# Identify long pass and 3pt shooting

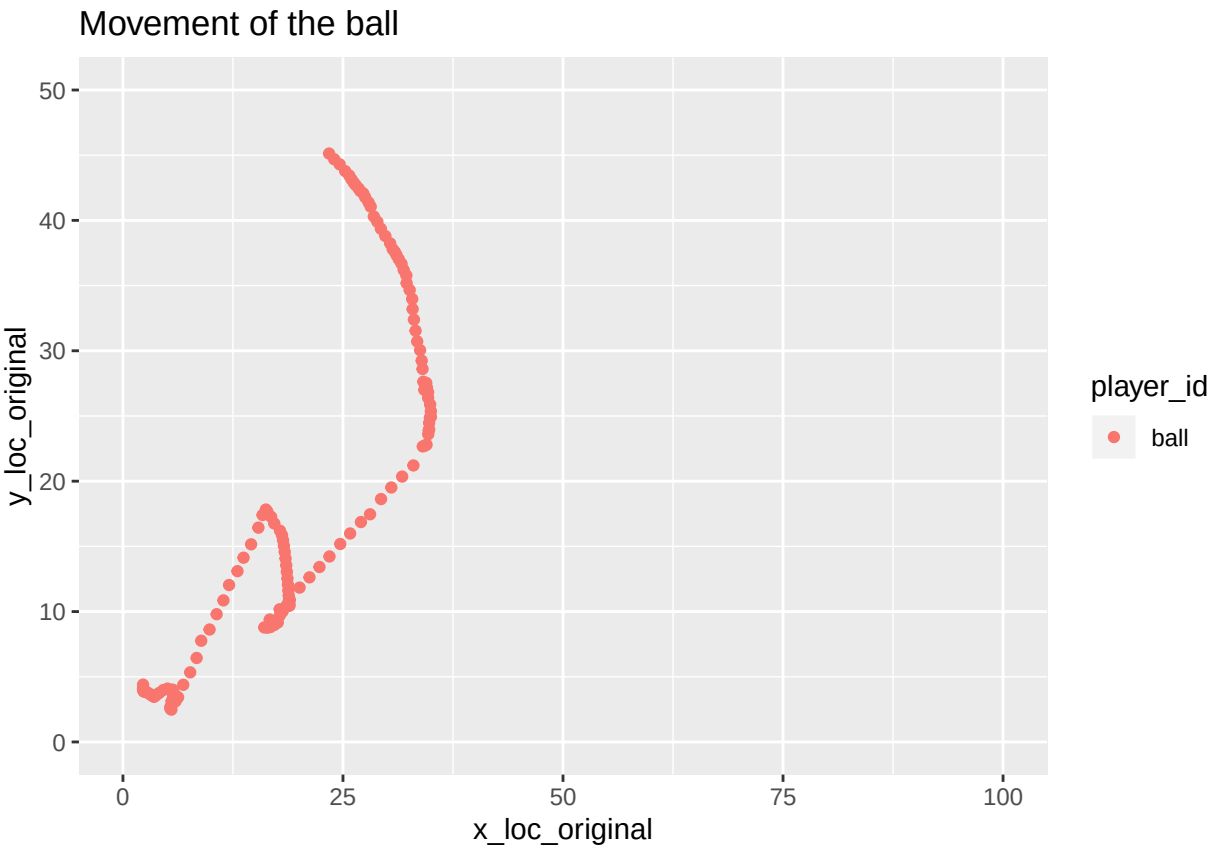
## Simply check the movement and event data

