# WBB PBP Analysis

December 21, 2022

```
#points scored on a possession minus the average points per possession for each game (summed per player
#players with a positive number lead to high value possessions when they shoot
#players with a negative number lead to low value possessions when they shoot
pbp %>%
 group_by(date) %>%
 filter(action == "GOOD" | action == "MISS") %>%
  mutate(points_scored = ifelse(action == "GOOD" & type == "LAYUP", 2, 0),
         points_scored = ifelse(action == "GOOD" & type == "JUMPER", 2, points_scored),
         points_scored = ifelse(action == "GOOD" & type == "3PTR", 3, points_scored),
         points_scored = ifelse(action == "GOOD" & type == "LAYUP", 2, points_scored),
         points_scored = ifelse(action == "GOOD" & type == "FT", 1, points_scored),
         possession = 1,
         total_game_points = sum(points_scored),
         total_game_possessions = sum(possession),
         exp_pos_pts = total_game_points/total_game_possessions,
         pts_over_exp = points_scored - exp_pos_pts) %>%
  ungroup() %>%
  group_by(checkname52) %>%
  summarise(total_points_added = sum(pts_over_exp),
            points added per = total points added/n()) %>%
  arrange(-points_added_per)
## # A tibble: 10 x 3
##
      checkname52
                       total_points_added points_added_per
##
      <chr>
                                    <dbl>
                                                     <dbl>
## 1 HEGLAND, GRACE
                                                    0.205
                                    14.1
## 2 BOTTEN, CAYDEN
                                     3.36
                                                    0.135
## 3 KELLY, RACHEL
                                    11.5
                                                    0.124
## 4 OLMEN, SOPHIE
                                    4.75
                                                    0.0779
## 5 YAGODINSKI, PAIGE
                                                    0.0511
                                    4.44
## 6 TANGEN, SARAH
                                    -4.52
                                                   -0.146
## 7 FIX, CASSIE
                                                   -0.259
                                   -17.8
## 8 HOWDESHELL, SARAH
                                    -3.33
                                                   -0.416
## 9 ZHEN, ZILING
                                    -5.99
                                                   -0.544
## 10 KORMANN, SOPHIA
                                    -6.54
                                                   -0.654
#same table as above, but by quarter
quarter_max <- pbp %>%
 group_by(date) %>%
  filter(action == "GOOD" | action == "MISS") %>%
  mutate(points_scored = ifelse(action == "GOOD" & type == "LAYUP", 2, 0),
         points_scored = ifelse(action == "GOOD" & type == "JUMPER", 2, points_scored),
         points scored = ifelse(action == "GOOD" & type == "3PTR", 3, points scored),
         points_scored = ifelse(action == "GOOD" & type == "LAYUP", 2, points_scored),
```

