DATA TAMING ASSIGNMENT 1

VINEETH MARIKUNTEMATHA RAVISHARADHYA-a1896845

23/02/2024

Setup

```
#Load the required packages
library(here)
library(tidyverse)
library(tidyr)
library(stringr)
library(forcats)
library(lubridate)
library(inspectdf)
library(ggplot2)
```

Q1. Loading the data

```
# Your student number goes here
ysn = 1896845
# Calculate your student number modulo 3
filenum <- ysn %% 3
filenum

## [1] 2

filename <- pasteO("D:/r studio files/Assignment_1/archery_2.csv")
filename

## [1] "D:/r studio files/Assignment_1/archery_2.csv"

# Read in the data
archery_2 <- read_csv("archery_2.csv")
# Display the first 10 lines of the data
archery_2

## # A tibble: 200 x 9
## name experienced started session_1 session_2 session_3 session_4 session_5</pre>
```

```
##
     <chr> <chr>
                              <chr>
                                      <chr>
                                               <chr>
                                                        <chr>
## 1 Eli
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
            Nο
## 2 Aria
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
                      15-Dec~ Target h~ Target h~ Target h~ Target h~
## 3 Isabel~ Yes
## 4 Sofia
           No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 5 Macken~ Yes
                      17-Apr~ Target h~ Target h~ Target h~ Target h~ Target h~
                     14-Apr~ Target h~ Target h~ Target h~ Target h~
## 6 Audrey Yes
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 7 Abigail No
## 8 Zoe
            Yes
                      15-Nov~ Target h~ Target h~ Target h~ Target h~
## 9 Claire No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 10 Floren~ No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## # i 190 more rows
## # i 1 more variable: session_6 <chr>
print(n=10, archery_2)
## # A tibble: 200 x 9
          experienced started session_1 session_2 session_3 session_4 session_5
     <chr> <chr> <chr>
##
                             <chr>
                                      <chr>
                                               <chr>
                                                        <chr>
## 1 Eli
           No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 2 Aria
            No
                     01-Feb~ Target h~ Target h~ Target h~ Target h~
## 3 Isabel~ Yes
                     15-Dec~ Target h~ Target h~ Target h~ Target h~
                     01-Feb~ Target h~ Target h~ Target h~ Target h~
## 4 Sofia No
## 5 Macken~ Yes
                     17-Apr~ Target h~ Target h~ Target h~ Target h~
## 6 Audrey Yes
                     14-Apr~ Target h~ Target h~ Target h~ Target h~ Target h~
## 7 Abigail No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 8 Zoe
            Yes
                      15-Nov~ Target h~ Target h~ Target h~ Target h~
## 9 Claire No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 10 Floren~ No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## # i 190 more rows
## # i 1 more variable: session_6 <chr>
```

Q2. Taking a random sample of archers

```
set.seed(1896845)
sample_archery_2 <- sample_n(archery_2, 100)</pre>
sample_archery_2
## # A tibble: 100 x 9
##
     name
            experienced started session_1 session_2 session_3 session_4 session_5
##
     <chr>
                                       <chr>
                                                <chr>
                                                          <chr>
## 1 River No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 2 Benjam~ No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
                       01-Feb~ Target h~ Target h~ Target h~ Target h~ Target h~
## 3 Chloe No
## 4 Lincoln Yes
                      15-Jul~ Target h~ Target h~ Target h~ Target h~
## 5 Charlo~ Yes
                       07-Dec~ Target h~ Target h~ Target h~ Target h~
## 6 Scarle~ No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 7 Poppy No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
## 8 Gabrie~ No
                      01-Feb~ Target h~ Target h~ Target h~ Target h~
```

```
## 9 Zayn No 01-Feb~ Target h~ Target h~ Target h~ Target h~ Target h~
## 10 Logan Yes 18-Sep~ Target h~ Target h~ Target h~ Target h~ Target h~
## # i 90 more rows
## # i 1 more variable: session_6 <chr>
```