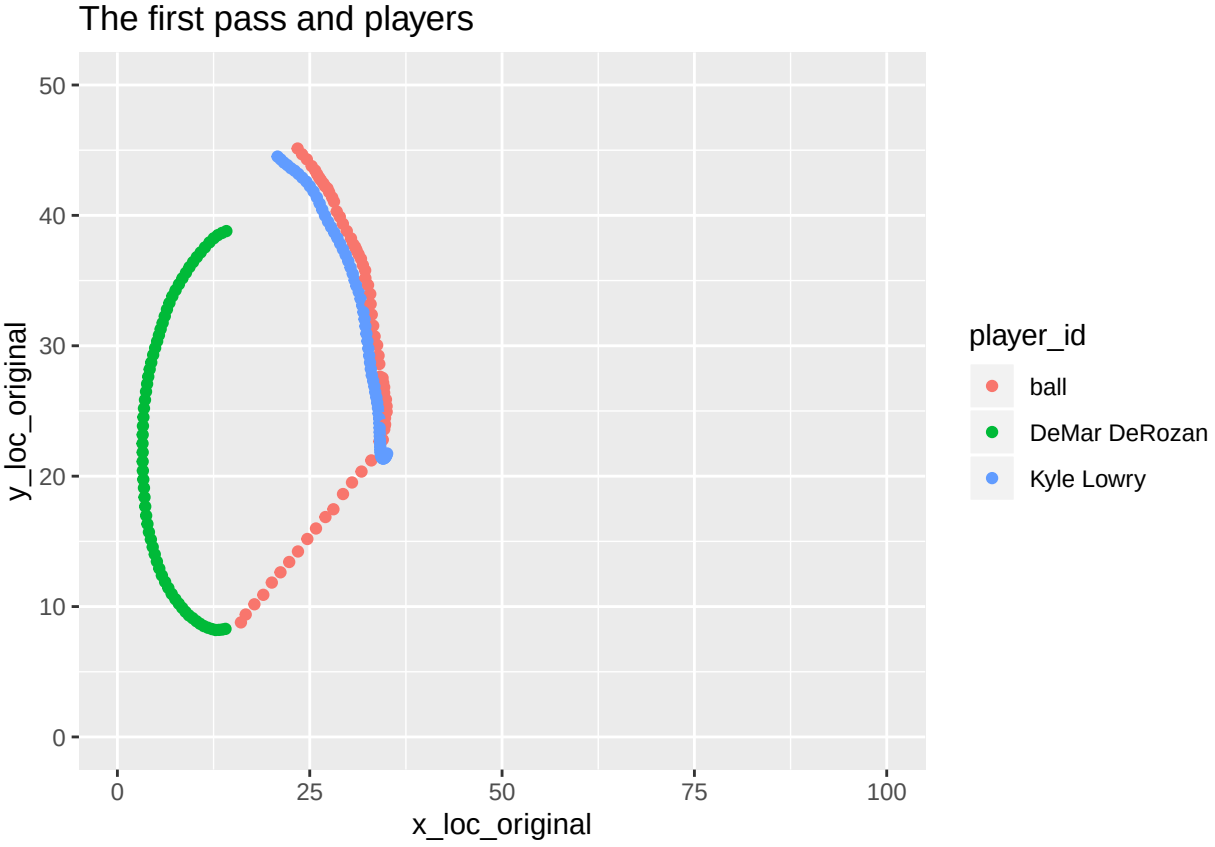
```
## # A tibble: 6 x 12
##   team_id player_id  x_loc y_loc radius game_clock shot_clock quarter game_id
##     <dbl>     <dbl>  <dbl> <dbl>  <dbl>      <dbl>      <dbl>   <dbl>   <dbl>
## 1 -1.00e0        -1 -201.  187.    3.64       711.       12.0       1  2.15e7
## 2  1.61e9      2449  111.  143.    0          711.       12.0       1  2.15e7
## 3  1.61e9    201960  115.   53.7   0          711.       12.0       1  2.15e7
## 4  1.61e9    200768 -195.  161.    0          711.       12.0       1  2.15e7
## 5  1.61e9    201942 -138.   94.2   0          711.       12.0       1  2.15e7
## 6  1.61e9    202685  -92.1 161.    0          711.       12.0       1  2.15e7
## # ... with 3 more variables: event_id <dbl>, x_loc_original <dbl>,
## #   y_loc_original <dbl>
```

