

# Advent of Code 2018: Some solutions

*Katie LeVan*

This is a narrative of my solutions to the [Advent of Code 2018](#) Challenge.

## — Day 1: Chronal Calibration —

“We’ve detected some temporal anomalies,” one of Santa’s Elves at the Temporal Anomaly Research and Detection Instrument Station tells you. She sounded pretty worried when she called you down here. “At 500-year intervals into the past, someone has been changing Santa’s history!”

“The good news is that the changes won’t propagate to our time stream for another 25 days, and we have a device” - she attaches something to your wrist - “that will let you fix the changes with no such propagation delay. It’s configured to send you 500 years further into the past every few days; that was the best we could do on such short notice.”

“The bad news is that we are detecting roughly fifty anomalies throughout time; the device will indicate fixed anomalies with stars. The other bad news is that we only have one device and you’re the best person for the job! Good lu—” She taps a button on the device and you suddenly feel like you’re falling. To save Christmas, you need to get all fifty stars by December 25th.

Collect stars by solving puzzles. Two puzzles will be made available on each day in the advent calendar; the second puzzle is unlocked when you complete the first. Each puzzle grants one star. Good luck!

After feeling like you’ve been falling for a few minutes, you look at the device’s tiny screen. “Error: Device must be calibrated before first use. Frequency drift detected. Cannot maintain destination lock.” Below the message, the device shows a sequence of changes in frequency (your puzzle input). A value like +6 means the current frequency increases by 6; a value like -3 means the current frequency decreases by 3.

For example, if the device displays frequency changes of +1, -2, +3, +1, then starting from a frequency of zero, the following changes would occur:

- Current frequency 0, change of +1; resulting frequency 1.
- Current frequency 1, change of -2; resulting frequency -1.
- Current frequency -1, change of +3; resulting frequency 2.
- Current frequency 2, change of +1; resulting frequency 3.

In this example, the resulting frequency is 3.

Here are other example situations:

- +1, +1, +1 results in 3
- +1, +1, -2 results in 0
- -1, -2, -3 results in -6

**Starting with a frequency of zero, what is the resulting frequency after all of the changes in frequency have been applied?**

```
# Input Data For Puzzle
chrg = read.delim(paste(wd, 'data/day-1-input.txt', sep='/'), header = F,
                  col.names = 'freq', encoding = 'UTF-8')
```

Because all additions and subtractions cancel each other out, the simple answer is to use the `sum()` function and return the net value.

Answer: `sum(chg$freq)` is 525

## — Part Two —

You notice that the device repeats the same frequency change list over and over. To calibrate the device, you need to find the first frequency it reaches twice.

For example, using the same list of changes above, the device would loop as follows:

- Current frequency 0, change of +1; resulting frequency 1.
- Current frequency 1, change of -2; resulting frequency -1.
- Current frequency -1, change of +3; resulting frequency 2.
- Current frequency 2, change of +1; resulting frequency 3. (At this point, the device continues from the start of the list.)
- Current frequency 3, change of +1; resulting frequency 4.
- Current frequency 4, change of -2; resulting frequency 2, which has already been seen.

In this example, the first frequency reached twice is 2. Note that your device might need to repeat its list of frequency changes many times before a duplicate frequency is found, and that duplicates might be found while in the middle of processing the list.

Here are other examples:

- +1, -1 first reaches 0 twice.
- +3, +3, +4, -2, -4 first reaches 10 twice.
- -6, +3, +8, +5, -6 first reaches 5 twice.
- +7, +7, -2, -7, -4 first reaches 14 twice.

**What is the first frequency your device reaches twice?**

My first instinct was to loop through the first list and see if a duplicate number emerged.

```
# Determine the list of increments
t = 0 # initialize t
add = vector() # create a container for each time increment

# Loop through the list of time steps
for(n in 1:nrow(chg)){
  # Current time step
  t = (t + chg$freq[n])
  # Log the time step
  add = c(add, t)
}

rm(n,t) # cleanup extra objects
```

Evaluating `any(duplicated(add))` yielded `FALSE` - therefore every number in the list was unique, with no duplicates. Next, I tried implementing a while loop to cycle through the `chg` list, keep track of each increment, and start over if the duplicate was not found. I quickly realized that this approach was taking minutes and no solution had been found.

Clearly, this wasn't the greatest solution from a [computational complexity](#) perspective.

