DS 2006 Midterm 1 Prep Questions

NAME: Olivia Park

NOTE: The `.rmd` version of the file is available here: (link)

Instructions

Please prepare reponses/solutions for the following questions. On the day of the exam, you will be given a new set of questions. You will use the solutions you've prepared for this exam during the exam.

During the exam, you will also be permitted to access the internet for publicly available content. You will not be allowed to communicate with anyone via the internet or any other means during the exam. This includes, but is not limited to:

- No messaging, emailing, or using social media to contact others.
- No posting questions or seeking answers on forums, chat rooms, chat bots (including large language models like ChatGPT), or any collaborative platforms.
- No sharing or discussing exam content with peers through any online or electronic medium.

You may **NOT** discuss any aspect of the exam or prep questions with anyone other than the instructor or TA. You may **NOT** share code or documents.

Submission instructions

1. Within your course repo, create a folder called `Midterm2`
2. Within the folder, create the script file `exam-prep.rmd` with your solutions. Create a rendered report in `.pdf` output.
3. Add, commit, and push to your repo on github.com.

Questions

Exam questions are organized into sections cooresponding to the learning outcomes of the course.

Section 1. Tools of the data scientists

> Learning objective: Use the tools of data scientists

> Learning objective: Implement best programming/coding practices

1.1 Write your name at the prompt above (line 6 in the script).

1.2 When you are done with the exam, please render this report as a pdf document.

1.3 Please explain the advantages of relative file paths over absolute file paths.

Absolute file paths make it easier to share and locate files. They are more portable and can be moved around because they are not attatched to 1 specific location. They also make working with Git and GitHub to collabroate on files much easier and more accessible to users.

Section 2. Probability

> Learning objective: Compare and contrast different definitions of probability, illustrating differences with simple examples

> Learning objective: Express the rules of probability verbally, mathematically, and computationally.

2.1 Machines A, B, and C generate widgets with a defect rate of 0.1, 0.01, and 0.001, respectively. Machine A generates twice as many widgets as B, and machine B generates twice as many widgets as machine C.

Please complete the cross table below based on the information above.

Hints:

- $P(A) = 2P(B) = 4P(B)$, which means $P(A) = 4/7$, $P(B) = 2/7$ and $P(C) = 1/7$.

	A	B	C	
Defective				
→ cell	.057	.0028	.00014	.0612
→ row	.950	.0467	.0023	
→ col	.1 .01	.001		
Functional				
→ cell	.514	.283	.1427	.939
→ row	.548	.301	.1520	
→ col	.9	.99	.999	
	4/7	2/7	1/7	1

2.2 If you did not know if a widget was functional or defective, what would be the probability that the widget is from machine A? Write the solution using mathematical notation, as in P(???) = ???.

$P(WidgetA) = 4/7 = .5714$

2.3 How would you update the probability from 2.3 if you knew the widget was defective? Write the solution using mathematical notation, as in P(???) = ???.

$P(A|Defective) = .950$

Section 3. Simulation

> Learning objective: Use probability models to build simulations of complex real world processes to answer research questions

3.1 In class and homework, you solved the birthday problem. The script we developed in class to solve the problem is below. Calculate the probability of a shared birthday for a class of size 20 with the code using only 30 replicates.

```
generate_class <- function(class_size){
  sample(1:365, class_size, replace=TRUE)
}

check_birthday <- function(class){
  class |>
    duplicated() |>
    any()
}

set.seed(230583)
R <- 30
replicates <- replicate(R, generate_class(20) |> check_birthday())
mean(replicates)
```

```
## [1] 0.3333333
```

3.2 Using the exact same number of replicates and class size, repeat the process of calculating the shared birthday probability 1000 times. Create a histogram of the 1000 estimates of the probability. The analytic solution to the birthday problem is

$$1 - \frac{365!}{365^k(365-k)!}$$

where k is the class size. Add the analytic solution as a vertical reference line using the **abline** command.

```
R <- 20
out <- rep(NA, 1000)
for(i in 1:length(out)){
  replicates <- replicate(R, generate_class(20) |> check_birthday())
  out[i] <- mean(replicates)
}
hist(out, breaks = 100, freq = FALSE, main = "Histogram of 1000 estimates of the shared birthday probabi
box()
analytic_solution <- function(k) {
  analytic_solution <- 1 - (exp(lfactorial(365) - lfactorial(365-k)) / (365^k))
}
k <- 20
analytic_prob <- analytic_solution(k)
abline(v = analytic_prob, col = "red")
```

Histogram of 1000 estimates of the shared birthday probability
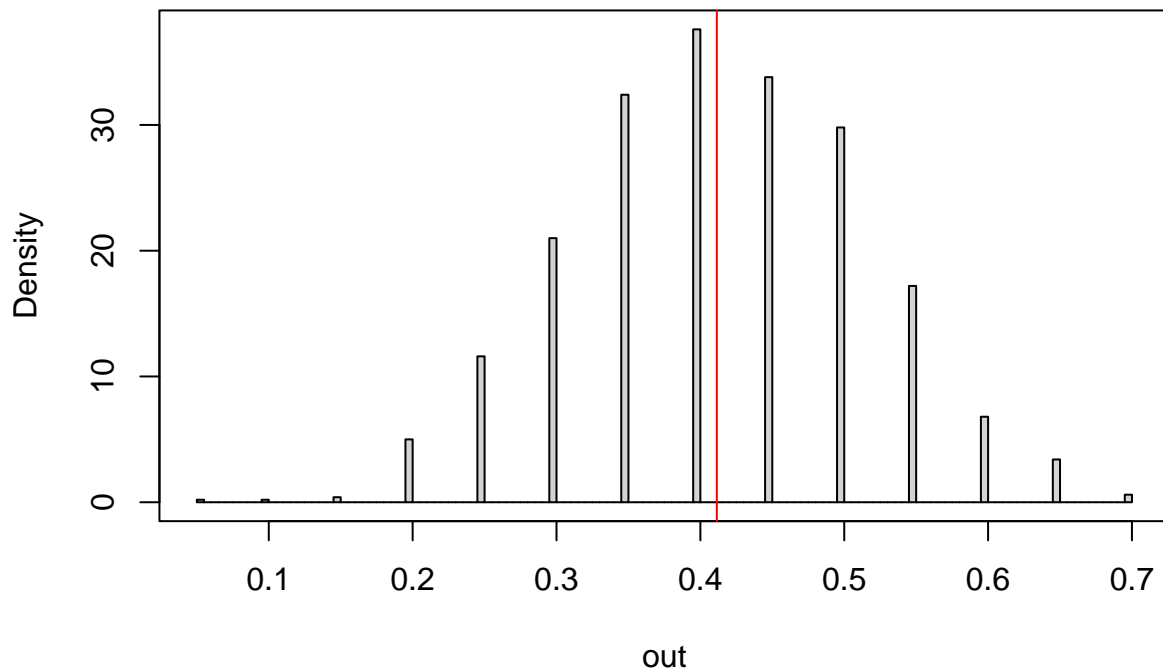

```r
#actual value of birthday(20)
print(analytic_prob)
```

```
## [1] 0.4114384
```