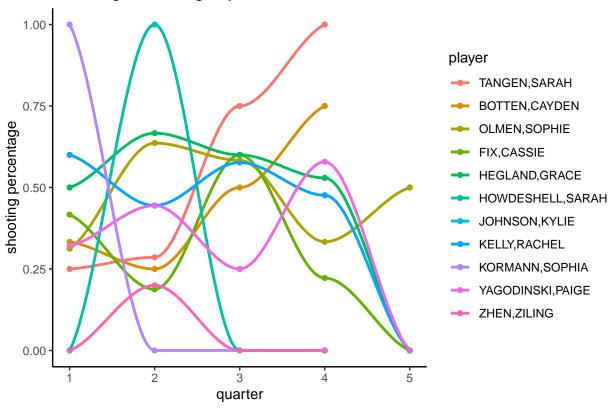
```
points_scored = ifelse(action == "GOOD" & type == "FT", 1, points_scored),
        possession = 1,
        total_game_points = sum(points_scored),
        total_game_possessions = sum(possession),
        exp_pos_pts = total_game_points/total_game_possessions,
        pts_over_exp = points_scored - exp_pos_pts) %>%
  ungroup() %>%
  group by(checkname52, quarter) %>%
  summarise(points_added = sum(pts_over_exp)) %>%
  arrange(quarter, -points_added)
## 'summarise()' has grouped output by 'checkname52'. You can override using the
## '.groups' argument.
pbp %>%
 filter(checkname52 != "TEAM") %>%
 filter(action == "ASSIST" | action == "TURNOVER") %>%
  group_by(checkname52) %>%
  summarise(assists = sum(action == "ASSIST"),
           turnovers = sum(action == "TURNOVER"),
           `a/t` = assists/turnovers) %>%
  arrange(-\a/t\)
## # A tibble: 10 x 4
                                          'a/t'
##
     checkname52 assists turnovers
##
      <chr>
                       <int> <int> <dbl>
## 1 HOWDESHELL, SARAH
                        4
                                   0 Inf
## 2 FIX, CASSIE
                         18
                                    18 1
## 3 YAGODINSKI,PAIGE 19
## 4 OLMEN,SOPHIE 11
                                    20 0.95
                                    12 0.917
                                   34 0.382
## 5 KELLY, RACHEL
                         13
## 6 HEGLAND, GRACE
                         9
                                   25 0.36
                          2
                                   8 0.25
## 7 ZHEN, ZILING
                                    8 0.125
## 8 BOTTEN, CAYDEN
                          1
## 9 TANGEN, SARAH
                          1
                                    9 0.111
## 10 KORMANN, SOPHIA
                                    2
#pbp %>%
 #filter(action == "TURNOVER" | action == "ASSIST") %>%
shooting_by_quarter <- pbp %>%
  filter(checkname52 != "TEAM") %>%
  filter(type != "FT") %>%
  group_by(checkname52, quarter) %>%
  mutate(points = ifelse(action == "GOOD" & type == "LAYUP", 2, 0),
        points = ifelse(action == "GOOD" & type == "JUMPER", 2, points),
        points = ifelse(action == "GOOD" & type == "3PTR", 3, points),
        points = ifelse(action == "GOOD" & type == "LAYUP", 2, points),
        points = ifelse(action == "GOOD" & type == "FT", 1, points)) %>%
  summarise(total_points = sum(points),
           shots = sum(action == "GOOD") + sum(action == "MISS"),
           shooting_pct = sum(action == "GOOD")/
```

```
(sum(action == "GOOD") + sum(action == "MISS")),
shooting_pct = ifelse(is.na(shooting_pct), 0, shooting_pct))
```

## 'summarise()' has grouped output by 'checkname52'. You can override using the
## '.groups' argument.

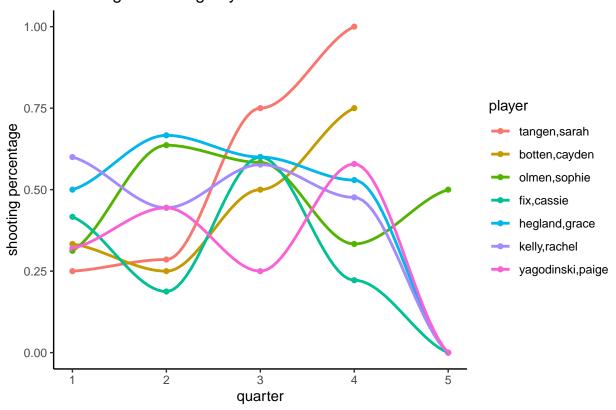
## 'geom\_smooth()' using method = 'loess' and formula 'y ~ x'

# Shooting Percentage by Quarter

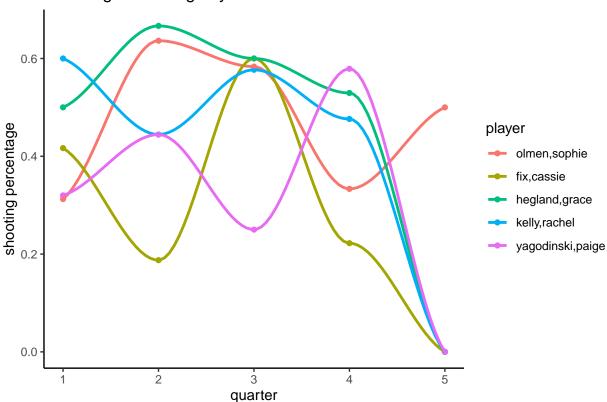


```
shooting_by_quarter <- shooting_by_quarter %>%
mutate(checkname52 = tolower(checkname52))
```