Q3. Tidying the data

Q3(a). Converting from wide form to long form

```
library(tidyr)
archery_wide_long <- gather(sample_archery_2, key = "session", value = "result", session_1:session_6)
sample archery 2
## # A tibble: 100 x 9
##
            experienced started session_1 session_2 session_3 session_4 session_5
##
     <chr>
            <chr>
                       <chr>
                              <chr>
                                       <chr>
                                                <chr>
                                                          <chr>
## 1 River No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 2 Benjam~ No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 3 Chloe
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
            No
## 4 Lincoln Yes
                       15-Jul~ Target h~ Target h~ Target h~ Target h~
## 5 Charlo~ Yes
                       07-Dec~ Target h~ Target h~ Target h~ Target h~
## 6 Scarle~ No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 7 Poppy
            No
## 8 Gabrie~ No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
## 9 Zayn
            No
                       01-Feb~ Target h~ Target h~ Target h~ Target h~
                       18-Sep~ Target h~ Target h~ Target h~ Target h~
## 10 Logan
            Yes
## # i 90 more rows
## # i 1 more variable: session_6 <chr>
```

Q3(b). Replacing result with 2 new columns

```
sample <- str_match(archery_wide_long$result,"Target hit (\\d+) times from (\\d+) shots")
archery_wide_long <- mutate(archery_wide_long,result = NULL)
archery_wide_long <-mutate(archery_wide_long,
shots = sample[,3],
hits = sample[,2]
)
archery_wide_long</pre>
```

```
## # A tibble: 600 x 6
##
               experienced started
     name
                                        session
                                                 shots hits
##
      <chr>
                <chr>
                           <chr>
                                        <chr>
                                                 <chr> <chr>
                           01-Feb-2024 session 1 37
##
  1 River
                                                       19
## 2 Benjamin No
                           01-Feb-2024 session 1 54
## 3 Chloe
               No
                           01-Feb-2024 session 1 41
                                                       24
## 4 Lincoln Yes
                           15-Jul-2023 session_1 62
                                                       53
## 5 Charlotte Yes
                           07-Dec-2022 session_1 66
                                                       59
```

```
## 6 Scarlett No
                           01-Feb-2024 session 1 55
                                                       28
                           01-Feb-2024 session_1 44
## 7 Poppy
                                                       23
               No
                           01-Feb-2024 session 1 57
## 8 Gabriella No
                                                       19
                           01-Feb-2024 session_1 38
## 9 Zayn
               No
                                                       19
## 10 Logan
               Yes
                           18-Sep-2022 session_1 61
## # i 590 more rows
archery_wide_long
```

```
## # A tibble: 600 x 6
                                               shots hits
     name
              experienced started
                                      session
##
                                               <chr> <chr>
     <chr>
               <chr>
                       <chr>
                                      <chr>
## 1 River
              No
                          01-Feb-2024 session 1 37
## 2 Benjamin No
                        01-Feb-2024 session 1 54
                                                     23
## 3 Chloe
                        01-Feb-2024 session 1 41
             No
## 4 Lincoln Yes
                        15-Jul-2023 session 1 62
                                                     53
## 5 Charlotte Yes
                          07-Dec-2022 session 1 66
                                                     59
## 6 Scarlett No
                          01-Feb-2024 session_1 55
                                                     28
## 7 Poppy
                          01-Feb-2024 session 1 44
                                                     23
              No
## 8 Gabriella No
                          01-Feb-2024 session_1 57
                                                     19
## 9 Zayn
               Nο
                          01-Feb-2024 session_1 38
                                                     19
                          18-Sep-2022 session_1 61
## 10 Logan
               Yes
                                                     53
## # i 590 more rows
```