3.3 What is the range (min and max) of the 1000 values you generated for the birthday probability estimate?

```r
#min and max
min(out)
```

```
## [1] 0.05
```

```r
max(out)
```

```
## [1] 0.7
```

3.4 What is the average absolute error of the 1000 estimates?

```r
absolute_error <- function(out, a_prob){
  sum_of_error <- 0
  for(i in 1:length(out)){
    sum_of_error <- sum_of_error + abs(out[i] - a_prob)
  }
```

```
  absolute_error <- sum_of_error / length(out)
}
abs_err <- absolute_error(out, analytic_prob)

print(abs_err)
```

```
## [1] 0.08261082
```

3.5 If you wanted to reduced the average absolute error by half, how many replicates (R) would you need to use?

To cut abosulte error in half, you can multiply R by 4. 4*20 = 80, so 80 should cut the absolute error in half. The exact solution can also be calcuated with this code:

```
for(i in 70:90){
  R <- i
  out_new <- rep(NA, 1000)
  for(i in 1:length(out_new)){
    replicates <- replicate(R, generate_class(20) |> check_birthday())
    out_new[i] <- mean(replicates)
  }
  analytic_prob <- analytic_solution(20)
  abs_err_new <- absolute_error(out_new, analytic_prob)
  if(abs_err_new <= abs_err / 2){
    print(abs_err_new)
    print(R)
    break
  }
}
```

```
## [1] 0.04113068
## [1] 85
```

3.6 Generate a plot of overlapping histograms to show the difference between R=30 and your R from the previous problem.

```
R2 <- 80 # Change this
out2 <- rep(NA, 1000)
for(i in 1:length(out2)){
  replicates <- replicate(R2, generate_class(20) |> check_birthday())
  out2[i] <- mean(replicates)
}

bins <- seq(0,1,by = 0.01)
hist(out, breaks = bins, col = "#ffabab50", freq = FALSE, main = "Histogram of 1000 estimates of the sha
hist(out2, breaks = bins, add=TRUE, col = "#6488ea59", freq = FALSE)
```

Histogram of 1000 estimates of the shared birthday probability
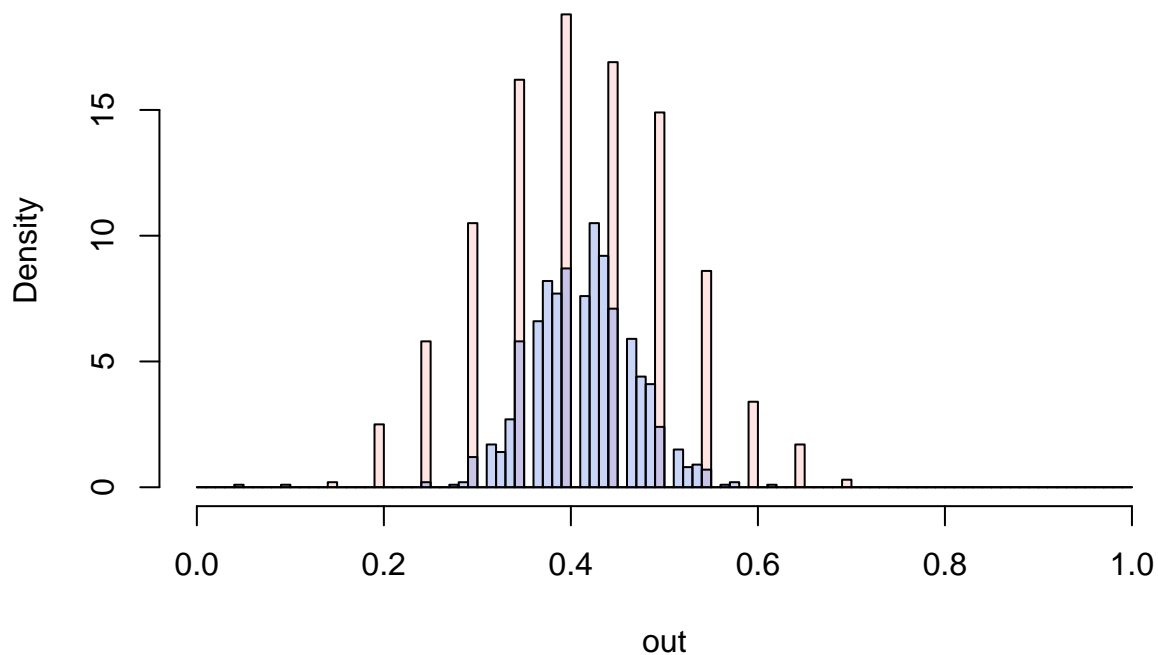
3.7 Calculate the average absolute error of the 1000 estimates generate with the new choice of R? Did it change as you expected it to?

Here is the calculations using R=80 for absulte error. It did change as I expected as the absolute error was cut from 0.08634973 to 0.0434552.

```
R <- 80
out <- rep(NA, 1000)
for(i in 1:length(out)){
  replicates <- replicate(R, generate_class(20) |> check_birthday())
  out[i] <- mean(replicates)
}

absolute_error <- function(out, a_prob){
  sum_of_error <- 0
  for(i in 1:length(out)){
    sum_of_error <- sum_of_error + abs(out[i] - a_prob)
  }
  absolute_error <- sum_of_error / length(out)
}
abs_err <- absolute_error(out, analytic_prob)

print(abs_err)
```

```
## [1] 0.04517973
```

Section 4. Diagnostics

Learning objective: apply cross table framework to the special case of binary outcomes

4.1 Suppose the prevalence of disease X is 0.1, and you are tasked with finding a diagnostic test so that the positive predictive value is 0.95. If the specificity of the test is 0.9, what must the sensitivity be?

	Cancer	Not Cancer	
Test = +	.005 ,row= .05, col=, .05	.09 , row= .95, col= .1	
Test = -	.095 , row=, col= .95	.81 , row=, col= .9	
	0.1	.9	1

Senistivity is a conditional column probability and is $P(+|C) = .05$

Section 5. Confounding vs Causal Pathway

Learning objective: define/describe confounding variables, Simpson's paradox, DAGs, and the causal pathway

5.1 The following function generates data from a cohort of individuals who were diagnosed with disease X. In the dataset, there is vaccination status, disease severity, and disease duration.

