Forest-goers GPS data

* Lon, lat, elevation, time and speed are recorded by the GPS logger that were configured to collect data every 30 minutes in 1st cycle and every 15 minutes in 2nd cycle
* GPS logger indicates the ID of the GPS logger. Note they can be reused between 2 cycles and some new ones were added in the 2nd cycle
* District is an ID for 1 of the 4 districts of data collection in southern Lao
* Cycle indicates whether this data is from 1st (~April to ~ august) or 2nd cycle (~September to ~ November)

To do:

* Start cleaning process:
  + Visualize data on interactive maps over google image (see attached preliminary R code for that Data\_Mapping\_GPS.R)
  + Explore when GPS loggers failed to record a data point (probably because was off)
  + Explore when GPS loggers recorded a point but coordinates are obviously off. This is most likely to happen when logger stayed under a roof for a long time. In preliminary analyses, I started cleaning this by removing points GPS points that where ridiculously far away from the previous and following GPS points whereas both of those were very close by
* Create time series of GPS tracks:
  + Distinguish destinations from transit locations
  + Identify home village (based on GPS frequency, potentially weighted by combination of landcover, proximity to road)
  + Create time series of “at-home”, “in transit”, “at travel destination”
  + Characterize “malaria at-risk” environment of destination and home village using landcover and environmental rasters (percent forest within Xkm, distance to closest primary roads, distance to water,…)
  + Characterize transit routes (via forest, roads, clustered, through way points,…)
* Descriptive statistics:
  + Distance traveled (per day, per trip, over the study period)
  + Travel frequency
  + Distribution of time spent at home, in villages, in transit, in forest, …
  + Define an “at-risk” environment, e.g at night in the forest away from home village and characterize distribution