Q3(c). Replace data in session column with session number

```
## # A tibble: 600 x 6
##
     name experienced started
                                      session shots hits
               <chr>
                                         <int> <chr> <chr>
##
     <chr>
                           <chr>
                                            1 37
## 1 River
                           01-Feb-2024
                                                    19
               No
                                            1 54
## 2 Benjamin No
                          01-Feb-2024
                                                    23
## 3 Chloe
                                            1 41
                                                    24
                          01-Feb-2024
## 4 Lincoln
               Yes
                          15-Jul-2023
                                            1 62
                                                    53
## 5 Charlotte Yes
                           07-Dec-2022
                                            1 66
                                                    59
## 6 Scarlett No
                                            1 55
                          01-Feb-2024
                                                    28
## 7 Poppy
                          01-Feb-2024
                                            1 44
                                                    23
## 8 Gabriella No
                          01-Feb-2024
                                            1 57
                                                    19
## 9 Zayn
               No
                           01-Feb-2024
                                            1 38
                                                    19
## 10 Logan
               Yes
                          18-Sep-2022
                                            1 61
                                                    53
## # i 590 more rows
```

Q3(d). Replace "Yes/No" with "TRUE/FALSE"

```
archery_wide_long$experienced[archery_wide_long$experienced == "Yes"] <- TRUE
archery_wide_long$experienced[archery_wide_long$experienced == "No"] <- FALSE
archery_wide_long</pre>
```

```
## # A tibble: 600 x 6
##
      name
                 experienced started
                                           session shots hits
##
      <chr>
                 <chr>>
                              <chr>
                                             <int> <chr> <chr>
    1 River
                 FALSE
                              01-Feb-2024
                                                 1 37
##
                                                          19
##
    2 Benjamin
                FALSE
                              01-Feb-2024
                                                 1 54
                                                          23
                                                 1 41
##
    3 Chloe
                 FALSE
                              01-Feb-2024
                                                          24
##
    4 Lincoln
                 TRUE
                              15-Jul-2023
                                                 1 62
                                                         53
   5 Charlotte TRUE
                                                 1 66
##
                              07-Dec-2022
                                                         59
##
    6 Scarlett FALSE
                              01-Feb-2024
                                                 1 55
                                                          28
                                                          23
##
   7 Poppy
                 FALSE
                              01-Feb-2024
                                                 1 44
    8 Gabriella FALSE
                              01-Feb-2024
                                                 1 57
                                                          19
    9 Zayn
                 FALSE
                              01-Feb-2024
                                                 1 38
                                                          19
##
                              18-Sep-2022
## 10 Logan
                 TRUE
                                                 1 61
                                                          53
## # i 590 more rows
```

```
#print first 10 outputs
print(n=10,archery_wide_long)
```

```
## # A tibble: 600 x 6
##
      name
                 experienced started
                                          session shots hits
##
      <chr>
                 <chr>
                              <chr>
                                             <int> <chr> <chr>
##
   1 River
                 FALSE
                             01-Feb-2024
                                                 1 37
                                                         19
##
    2 Benjamin FALSE
                             01-Feb-2024
                                                 1 54
                                                         23
                                                 1 41
##
    3 Chloe
                 FALSE
                             01-Feb-2024
                                                         24
##
    4 Lincoln
                 TRUE
                             15-Jul-2023
                                                 1 62
                                                         53
##
   5 Charlotte TRUE
                             07-Dec-2022
                                                 1 66
                                                         59
##
   6 Scarlett FALSE
                             01-Feb-2024
                                                 1 55
                                                         28
##
    7 Poppy
                 FALSE
                              01-Feb-2024
                                                 1 44
                                                         23
##
    8 Gabriella FALSE
                              01-Feb-2024
                                                 1 57
                                                         19
    9 Zayn
                 FALSE
                              01-Feb-2024
                                                 1 38
                                                         19
## 10 Logan
                              18-Sep-2022
                                                 1 61
                 TRUE
                                                         53
## # i 590 more rows
```

Q4. Identifying data types

name = It is a "Categorical Nominal", as they are not categorized on the basis of hierarchy , making it as a nominal variable.