Generate 1000 draws from the function and create the cross table for vaccination status and disease duration. Calculate a summary of the effect of vaccination by calculating the difference is conditional probabilities:

$$\Delta = P(\text{Short recovery}|\text{vaccinated}) - P(\text{Short recovery}|\text{unvaccinated})$$

```r
vax_data <- function(R){
    vs <- rbinom(R,1,.5)
    ds <- rbinom(R,1,.25*(vs==1)+.75*(vs==0))
    rt <- rbinom(R,1,.7*(ds==1)+.5*(ds==0))
    data.frame(
        vaccination_status = c("vaccinated","unvaccinated")[vs+1]
      , disease_severity = c("mild","severe")[ds+1]
      , recovery_time = c("short","long")[rt+1]
    )
}

calc_delta <- function(df){
  num_vax <- 0
  vax_and_recov <- 0

  for(i in 1:nrow(df)){
    if(df[i,1] == "vaccinated"){
      num_vax <- num_vax + 1
      if(df[i,3] == "short"){
        vax_and_recov <- vax_and_recov + 1
      }
    }
  }
  p_short_given_vax = vax_and_recov / num_vax
```

```
  num_unvax <- 0
  unvax_and_recov <- 0

  for(i in 1:nrow(df)){
    if(df[i,1] == "unvaccinated"){
      num_unvax <- num_unvax + 1
      if(df[i,3] == "short"){
        unvax_and_recov <- unvax_and_recov + 1
      }
    }
  }
  p_short_given_unvax = unvax_and_recov / num_unvax

  delta = p_short_given_vax - p_short_given_unvax
}

Ra <- 1000
vax <- vax_data(Ra)

print(vax)
```

```
##       vaccination_status disease_severity recovery_time
## 1           unvaccinated           severe          long
## 2           unvaccinated           severe          long
## 3           unvaccinated             mild         short
## 4             vaccinated           severe          long
## 5           unvaccinated           severe          long
## 6           unvaccinated             mild          long
## 7             vaccinated           severe          long
## 8           unvaccinated             mild         short
## 9             vaccinated           severe         short
## 10            vaccinated           severe         short
## 11          unvaccinated             mild          long
## 12            vaccinated           severe         short
## 13          unvaccinated             mild         short
## 14            vaccinated           severe          long
## 15          unvaccinated             mild         short
## 16          unvaccinated             mild          long
## 17            vaccinated           severe          long
## 18            vaccinated           severe          long
## 19            vaccinated           severe          long
## 20          unvaccinated             mild          long
## 21            vaccinated           severe          long
## 22            vaccinated           severe          long
## 23          unvaccinated             mild         short
## 24            vaccinated             mild         short
## 25            vaccinated             mild         short
## 26          unvaccinated             mild          long
## 27            vaccinated           severe         short
## 28            vaccinated           severe          long
## 29            vaccinated           severe          long
## 30            vaccinated             mild         short
## 31          unvaccinated             mild          long
```

