

Survey on the R packages

Rocio Joo

October 14, 2018

Survey Questions

The purpose of this anonymous survey was an assessment about the current state of the art of movement-related R packages about three aspects:

1. How popular those packages are;
2. How well documented they are;
3. How relevant they are for users.

The survey got exemption from the Institutional Review Board at University of Florida. IRB02 Office, Box 112250, University of Florida, Gainesville, FL 32611-2250.

Since the only question analyzed for the purpose of this review concerns package documentation, we only detail that question here.

Package documentation

How helpful is the documentation provided for each of the packages you've used for your work? Documentation includes what is contained in the manual and help pages, vignettes, published manuscripts, and other material about the package provided by the authors. Please answer using one of the following options:

- Not enough: It's not enough to let me know how to do what I need;
- Basic: It's enough to let me get started with simple use of the functions but not to go further (e.g. use all arguments in the functions, or put extra variables);
- Good: I did everything I wanted and needed to do with it;
- Excellent: I ended up doing even more than what I planned because of the excellent information in the documentation.
- Don't remember: I honestly can't remember...

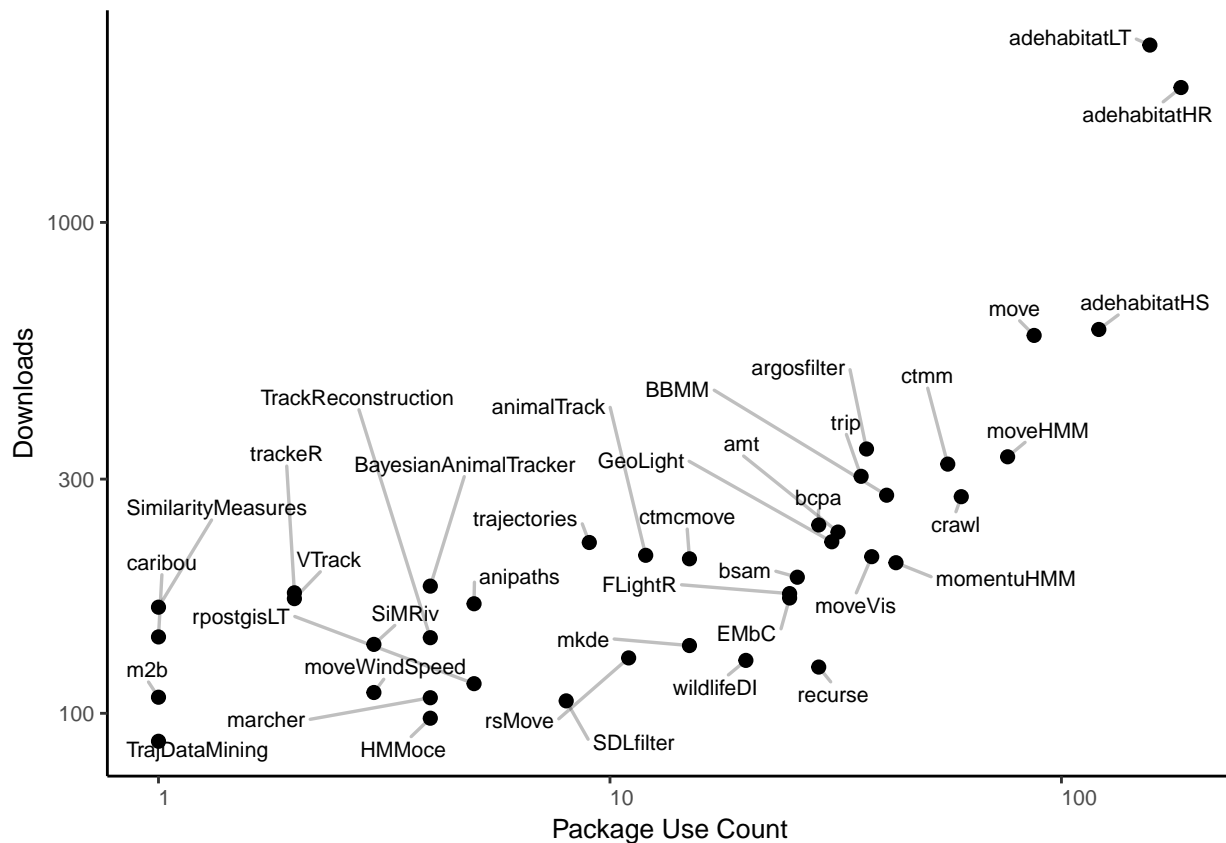
Survey representativity

There was no previous selection of the participants and no probabilistic sampling was involved. The survey was advertised by Twitter, email lists and the mablab website: <http://matlab.org>. We analyzed only completed surveys (all questions answered); 225 participants completed the survey.