"experienced" = "Categorical Nominal", as it appears to be the status of the archers ,with true and false and also its a nominal variable as it has no inherent order from all the categories.

"started" = it is a "Quantitative Continuous" as it represents continuous range of values and also the this variable holds the date of each athlete or archer started

"session" = it is a "Quantitative Discrete" as this variable is countable and distinct value as it doesnt have fractions or decimals

"shots" = here this variable has the number of shots taken by every archer and which tends to have countable and distinct values, so it is "Quantitative Discrete" .

"hits" = it is similar to shots as it is countable and has distinct value so it is "Quantitative Discrete"

"days_experience" = it is a "Quantitative Continuous" as this variable can take any non variable non negative real value within a continuous range.

"accuracy" = it is a "Quantitative Continuous, as it represents the ratio of hits to shots along with the accuracy of every archer.

Q5(a). Taming the data

```
archery_wide_long$hits <- as.integer(archery_wide_long$hits)
archery_wide_long$shots <- as.integer(archery_wide_long$shots)
archery_wide_long</pre>
```

```
## # A tibble: 600 x 6
##
            experienced started
                                  session shots hits
    name
##
     <chr>
             <chr>
                       <chr>
                                    <int> <int> <int>
                        01-Feb-2024
                                           37
##
  1 River
             FALSE
                                                 19
                                      1
  2 Benjamin FALSE
                       01-Feb-2024
                                       1
                                           54
                                                 23
  3 Chloe
                      01-Feb-2024
                                            41
##
             FALSE
                                       1
                                                 24
##
   4 Lincoln
            TRUE
                       15-Jul-2023
                                       1
                                            62
                                                 53
                                       1 66
## 5 Charlotte TRUE
                       07-Dec-2022
                                                 59
## 6 Scarlett FALSE
                       01-Feb-2024
                                                 28
                                       1 44
                                                 23
## 7 Poppy
             FALSE
                       01-Feb-2024
## 8 Gabriella FALSE
                       01-Feb-2024
                                       1
                                           57
                                                 19
                                       1
## 9 Zayn
             FALSE
                       01-Feb-2024
                                           38
                                                 19
                       18-Sep-2022
## 10 Logan
             TRUE
                                       1
                                           61
                                                 53
## # i 590 more rows
```

Q5(b). Taming the data

•

```
archery_wide_long$started = dmy(archery_wide_long$started)
archery_wide_long$experienced = as.logical(archery_wide_long$experienced)
archery_wide_long
```

```
## # A tibble: 600 x 6
             experienced started
                                  session shots hits
##
     name
##
     <chr>
             <lgl> <date> <int> <int> <int><</pre>
## 1 River FALSE
                       2024-02-01
                                     1
                                           37
                                                 19
## 2 Benjamin FALSE
                       2024-02-01
                                       1
                                           54
                                                 23
## 3 Chloe FALSE
                                           41
                        2024-02-01
                                       1
                                                 24
```

```
## 4 Lincoln TRUE
                          2023-07-15
                                                    53
## 5 Charlotte TRUE
                          2022-12-07
                                               66
                                                    59
## 6 Scarlett FALSE
                                               55
                          2024-02-01
                                                    28
                                               44
                                                    23
## 7 Poppy
              FALSE
                          2024-02-01
## 8 Gabriella FALSE
                          2024-02-01
                                                    19
## 9 Zayn
              FALSE
                          2024-02-01
                                          1
                                               38
                                                    19
## 10 Logan
               TRUE
                          2022-09-18
## # i 590 more rows
#print first 10 outputs
print(n=10, archery_wide_long)
## # A tibble: 600 x 6
##
              experienced started
                                    session shots hits
     name
     <chr>
              <lgl>
                                      <int> <int> <int>
                          <date>
## 1 River
              FALSE
                          2024-02-01
                                          1
                                               37
## 2 Benjamin FALSE
                                          1
                                               54
                        2024-02-01
## 3 Chloe
            FALSE
                        2024-02-01
                                          1
                                               41
                                                    24
## 4 Lincoln TRUE
                         2023-07-15
                                                    53
## 5 Charlotte TRUE
                         2022-12-07
                                               66
                                                    59
                                          1
## 6 Scarlett FALSE
                          2024-02-01
                                          1
                                                    28
## 7 Poppy
             FALSE
                                          1 44 23
                         2024-02-01
## 8 Gabriella FALSE
                          2024-02-01
                                          1 57 19
                                            38
## 9 Zayn
              FALSE
                          2024-02-01
                                          1
                                                    19
               TRUE
                          2022-09-18
                                               61
                                                    53
## 10 Logan
```