Looking at the patterns in the example solutions more closely, I realized that the second round of numbers in the list was just a repeat of the first set, where each number was added to the sum of all numbers in the original list.

For example, if the original list of offsets +1, -2, +3, +1 had a sum of 3 and the increments at each time step were +1, -1, +2, +3 where the offsets were applied to zero, then the next time steps would be:

- +4, +2, +5, +6
- +7, +5, +8, +9
- +10, +8, +11, +12 etc.

```
# Part 2: Determine first duplicate number
# when duplicate is in existing list, can use 'duplicated' function to find dup
# when many iterations required, can use the fact that the sum() of the list
# is the new starting number whereby all list items are incremented

# If there is a duplicate in the first try
if(any(duplicated(add))){
  # store that duplicate
  dup = (add[duplicated(add)][1])
} else {
  # Increment entire list by the sum
  # until duplicate is found
  final = add;
  while(!any(duplicated(final))){
    # Log the time step
    # If a duplicate is found
    # then the while loop ends
    add = (sum(chg$freq) + add)
    final = c(final, add)
  }
  dup = (final[duplicated(final)][1]) # Else, this is the first duplicate
}
```

When this logic is applied, the first duplicate seen is at 75749 (requiring looping over the list 145 times), which explains in part why the slower one-by-one evaluation method was taking eons to complete (since the first duplicate is at position 146725 in the list).

## — Day 2: Inventory Management System —

You stop falling through time, catch your breath, and check the screen on the device. “Destination reached. Current Year: 1518. Current Location: North Pole Utility Closet 83N10.” You made it! Now, to find those anomalies.

Outside the utility closet, you hear footsteps and a voice. “...I’m not sure either. But now that so many people have chimneys, maybe he could sneak in that way?” Another voice responds, “Actually, we’ve been working on a new kind of suit that would let him fit through tight spaces like that. But, I heard that a few days ago, they lost the prototype fabric, the design plans, everything! Nobody on the team can even seem to remember important details of the project!”

“Wouldn’t they have had enough fabric to fill several boxes in the warehouse? They’d be stored together, so the box IDs should be similar. Too bad it would take forever to search the warehouse for two similar box IDs...” They walk too far away to hear any more.

Late at night, you sneak to the warehouse - who knows what kinds of paradoxes you could cause if you were discovered - and use your fancy wrist device to quickly scan every box and produce a list of the likely candidates (your puzzle input).

To make sure you didn’t miss any, you scan the likely candidate boxes again, counting the number that have an ID containing exactly two of any letter and then separately counting those with exactly three of any letter. You can multiply those two counts together to get a rudimentary checksum and compare it to what your device predicts.

For example, if you see the following box IDs:

- abcdef contains no letters that appear exactly two or three times.
- bababc contains two a and three b, so it counts for both.
- abbccde contains two b, but no letter appears exactly three times.
- abcccd contains three c, but no letter appears exactly two times.
- aabccd contains two a and two d, but it only counts once.
- abcdee contains two e.
- ababab contains three a and three b, but it only counts once.

Of these box IDs, four of them contain a letter which appears exactly twice, and three of them contain a letter which appears exactly three times. Multiplying these together produces a checksum of  $4 * 3 = 12$ .

**What is the checksum for your list of box IDs?**

```
# Input Data
boxes = read.delim(paste(wd, 'data/day-2-input.txt', sep='/'), header = F,
                  col.names = 'IDs', encoding = 'UTF-8')
```

Using the `strsplit()` function, you can get a list of all single elements from the vector of boxes identifiers. From there, it’s trivial to find duplicates within the list (using the `duplicated()` function). Any duplicates that are exactly double or triple repeats in a string can be logged.

```
# Part 1: Checksum; look for doubles & triples
# First create a list that is easier to work with
boxes.split = strsplit(boxes$IDs, split = '')

# make places to store info about detected doubles and triples
boxes[, c('dbl', 'tpl')] = NA
# For each box
for(n in 1:nrow(boxes)){
  # Get all the letters
  b = boxes.split[[n]]
  # Get the duplicates
  dups = b[duplicated(b)]

  if(length(b)==0){
    boxes$dbl[n] = boxes$tpl[n] = 0
  } else {
    # For each detected duplicate
```

```

for(m in dups){
  # Was it repeated exactly twice?
  if(length(b[b%in%m])%in%2){
    boxes$dbl[n] = 1 # log the answer
  }
  # Was it repeated exactly three times?
  if(length(b[b%in%m])%in%3){
    boxes$tpl[n] = 1 # log the answer
  }
}
}
}

```

The the sums of those logs can be multiplied as the puzzle requires.