8

```
## 32       vaccinated      severe    short
## 33       vaccinated      severe     long
## 34       vaccinated      severe     long
## 35     unvaccinated        mild     long
## 36       vaccinated      severe     long
## 37       vaccinated      severe    short
## 38     unvaccinated        mild    short
## 39     unvaccinated        mild     long
## 40       vaccinated      severe     long
## 41       vaccinated      severe     long
## 42     unvaccinated        mild    short
## 43     unvaccinated        mild     long
## 44       vaccinated      severe    short
## 45     unvaccinated        mild     long
## 46     unvaccinated        mild    short
## 47       vaccinated        mild     long
## 48       vaccinated      severe     long
## 49       vaccinated        mild     long
## 50       vaccinated      severe     long
## 51       vaccinated      severe    short
## 52       vaccinated        mild     long
## 53     unvaccinated      severe     long
## 54       vaccinated      severe     long
## 55       vaccinated      severe     long
## 56     unvaccinated        mild    short
## 57       vaccinated      severe     long
## 58     unvaccinated        mild    short
## 59       vaccinated        mild     long
## 60       vaccinated      severe     long
## 61     unvaccinated        mild    short
## 62       vaccinated        mild    short
## 63       vaccinated      severe    short
## 64     unvaccinated      severe     long
## 65     unvaccinated        mild    short
## 66     unvaccinated        mild    short
## 67       vaccinated        mild    short
## 68     unvaccinated        mild     long
## 69       vaccinated        mild    short
## 70     unvaccinated        mild     long
## 71     unvaccinated      severe    short
## 72       vaccinated      severe    short
## 73       vaccinated      severe    short
## 74       vaccinated      severe    short
## 75       vaccinated        mild     long
## 76     unvaccinated        mild     long
## 77     unvaccinated        mild    short
## 78     unvaccinated        mild     long
## 79     unvaccinated        mild     long
## 80       vaccinated        mild    short
## 81       vaccinated      severe     long
## 82       vaccinated      severe     long
## 83       vaccinated        mild     long
## 84     unvaccinated      severe    short
## 85       vaccinated        mild    short
```

```
## 86         vaccinated        mild      long
## 87         vaccinated        severe     short
## 88       unvaccinated        mild      short
## 89       unvaccinated        severe      long
## 90         vaccinated        severe      long
## 91         vaccinated        mild       long
## 92       unvaccinated        mild       long
## 93         vaccinated        severe      long
## 94         vaccinated        severe      long
## 95       unvaccinated        mild      short
## 96         vaccinated        severe      long
## 97       unvaccinated        mild       long
## 98       unvaccinated        severe      long
## 99         vaccinated        mild      short
## 100        vaccinated        severe     short
## 101      unvaccinated        mild      short
## 102      unvaccinated        mild      short
## 103      unvaccinated        severe     short
## 104        vaccinated        severe      long
## 105      unvaccinated        mild      short
## 106      unvaccinated        mild      short
## 107      unvaccinated        mild      short
## 108      unvaccinated        severe      long
## 109      unvaccinated        mild       long
## 110      unvaccinated        severe      long
## 111      unvaccinated        mild      short
## 112        vaccinated        mild      short
## 113      unvaccinated        severe     short
## 114      unvaccinated        mild      short
## 115      unvaccinated        severe      long
## 116      unvaccinated        mild      short
## 117      unvaccinated        mild      short
## 118      unvaccinated        mild       long
## 119        vaccinated        severe      long
## 120        vaccinated        mild      short
## 121        vaccinated        mild       long
## 122      unvaccinated        mild      short
## 123      unvaccinated        mild       long
## 124        vaccinated        severe      long
## 125        vaccinated        severe      long
## 126      unvaccinated        severe     short
## 127        vaccinated        severe      long
## 128        vaccinated        severe      long
## 129        vaccinated        mild      short
## 130      unvaccinated        mild       long
## 131      unvaccinated        mild      short
## 132        vaccinated        severe      long
## 133        vaccinated        severe      long
## 134        vaccinated        severe      long
## 135      unvaccinated        mild       long
## 136        vaccinated        mild       long
## 137        vaccinated        severe      long
## 138      unvaccinated        severe      long
## 139        vaccinated        severe     short
```

```
## 140           vaccinated          severe          short
## 141         unvaccinated            mild          short
## 142           vaccinated          severe          short
## 143           vaccinated          severe           long
## 144         unvaccinated            mild          short
## 145           vaccinated            mild          short
## 146           vaccinated            mild           long
## 147         unvaccinated          severe           long
## 148           vaccinated            mild          short
## 149           vaccinated          severe           long
## 150         unvaccinated            mild           long
## 151           vaccinated          severe           long
## 152           vaccinated          severe          short
## 153           vaccinated          severe          short
## 154           vaccinated          severe           long
## 155           vaccinated            mild          short
## 156         unvaccinated          severe           long
## 157           vaccinated          severe          short
## 158         unvaccinated            mild          short
## 159           vaccinated          severe           long
## 160           vaccinated            mild          short
## 161           vaccinated          severe           long
## 162         unvaccinated            mild           long
## 163         unvaccinated            mild           long
## 164           vaccinated          severe           long
## 165         unvaccinated          severe           long
## 166           vaccinated          severe           long
## 167           vaccinated          severe           long
## 168         unvaccinated          severe          short
## 169           vaccinated            mild          short
## 170           vaccinated            mild          short
## 171         unvaccinated            mild          short
## 172         unvaccinated          severe          short
## 173         unvaccinated            mild           long
## 174           vaccinated          severe           long
## 175         unvaccinated            mild          short
## 176           vaccinated          severe           long
## 177         unvaccinated            mild          short
## 178           vaccinated            mild          short
## 179           vaccinated          severe           long
## 180         unvaccinated            mild          short
## 181           vaccinated          severe          short
## 182           vaccinated          severe           long
## 183         unvaccinated            mild          short
## 184           vaccinated          severe           long
## 185         unvaccinated            mild          short
## 186           vaccinated            mild           long
## 187           vaccinated          severe           long
## 188         unvaccinated          severe           long
## 189         unvaccinated            mild           long
## 190         unvaccinated            mild          short
## 191         unvaccinated            mild          short
## 192           vaccinated          severe           long
## 193           vaccinated          severe          short
```

```
## 194       unvaccinated       mild       long
## 195         vaccinated     severe       long
## 196       unvaccinated       mild      short
## 197       unvaccinated       mild       long
## 198       unvaccinated       mild       long
## 199       unvaccinated     severe       long
## 200       unvaccinated       mild      short
## 201         vaccinated       mild       long
## 202       unvaccinated     severe       long
## 203       unvaccinated       mild      short
## 204       unvaccinated       mild      short
## 205         vaccinated     severe       long
## 206       unvaccinated       mild      short
## 207         vaccinated     severe      short
## 208         vaccinated       mild      short
## 209         vaccinated     severe       long
## 210         vaccinated       mild      short
## 211         vaccinated     severe       long
## 212       unvaccinated       mild       long
## 213       unvaccinated       mild      short
## 214         vaccinated     severe       long
## 215         vaccinated     severe       long
## 216       unvaccinated       mild      short
## 217         vaccinated     severe       long
## 218       unvaccinated       mild      short
## 219         vaccinated       mild      short
## 220       unvaccinated       mild       long
## 221         vaccinated     severe       long
## 222       unvaccinated       mild      short
## 223         vaccinated       mild       long
## 224       unvaccinated       mild      short
## 225         vaccinated     severe       long
## 226         vaccinated     severe       long
## 227         vaccinated     severe       long
## 228       unvaccinated       mild       long
## 229         vaccinated     severe      short
## 230         vaccinated       mild       long
## 231       unvaccinated       mild      short
## 232         vaccinated       mild      short
## 233         vaccinated     severe       long
## 234       unvaccinated       mild       long
## 235       unvaccinated     severe       long
## 236         vaccinated     severe       long
## 237       unvaccinated     severe       long
## 238       unvaccinated       mild       long
## 239         vaccinated     severe       long
## 240       unvaccinated       mild      short
## 241       unvaccinated       mild       long
## 242       unvaccinated     severe      short
## 243       unvaccinated       mild      short
## 244         vaccinated       mild       long
## 245       unvaccinated       mild       long
## 246       unvaccinated       mild       long
## 247         vaccinated       mild       long
```

```