Q6. Adding two new colums to dataset

adding days_experience column

i 590 more rows

```
## # A tibble: 600 x 7
##
     name
            experienced started
                                    session shots hits days_experience
##
     <chr>
              <lgl>
                         <date>
                                      <int> <int> <int>
## 1 River
              FALSE
                          2024-02-01
                                          1
                                                    19
                                                                    0
   2 Benjamin FALSE
##
                          2024-02-01
                                          1
                                              54
                                                    23
                                                                    0
## 3 Chloe
              FALSE
                         2024-02-01
                                         1
                                             41
                                                    24
                                                                    0
## 4 Lincoln TRUE
                         2023-07-15
                                                    53
                                                                  201
## 5 Charlotte TRUE
                         2022-12-07
                                         1 66 59
                                                                  421
## 6 Scarlett FALSE
                         2024-02-01
```

```
## 7 Poppy
                            2024-02-01
                                                         23
                FALSE
                                                   44
                                                                          0
## 8 Gabriella FALSE
                            2024-02-01
                                              1
                                                   57
                                                         19
                                                                          0
## 9 Zayn
                                              1
                                                   38
                FALSE
                            2024-02-01
                                                         19
                                                                          0
                            2022-09-18
                                                   61
                                                         53
                                                                        501
## 10 Logan
                TRUE
## # i 590 more rows
```

adding accuracy column

```
## # A tibble: 600 x 8
##
      name
                experienced started
                                       session shots hits days_experience accuracy
##
                                         <int> <int> <int>
                                                                      <int>
      <chr>
                <1g1>
                            <date>
                                                                               <dbl>
##
   1 River
                FALSE
                            2024-02-01
                                             1
                                                  37
                                                                          0
                                                                               0.514
                                                                               0.426
##
   2 Benjamin FALSE
                            2024-02-01
                                             1
                                                   54
                                                         23
                                                                          0
## 3 Chloe
                FALSE
                            2024-02-01
                                                         24
                                                                               0.585
                                             1
## 4 Lincoln
                TRUE
                            2023-07-15
                                                  62
                                                                        201
                                                                               0.855
                                             1
                                                         53
## 5 Charlotte TRUE
                            2022-12-07
                                                  66
                                                         59
                                                                        421
                                                                               0.894
                                             1
## 6 Scarlett FALSE
                            2024-02-01
                                             1
                                                  55
                                                         28
                                                                          0
                                                                               0.509
## 7 Poppy
                            2024-02-01
                                             1
                                                         23
                                                                          0
                                                                               0.523
                FALSE
## 8 Gabriella FALSE
                            2024-02-01
                                             1
                                                  57
                                                         19
                                                                          0
                                                                               0.333
                            2024-02-01
                                             1
                                                  38
                                                                               0.5
## 9 Zayn
                FALSE
                                                         19
                                                                          0
## 10 Logan
                TRUE
                            2022-09-18
                                             1
                                                  61
                                                         53
                                                                        501
                                                                               0.869
## # i 590 more rows
```

```
#print first 10 outputs
print(n=10,archery_wide_long)
```

```
## # A tibble: 600 x 8
##
     name
                experienced started
                                       session shots hits days_experience accuracy
##
      <chr>
                <lgl>
                            <date>
                                         <int> <int> <int>
                                                                     <int>
                                                                              <dbl>
## 1 River
               FALSE
                            2024-02-01
                                             1
                                                  37
                                                                         0
                                                                              0.514
                                                        19
## 2 Benjamin FALSE
                            2024-02-01
                                             1
                                                  54
                                                        23
                                                                         0
                                                                              0.426
## 3 Chloe
               FALSE
                            2024-02-01
                                             1
                                                  41
                                                                         0
                                                                              0.585
                                                        24
## 4 Lincoln
               TRUE
                            2023-07-15
                                                                       201
                                                                              0.855
## 5 Charlotte TRUE
                                                                       421
                                                                              0.894
                           2022-12-07
                                             1
                                                  66
                                                        59
## 6 Scarlett FALSE
                            2024-02-01
                                             1
                                                  55
                                                        28
                                                                         0
                                                                              0.509
                                                  44
                                                        23
                                                                              0.523
## 7 Poppy
                            2024-02-01
                                             1
                                                                         0
                FALSE
## 8 Gabriella FALSE
                            2024-02-01
                                                  57
                                                        19
                                                                         0
                                                                              0.333
## 9 Zayn
                FALSE
                            2024-02-01
                                             1
                                                  38
                                                        19
                                                                         0
                                                                              0.5
                TRUE
                            2022-09-18
                                                        53
                                                                       501
                                                                              0.869
## 10 Logan
## # i 590 more rows
```