Answer: `sum(boxes$dbl, na.rm = T) * sum(boxes$tpl, na.rm = T)` is 5434.

## — Part Two —

Confident that your list of box IDs is complete, you're ready to find the boxes full of prototype fabric.

The boxes will have IDs which differ by exactly one character at the same position in both strings. For example, given the following box IDs:

abcde fghij klmno pqrst fguij axcye wvxyz

The IDs abcde and axcye are close, but they differ by two characters (the second and fourth). However, the IDs fghij and fguij differ by exactly one character, the third (h and u). Those must be the correct boxes.

**What letters are common between the two correct box IDs? (In the example above, this is found by removing the differing character from either ID, producing fgij.)**

Using the `adist()` function from the `utils` library makes this solution super simple. `adist()` will evaluate the approximate Levenshtein string distances between two strings `s` and `t` such that “the minimal possibly weighted number of insertions, deletions and substitutions needed to transform `s` into `t` (so that the transformation exactly matches `t`)” is returned.

```

# Part 2: Figure out which box IDs are only one letter apart
# Use the distance function to assess differences

# A vector of all IDs
boxes = boxes$IDs

# a data frame of distances
# we are looking for the case where adist returns value of 1
mat = as.data.frame(adist(boxes))

# assign names of boxes to the rows
rownames(mat) = boxes

# Return rownames where the value of the data.frame was 1
box.list = (rownames(which(mat == 1, arr.ind = TRUE)))

```

For me, the `box.list` of similar names returns “agitmjdjvlhedpsyoqfzuknpjwt” and “agirmjdjvlhedpsyoqfzuknpjwt”, which differ in the 4th letter. Thus, my puzzle answer was agimjdjvlhedpsyoqfzuknpjwt (removing that 4th different letter).

## — Day 3: No Matter How You Slice It —

The Elves managed to locate the chimney-squeeze prototype fabric for Santa’s suit (thanks to someone who helpfully wrote its box IDs on the wall of the warehouse in the middle of the night). Unfortunately, anomalies are still affecting them - nobody can even agree on how to cut the fabric.

The whole piece of fabric they’re working on is a very large square - at least 1000 inches on each side.

Each Elf has made a claim about which area of fabric would be ideal for Santa’s suit. All claims have an ID and consist of a single rectangle with edges parallel to the edges of the fabric. Each claim’s rectangle is defined as follows:

The number of inches between the left edge of the fabric and the left edge of the rectangle. The number of inches between the top edge of the fabric and the top edge of the rectangle. The width of the rectangle in inches. The height of the rectangle in inches. A claim like `#123 @ 3,2: 5x4` means that claim ID 123 specifies a rectangle 3 inches from the left edge, 2 inches from the top edge, 5 inches wide, and 4 inches tall. Visually, it claims the square inches of fabric represented by `#` (and ignores the square inches of fabric represented by `.`) in the diagram below:

```

.....
.....
...#####...
...#####...
...#####...
...#####...
.....
.....
.....

```

The problem is that many of the claims overlap, causing two or more claims to cover part of the same areas. For example, consider the following claims:

- `#1 @ 1,3: 4x4`
- `#2 @ 3,1: 4x4`
- `#3 @ 5,5: 2x2`

Visually, these claim the following areas:

```

.....
...2222.
...2222.
.11XX22.
.11XX22.
.111133.

```

.111133.

.....

The four square inches marked with X are claimed by both 1 and 2. (Claim 3, while adjacent to the others, does not overlap either of them.)

If the Elves all proceed with their own plans, none of them will have enough fabric. How many square inches of fabric are within two or more claims?

```
library(sp)

# Input Data
suit = read.delim(paste(wd,'data/day-3-input.txt',sep='/'), header = F,
                  col.names = 'loc', encoding = 'UTF-8')
```