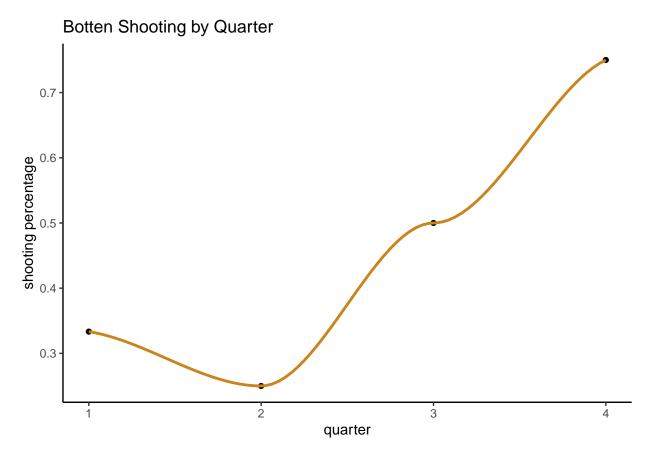
### Shooting Percentage by Quarter



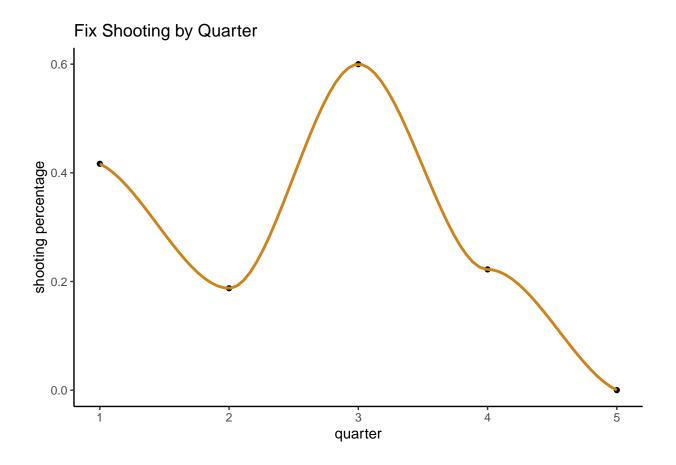
## Shooting Percentage by Quarter



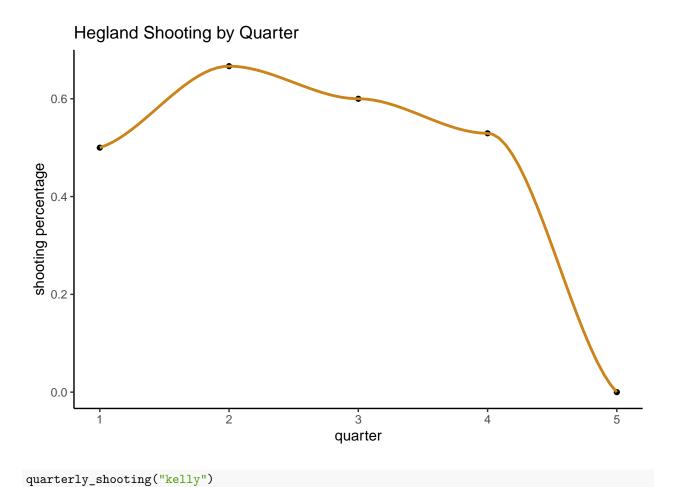
## 'geom\_smooth()' using method = 'loess' and formula 'y ~ x'