Q7. Display the sample statistics for the numerical values in your dataset

```
inspect_num(archery_wide_long)
## # A tibble: 5 x 10
##
     col name
                           q1 median
                                                     q3
                  min
                                                                   sd pcnt_na hist
                                          mean
                                                          max
##
     <chr>>
                 <dbl> <dbl>
                                <dbl>
                                         <dbl>
                                                 <dbl> <dbl>
                                                                <dbl>
                                                                         <dbl> <named >
## 1 session
                1
                        2
                                3.5
                                         3.5
                                                 5
                                                            6
                                                                1.71
                                                                             0 <tibble>
## 2 shots
               22
                       43
                               51
                                        52.6
                                                62
                                                           96
                                                               12.2
                                                                             0 <tibble>
                                                           83 16.6
## 3 hits
               10
                       24
                               39.5
                                        39.6
                                                                             0 <tibble>
                                                54
## 4 days exp~
                0
                        0
                              206.
                                       269.
                                               497.
                                                          910 297.
                                                                             0 <tibble>
                                                                             0 <tibble>
## 5 accuracy
                0.333 0.560
                                0.762
                                         0.726
                                                 0.898
                                                            1
                                                                0.181
#q8. Subset of dataset to contain experienced ones
archery_wide_long <- filter(archery_wide_long, experienced == TRUE)
archery_wide_long
## # A tibble: 318 x 8
##
                                         session shots hits days_experience accuracy
      name
                 experienced started
##
      <chr>
                 <lgl>
                             <date>
                                           <int> <int> <int>
                                                                         <int>
                                                                                   <dbl>
##
   1 Lincoln
                TRUE
                             2023-07-15
                                               1
                                                     62
                                                           53
                                                                           201
                                                                                  0.855
   2 Charlotte TRUE
                             2022-12-07
                                                           59
                                                                           421
                                                                                  0.894
##
   3 Logan
                 TRUE
                             2022-09-18
                                               1
                                                     61
                                                           53
                                                                           501
                                                                                  0.869
##
   4 Sophia
                 TRUE
                             2022-04-20
                                               1
                                                           43
                                                                           652
                                                                                  0.843
##
  5 Bella
                TRUE
                             2021-10-09
                                               1
                                                     64
                                                           62
                                                                           845
                                                                                  0.969
   6 Lavla
                TRUE
                             2022-09-09
                                               1
                                                     66
                                                           58
                                                                           510
                                                                                  0.879
    7 Charlie
                             2023-04-19
                                                     67
                                                           48
                                                                           288
                                                                                  0.716
##
                TRUE
                                               1
    8 Millie
                                                                           349
##
                 TRUE
                             2023-02-17
                                               1
                                                           69
                                                                                  0.821
##
  9 Zara
                 TRUE
                             2022-05-26
                                               1
                                                     43
                                                           40
                                                                           616
                                                                                  0.930
## 10 Emily
                 TRUE
                             2022-09-02
                                                     54
                                                           49
                                                                           517
                                                                                  0.907
## # i 308 more rows
#print first 10 outputs
print(n=10, archery_wide_long)
## # A tibble: 318 x 8
##
                experienced started
                                         session shots hits days_experience accuracy
      name
##
      <chr>
                 <lgl>
                             <date>
                                           <int> <int> <int>
                                                                         <int>
                                                                                  <dbl>
##
    1 Lincoln
                TRUE
                             2023-07-15
                                               1
                                                     62
                                                           53
                                                                           201
                                                                                  0.855
##
   2 Charlotte TRUE
                             2022-12-07
                                               1
                                                     66
                                                           59
                                                                           421
                                                                                  0.894
                TRUE
                             2022-09-18
                                                     61
                                                           53
                                                                           501
                                                                                  0.869
##
   3 Logan
                                               1
##
    4 Sophia
                 TRUE
                             2022-04-20
                                               1
                                                     51
                                                           43
                                                                           652
                                                                                  0.843
##
   5 Bella
                TRUE
                             2021-10-09
                                               1
                                                     64
                                                           62
                                                                           845
                                                                                  0.969
    6 Lavla
                 TRUE
                             2022-09-09
                                                           58
                                                                           510
                                                                                  0.879
   7 Charlie
                                                     67
                                                                           288
                                                                                  0.716
##
                TRUE
                             2023-04-19
                                               1
                                                           48
##
    8 Millie
                 TRUE
                             2023-02-17
                                               1
                                                     84
                                                           69
                                                                           349
                                                                                  0.821
```