## 248       unvaccinated        severe       short
## 249       unvaccinated          mild       short
## 250         vaccinated        severe       short
## 251         vaccinated        severe       short
## 252         vaccinated          mild        long
## 253       unvaccinated          mild        long
## 254         vaccinated        severe       short
## 255       unvaccinated          mild       short
## 256       unvaccinated          mild        long
## 257         vaccinated        severe        long
## 258         vaccinated          mild       short
## 259       unvaccinated          mild       short
## 260         vaccinated          mild        long
## 261       unvaccinated          mild       short
## 262         vaccinated        severe        long
## 263       unvaccinated          mild       short
## 264       unvaccinated        severe       short
## 265       unvaccinated          mild       short
## 266       unvaccinated          mild       short
## 267       unvaccinated        severe        long
## 268         vaccinated          mild       short
## 269         vaccinated          mild        long
## 270         vaccinated        severe       short
## 271         vaccinated        severe        long
## 272       unvaccinated          mild       short
## 273         vaccinated        severe        long
## 274       unvaccinated        severe        long
## 275       unvaccinated          mild        long
## 276         vaccinated        severe        long
## 277       unvaccinated          mild       short
## 278         vaccinated        severe       short
## 279       unvaccinated          mild        long
## 280       unvaccinated        severe        long
## 281         vaccinated        severe        long
## 282       unvaccinated          mild        long
## 283       unvaccinated          mild       short
## 284         vaccinated          mild       short
## 285       unvaccinated        severe        long
## 286       unvaccinated          mild        long
## 287       unvaccinated          mild        long
## 288         vaccinated        severe        long
## 289         vaccinated        severe        long
## 290       unvaccinated          mild       short
## 291       unvaccinated        severe        long
## 292         vaccinated        severe        long
## 293       unvaccinated          mild        long
## 294       unvaccinated          mild       short
## 295       unvaccinated          mild       short
## 296         vaccinated        severe       short
## 297       unvaccinated          mild        long
## 298       unvaccinated        severe       short
## 299         vaccinated        severe        long
## 300         vaccinated        severe       short
## 301       unvaccinated          mild        long
```

```
## 302      unvaccinated          mild       long
## 303        vaccinated        severe      short
## 304        vaccinated        severe       long
## 305        vaccinated        severe      short
## 306        vaccinated          mild      short
## 307        vaccinated        severe      short
## 308        vaccinated          mild       long
## 309        vaccinated        severe       long
## 310        vaccinated        severe      short
## 311      unvaccinated          mild       long
## 312        vaccinated        severe       long
## 313      unvaccinated          mild      short
## 314      unvaccinated          mild       long
## 315        vaccinated          mild      short
## 316        vaccinated        severe      short
## 317        vaccinated        severe       long
## 318        vaccinated          mild       long
## 319        vaccinated        severe       long
## 320        vaccinated        severe       long
## 321      unvaccinated          mild       long
## 322      unvaccinated          mild       long
## 323        vaccinated          mild      short
## 324      unvaccinated        severe       long
## 325      unvaccinated        severe      short
## 326      unvaccinated          mild       long
## 327      unvaccinated          mild       long
## 328        vaccinated        severe       long
## 329      unvaccinated          mild      short
## 330      unvaccinated          mild      short
## 331        vaccinated        severe       long
## 332      unvaccinated          mild       long
## 333      unvaccinated        severe       long
## 334      unvaccinated          mild      short
## 335      unvaccinated          mild       long
## 336        vaccinated        severe       long
## 337      unvaccinated          mild      short
## 338      unvaccinated        severe      short
## 339        vaccinated        severe       long
## 340      unvaccinated          mild       long
## 341      unvaccinated          mild      short
## 342        vaccinated        severe       long
## 343        vaccinated        severe       long
## 344        vaccinated        severe      short
## 345      unvaccinated          mild      short
## 346        vaccinated        severe       long
## 347        vaccinated        severe      short
## 348        vaccinated        severe       long
## 349      unvaccinated        severe       long
## 350      unvaccinated        severe       long
## 351        vaccinated        severe       long
## 352        vaccinated        severe      short
## 353        vaccinated        severe       long
## 354      unvaccinated          mild       long
## 355      unvaccinated        severe       long
```

```
## 356        vaccinated        severe        long
## 357      unvaccinated        severe       short
## 358        vaccinated        severe        long
## 359      unvaccinated          mild        long
## 360        vaccinated        severe        long
## 361      unvaccinated        severe        long
## 362      unvaccinated          mild       short
## 363      unvaccinated        severe       short
## 364      unvaccinated          mild       short
## 365        vaccinated        severe        long
## 366      unvaccinated          mild        long
## 367      unvaccinated          mild       short
## 368      unvaccinated          mild       short
## 369      unvaccinated          mild        long
## 370        vaccinated        severe        long
## 371        vaccinated        severe        long
## 372        vaccinated        severe       short
## 373        vaccinated        severe        long
## 374        vaccinated          mild        long
## 375      unvaccinated          mild        long
## 376      unvaccinated          mild        long
## 377      unvaccinated          mild        long
## 378      unvaccinated          mild        long
## 379      unvaccinated        severe        long
## 380      unvaccinated          mild        long
## 381        vaccinated          mild       short
## 382      unvaccinated          mild       short
## 383      unvaccinated          mild        long
## 384        vaccinated        severe       short
## 385      unvaccinated          mild       short
## 386      unvaccinated          mild        long
## 387      unvaccinated          mild       short
## 388      unvaccinated        severe        long
## 389      unvaccinated          mild       short
## 390        vaccinated        severe        long
## 391      unvaccinated          mild        long
## 392      unvaccinated        severe        long
## 393        vaccinated        severe       short
## 394        vaccinated          mild        long
## 395      unvaccinated          mild       short
## 396        vaccinated          mild        long
## 397      unvaccinated        severe        long
## 398        vaccinated        severe       short
## 399      unvaccinated          mild       short
## 400        vaccinated        severe        long
## 401      unvaccinated          mild       short
## 402      unvaccinated        severe       short
## 403        vaccinated          mild        long
## 404        vaccinated        severe       short
## 405        vaccinated        severe        long
## 406      unvaccinated        severe        long
## 407      unvaccinated          mild        long
## 408        vaccinated        severe        long
## 409      unvaccinated          mild        long
```

```
## 410          unvaccinated          mild      short
## 411          unvaccinated        severe       long
## 412          unvaccinated          mild       long
## 413          unvaccinated        severe       long
## 414            vaccinated        severe      short
## 415          unvaccinated          mild      short
## 416          unvaccinated          mild      short
## 417          unvaccinated          mild       long
## 418          unvaccinated        severe       long
## 419            vaccinated        severe      short
## 420            vaccinated          mild       long
## 421          unvaccinated          mild       long
## 422            vaccinated        severe       long
## 423          unvaccinated          mild      short
## 424          unvaccinated        severe       long
## 425            vaccinated        severe       long
## 426            vaccinated          mild      short
## 427            vaccinated        severe       long
## 428          unvaccinated          mild      short
## 429            vaccinated        severe       long
## 430          unvaccinated          mild      short
## 431            vaccinated          mild      short
## 432            vaccinated        severe       long
## 433          unvaccinated          mild      short
## 434          unvaccinated          mild      short
## 435          unvaccinated          mild      short
## 436            vaccinated          mild      short
## 437          unvaccinated          mild       long
## 438          unvaccinated          mild      short
## 439          unvaccinated          mild       long
## 440            vaccinated        severe      short
## 441          unvaccinated          mild       long
## 442            vaccinated        severe       long
## 443          unvaccinated          mild      short
## 444          unvaccinated          mild      short
## 445          unvaccinated          mild      short
## 446          unvaccinated        severe      short
## 447            vaccinated        severe       long
## 448            vaccinated        severe      short
## 449          unvaccinated          mild      short
## 450          unvaccinated          mild       long
## 451            vaccinated        severe       long
## 452          unvaccinated          mild      short
## 453            vaccinated        severe       long
## 454            vaccinated        severe       long
## 455          unvaccinated        severe      short
## 456            vaccinated        severe       long
## 457            vaccinated        severe      short
## 458            vaccinated        severe       long
## 459            vaccinated        severe       long
## 460            vaccinated        severe       long
## 461            vaccinated        severe       long
## 462          unvaccinated          mild       long
## 463            vaccinated          mild      short
```

```