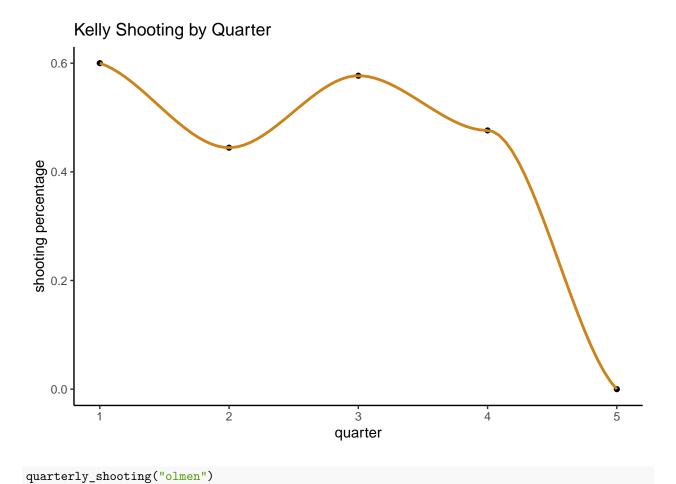


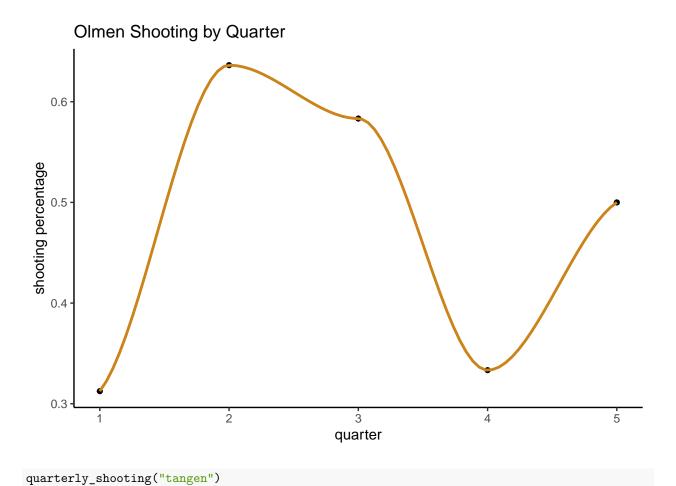
quarterly\_shooting("fix")

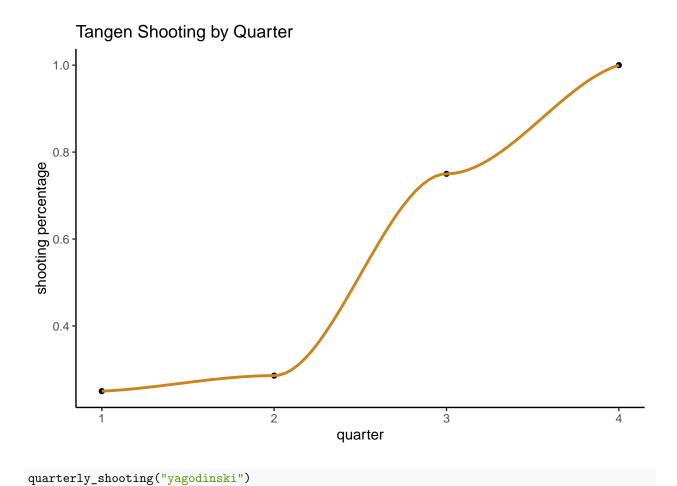


quarterly\_shooting("hegland")