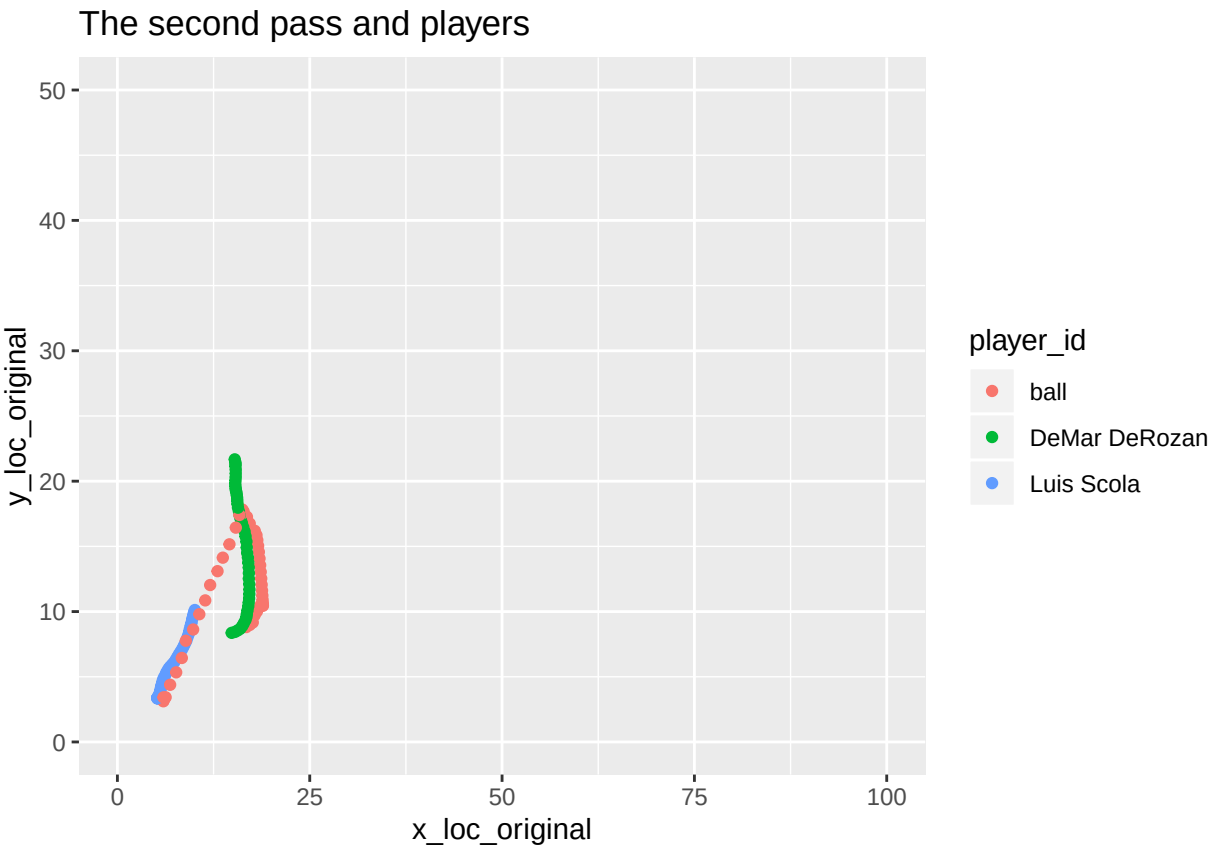
```
## # A tibble: 6 x 33
##   GAME_ID EVENTNUM EVENTMSGTYPE EVENTMSGACTIONT~ PERIOD WCTIMESTRING
##   <chr>      <dbl>        <dbl>            <dbl>  <dbl> <time>
## 1 002150~        0           12                0      1 19:41
## 2 002150~        1           10                0      1 19:41
## 3 002150~        2            2              101      1 19:42
## 4 002150~        3            4                0      1 19:42
## 5 002150~        4            1               87      1 19:42
## 6 002150~        5            1               48      1 19:42
## # ... with 27 more variables: PCTIMESTRING <time>, HOMEDESCRIPTION <chr>,
## #   NEUTRALDESCRIPTION <lgl>, VISITORDESCRIPTION <chr>, SCORE <chr>,
## #   SCOREMARGIN <chr>, PERSON1TYPE <dbl>, PLAYER1_ID <dbl>, PLAYER1_NAME <chr>,
## #   PLAYER1_TEAM_ID <dbl>, PLAYER1_TEAM_CITY <chr>,
## #   PLAYER1_TEAM_NICKNAME <chr>, PLAYER1_TEAM_ABBREVIATION <chr>,
## #   PERSON2TYPE <dbl>, PLAYER2_ID <dbl>, PLAYER2_NAME <chr>,
## #   PLAYER2_TEAM_ID <dbl>, PLAYER2_TEAM_CITY <chr>,
## #   PLAYER2_TEAM_NICKNAME <chr>, PLAYER2_TEAM_ABBREVIATION <chr>,
## #   PERSON3TYPE <dbl>, PLAYER3_ID <dbl>, PLAYER3_NAME <chr>,
## #   PLAYER3_TEAM_ID <dbl>, PLAYER3_TEAM_CITY <chr>,
## #   PLAYER3_TEAM_NICKNAME <chr>, PLAYER3_TEAM_ABBREVIATION <chr>
```