The first thing I did was visualize the problem, using a function I wrote called `makePoly` to convert the claim format (i.e., #1 @ 1,3: 4x4) into polygons.

```
# Make a Polygon for each claim
# makePoly will convert each claim into a Polygon
makePoly <- function(input){
  # Get the values
  vals = unlist(strsplit(input, split = ' '))
  # Set the beginning point
  beg = as.numeric(gsub(':', '', unlist(strsplit(vals[3], split = ', ')))) # x, y
  # What is the furthest point?
  inc = as.numeric(unlist(strsplit(vals[4], split = 'x')))) # x, y
  # Make a matrix
  m <- matrix(c(beg, # start corner
                beg[1]+inc[1], beg[2], # X corner
                beg[1]+inc[1], beg[2]+inc[2], # far corner
                beg[1], beg[2]+inc[2]), # Y corner
              ncol = 2, byrow = TRUE)

  out <- Polygons(list(Polygon(m)), gsub('#', '', vals[1]))
  return(out)
}

# Covert each suit location into a polygon
locs = sapply(suit$loc, makePoly)

# Then the SpatialPolygons function can format those polygons
# so they are ready to plot
polys = SpatialPolygons((locs))
```

These are then graphable, for a nice visual of the problem.

```
# Plotting the polygons
plot(polys, border = rainbow(n = 30),
     xlab = "X", ylab = "Y",
     main = "Claims on the Santa Suit material",
```

```
xlim=c(0,1000),ylim=c(0,1000))
axis(1);axis(2); box()
```

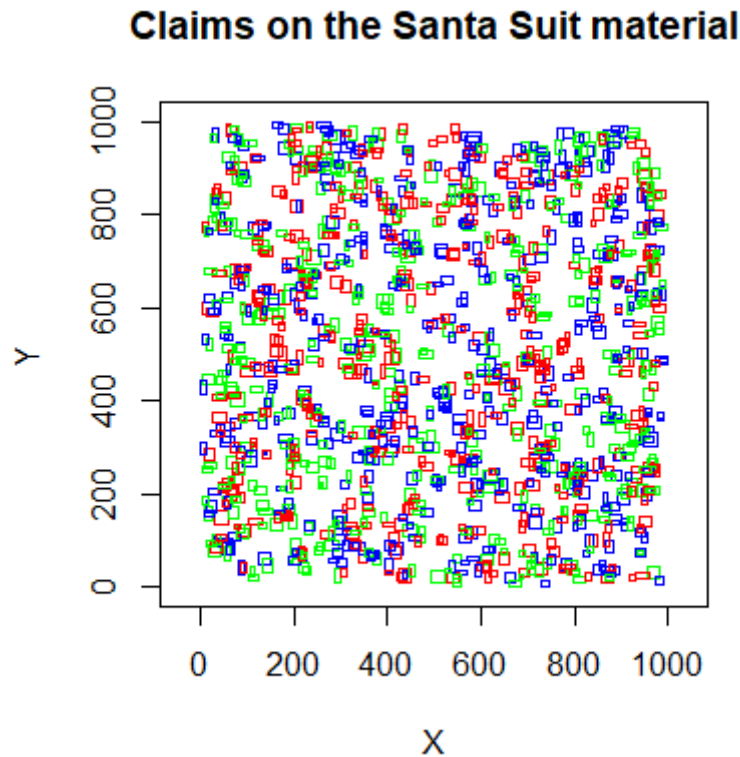


Figure 1: Graphical representation of the problem.

It seemed computationally easier to just convert each claim into a list of coordinate locations it represented, using a function I wrote called `makeList`. From there it's as simple as applying the function and getting the list of duplicate locations.

```
# Convert each item into it's list of addresses,
# then look for ones with > 1 instance
makeList <- function(input){
  # Get the values
  vals = unlist(strsplit(input, split = ' '))
  # Set the beginning point
  beg = as.numeric(gsub(':', '', unlist(strsplit(vals[3], split = ',')))) # x, y
  # What is the furthest point?
  inc = as.numeric(unlist(strsplit(vals[4], split = 'x')))) # x, y

  mini = (beg + 1) # x, y start corner
  maxi = (beg + inc) # x, y max corner
  coord = vector()
  for(x in mini[1]:maxi[1]){
    for (y in mini[2]:maxi[2]){
      coord = c(coord, paste(x,y,sep=','))
    }
  }
}
```



```

    }
  }
  return(coord)
}

# Use the makeList function to transform each suit location
# into the list of coordinates is comprises
coords = (sapply(suit$loc, makeList))
# unlist for ease of use
cor = unlist(coords)
# determine which claim locations are repeated
dup = unique(cor[duplicated(cor)])

length(dup) # area of duplicate claims

```

The area is easy to calculate by looking at the length of the list of unique names of duplicate claim locations, since each unique swatch location is area of size 1. The `length(dup)` is 111266 and the solution to the puzzle.