## 464         vaccinated              mild        short
## 465         vaccinated            severe         long
## 466         vaccinated            severe         long
## 467         vaccinated            severe         long
## 468         vaccinated            severe         long
## 469         vaccinated              mild        short
## 470         vaccinated            severe         long
## 471       unvaccinated              mild         long
## 472         vaccinated            severe        short
## 473       unvaccinated              mild        short
## 474       unvaccinated              mild         long
## 475       unvaccinated            severe         long
## 476       unvaccinated              mild         long
## 477       unvaccinated              mild         long
## 478       unvaccinated              mild        short
## 479       unvaccinated            severe         long
## 480       unvaccinated              mild         long
## 481       unvaccinated            severe         long
## 482         vaccinated            severe        short
## 483       unvaccinated            severe         long
## 484         vaccinated            severe        short
## 485         vaccinated              mild        short
## 486       unvaccinated              mild         long
## 487       unvaccinated              mild         long
## 488       unvaccinated            severe         long
## 489       unvaccinated            severe         long
## 490       unvaccinated              mild        short
## 491         vaccinated            severe         long
## 492         vaccinated              mild         long
## 493       unvaccinated              mild         long
## 494         vaccinated            severe        short
## 495       unvaccinated              mild         long
## 496       unvaccinated            severe        short
## 497       unvaccinated            severe         long
## 498       unvaccinated            severe        short
## 499       unvaccinated            severe         long
## 500         vaccinated            severe         long
## 501         vaccinated            severe        short
## 502         vaccinated            severe        short
## 503         vaccinated            severe        short
## 504       unvaccinated              mild         long
## 505       unvaccinated              mild         long
## 506         vaccinated            severe         long
## 507         vaccinated            severe         long
## 508         vaccinated            severe         long
## 509         vaccinated            severe         long
## 510       unvaccinated            severe         long
## 511       unvaccinated              mild         long
## 512       unvaccinated              mild         long
## 513         vaccinated              mild         long
## 514       unvaccinated              mild        short
## 515         vaccinated            severe        short
## 516       unvaccinated            severe        short
## 517       unvaccinated              mild        short
```

```
## 518     unvaccinated      severe       long
## 519     unvaccinated        mild       long
## 520       vaccinated      severe       long
## 521       vaccinated      severe       long
## 522       vaccinated        mild      short
## 523       vaccinated      severe      short
## 524       vaccinated      severe      short
## 525     unvaccinated        mild       long
## 526       vaccinated      severe       long
## 527     unvaccinated      severe       long
## 528     unvaccinated        mild      short
## 529     unvaccinated        mild       long
## 530       vaccinated      severe       long
## 531     unvaccinated        mild       long
## 532       vaccinated      severe       long
## 533       vaccinated      severe      short
## 534       vaccinated      severe       long
## 535     unvaccinated      severe       long
## 536       vaccinated        mild       long
## 537     unvaccinated        mild       long
## 538     unvaccinated      severe       long
## 539     unvaccinated        mild      short
## 540     unvaccinated        mild      short
## 541       vaccinated        mild      short
## 542     unvaccinated        mild       long
## 543       vaccinated        mild      short
## 544       vaccinated      severe       long
## 545       vaccinated      severe      short
## 546       vaccinated        mild      short
## 547     unvaccinated        mild       long
## 548     unvaccinated        mild      short
## 549       vaccinated      severe       long
## 550     unvaccinated        mild       long
## 551       vaccinated      severe       long
## 552     unvaccinated      severe       long
## 553       vaccinated      severe      short
## 554       vaccinated      severe       long
## 555     unvaccinated      severe      short
## 556       vaccinated      severe       long
## 557       vaccinated      severe       long
## 558       vaccinated        mild      short
## 559       vaccinated      severe      short
## 560     unvaccinated        mild       long
## 561       vaccinated      severe       long
## 562     unvaccinated        mild      short
## 563       vaccinated      severe       long
## 564     unvaccinated        mild       long
## 565       vaccinated      severe       long
## 566       vaccinated      severe       long
## 567     unvaccinated        mild      short
## 568     unvaccinated        mild      short
## 569     unvaccinated        mild      short
## 570       vaccinated      severe      short
## 571     unvaccinated        mild       long
```

```
## 572          unvaccinated          severe          long
## 573            vaccinated          severe          long
## 574            vaccinated          severe          long
## 575          unvaccinated          severe         short
## 576            vaccinated          severe          long
## 577            vaccinated          severe         short
## 578            vaccinated            mild         short
## 579          unvaccinated            mild         short
## 580            vaccinated          severe          long
## 581            vaccinated          severe          long
## 582          unvaccinated            mild          long
## 583          unvaccinated          severe         short
## 584          unvaccinated            mild         short
## 585            vaccinated          severe         short
## 586            vaccinated          severe         short
## 587            vaccinated            mild         short
## 588          unvaccinated            mild         short
## 589          unvaccinated            mild         short
## 590          unvaccinated            mild         short
## 591          unvaccinated          severe          long
## 592            vaccinated          severe          long
## 593          unvaccinated            mild         short
## 594            vaccinated          severe          long
## 595            vaccinated            mild          long
## 596          unvaccinated            mild         short
## 597          unvaccinated          severe          long
## 598            vaccinated            mild          long
## 599            vaccinated          severe         short
## 600            vaccinated          severe          long
## 601          unvaccinated          severe          long
## 602            vaccinated            mild          long
## 603            vaccinated          severe          long
## 604            vaccinated          severe         short
## 605            vaccinated          severe          long
## 606            vaccinated            mild          long
## 607          unvaccinated          severe          long
## 608            vaccinated          severe          long
## 609          unvaccinated            mild         short
## 610          unvaccinated          severe          long
## 611          unvaccinated            mild          long
## 612          unvaccinated          severe          long
## 613          unvaccinated            mild         short
## 614          unvaccinated            mild         short
## 615          unvaccinated            mild         short
## 616            vaccinated          severe          long
## 617          unvaccinated          severe          long
## 618            vaccinated          severe          long
## 619          unvaccinated            mild         short
## 620          unvaccinated            mild         short
## 621            vaccinated          severe          long
## 622            vaccinated          severe          long
## 623            vaccinated          severe         short
## 624            vaccinated            mild          long
## 625          unvaccinated            mild          long
```

```
## 626        vaccinated        severe        short
## 627      unvaccinated          mild         long
## 628        vaccinated        severe         long
## 629        vaccinated        severe         long
## 630      unvaccinated        severe         long
## 631      unvaccinated          mild        short
## 632      unvaccinated        severe         long
## 633      unvaccinated          mild         long
## 634        vaccinated        severe         long
## 635      unvaccinated          mild         long
## 636        vaccinated        severe         long
## 637      unvaccinated        severe        short
## 638        vaccinated        severe        short
## 639        vaccinated          mild         long
## 640        vaccinated        severe         long
## 641      unvaccinated          mild        short
## 642        vaccinated        severe        short
## 643        vaccinated        severe         long
## 644      unvaccinated          mild         long
## 645        vaccinated        severe         long
## 646      unvaccinated        severe         long
## 647      unvaccinated          mild        short
## 648        vaccinated        severe        short
## 649        vaccinated        severe         long
## 650      unvaccinated        severe         long
## 651        vaccinated        severe         long
## 652        vaccinated          mild        short
## 653      unvaccinated          mild        short
## 654        vaccinated        severe        short
## 655        vaccinated        severe        short
## 656      unvaccinated          mild         long
## 657        vaccinated        severe         long
## 658        vaccinated        severe        short
## 659      unvaccinated          mild        short
## 660      unvaccinated        severe         long
## 661        vaccinated        severe         long
## 662      unvaccinated          mild        short
## 663        vaccinated        severe         long
## 664      unvaccinated          mild        short
## 665        vaccinated          mild         long
## 666        vaccinated        severe         long
## 667      unvaccinated        severe        short
## 668      unvaccinated          mild         long
## 669      unvaccinated          mild         long
## 670      unvaccinated        severe        short
## 671        vaccinated        severe         long
## 672        vaccinated        severe         long
## 673        vaccinated        severe        short
## 674      unvaccinated          mild        short
## 675        vaccinated        severe         long
## 676        vaccinated        severe        short
## 677        vaccinated        severe         long
## 678      unvaccinated        severe        short
## 679      unvaccinated          mild        short
```

```
