

A complete example

Tom Stewart

2018-01-12

This article will demonstrate a real world example with a 191-participant dataset.

Reading data

All data should be loaded at the beginning of the script. This includes the PALMS dataset, and any other shapefiles and csv files. Any preprocessing, such as buffering can also be done here.

```
library(palmsplusr)
library(readr)

# Load PALMS dataset
palms <- read_palms("F:/data/csv/palms_output.csv")

# Load other csvs
participant_basis <- read_csv("F:/data/csv/participant_basis.csv")
class_timetable <- read_csv("F:/data/csv/class_timetable.csv")

# Load shapefiles (home points & schoolyard polygons)
home <- read_sf("F:/data/shapefiles/home.shp")
school <- read_sf("F:/data/shapefiles/school.shp")

# Buffer home points
home.buffer <- palms_buffer(point = home, distance = 100, crs = 2193)
```

The participant_basis file is used to link participant identifier with the correct school and classroom:

```
head(participant_basis, n = 5)
```

```
## # A tibble: 5 x 3
##   identifier school_id class_id
##   <chr>         <int>   <int>
## 1 RC0400         2       2
## 2 RC0402         2       2
## 3 RC0404         2       2
## 4 RC0406         2       2
## 5 RC0414         2       2
```

The class_timetable file details start and end times for school and recess periods, for each school/classroom:

```
head(class_timetable, n = 5)
```

```
## # A tibble: 5 x 9
##   school_id class_id date      school_start school_end
##   <int>    <int> <date>      <dtm>         <dtm>
## 1      2      2 2013-06-10 2013-06-10 08:45:00 2013-06-10 15:15:00
## 2      2      2 2013-06-11 2013-06-11 08:45:00 2013-06-11 15:15:00
## 3      2      2 2013-06-12 2013-06-12 08:45:00 2013-06-12 15:15:00
## 4      2      2 2013-06-13 2013-06-13 08:45:00 2013-06-13 15:15:00
## 5      2      2 2013-06-14 2013-06-14 08:45:00 2013-06-14 15:15:00
## # ... with 4 more variables: recess1_start <dtm>, recess1_end <dtm>,
## #   recess2_start <dtm>, recess2_end <dtm>
```

Creating the five field tables

All formulas for building `palmsplus`, `days`, `trajectories`, and `multimodal` are created. This is done by creating the five field tables:

- `palmsplus_fields`
- `palmsplus_domains`
- `trajectory_fields`
- `trajectory_locations`
- `multimodal_fields`

Most of these tables can be populated using the `palms_load_defaults()`. I use this function as a base, but then add additional fields that are specific to my dataset and research questions:

```
# Remove all field tables that exist
palms_remove_tables()

# load defaults for:
# * palmsplus_fields
# * trajectory_fields
# * multimodal_fields
palms_load_defaults(palms_epoch(palms))

# palmsplus_fields (from other files)
palms_add_field("in_school_time", "palms_in_time(., i, class_timetable,
    participant_basis, school_start, school_end)")
palms_add_field("in_recess1", "palms_in_time(., i, class_timetable,
    participant_basis, recess1_start, recess1_end)")
palms_add_field("in_recess2", "palms_in_time(., i, class_timetable,
    participant_basis, recess2_start, recess2_end)")
palms_add_field("at_home", "palms_in_polygon(., filter(home.buffer, identifier == i), identifier)")
palms_add_field("at_school", "palms_in_polygon(., filter(school,
    school_id == participant_basis %>% filter(identifier == i) %>%
    pull(school_id)))")

# palmsplus_domains
palms_add_domain("home", "at_home")
palms_add_domain("school", "(!at_home & at_school & in_school_time)")
palms_add_domain("transport", "!at_home & !(at_school & in_school_time) &
    (pedestrian | bicycle | vehicle)")
palms_add_domain("leisure", "!at_home & !(at_school & in_school_time) &
    !pedestrian & !bicycle & !vehicle")

# trajectory_locations
palms_add_trajectory_location("home_school", "at_home", "at_school")
palms_add_trajectory_location("school_home", "at_school", "at_home")
palms_add_trajectory_location("home_home", "at_home", "at_home")
palms_add_trajectory_location("school_school", "at_school", "at_school")
```

I didn't specify any additional `trajectory_fields` or `multimodal_fields` as the default fields will suffice.