## — Part Two —

Amidst the chaos, you notice that exactly one claim doesn't overlap by even a single square inch of fabric with any other claim. If you can somehow draw attention to it, maybe the Elves will be able to make Santa's suit after all!

For example, in the claims above, only claim 3 is intact after all claims are made.

### What is the ID of the only claim that doesn't overlap?

Since the suit location list wasn't long, it was quick to iterate over the list and find the claim that didn't have a value in the duplicate claim location list.

```

# Part 2: Find the one square that doesn't overlap
# unduplicated coordinates
for (i in 1:nrow(suit)){
  if(length(coords[[i]][coords[[i]]%in%dup])==0){
    print(suit[i,])
  }
}

```

Revealing that claim # 266 had no overlap.

## — Day 4: Repose Record —

You've sneaked into another supply closet - this time, it's across from the prototype suit manufacturing lab. You need to sneak inside and fix the issues with the suit, but there's a guard stationed outside the lab, so this is as close as you can safely get.

As you search the closet for anything that might help, you discover that you're not the first person to want to sneak in. Covering the walls, someone has spent an hour starting every midnight for the past few months secretly observing this guard post! They've been writing down the ID of the one guard on duty that night

- the Elves seem to have decided that one guard was enough for the overnight shift - as well as when they fall asleep or wake up while at their post (your puzzle input).

For example, consider the following records, which have already been organized into chronological order:

[1518-11-01 00:00] Guard #10 begins shift [1518-11-01 00:05] falls asleep [1518-11-01 00:25] wakes up [1518-11-01 00:30] falls asleep [1518-11-01 00:55] wakes up [1518-11-01 23:58] Guard #99 begins shift [1518-11-02 00:40] falls asleep [1518-11-02 00:50] wakes up [1518-11-03 00:05] Guard #10 begins shift [1518-11-03 00:24] falls asleep [1518-11-03 00:29] wakes up [1518-11-04 00:02] Guard #99 begins shift [1518-11-04 00:36] falls asleep [1518-11-04 00:46] wakes up [1518-11-05 00:03] Guard #99 begins shift [1518-11-05 00:45] falls asleep [1518-11-05 00:55] wakes up

Timestamps are written using year-month-day hour:minute format. The guard falling asleep or waking up is always the one whose shift most recently started. Because all asleep/awake times are during the midnight hour (00:00 - 00:59), only the minute portion (00 - 59) is relevant for those events.

Visually, these records show that the guards are asleep at these times:

```

Date ID Minute 00000000000111111111222222222333333333444444444555555555 0123456789012345678901234567890123
11-01 #10 .....#####.....#####.....
11-02 #99 .....#####..... 11-03 #10 .....#####
11-04 #99 .....#####..... 11-05 #99 .....#####

```

The columns are Date, which shows the month-day portion of the relevant day; ID, which shows the guard on duty that day; and Minute, which shows the minutes during which the guard was asleep within the midnight hour. (The Minute column's header shows the minute's ten's digit in the first row and the one's digit in the second row.) Awake is shown as ., and asleep is shown as #.

Note that guards count as asleep on the minute they fall asleep, and they count as awake on the minute they wake up. For example, because Guard #10 wakes up at 00:25 on 1518-11-01, minute 25 is marked as awake.

If you can figure out the guard most likely to be asleep at a specific time, you might be able to trick that guard into working tonight so you can have the best chance of sneaking in. You have two strategies for choosing the best guard/minute combination.

Strategy 1: Find the guard that has the most minutes asleep. What minute does that guard spend asleep the most?

In the example above, Guard #10 spent the most minutes asleep, a total of 50 minutes (20+25+5), while Guard #99 only slept for a total of 30 minutes (10+10+10). Guard #10 was asleep most during minute 24 (on two days, whereas any other minute the guard was asleep was only seen on one day).

While this example listed the entries in chronological order, your entries are in the order you found them. You'll need to organize them before they can be analyzed.

What is the ID of the guard you chose multiplied by the minute you chose?

(In the above example, the answer would be  $10 * 24 = 240$ .)

