%
  filter(gender!="total")
```

```
##             discipline gender applications awards success
## 1      Social sciences    men          425     65    15.3
## 2      Social sciences  women          409     47    11.5
## 3     Medical sciences    men          245     46    18.8
## 4     Medical sciences  women          260     29    11.2
## 5    Interdisciplinary    men          105     12    11.4
## 6    Interdisciplinary  women           78     17    21.8
## 7           Humanities    men          230     33    14.3
## 8           Humanities  women          166     32    19.3
## 9   Technical sciences    men          189     30    15.9
## 10  Technical sciences  women           62     13    21.0
## 11 Earth/life sciences    men          156     38    24.4
## 12 Earth/life sciences  women          126     18    14.3
```

```
## 13   Physical sciences    men          135     26     19.3
## 14   Physical sciences  women           39      9     23.1
## 15   Chemical sciences    men           83     22     26.5
## 16   Chemical sciences  women           39     10     25.6
## 17             Physics    men           67     18     26.9
## 18             Physics  women            9      2     22.2
```

(5)

```r
research_funding_rates%>%
  mutate(discipline=reorder(discipline, success_rates_total))%>%
  rename(success_total=success_rates_total,
         success_men=success_rates_men,
         success_women=success_rates_women)%>%
  gather(key, value, -discipline)%>%
  separate(key, into = c("status", "gender"))%>%
  spread(status, value)%>%
  filter(gender!="total")%>%
  ggplot(aes(discipline, success,
             size=applications, color=gender))+
  geom_point()+
  theme(axis.text.x = element_text(angle=90, vjust = 1))
```