1

43

40

49

616

517

0.930

0.907

2022-05-26

2022-09-02

##

9 Zara

i 308 more rows

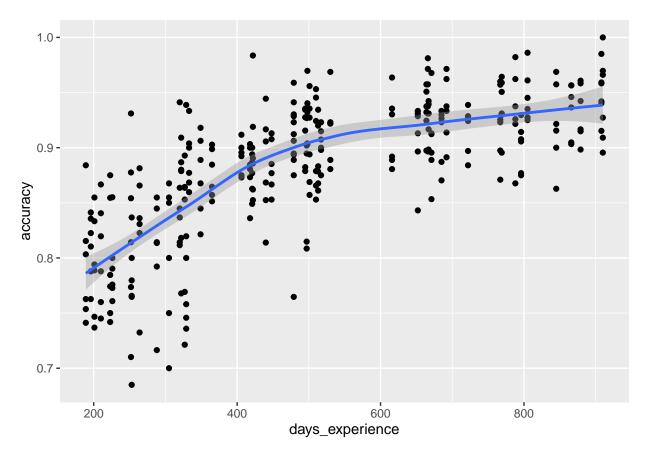
10 Emily

TRUE

TRUE

#Q9. Scatter plot of accuracy and days_experience

```
ggplot(archery_wide_long, aes(x = days_experience, y = accuracy)) +
  geom_point() +
  geom_smooth()
```



q10. Since there is a rise in the days_experience , its a sign of positive trend, where it implies that the archers are getting more experience and their accuracy level tends to improve. However at the curve over 750 days becaomes more stable, so to be the better archer it takes anywhere around 750 to 800 days by taking plot as the reference.