## Look at Event 1 and the movement of ball

There seems to exist two pass in the ball movement. Kyle Lowry initially dribbles the ball and pass it to others


Movement of the ball
y_loc_original
x_loc_original
0
10
20
30
40
50
0
25
50
75
100
player_id
ball

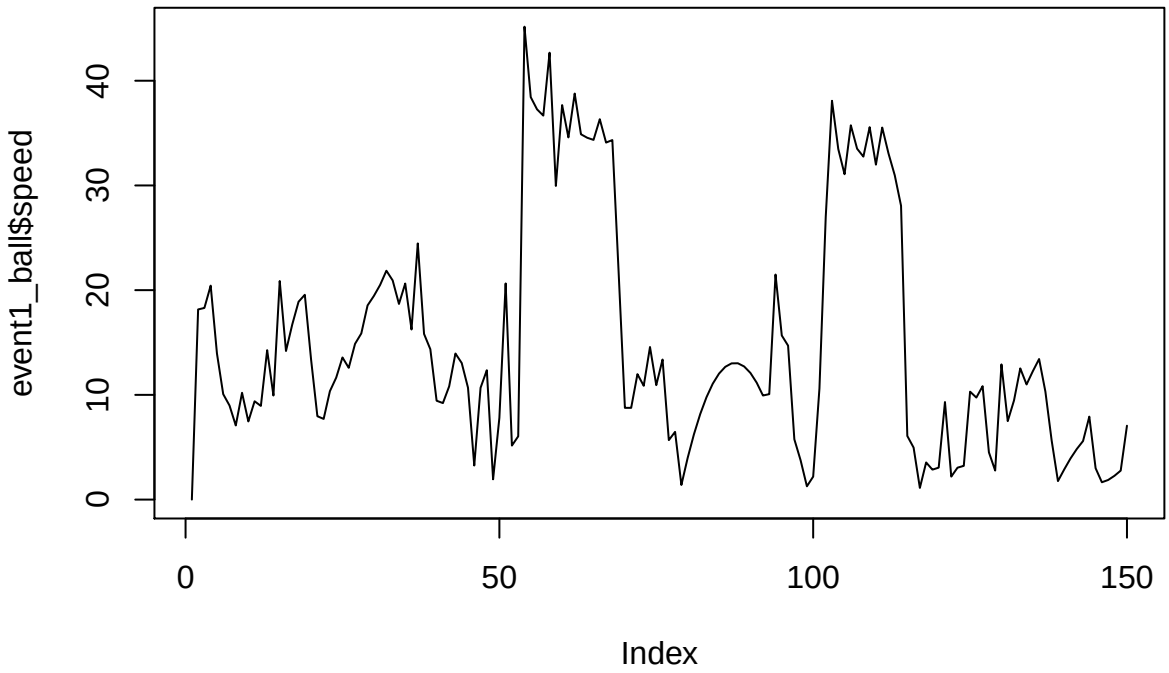

The first pass and players
y_loc_original
x_loc_original
0
10
20
30
40
50
0
25
50
75
100
player_id
ball
DeMar DeRozan
Kyle Lowry