## 680    vaccinated     severe    long
## 681    vaccinated       mild    long
## 682    vaccinated     severe   short
## 683  unvaccinated       mild    long
## 684  unvaccinated       mild   short
## 685  unvaccinated     severe    long
## 686    vaccinated     severe    long
## 687  unvaccinated       mild    long
## 688  unvaccinated       mild   short
## 689    vaccinated     severe   short
## 690    vaccinated     severe    long
## 691  unvaccinated       mild   short
## 692    vaccinated       mild    long
## 693    vaccinated       mild   short
## 694  unvaccinated       mild    long
## 695  unvaccinated       mild    long
## 696  unvaccinated       mild   short
## 697    vaccinated     severe    long
## 698  unvaccinated       mild    long
## 699  unvaccinated     severe   short
## 700    vaccinated     severe    long
## 701  unvaccinated     severe   short
## 702  unvaccinated       mild    long
## 703  unvaccinated     severe    long
## 704  unvaccinated       mild   short
## 705  unvaccinated       mild   short
## 706    vaccinated     severe    long
## 707    vaccinated     severe    long
## 708  unvaccinated       mild    long
## 709  unvaccinated       mild    long
## 710  unvaccinated       mild    long
## 711  unvaccinated       mild   short
## 712    vaccinated     severe   short
## 713  unvaccinated       mild   short
## 714  unvaccinated       mild   short
## 715    vaccinated     severe    long
## 716    vaccinated       mild    long
## 717    vaccinated     severe    long
## 718    vaccinated     severe    long
## 719  unvaccinated       mild   short
## 720  unvaccinated       mild    long
## 721    vaccinated       mild    long
## 722    vaccinated       mild   short
## 723  unvaccinated       mild    long
## 724  unvaccinated       mild   short
## 725  unvaccinated       mild    long
## 726  unvaccinated       mild    long
## 727  unvaccinated       mild    long
## 728    vaccinated       mild    long
## 729  unvaccinated     severe    long
## 730    vaccinated     severe    long
## 731  unvaccinated       mild    long
## 732    vaccinated       mild    long
## 733  unvaccinated     severe   short
```

```
## 734       vaccinated       mild       long
## 735     unvaccinated       mild       long
## 736       vaccinated     severe       long
## 737     unvaccinated       mild       long
## 738     unvaccinated       mild      short
## 739     unvaccinated     severe       long
## 740     unvaccinated       mild      short
## 741     unvaccinated       mild       long
## 742     unvaccinated       mild       long
## 743     unvaccinated       mild       long
## 744       vaccinated       mild      short
## 745     unvaccinated     severe       long
## 746     unvaccinated     severe       long
## 747     unvaccinated       mild      short
## 748     unvaccinated       mild      short
## 749       vaccinated     severe       long
## 750       vaccinated     severe      short
## 751       vaccinated       mild       long
## 752     unvaccinated       mild      short
## 753       vaccinated       mild       long
## 754       vaccinated     severe      short
## 755     unvaccinated       mild      short
## 756     unvaccinated       mild       long
## 757       vaccinated     severe       long
## 758       vaccinated     severe      short
## 759       vaccinated     severe       long
## 760       vaccinated     severe       long
## 761       vaccinated     severe      short
## 762       vaccinated     severe      short
## 763       vaccinated     severe       long
## 764       vaccinated     severe       long
## 765     unvaccinated     severe       long
## 766       vaccinated     severe       long
## 767       vaccinated     severe       long
## 768     unvaccinated       mild       long
## 769     unvaccinated       mild      short
## 770     unvaccinated       mild       long
## 771       vaccinated     severe       long
## 772     unvaccinated     severe       long
## 773       vaccinated     severe      short
## 774     unvaccinated     severe       long
## 775       vaccinated     severe       long
## 776       vaccinated       mild      short
## 777     unvaccinated       mild      short
## 778       vaccinated       mild       long
## 779       vaccinated     severe       long
## 780       vaccinated       mild      short
## 781       vaccinated       mild       long
## 782     unvaccinated       mild      short
## 783     unvaccinated       mild       long
## 784     unvaccinated     severe       long
## 785     unvaccinated       mild      short
## 786     unvaccinated       mild      short
## 787       vaccinated     severe       long
```

```
## 788       vaccinated        mild      short
## 789       vaccinated      severe      short
## 790     unvaccinated        mild      short
## 791     unvaccinated        mild      short
## 792     unvaccinated      severe       long
## 793       vaccinated      severe      short
## 794     unvaccinated        mild      short
## 795       vaccinated        mild      short
## 796       vaccinated      severe       long
## 797       vaccinated        mild      short
## 798       vaccinated      severe      short
## 799       vaccinated        mild      short
## 800       vaccinated      severe       long
## 801     unvaccinated        mild       long
## 802     unvaccinated        mild      short
## 803     unvaccinated      severe      short
## 804     unvaccinated        mild       long
## 805       vaccinated      severe       long
## 806       vaccinated      severe      short
## 807     unvaccinated        mild      short
## 808     unvaccinated        mild       long
## 809       vaccinated      severe       long
## 810       vaccinated      severe      short
## 811     unvaccinated        mild       long
## 812     unvaccinated        mild       long
## 813     unvaccinated        mild      short
## 814     unvaccinated        mild      short
## 815     unvaccinated        mild      short
## 816     unvaccinated        mild      short
## 817       vaccinated      severe       long
## 818       vaccinated        mild      short
## 819       vaccinated      severe      short
## 820     unvaccinated      severe       long
## 821       vaccinated        mild       long
## 822       vaccinated      severe      short
## 823     unvaccinated      severe       long
## 824     unvaccinated        mild       long
## 825       vaccinated        mild      short
## 826     unvaccinated        mild      short
## 827       vaccinated      severe       long
## 828     unvaccinated        mild      short
## 829       vaccinated      severe       long
## 830     unvaccinated      severe      short
## 831       vaccinated        mild       long
## 832       vaccinated        mild      short
## 833       vaccinated      severe       long
## 834     unvaccinated        mild      short
## 835     unvaccinated        mild       long
## 836     unvaccinated        mild      short
## 837     unvaccinated        mild       long
## 838     unvaccinated        mild      short
## 839     unvaccinated        mild       long
## 840     unvaccinated        mild       long
## 841     unvaccinated        mild      short
```

```
## 842     vaccinated     mild    short
## 843   unvaccinated     mild     long
## 844   unvaccinated     mild     long
## 845     vaccinated   severe     long
## 846     vaccinated   severe     long
## 847   unvaccinated     mild     long
## 848   unvaccinated     mild     long
## 849   unvaccinated     mild     long
## 850     vaccinated     mild    short
## 851     vaccinated   severe     long
## 852   unvaccinated   severe     long
## 853     vaccinated   severe     long
## 854     vaccinated     mild     long
## 855     vaccinated   severe     long
## 856     vaccinated   severe    short
## 857   unvaccinated   severe    short
## 858   unvaccinated     mild    short
## 859     vaccinated   severe     long
## 860     vaccinated   severe     long
## 861     vaccinated     mild     long
## 862     vaccinated   severe     long
## 863     vaccinated   severe     long
## 864     vaccinated     mild     long
## 865   unvaccinated     mild     long
## 866   unvaccinated   severe     long
## 867   unvaccinated     mild     long
## 868     vaccinated   severe    short
## 869   unvaccinated   severe     long
## 870   unvaccinated   severe    short
## 871   unvaccinated     mild    short
## 872     vaccinated   severe    short
## 873     vaccinated     mild     long
## 874   unvaccinated   severe     long
## 875   unvaccinated     mild     long
## 876     vaccinated   severe     long
## 877     vaccinated   severe     long
## 878     vaccinated     mild     long
## 879     vaccinated   severe    short
## 880   unvaccinated     mild     long
## 881   unvaccinated     mild    short
## 882     vaccinated   severe    short
## 883     vaccinated     mild    short
## 884   unvaccinated     mild    short
## 885   unvaccinated   severe     long
## 886     vaccinated   severe     long
## 887   unvaccinated   severe     long
## 888   unvaccinated     mild     long
## 889   unvaccinated   severe     long
## 890   unvaccinated   severe     long
## 891   unvaccinated   severe     long
## 892     vaccinated   severe     long
## 893   unvaccinated     mild     long
## 894   unvaccinated     mild    short
## 895   unvaccinated     mild    short
```

```
## 896    unvaccinated      mild     long
## 897    unvaccinated      mild    short
## 898    unvaccinated      mild    short
## 899    unvaccinated      mild     long
## 900    unvaccinated      mild    short
## 901    unvaccinated    severe     long
## 902      vaccinated    severe     long
## 903      vaccinated    severe     long
## 904    unvaccinated      mild    short
## 905      vaccinated    severe     long
## 906      vaccinated    severe    short
## 907    unvaccinated      mild    short
## 908      vaccinated    severe     long
## 909      vaccinated      mild    short
## 910      vaccinated    severe    short
## 911      vaccinated    severe     long
## 912    unvaccinated    severe     long
## 913      vaccinated    severe    short
## 914      vaccinated    severe     long
## 915      vaccinated    severe     long
## 916    unvaccinated      mild    short
## 917    unvaccinated      mild     long
## 918    unvaccinated    severe     long
## 919    unvaccinated      mild    short
## 920      vaccinated    severe     long
## 921    unvaccinated    severe    short
## 922      vaccinated    severe    short
## 923    unvaccinated      mild     long
## 924    unvaccinated      mild    short
## 925      vaccinated    severe     long
## 926    unvaccinated      mild    short
## 927    unvaccinated      mild    short
## 928    unvaccinated      mild     long
## 929      vaccinated      mild    short
## 930      vaccinated      mild     long
## 931    unvaccinated      mild     long
## 932    unvaccinated      mild    short
## 933    unvaccinated    severe     long
## 934      vaccinated    severe     long
## 935    unvaccinated    severe    short
## 936      vaccinated    severe    short
## 937    unvaccinated      mild     long
## 938      vaccinated    severe     long
## 939    unvaccinated      mild    short
## 940      vaccinated      mild     long
## 941    unvaccinated      mild    short
## 942    unvaccinated      mild     long
## 943      vaccinated    severe     long
## 944    unvaccinated      mild    short
## 945    unvaccinated      mild     long
## 946    unvaccinated      mild    short
## 947    unvaccinated      mild     long
## 948      vaccinated      mild    short
## 949    unvaccinated      mild    short
```