All five of the field tables look like:

palmsplus_fields

Table 1: palmsplus_fields

name	formula	domain_field
weekday	dow < 6	FALSE
weekend	dow > 5	FALSE
indoors	iov == 3	FALSE
outdoors	iov == 1	FALSE
in_vehicle	iov == 2	FALSE
inserted	fixtypecode == 6	FALSE
pedestrian	tripmot == 1	FALSE
bicycle	tripmot == 2	FALSE
vehicle	tripmot == 3	FALSE
nonwear	activityintensity < 0	TRUE
wear	activityintensity >= 0	TRUE
sedentary	activityintensity == 0	TRUE
light	activityintensity == 1	TRUE
moderate	activityintensity == 2	TRUE
vigorous	activityintensity == 3	TRUE
mvpa	moderate + vigorous	TRUE
in_school_time	palms_in_time(., i, class_timetable, participant_basis, school_start, school_end) FALSE	
in_recess1	palms_in_time(., i, class_timetable, participant_basis, recess1_start, recess1_end) FALSE	
in_recess2	palms_in_time(., i, class_timetable, participant_basis, recess2_start, recess2_end) FALSE	
at_home	palms_in_polygon(., filter(home.buffer, identifier == i), identifier)	FALSE
at_school	palms_in_polygon(., filter(school, school_id == participant_basis %>% filter(identifier == i) %>% pull(school_id))) FALSE	

palmsplus_domains

Table 2: palmsplus_domains

name	formula
home	at_home
school	(!at_home & at_school & in_school_time)
transport	!at_home & !(at_school & in_school_time) & (pedestrian bicycle vehicle)
leisure	!at_home & !(at_school & in_school_time) & !pedestrian & !bicycle & !vehicle

trajectory_fields

Table 3: trajectory_fields

name	formula	after_conversion
mot	first(tripmot)	FALSE
date	first(as.Date(datetime))	FALSE
start	datetime[triptype==1]	FALSE
end	datetime[triptype==4]	FALSE
duration	as.numeric(difftime(end, start, units = "secs") + 15)	FALSE
nonwear	sum(activityintensity < 0) * 15	FALSE
wear	sum(activityintensity >= 0) * 15	FALSE
sedentary	sum(activityintensity == 0) * 15	FALSE
light	sum(activityintensity == 1) * 15	FALSE
moderate	sum(activityintensity == 2) * 15	FALSE
vigorous	sum(activityintensity == 3) * 15	FALSE
mvp	moderate + vigorous	FALSE
length	as.numeric(st_length(.))	TRUE
speed	(length / duration) * 3.6	TRUE

trajectory_locations

Table 4: trajectory_locations

name	start_criteria	end_criteria
home_school	at_home	at_school
school_home	at_school	at_home
home_home	at_home	at_home
school_school	at_school	at_school

multimodal_fields

Table 5: multimodal_fields

name	func
duration	sum
nonwear	sum
wear	sum
sedentary	sum
light	sum
moderate	sum
vigorous	sum
mvp	sum
length	sum
speed	mean

Building palmsplus, days, trajectories, and multimodal datasets

Once all of the formula tables are created, it is simple to build all of the datasets. I've included a timer here to see how long it takes to process:

```
# Start the timer
t <- proc.time()

# Build palmsplus
palmsplus <- palms_build_palmsplus(palms, verbose = FALSE)

# Build days
days <- palms_build_days(palmsplus)

# Build trajectories
trajectories <- palms_build_trajectories(palmsplus)

# Build multimodal trips
multimodal <- palms_build_multimodal(trajectories, 200, 10)

## Calculating multimodal eligibility...done
## Assigning trip numbers...done
## Calculating fields...done
cat("Run time:", (proc.time() - t)[[1]]/60, "minutes")
```

```
## Run time: 20.18233 minutes
```

Just over 15 minutes for 191 participants isn't too bad!

Saving results

```
write_csv(palmsplus, "palmsplus.csv")
write_csv(days, "days.csv")
write_csv(trajectories, "trajectories.csv")
write_csv(multimodal, "multimodal.csv")

st_write(palmsplus, "palmsplus.shp")
st_write(trajectories, "trajectories.shp")
st_write(multimodal, "multimodal.shp")
```