## Fit the passing

We preprocess the data (using python) and calculate the speed for each player and the ball at each timestep.

```
## # A tibble: 6 x 14
##   index team_id player_id x_loc y_loc radius game_clock shot_clock quarter
##   <dbl> <chr>   <chr>     <dbl> <dbl>  <dbl>      <dbl>      <dbl>   <dbl>
## 1     0 ball    ball      -201.  187.  3.64        711.       12.0       1
## 2    21 ball    ball      -197.  193.  3           711.       12.0       1
## 3    31 ball    ball      -193.  199.  2.25        711.       12.0       1
## 4    43 ball    ball      -188.  205.  1.49        711.       11.9       1
## 5    51 ball    ball      -184.  210.  0.507       711.       11.9       1
## 6    65 ball    ball      -181.  212.  0.956       711.       11.9       1
## # ... with 5 more variables: game_id <dbl>, event_id <dbl>,
## #   x_loc_original <dbl>, y_loc_original <dbl>, speed <dbl>
```

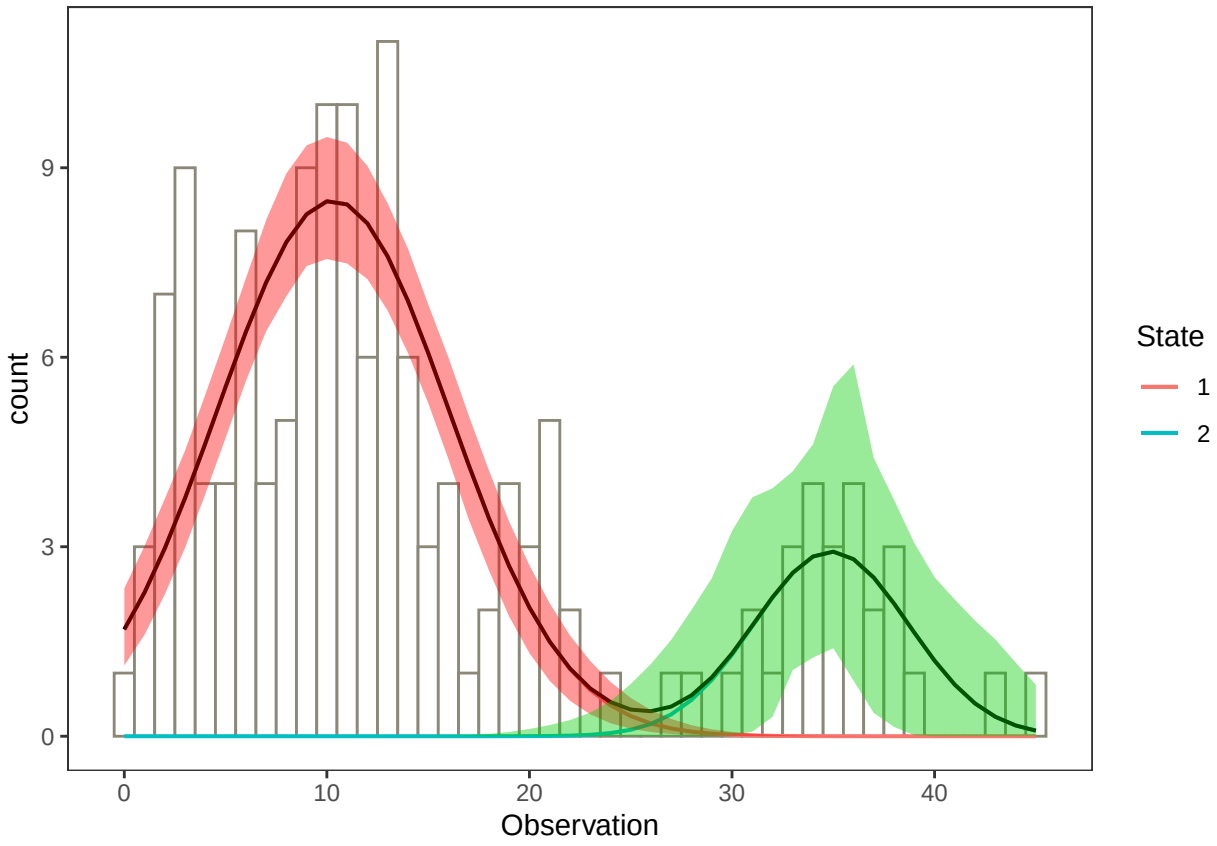


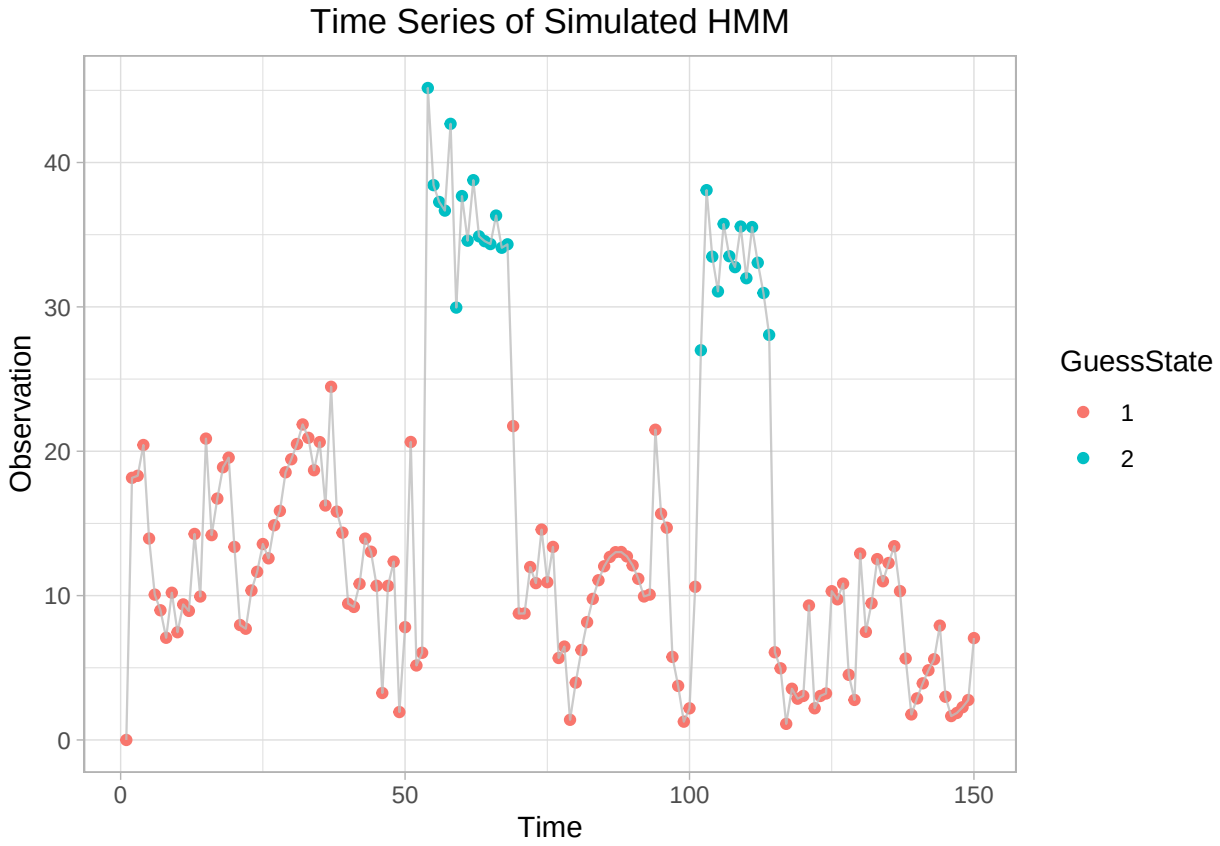

After investigating the speed data for these two potential state, we assign the initial value and starting fitting two states, a pass occurs or a pass doesn't occur.

```
# number of states, 2
m <- 2
# the mean of two state
mu <- c(9, 26)
# standard deviation of two states
sigma <- c(12, 4)
# transition probability
gamma <- matrix(c(0.99, 0.01, 0.02, 0.98), m, m, byrow = TRUE)
# Initial state probability
delta <- c(1, 0)
```

```
## $m
## [1] 2
##
## $mu
## [1] 10.30614 34.89330
##
## $sigma
## [1] 5.738578 3.821688
##
## $gamma
##            [,1]       [,2]
## [1,] 0.98481175 0.01518825
## [2,] 0.07711296 0.92288704
##
## $delta
## [1] 0.8354491 0.1645509
##
## $code
## [1] 1
```

```
##
## $mllk
## [1] 480.6185
##
## $AIC
## [1] 973.2369
##
## $BIC
## [1] 991.3007
```





