

Airbnb data ER

age_gender_bkts
age_bucket = [0-4,5-9,10-14,...,95-99,100+]
country_destination = [AU,CA,DE,ES,FR,GB,IT,NL,PT,US]
gender = [male,female]
population_in_thousands = type_int
year = [2015]

sample_submission_NDF	
PK	<u>id</u>
	country = [AU,CA,DE,ES,FR,GB,IT,NL,PT,US,NDF,other] (NDF = no destination found, other = other countries)

train_users / test_users	
PK	<u>id</u>
	date_account_created = type_date timestamp_first_active = type_int (timestamp) date_first_booking = type_date (included missing val) (N/A in test_users) gender = [-unknown-,MALE,FEMALE,OTHER] age = type_int (included missing val) signup_method = [basic,facebook,google] signup_flow = [0,1,2,3,6,24] (not understand) language = [ca,cs,da,de,el,en,...,ru,sv,th,tr,zh] affiliate_channel = [api,content,direct...,seo] (direct same as below) affiliate_provider = [baidu,bing,google,craigslist,...,direct,facebook] first_affiliate_tracked = [linked,marketing,untracked,...] (included missing val) signup_app = [Android,iOS,Moweb,Web] first_device_type = [Tablet,iPhone,Android,Mac,Win,Desktop...,other] first_browser = [Browsers...] country_destination = [AU,CA,DE,ES,FR,GB,IT,NL,PT,US,NDF,other] (result)

countries
country_destination = [AU,CA,DE,ES,FR,GB,IT,NL,PT,US]
lat_destination = type_float
lng_destination = type_float
distance_km = type_float (compared to US)
destination_km2 = type_int (not understand)
destination_language = [deu,eng,fra,ita,nld,por,spa]
language_levenshtein_distance = type_float (edit distance compare to ENG)

sessions
user_id
action = [lookup,search_result,personal,...] (action_location_module_list)
action_type = [click,view,data,...] (action_list) (included missing)
action_detail = [view_result,wishlist_update...] (action_object_location) (included missing)
device_type = [Windows,Tablet,Android,...]
secs_elapsed = type_int (included missing)