To get an idea of how representative the survey was of the population of R-tracking-packages users, we compared the number of participants that used each package to the number of monthly downloads that each package has.

The number of downloads were calculated using the R package `cran.stats`. It calculates the number of independent downloads by each package (subtracting downloads by dependencies) by day. It only works for downloads using Rstudio, and for packages on CRAN. So for the tracking packages on CRAN, we computed the average of downloads per month, from September 2017 to August 2018; less months were considered for packages that were less than one year old.

There is no perfect match between the number of users and the number of downloads per package, but a correlation 0.88 for the 42 packages on CRAN, provides evidence of a good representativity of the users of tracking packages in the survey. A log-log plot for both metrics is shown in the figure below.



Survey analysis

For this work, we only analyzed the question about package documentation; only responses of participants who remembered the documentation were considered in the analysis.

We counted the number of participants who expressed that the documentation was either good or excellent, and divided by the total of answers on documentation given about the package. The packages that had more than 0.75 (or 75%) and at least 10 respondents, were considered to have “great documentation”.

Summarized results

Here are the results of the survey regarding documentation as well as the count of participants that use each package.

package	Not.enough	Basic	Good	Excellent	Don.t.remember	Use_Counts	Great_Doc
adehabitatHR	2	26	76	62	18	184	0.83
adehabitatHS	0	14	52	35	20	121	0.86
adehabitatLT	2	14	66	58	17	157	0.89
amt	3	8	11	2	8	32	0.54
animalTrack	1	3	3	0	5	12	0.43
anipaths	0	2	2	0	1	5	0.50
argosfilter	1	8	9	3	16	37	0.57
argosTrack	1	5	1	1	6	14	0.25
BayesianAnimalTracker	1	2	0	1	0	4	0.25
BBMM	3	10	12	1	15	41	0.50

package	Not.enough	Basic	Good	Excellent	Don.t.remember	Use_Counts	Great_Doc
bcpa	1	4	12	3	9	29	0.75
bsam	0	5	13	3	5	26	0.76
caribou	0	0	0	0	1	1	NaN
crawl	4	15	12	3	26	60	0.44
ctmcmove	0	6	2	0	7	15	0.25
ctmm	1	7	20	12	16	56	0.80
EMbC	1	3	11	7	3	25	0.82
feedr	0	0	0	0	0	0	NaN
FLightR	0	5	9	3	8	25	0.71
GeoLight	0	6	16	4	5	31	0.77
hab	3	3	5	1	2	14	0.50
HMMoce	0	0	1	1	2	4	1.00
kftrack	1	0	2	1	3	7	0.75
m2b	0	0	0	0	1	1	NaN
marcher	0	0	2	1	1	4	1.00
migrateR	0	6	3	4	6	19	0.54
mkde	0	2	2	0	11	15	0.50
momentuHMM	0	2	11	19	11	43	0.94
move	0	15	35	14	23	87	0.77
moveHMM	2	4	26	25	19	76	0.89
movementAnalysis	1	5	1	0	2	9	0.14
moveNT	0	1	0	0	0	1	0.00
moveVis	4	6	17	5	6	38	0.69
moveWindSpeed	1	1	0	0	1	3	0.00
probGLS	0	1	2	2	1	6	0.80
recurse	1	3	8	5	12	29	0.76
rhr	0	6	4	1	2	13	0.45
rpostgisLT	0	1	2	0	2	5	0.67
rsMove	2	0	2	1	6	11	0.60
SDLfilter	0	0	2	2	4	8	1.00
SGAT/TripEstimation	6	4	6	2	6	24	0.44
SimilarityMeasures	0	0	1	0	0	1	1.00
SiMRiv	0	0	2	1	0	3	1.00
smam	0	0	0	0	0	0	NaN
T-LoCoH	4	5	10	7	17	43	0.65
telemetr	0	2	0	0	0	2	0.00
trackeR	0	1	0	0	1	2	0.00
trackit	2	0	1	1	2	6	0.50
TrackReconstruction	0	1	1	0	2	4	0.50
TrajDataMining	0	0	1	0	0	1	1.00
trajectories	0	2	1	0	6	9	0.33
trip	1	7	9	4	15	36	0.62
TwilightFree	0	3	4	0	2	9	0.57
ukfsst/kfsst	0	2	0	1	4	7	0.33
VTrack	0	0	2	0	0	2	1.00
wildlifeDI	0	3	8	5	4	20	0.81