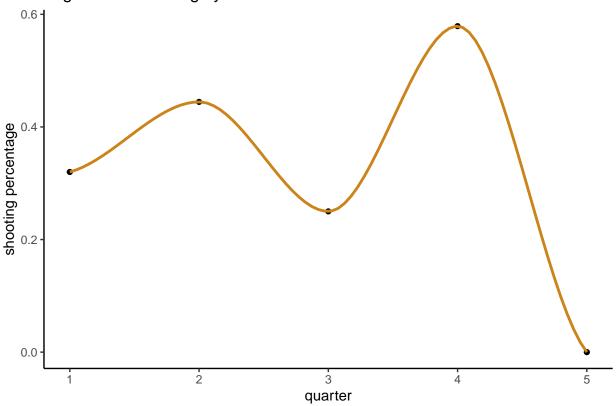








### Yagodinski Shooting by Quarter



```
#strong first quarter shooters: hegland, kelly
#strong second quarter shooters: hegland, olmen
#strong third quarter shooters: fix, hegland, kelly, olmen, tangen
#strong fourth quarter shooters: botten, hegland, tangen, yagodinski
```

```
#points per shot (excluding free throws)
pbp %>%
  filter(checkname52 != "TEAM") %>%
  filter(type != "FT") %>%
  group_by(checkname52) %>%
  mutate(points = ifelse(action == "GOOD" & type == "LAYUP", 2, 0),
        points = ifelse(action == "GOOD" & type == "JUMPER", 2, points),
        points = ifelse(action == "GOOD" & type == "3PTR", 3, points),
        points = ifelse(action == "GOOD" & type == "LAYUP", 2, points)) %>%
  summarise(total_points = sum(points),
        shots = sum(action == "GOOD") + sum(action == "MISS"),
        pts_per_shot = total_points/shots) %>%
  arrange(-pts_per_shot)
```

```
## # A tibble: 11 x 4
##
      checkname52
                        total_points shots pts_per_shot
##
      <chr>
                               <dbl> <int>
                                                    <dbl>
                                         55
                                                    1.22
##
   1 HEGLAND, GRACE
                                  67
   2 KELLY, RACHEL
                                  86
                                         77
                                                    1.12
   3 OLMEN, SOPHIE
                                  54
                                         50
                                                    1.08
```

```
## 4 BOTTEN, CAYDEN
                               26
                                     25
                                              1.04
## 5 YAGODINSKI, PAIGE
                               76 76
                                              1
                              20 23
## 6 TANGEN, SARAH
                                              0.870
## 7 FIX, CASSIE
                              33 49
                                              0.673
                               2
                                    6
## 8 HOWDESHELL, SARAH
                                              0.333
                                    9
## 9 KORMANN, SOPHIA
                              2
                                              0.222
## 10 ZHEN.ZILING
                               2
                                    9
                                              0.222
## 11 JOHNSON, KYLIE
                               0 0
                                            NaN
#shot distribution by player
pbp %>%
 filter(checkname52 != "TEAM") %>%
  filter(action == "GOOD" | action == "MISS") %>%
  group_by(checkname52) %>%
  summarise(layups = sum(type == "LAYUP"),
           mid_range = sum(type == "JUMPER"),
           threes = sum (type == "3PTR")
           )
## # A tibble: 10 x 4
     checkname52
                     layups mid_range threes
##
     <chr>>
                      <int> <int> <int>
## 1 BOTTEN, CAYDEN
                         8
                                         10
## 2 FIX, CASSIE
                         8
                                   31
                                         10
## 3 HEGLAND, GRACE
                         35
                                  8
                                         12
## 4 HOWDESHELL, SARAH
                         2
                                         0
                                   4
                                   21
                                         10
## 5 KELLY, RACHEL
                         46
## 6 KORMANN, SOPHIA
                        8
                                  1
                                         0
## 7 OLMEN, SOPHIE
                        6
                                  22
                                        22
                         6
## 8 TANGEN, SARAH
                                   10
                                         7
## 9 YAGODINSKI,PAIGE
                         16
                                   11
                                         49
## 10 ZHEN,ZILING
                                   7
pbp %>%
 filter(checkname52 != "TEAM") %>%
 filter(action == "GOOD" | action == "MISS") %>%
  group_by(checkname52) %>%
  summarise(layups = sum(type == "LAYUP"),
           layup_pct = sum(action == "GOOD" & type == "LAYUP")/
             sum(type == "LAYUP")
           ) %>%
  filter(layups > 5) %>%
  arrange(-layup_pct)
## # A tibble: 8 x 3
##
    checkname52
                    layups layup_pct
##
    <chr>
                     <int>
                               <dbl>
## 1 YAGODINSKI, PAIGE
                       16
                               0.75
## 2 OLMEN, SOPHIE
                        6
                               0.667
## 3 KELLY, RACHEL
                        46
                              0.630
## 4 FIX, CASSIE
                        8
                              0.625
## 5 HEGLAND, GRACE
                       35
                               0.6
```

