kaggle

Host   Competitions   Scripts   Jobs   Community ▾   Rupak Chakraborty   Logout

Thank you for accepting the rules.

Jobs • 188 teams

Airbnb Recruiting: New User Bookings

Wed 25 Nov 2015   Thu 11 Feb 2016 (2 months to go)

Dashboard

Home
  Data
  Make a submission

Information
  Description
  Evaluation
  Rules
  Timeline

Forum

Scripts
  New Script
  New Notebook

Leaderboard

My Team
  Upload your model

My Submissions

Leaderboard

1. jil663
2. Just_a_try
3. zhiminzhou
4. Anonymous 15201
5. Adil
6. Sandro
7. D3PO
8. Vladimir Larin
9. Josh Stone
10. CamdenClark

148 Scripts

NDCG example
5 Votes / yesterday / Python

FisrtScript
8 Votes / 4 days ago / Python

Competition Details   »   Get the Data   »   Make a submission

Data Files

File Name	Available Formats
countries.csv	.zip (546 b)
age_gender_bkts.csv	.zip (2.46 kb)
sample_submission.csv	.zip (387.29 kb)
test_users.csv	.zip (804.86 kb)
train_users.csv	.zip (3.28 mb)
sessions.csv	.zip (31.34 mb)

In this challenge, you are given a list of users along with their demographics, web session records, and some summary statistics. You are asked to predict which country a new user's first booking destination will be. All the users in this dataset are from the USA.

There are 12 possible outcomes of the destination country: 'US', 'FR', 'CA', 'GB', 'ES', 'IT', 'PT', 'NL','DE', 'AU', 'NDF' (no destination found), and 'other'. Please note that 'NDF' is different from 'other' because 'other' means there was a booking, but is to a country not included in the list, while 'NDF' means there wasn't a booking.

The training and test sets are split by dates. In the test set, you will predict all the new users with first activities after 4/1/2014. In the sessions dataset, the data only dates back to 1/1/2014, while the users dataset dates back to 2010.

File descriptions

- **train_users.csv** - the training set of users
- **test_users.csv** - the test set of users

  - id: user id
  - date_account_created: the date of account creation

Forum (19 topics)

teams

players

entries

- timestamp_first_active: timestamp of the first activity, note that it can be earlier than date_account_created or date_first_booking because a user can search before signing up
- date_first_booking: date of first booking
- gender
- age
- signup_method
- signup_flow: the page a user came to signup up from
- language: international language preference
- affiliate_channel: what kind of paid marketing
- affiliate_provider: where the marketing is e.g. google, craigslist, other
- first_affiliate_tracked: whats the first marketing the user interacted with before the signing up
- signup_app
- first_device_type
- first_browser
- country_destination: this is the **target variable** you are to predict

- **sessions.csv** - web sessions log for users

  - user_id: to be joined with the column 'id' in users table
  - action
  - action_type
  - action_detail
  - device_type
  - secs_elapsed

- **countries.csv** - summary statistics of destination countries in this dataset and their locations
- **age_gender_bkts.csv** - summary statistics of users' age group, gender, country of destination
- **sample_submission.csv** - correct format for submitting your predictions

About   Our Team   Careers   Terms   Privacy   Contact/Support