I loaded the data into my session and imported a quick function to determine mode. For some reason, R has default functions for calculation of mean and median, but not mode. Mode is useful for determining which minute the guard is most commonly asleep.

```
# Input Data
duty = read.delim(paste(wd, 'data/day-4-input.txt', sep='/'), header = F,
                  col.names = 'log', encoding = 'UTF-8')

# Get the mode
getmode <- function(v) {
```

```

uniqv <- unique(v)
uniqv[which.max(tabulate(match(v, univq)))]
}

```

But the data are not immediately useable, with some cleanup required. First I isolated the timestamps, descriptive text and guard identifier from the duty log. Using some date handling, I was able to sort the dataset pretty easily. Relying on that ordering, I determined which records pertained to specific guards where not otherwise indicated. With that preprocessing complete, I am ready to calculate sleep time.

```

# Part 1: Find the guard that has the most minutes asleep.
# What minute does that guard spend asleep the most?

# Clean up the duty log
duty$date = substr(duty$log,2,17) # Split time stamps from file
duty$info = gsub('\\[[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}\\]',
                 '',duty$log) # get info separate from time stamp
duty$guardID = gsub('[A-z# ]','',duty$info) # find Guard ID from shift starts

# Sort by date
duty$date = as.POSIXct(duty$date)
duty = duty[order(duty$date), ]

# Determine guard on duty
n = 1; glist = which(!duty$guardID%in%'')
while (n<nrow(duty)){
  # What's the current guard's ID?
  gID = duty$guardID[n]
  # When is the next shift?
  if(which(glist%in%n)<length(glist)){
    nxt.shift = glist[which(glist%in%n)+1]
    # Assign all current time stamps to current guard
    duty$guardID[n:(nxt.shift-1)] = gID

    n = nxt.shift
  } else {
    duty$guardID[n:nrow(duty)] = gID
    break
  }
}
}

```

For each guard, the difftime function will yield the time difference between two times. difftime just outputs the single value of endtime - starttime, so i separately generated a string of minutes and ran the 'getmode' function over that to determine the most prevalent minute.

```

# Which guard spent the most time asleep?
duty.summary = data.frame(guardID = unique(duty$guardID))
for(g in 1:nrow(duty.summary)){
  # Isolate that guard's schedule
  sch = duty[duty$guardID%in%duty.summary$guardID[g], ]
  sch$min = as.numeric(substr(sch$date, 15,16))

  sleep = mins = vector()
  for(time in 1:nrow(sch)){

```

```

    if(grepl('sleep', sch$info[time])){
      # Amount of time sleeping
      sleep = c(sleep, difftime(sch$date[time+1],sch$date[time], units = 'mins'))
      mins = c(mins, sch$min[time]:((sch$min[time+1])-1))
    }
  }

  duty.summary$totalMinutesAsleep[g] = sum(sleep)
  duty.summary$bestMin[g] = paste(getmode(mins), collapse = '|')

  # Use for Part 2
  if(!is.na(duty.summary$bestMin[g])){
    duty.summary$numInMode[g] = length(mins[mins%in%duty.summary$bestMin[g]])
  } else {
    duty.summary$numInMode[g] = NA
  }
}

```

From there, the guardID \* most common minute is easy, and gave me an answer of 95199.

```

# What is the ID of the guard you chose multiplied by the minute you chose?
as.numeric(duty.summary$guardID[duty.summary$totalMinutesAsleep%in%max(duty.summary$totalMinutesAsleep)])
as.numeric(duty.summary$bestMin[duty.summary$totalMinutesAsleep%in%max(duty.summary$totalMinutesAsleep)])

```

## — Part Two —

Strategy 2: Of all guards, which guard is most frequently asleep on the same minute?

In the example above, Guard #99 spent minute 45 asleep more than any other guard or minute - three times in total. (In all other cases, any guard spent any minute asleep at most twice.)

**What is the ID of the guard you chose multiplied by the minute you chose?**

(In the above example, the answer would be  $99 * 45 = 4455$ .)

When iterating over the first list, I also evaluated `length(mins[mins%in%duty.summary$bestMin[g]])`; or, in other words, if the best minute is the “mode” of the data for each guard, how many instances did that minute occur for each guard? From there its just multiplying the guardID \* “mode” minute, for the guard with the largest number of sleep times occurring in that minute.

```

# Part 2: Of all guards, which guard is most frequently asleep on the same minute?
# What is the ID of the guard you chose multiplied by the minute you chose?
as.numeric(duty.summary$guardID[duty.summary$numInMode%in%max(duty.summary$numInMode)]) *
as.numeric(duty.summary$bestMin[duty.summary$numInMode%in%max(duty.summary$numInMode)])

```

My puzzle answer was 7887.