8

0.5

## 6 BOTTEN, CAYDEN

```
## 7 TANGEN,SARAH 6
## 8 KORMANN,SOPHIA 8
                                 0.333
                                 0.125
pbp %>%
  filter(checkname52 != "TEAM") %>%
  filter(action == "GOOD" | action == "MISS") %>%
  group_by(checkname52) %>%
  summarise(mid_range = sum(type == "JUMPER"),
            mid_pct = sum(action == "GOOD" & type == "JUMPER")/
              sum(type == "JUMPER")
            ) %>%
  filter(mid_range > 5) %>%
  arrange(-mid_pct)
## # A tibble: 8 x 3
##
     checkname52
                      mid_range mid_pct
##
     <chr>
                         <int> <dbl>
                            8 0.625
## 1 HEGLAND, GRACE
## 2 OLMEN, SOPHIE
                             22 0.5
                            10 0.5
## 3 TANGEN, SARAH
## 4 BOTTEN, CAYDEN
                            7 0.429
## 5 KELLY, RACHEL
                           21 0.381
## 6 FIX, CASSIE
                            31 0.323
## 7 YAGODINSKI, PAIGE
                           11 0.182
## 8 ZHEN,ZILING
                            7 0.143
pbp %>%
  filter(checkname52 != "TEAM") %>%
  filter(action == "GOOD" | action == "MISS") %>%
  group_by(checkname52) %>%
  summarise(threes = sum (type == "3PTR"),
            three_pct = sum(action == "GOOD" & type == "3PTR")/
              sum(type == "3PTR")
            ) %>%
  filter(threes > 5) %>%
  arrange(-three_pct)
## # A tibble: 7 x 3
##
     checkname52
                      threes three_pct
##
     <chr>
                      <int>
                                 <dbl>
## 1 HEGLAND, GRACE
                                 0.417
                         12
## 2 BOTTEN, CAYDEN
                         10
                                 0.4
## 3 KELLY, RACHEL
                         10
                                 0.4
## 4 OLMEN, SOPHIE
                         22
                                0.364
## 5 YAGODINSKI,PAIGE
                         49
                               0.327
                          7
## 6 TANGEN, SARAH
                                0.286
## 7 FIX, CASSIE
                         10
                                0.1
pbp %>%
  group_by(checkname52) %>%
  filter(action == "GOOD" | action == "MISS") %>%
  summarise(fts = sum(type == "FT"),
```

```
ft_pct = sum(action == "GOOD" & type == "FT")/
              sum(type == "FT")
            ) %>%
  filter(fts > 5) %>%
  arrange(-ft_pct)
## # A tibble: 6 x 3
##
     checkname52
                        fts ft_pct
##
     <chr>
                      <int> <dbl>
## 1 HEGLAND, GRACE
                        14 0.714
## 2 OLMEN, SOPHIE
                         11 0.636
## 3 KELLY, RACHEL
                         16 0.625
## 4 FIX, CASSIE
                        20 0.6
## 5 YAGODINSKI,PAIGE
                       11 0.545
## 6 TANGEN, SARAH
                         8 0.5
pbp %>%
  filter(action == "GOOD" | action == "MISS") %>%
  mutate(points = ifelse(action == "GOOD" & type == "LAYUP", 2, 0),
         points = ifelse(action == "GOOD" & type == "JUMPER", 2, points),
         points = ifelse(action == "GOOD" & type == "3PTR", 3, points),
         points = ifelse(action == "GOOD" & type == "LAYUP", 2, points),
         points = ifelse(action == "GOOD" & type == "FT", 1, points))%>%
  group_by(checkname52) %>%
  summarise(`TS%` = 100*sum(points)/
              (2*(sum(type == "LAYUP" | type == "JUMPER" | type == "3PTR") +
                    0.44*sum(type == "FT")
                  )
                                   )
            ) %>%
  arrange(-`TS%`)
## # A tibble: 10 x 2
##
      checkname52
                       'TS%'
      <chr>
                       <dbl>
## 1 HEGLAND, GRACE
                        62.9
## 2 KELLY, RACHEL
                        57.1
## 3 OLMEN, SOPHIE
                        55.6
## 4 BOTTEN, CAYDEN
                        52
## 5 YAGODINSKI, PAIGE 50.7
## 6 TANGEN, SARAH
                        45.2
## 7 FIX, CASSIE
                        38.9
## 8 HOWDESHELL, SARAH 29.1
## 9 ZHEN, ZILING
                        20.2
## 10 KORMANN, SOPHIA
                        10.6
usage <- function(player){</pre>
  lineups %>%
    mutate(Lineup = tolower(Lineup)) %>%
    mutate(total_mins = sum(game_time)) %>%
    filter(str_detect(Lineup, {{player}})) %>%
    full_join(pbp) %>%
```

```
mutate(checkname52 = tolower(checkname52)) %>%
    summarise(usage_rate = sum(100 * (sum(action == "MISS" | action == "GOOD" & type != "FT" & str_dete
                                  ) * (total_mins/5) /
                sum(game_time) * (sum(action == "MISS" | action == "GOOD" & type != "FT") + 0.44 * sum(
    )
}
usage("kelly")
## Error in mutate(., Lineup = tolower(Lineup)): object 'lineups' not found
lineups %>%
  filter(str_detect(Lineup, "Botten")) %>%
    summarise(total_mins = sum(game_time))
## Error in filter(., str_detect(Lineup, "Botten")): object 'lineups' not found
lineups %>%
    summarise(total_mins = sum(game_time))
## Error in summarise(., total_mins = sum(game_time)): object 'lineups' not found
rebounding <- function(player){</pre>
  lineups %>%
    mutate(Lineup = tolower(Lineup)) %>%
    mutate(total_mins = sum(game_time)) %>%
    filter(str_detect(Lineup, {{player}})) %>%
    left_join(pbp2) %>%
    mutate(checkname52 = tolower(checkname52)) %>%
    summarise(rebound_rate = 100 * (sum(action == "REBOUND" & str_detect(checkname52, {{player}})) * (t
              /(sum(game_time)*sum(action == "REBOUND"))
              ) %>%
    head(1)
}
rebounding("botten")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
rebounding("fix")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
rebounding("hegland")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
```

```
rebounding("kelly")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
rebounding("kormann")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
rebounding("olmen")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
rebounding("yagodinski")
## Error in h(simpleError(msg, call)): error in evaluating the argument 'x' in selecting a method for f
pbp %>%
 filter(action == "STEAL" | action == "BLOCK" | action == "FOUL") %>%
 group_by(checkname52) %>%
 summarise(steals = sum(action == "STEAL"),
           blocks = sum(action == "BLOCK"),
           fouls = sum(action == "FOUL"),
           contests = n(),
           foul rate = fouls/contests) %>%
 arrange(foul_rate)
## # A tibble: 10 x 6
##
     ##
     <chr>
                      <int> <int> <int>
                                            <int>
                                                      <dbl>
## 1 FIX, CASSIE
                                                      0.469
                         14
                                 3
                                      15
                                               32
## 2 HEGLAND, GRACE
                          9
                                      16
                                               31
                                                      0.516
                                  6
## 3 KELLY, RACHEL
                           6
                                      25
                                               47
                                                      0.532
                                16
                          2
## 4 BOTTEN, CAYDEN
                                 1
                                       5
                                                8
                                                      0.625
## 5 YAGODINSKI, PAIGE
                          6
                                 2
                                      17
                                               25
                                                      0.68
## 6 OLMEN, SOPHIE
                                 2
                                               18
                                                      0.833
                           1
                                      15
## 7 TANGEN, SARAH
                           0
                                       7
                                                8
                                                      0.875
                                 1
                                                2
## 8 HOWDESHELL, SARAH
                           0
                                 0
                                       2
                                                     1
## 9 KORMANN, SOPHIA
                           0
                                       4
                                                4
                                                      1
                                 0
## 10 ZHEN, ZILING
                           0
                                 0
                                       4
                                                4
#proportions of players rebounds that are offensive rebounds
pbp %>%
 filter(action == "REBOUND") %>%
 filter(checkname52 != "TEAM") %>%
 group_by(checkname52) %>%
 summarise(offensive = sum(type == "OFF"),
           defensive = sum(type == "DEF"),
           total = n(),
           off reb rate = offensive/total) %>%
```