25

```
## 950      unvaccinated         mild      short
## 951        vaccinated         mild      short
## 952      unvaccinated         mild       long
## 953      unvaccinated         mild      short
## 954      unvaccinated       severe       long
## 955        vaccinated         mild       long
## 956        vaccinated         mild       long
## 957      unvaccinated       severe      short
## 958        vaccinated         mild       long
## 959      unvaccinated         mild      short
## 960        vaccinated       severe       long
## 961        vaccinated       severe       long
## 962      unvaccinated         mild       long
## 963      unvaccinated         mild       long
## 964      unvaccinated         mild       long
## 965      unvaccinated       severe       long
## 966        vaccinated       severe       long
## 967      unvaccinated       severe       long
## 968        vaccinated       severe       long
## 969      unvaccinated         mild      short
## 970        vaccinated       severe      short
## 971        vaccinated       severe       long
## 972        vaccinated         mild      short
## 973        vaccinated         mild      short
## 974      unvaccinated         mild      short
## 975        vaccinated       severe      short
## 976      unvaccinated         mild      short
## 977      unvaccinated         mild       long
## 978        vaccinated       severe       long
## 979      unvaccinated       severe       long
## 980      unvaccinated         mild       long
## 981      unvaccinated       severe       long
## 982      unvaccinated         mild      short
## 983        vaccinated       severe      short
## 984      unvaccinated         mild       long
## 985      unvaccinated         mild      short
## 986        vaccinated       severe       long
## 987      unvaccinated         mild      short
## 988        vaccinated       severe       long
## 989        vaccinated       severe       long
## 990        vaccinated       severe      short
## 991        vaccinated       severe      short
## 992      unvaccinated       severe       long
## 993        vaccinated       severe       long
## 994        vaccinated         mild      short
## 995        vaccinated       severe       long
## 996        vaccinated         mild       long
## 997      unvaccinated       severe       long
## 998        vaccinated       severe      short
## 999      unvaccinated         mild      short
## 1000     unvaccinated         mild       long
```

```r
delta <- calc_delta(vax)
print(delta)
```

```
## [1] -0.09706736
```

5.2 Stratify the table by disease severity. As in the previous problem, calculate the same treatment effect in each strata.

```
dfMild <- vax[which(vax$disease_severity == "mild"),]

delta_mild <- calc_delta(dfMild)
print(delta_mild)
```

```
## [1] -0.008308532
```

```
dfSevere <- vax[which(vax$disease_severity == "severe"),]

delta_severe <- calc_delta(dfSevere)
print(delta_severe)
```

```
## [1] 0.03355211
```

5.3 Based on the summary of the treatment effect that you observe in the combined and stratified tables, does vaccination help shorten recovery?

Based on the combined table, vacciaation does not help shorten recovery. Based on the combined tables, it seems as though vaccination can help if the disease is severe but not if it is a mild case.

5.4 Which measure of treatment effect is most persuasive? The combined estimate or the stratified estimates? Which estimate should you rely on? Explain why, creating a DAG to represent relationship between the variables.

The most persuasive measure is the startified data because it shows that vaccination has a different impact on recovery. Therefore, we should rely on the stratified treatment effect estimate rather than the unstratified data.

Vaccination Status (V) –> Disease Severity (D) –> Recovery Time (R) | +————————————————— ——> Recovery Time (R)