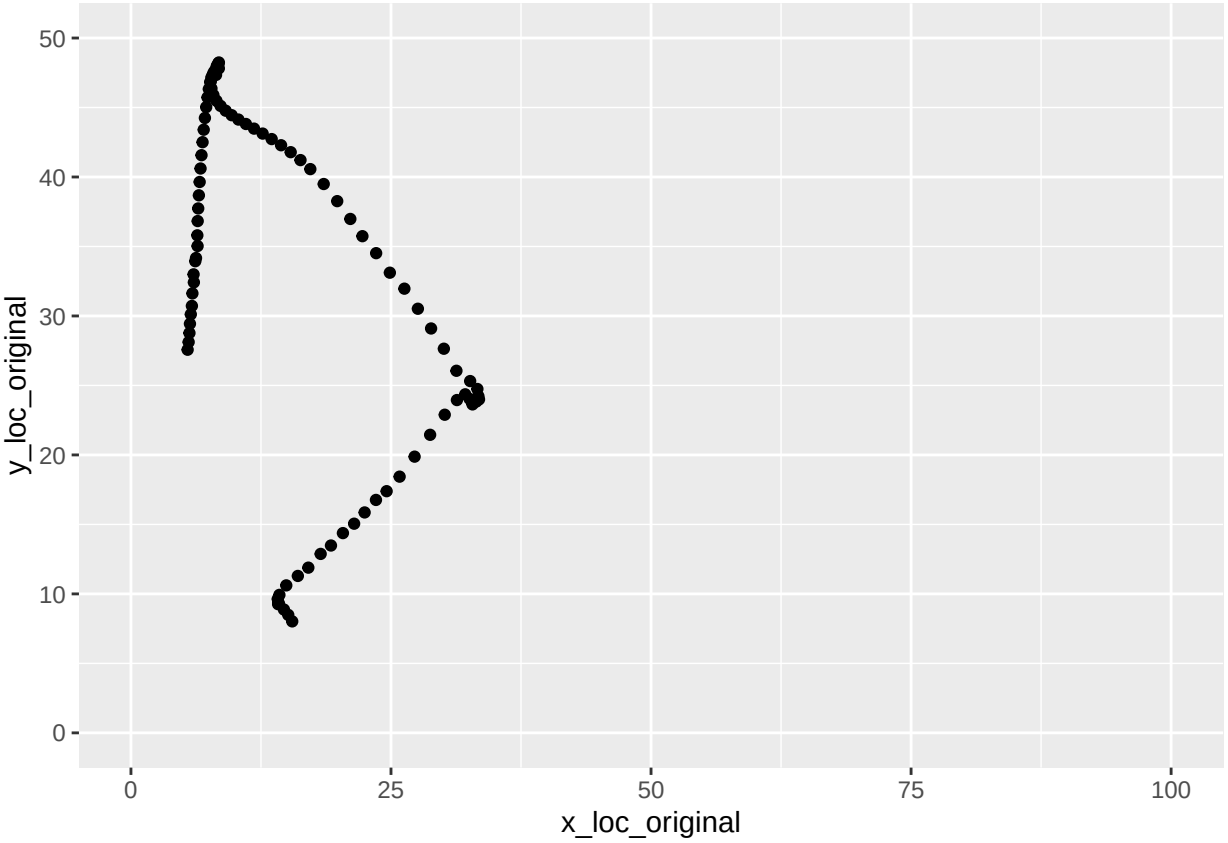

## 3-pt shooting

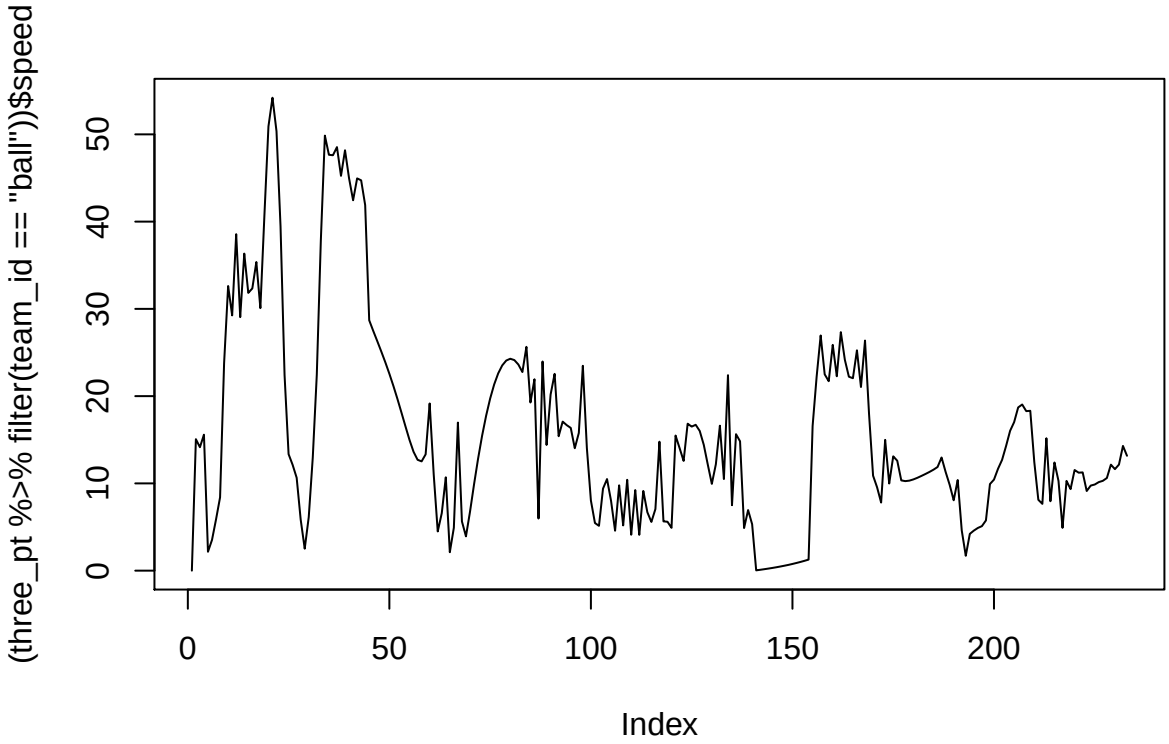
We also prepare the data for three point shooting.

```
## # A tibble: 6 x 14
##   index team_id player_id  x_loc y_loc radius game_clock shot_clock quarter
##   <dbl> <chr>   <chr>      <dbl> <dbl>  <dbl>      <dbl>      <dbl>   <dbl>
## 1     0 ball    ball       170.  108.    3.15       558.       15.0       1
## 2     1 TOR     DeMarre ~  164.  115.    0          558.       15.0       1
## 3     2 TOR     Kyle Low~  -69.0 280.    0          558.       15.0       1
## 4     3 TOR     DeMar De~ -237.   49.2   0          558.       15.0       1
## 5     4 TOR     Jonas Va~   28.5 104.    0          558.       15.0       1
## 6     5 CHA     Marvin W~  159.   43.3   0          558.       15.0       1
## # ... with 5 more variables: game_id <dbl>, event_id <dbl>,
## #   x_loc_original <dbl>, y_loc_original <dbl>, speed <dbl>
```

I want to fit it together with passing/dribbling, but right now I think the speed enough for long passing, but it is not enough to identify the dribble by player and shoot by player (since they are close).

I believe distance to the hoop could be another criteria for us to identify a shooting.




Lastly, I create two gif plots.