arrange(-off\_reb\_rate)

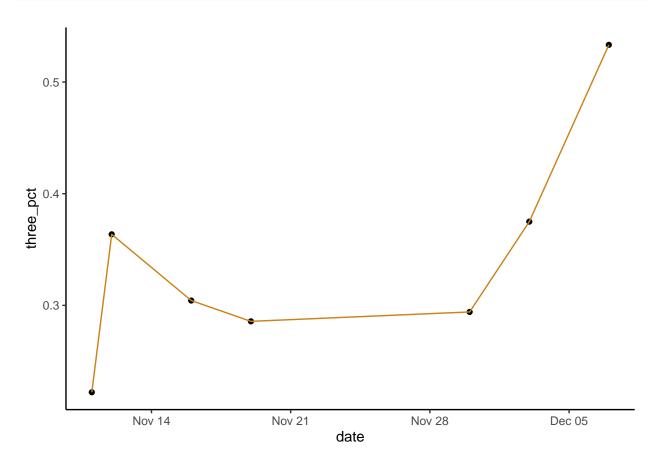
```
## # A tibble: 10 x 5
##
     <chr>
                    <int> <int> <int>
##
## 1 KORMANN, SOPHIA
                        6
                                               0.545
                                 5
                                       11
## 2 HOWDESHELL, SARAH
                         2
                                  3
                                       5
                                                0.4
## 3 OLMEN, SOPHIE
                         7
                                  12
                                       19
                                               0.368
## 4 KELLY.RACHEL
                         17
                                  37
                                               0.315
## 5 HEGLAND, GRACE
                        17
                                  41
                                               0.293
                                       58
                                  10
## 6 TANGEN.SARAH
                         4
                                       14
                                                0.286
## 7 BOTTEN, CAYDEN
                         3
                                  11
                                      14
                                               0.214
## 8 YAGODINSKI, PAIGE
                        11
                                  48 59
                                                0.186
## 9 FIX, CASSIE
                          5
                                  25
                                       30
                                                0.167
## 10 ZHEN, ZILING
                          0
pbp_full %>%
 filter(action == "REBOUND") %>%
 group_by(date) %>%
 summarise(sto off = sum(team == "STO" & type == "OFF"),
          sto_def = sum(team == "STO" & type == "DEF"),
          opp_off = sum(team != "STO" & type == "OFF"),
          opp_def = sum(team != "STO" & type == "DEF"),
          orb_pct = sto_off/(sto_off + opp_def),
          drb_pct = sto_def/(sto_def + opp_off)) %>%
 add_column(result = c(1,1,1,0,0,0,0))
## # A tibble: 7 x 8
           sto_off sto_def opp_off opp_def orb_pct drb_pct result
    date
##
    <date>
             <int> <int> <int> <int>
                                          <dbl>
                                                  <dbl> <dbl>
                                      25 0.390
                                                  0.696
## 1 2022-11-11
                16
                        32
                              14
## 2 2022-11-12
                               20
                                      17 0.370
                 10
                        31
                                                 0.608
## 3 2022-11-16
                  23
                        32
                              10
                                      24 0.489
                                                 0.762
## 4 2022-11-19
                              19
                                     21 0.382
                13
                        35
                                                 0.648
                                     23 0.258
## 5 2022-11-30
                 8
                       26
                              16
                                                 0.619
                 8
## 6 2022-12-03
                        28
                               13
                                      14 0.364
                                                 0.683
                                                           0
## 7 2022-12-07
                        29
                               5
                                      19 0.136
                                                 0.853
```