## — Day 5: Alchemical Reduction —

You’ve managed to sneak in to the prototype suit manufacturing lab. The Elves are making decent progress, but are still struggling with the suit’s size reduction capabilities.

While the very latest in 1518 alchemical technology might have solved their problem eventually, you can do better. You scan the chemical composition of the suit's material and discover that it is formed by extremely long polymers (one of which is available as your puzzle input).

The polymer is formed by smaller units which, when triggered, react with each other such that two adjacent units of the same type and opposite polarity are destroyed. Units' types are represented by letters; units' polarity is represented by capitalization. For instance, r and R are units with the same type but opposite polarity, whereas r and s are entirely different types and do not react.

For example:

In aA, a and A react, leaving nothing behind. In abBA, bB destroys itself, leaving aA. As above, this then destroys itself, leaving nothing. In abAB, no two adjacent units are of the same type, and so nothing happens. In aabAAB, even though aa and AA are of the same type, their polarities match, and so nothing happens. Now, consider a larger example, dabAcCaCBACcCaDA:

dabAcCaCBACcCaDA The first 'cC' is removed. dabAaCBACcCaDA This creates 'Aa', which is removed. dabCBACcCaDA Either 'cC' or 'Cc' are removed (the result is the same). dabCBACaDA No further actions can be taken. After all possible reactions, the resulting polymer contains 10 units.

**How many units remain after fully reacting the polymer you scanned?**

```
# Input Data
polymer = as.character(read.delim(paste(wd, 'data/day-5-input.txt', sep='/'),
                                   header = F, encoding = 'UTF-8'))
```

This one is pretty straightforward to do using regular expressions and substitution. The gsub() function does both, and when implemented under a while loop, i can iterate through the alphabet to assure that all valid removals have been implemented and break the loop when I've gone through the whole list without removing any additional characters.

```
# Part 1: How many units remain after fully reacting the polymer you scanned?
# Setup; so we can track progress
x = t = 1; start = nchar(polymer)
while(x<3){
  L = LETTERS[t]
  l = letters[t]
  # Remove Aa & aA, and repeat for all letters
  polymer = gsub(paste0(L,l), '', polymer)
  polymer = gsub(paste0(l,L), '', polymer)

  n = nchar(polymer) # is this string getting shorter?
  print(n)
  if(t<26){
    t = t+1
  } else {
    t = 1
    if(start>n){
      start = n
    } else{
      x = x + 1
    }
  }
}
```

My puzzle answer (n) was 11754.

## — Part Two —

Time to improve the polymer.

One of the unit types is causing problems; it's preventing the polymer from collapsing as much as it should. Your goal is to figure out which unit type is causing the most problems, remove all instances of it (regardless of polarity), fully react the remaining polymer, and measure its length.

For example, again using the polymer `dabAcCaCBACcCaDA` from above:

Removing all A/a units produces `dbcCCBcCcD`. Fully reacting this polymer produces `dbCBcD`, which has length 6. Removing all B/b units produces `daAcCaCAcCcaDA`. Fully reacting this polymer produces `daCAcaDA`, which has length 8. Removing all C/c units produces `dabAaBAaDA`. Fully reacting this polymer produces `daDA`, which has length 4. Removing all D/d units produces `abAcCaCBACcCaA`. Fully reacting this polymer produces `abCBAC`, which has length 6. In this example, removing all C/c units was best, producing the answer 4.

**What is the length of the shortest polymer you can produce by removing all units of exactly one type and fully reacting the result?**

This required the same implementation as before with the twist that I remove (via regex) one letter of any case before starting the process.

```
# Part 2: If one letter (of both cases) were removed,
# which one should be removed to make the most compact polymer?
poly = vector()
for(alpha in LETTERS){
  polymer = as.character(read.delim('data/day-5-input.txt',
                                     header = F, encoding = 'UTF-8'))
  polymer = gsub(paste0('[',alpha,tolower(alpha),']{1}'), '', polymer)
  x = t = 1; start = nchar(polymer)
  while(x<3){
    L = LETTERS[t]
    l = letters[t]
    # Remove Aa & aA, and repeat for all letters
    polymer = gsub(paste0(L,l), '', polymer)
    polymer = gsub(paste0(l,L), '', polymer)

    n = nchar(polymer) # is this string getting shorter?
    print(n)
    if(t<26){
      t = t+1
    } else {
      t = 1
      if(start>n){
        start = n
      } else{
        x = x + 1
      }
    }
  }
}
```

```
poly = c(poly, n)
}
```

In the end, removing the letter T to start made for the most compact list `LETTERS[which(poly%in%min(poly))]` and for me was of length 4098.

