

# peakPerformR: Athlete Prime Analysis Package

Jacob Berlin and Eli Vatsaas

2025-03-30

## Introduction

### Building Sports Data

First, let's build data for all sports using the cached data available in the package:

```
# Build all sports data
all_sports <- build_all_sports()

# Display summary of available sports data
sapply(all_sports, nrow)
```

```
## $player_id
## NULL
##
## $player_name
## NULL
##
## $season
## NULL
##
## $age
## NULL
##
## $position
## NULL
##
## $value
## NULL
##
## $pct.season.played
## NULL
##
## $games_played
## NULL
##
## $sport
## NULL
##
## $slug
## NULL
```

```
##
## $gender
## NULL
##
## $league
## NULL
##
## $id
## NULL
##
## $player_value
## NULL
##
## $career_games
## NULL
##
## $career_seasons
## NULL
##
## $career_mean
## NULL
##
## $z_score
## NULL
```

```
players_with_sufficient_data <- all_sports %>%
  group_by(id) %>%
  summarize(num_seasons = n_distinct(season),
            player_name = first(player_name)) %>%
  filter(num_seasons >= 5) %>%
  select(player_name, id)

print(paste("Number of players with 5+ seasons:", length(players_with_sufficient_data$player_name)))
```

```
## [1] "Number of players with 5+ seasons: 10931"
```

```
filtered_data <- all_sports %>%
  filter(id %in% players_with_sufficient_data$id)
```

## Processing Player Trajectories

Now, let's process player trajectories with specific filters:

```
# Apply filters for player trajectories
player_trajectories <- process_player_trajectories(filtered_data)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      3.000   4.000   5.000   5.214   6.000   9.000
```

```
# View the first few processed player trajectories
head(player_trajectories)
```

```
## $models
## # A tibble: 10,931 x 8
## # Groups:   id [10,931]
##   id          data      player_name model_info  fit_success model      knots
##   <chr>      <list>    <chr>      <list>      <lgl>      <list>    <dbl>
## 1 MLB-jeff-bagw~ <tibble> Jeff Bagwe~ <named list> TRUE      <smth.spl> 3
## 2 MLB-frank-tho~ <tibble> Frank Thom~ <named list> TRUE      <smth.spl> 8
## 3 MLB-ken-griff~ <tibble> Ken Griff~ <named list> TRUE      <smth.spl> 5
## 4 MLB-kenny-lof~ <tibble> Kenny Loft~ <named list> TRUE      <smth.spl> 8
## 5 MLB-barry-bon~ <tibble> Barry Bonds <named list> TRUE      <smth.spl> 5
## 6 MLB-albert-be~ <tibble> Albert Bel~ <named list> TRUE      <smth.spl> 6
## 7 MLB-fred-mcgr~ <tibble> Fred McGri~ <named list> TRUE      <smth.spl> 8
## 8 MLB-moises-al~ <tibble> Moises Alou <named list> TRUE      <smth.spl> 3
## 9 MLB-cal-ripke~ <tibble> Cal Ripken~ <named list> TRUE      <smth.spl> 7
## 10 MLB-rafael-pa~ <tibble> Rafael Pal~ <named list> TRUE      <smth.spl> 3
## # i 10,921 more rows
## # i 1 more variable: fit_method <chr>
##
## $trajectories
## # A tibble: 1,093,100 x 3
## # Groups:   id [10,931]
##   id          age predicted_value
##   <chr>      <dbl>      <dbl>
## 1 MLB-jeff-bagwell-00 26          4.93
## 2 MLB-jeff-bagwell-00 26.1         5.01
## 3 MLB-jeff-bagwell-00 26.2         5.08
## 4 MLB-jeff-bagwell-00 26.3         5.16
## 5 MLB-jeff-bagwell-00 26.4         5.24
## 6 MLB-jeff-bagwell-00 26.6         5.31
## 7 MLB-jeff-bagwell-00 26.7         5.39
## 8 MLB-jeff-bagwell-00 26.8         5.47
## 9 MLB-jeff-bagwell-00 26.9         5.55
## 10 MLB-jeff-bagwell-00 27          5.62
## # i 1,093,090 more rows
##
## $knot_info
## # A tibble: 10,931 x 5
## # Groups:   id [10,931]
##   id          player_name      knots fit_method fit_success
##   <chr>      <chr>      <dbl> <chr>      <lgl>
## 1 MLB-jeff-bagwell-00 Jeff Bagwell      3 method_1 TRUE
## 2 MLB-frank-thomas-00 Frank Thomas      8 method_4 TRUE
## 3 MLB-ken-griffey-jr-00 Ken Griffey Jr.    5 method_1 TRUE
## 4 MLB-kenny-lofton-00 Kenny Lofton      8 method_4 TRUE
## 5 MLB-barry-bonds-00 Barry Bonds       5 method_1 TRUE
## 6 MLB-albert-belle-00 Albert Belle      6 method_4 TRUE
## 7 MLB-fred-mcgriff-00 Fred McGriff      8 method_4 TRUE
## 8 MLB-moises-alou-00 Moises Alou       3 method_1 TRUE
## 9 MLB-cal-ripken-jr-00 Cal Ripken Jr.    7 method_4 TRUE
## 10 MLB-rafael-palmeiro-00 Rafael Palmeiro   3 method_1 TRUE
```

```
## # i 10,921 more rows
```

```
#  
# # Count players by sport  
# player_counts <- player_trajectories %>%  
#   left_join(filtered_data %>% select(id, league), by="id") %>%  
#   group_by(league) %>%  
#   summarize(player_count = n_distinct(player_id))  
#  
# print(player_counts)
```