#### #games where rebounding is higher generally corellate strongly with wins

## # A tibble: 7 x 7

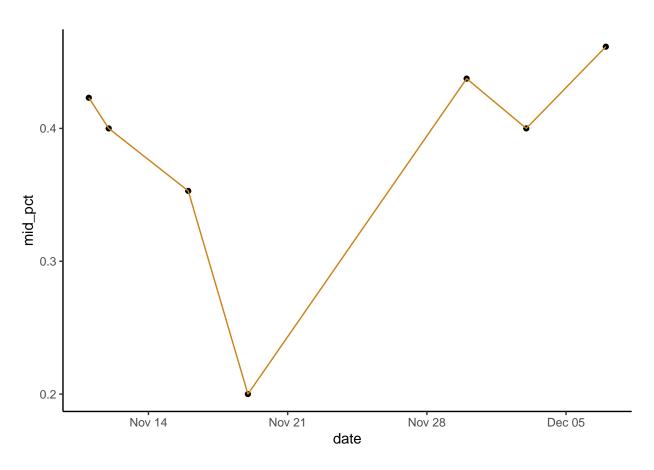
```
##
     date
                 layups mid_range threes layup_pct mid_pct three_pct
##
     <date>
                  <int>
                             <int>
                                    <int>
                                               <dbl>
                                                        <dbl>
                                                                   <dbl>
## 1 2022-11-11
                                               0.682
                                                        0.423
                                                                   0.222
                     22
                                26
                                       18
## 2 2022-11-12
                     17
                                15
                                       11
                                               0.706
                                                        0.4
                                                                  0.364
## 3 2022-11-16
                     29
                                17
                                       23
                                               0.379
                                                        0.353
                                                                  0.304
## 4 2022-11-19
                     18
                                20
                                       21
                                               0.722
                                                        0.2
                                                                  0.286
## 5 2022-11-30
                     22
                                16
                                       17
                                               0.591
                                                        0.438
                                                                  0.294
## 6 2022-12-03
                                                        0.4
                     13
                                               0.462
                                                                  0.375
                                15
                                       16
## 7 2022-12-07
                     15
                                13
                                       15
                                               0.533
                                                        0.462
                                                                  0.533
```

```
game_shooting %>%
  ggplot(aes(x = date, y = three_pct)) +
  geom_point() +
  geom_line(color = "#CB831A") +
  theme_classic()
```



```
#three pointers have recently gotten significantly better

game_shooting %>%
    ggplot(aes(x = date, y = mid_pct)) +
    geom_point() +
    geom_line(color = "#CB831A") +
    theme_classic()
```



```
#mid range shots have been relatively consistent outside of the 11/19 game
game_shooting %>%
   ggplot(aes(x = date, y = layup_pct)) +
   geom_point() +
   geom_line(color = "#CB831A") +
   theme_classic()
```