## — Day 6: Chronal Coordinates —

The device on your wrist beeps several times, and once again you feel like you’re falling.

“Situation critical,” the device announces. “Destination indeterminate. Chronal interference detected. Please specify new target coordinates.”

The device then produces a list of coordinates (your puzzle input). Are they places it thinks are safe or dangerous? It recommends you check manual page 729. The Elves did not give you a manual.

If they’re dangerous, maybe you can minimize the danger by finding the coordinate that gives the largest distance from the other points.

Using only the Manhattan distance, determine the area around each coordinate by counting the number of integer X,Y locations that are closest to that coordinate (and aren’t tied in distance to any other coordinate).

Your goal is to find the size of the largest area that isn’t infinite. For example, consider the following list of coordinates:

- 1, 1
- 1, 6
- 8, 3
- 3, 4
- 5, 5
- 8, 9

If we name these coordinates A through F, we can draw them on a grid, putting 0,0 at the top left:

```
.....
.A.....
.....
.....C.
...D.....
....E....
.B.....
.....
.....
.....F.
```

This view is partial - the actual grid extends infinitely in all directions. Using the Manhattan distance, each location's closest coordinate can be determined, shown here in lowercase:

```
aaaaa.cccc
aAaaa.cccc
aaaddecccc
aadddeccCc
..dDdecccc
bb.deEeccc
bBb.eeee..
bbb.eeefff
bbb.eeffff
bbb.ffffFf
```

Locations shown as . are equally far from two or more coordinates, and so they don't count as being closest to any.

In this example, the areas of coordinates A, B, C, and F are infinite - while not shown here, their areas extend forever outside the visible grid. However, the areas of coordinates D and E are finite: D is closest to 9 locations, and E is closest to 17 (both including the coordinate's location itself). Therefore, in this example, the size of the largest area is 17.

**What is the size of the largest area that isn't infinite?**

## — Day 7: The Sum of Its Parts —

You find yourself standing on a snow-covered coastline; apparently, you landed a little off course. The region is too hilly to see the North Pole from here, but you do spot some Elves that seem to be trying to unpack something that washed ashore. It's quite cold out, so you decide to risk creating a paradox by asking them for directions.

"Oh, are you the search party?" Somehow, you can understand whatever Elves from the year 1018 speak; you assume it's Ancient Nordic Elvish. Could the device on your wrist also be a translator? "Those clothes don't look very warm; take this." They hand you a heavy coat.

"We do need to find our way back to the North Pole, but we have higher priorities at the moment. You see, believe it or not, this box contains something that will solve all of Santa's transportation problems - at least, that's what it looks like from the pictures in the instructions." It doesn't seem like they can read whatever language it's in, but you can: "Sleigh kit. Some assembly required."

"'Sleigh'? What a wonderful name! You must help us assemble this 'sleigh' at once!" They start excitedly pulling more parts out of the box.

The instructions specify a series of steps and requirements about which steps must be finished before others can begin (your puzzle input). Each step is designated by a single letter. For example, suppose you have the following instructions:

Step C must be finished before step A can begin. Step C must be finished before step F can begin. Step A must be finished before step B can begin. Step A must be finished before step D can begin. Step B must be finished before step E can begin. Step D must be finished before step E can begin. Step F must be finished before step E can begin. Visually, these requirements look like this:



```

-->A--->B--
/      \      \
C        -->D----->E
\              /
----->F-----

```

Your first goal is to determine the order in which the steps should be completed. If more than one step is ready, choose the step which is first alphabetically. In this example, the steps would be completed as follows:

Only C is available, and so it is done first. Next, both A and F are available. A is first alphabetically, so it is done next. Then, even though F was available earlier, steps B and D are now also available, and B is the first alphabetically of the three. After that, only D and F are available. E is not available because only some of its prerequisites are complete. Therefore, D is completed next. F is the only choice, so it is done next. Finally, E is completed. So, in this example, the correct order is CABDFE.

**In what order should the steps in your instructions be completed?**