

Trusted Housesitters

Principal Data Scientist Technical Task Dataset Documentation

This document explains the tables available for the Principal Data Scientist technical task. It's designed to save you time by clearly outlining the information in each field, so you can focus on the core clustering task. You are welcome to use any or all of these tables as needed.

Table: assignments

Description: Each row in the assignments model represents a sit (or assignment) created on the platform, detailing metrics and dimensions such as start/end dates, locations, and sitter, owner and sit details. Both sitters and owners (or listings) can have multiple assignments.

Field	Field Description
id	Unique identifier for each assignment/sit
created_dt	Date the assignment/sit was created
last_modified_dt	Date the assignment/sit was modified
start_dt	Date the sit is due to start
end_dt	Date the sit is due to end
private	Is true if the sit is private
approximate_dates	Is true if the sit start/end dates are approximate dates
sitter_found	Is true if the sit has been filled
listing_id	ID of the listing associated with the assignment/sit
owner_user_id	ID of the user who created the assignment/sit
ocountry	Country of the listing
oregion	Region of the listing
ocity	City of the listing
profile_id	ID of the sitter's profile, if the sit has been filled
sitter_user_id	User ID of the sitter who has filled the sit
scountry	Country of the sitter
sregion	Region of the sitter
scity	City of the sitter
is_new	Is true if the assignment/sit is new
reviewing	Is true if the assignment/sit is in review
welcome_guide_shared_ts	Date and time the owner shared their welcome guide with the sitter. This contains details about their home and pets, and is optional.
deleted	Is true if the assignment/sit was deleted

Table: profiles

Description: Each record in profiles represents an approved sitter profile that has been created. It brings together dimensions and metrics associated with each profile relating to the sitter details, photos, assignments, reviews and applications.

Field	Field Description
id	Unique identifier for each sitter profile
user_id	The user ID of the sitter
modified_dt	The date the profile was last modified
latitude	The latitude of the sitter
longitude	The longitude of the sitter
pct_complete	The percentage of the profile that is complete
travelling_as	The sex of the sitter and their travel companion (where applicable) <ul style="list-style-type: none">• F for Female• M for Male• N for Non-binary• T for Trans• U for Unspecified• X for Blank
with_children	Is true if the sitter is travelling with children
child_ages	The ages of the children, where applicable
sitting_with_another	Is true if the sitter is travelling with someone else e.g. a partner
birth_date	The sitter's date of birth
partner_birth_date	The sitter's travel companion's date of birth
occupation_type	The sitter's occupation type
prev_sitting_experience	Is true if the sitter has previous sitting experience
other_animals	A list of other animals the sitter has experience with
has_dog_experience	Is true if the sitter has experience with dogs
has_cat_experience	Is true if the sitter has experience with cats
has_reptile_experience	Is true if the sitter has experience with reptiles
has_horse_experience	Is true if the sitter has experience with horses
has_fish_experience	Is true if the sitter has experience with fish
has_poultry_experience	Is true if the sitter has experience with poultry
has_farm_animal_experience	Is true if the sitter has experience with farm animals
has_bird_experience	Is true if the sitter has experience with birds
has_small_pet_experience	Is true if the sitter has experience with small pets
prefers_all_countries	Is true if the sitter has selected at least 250 countries
preferred_countries	List of preferred countries selected by the user

location_wish_list	List of preferred locations selected by the user
nb_reviews	The number of reviews received, given the assignment has not been deleted
nb_5s_reviews	The number of 5 star reviews received, given the assignment has not been deleted. The sitter must have received 5 stars for all categories (organised, reliable, self sufficient, tidy and pet care)
nb_applications	The number of unique sits the sitter has applied for or been invited to
nb_sits_completed	The number of sits completed by the sitter, given the assignment has not been deleted and has already ended
nb_domestic_sits	The number of domestic (same country) sits completed by the sitter, given the assignment has not been deleted and has already ended
nb_local_sits	The number of local (same city) sits completed by the sitter, given the assignment has not been deleted and has already ended
nb_sits_booked	The number of sits booked for the future, given the assignment has not been deleted
years_of_experience	The number of years of experience the sitter has
skills	A list of animal related skills the sitter has
nb_of_sitter_pets	The number of sitter pets associated with the profile
daily_minutes_willing_to_walk_dogs	How many daily minutes the sitter is willing to walk dogs
dog_size_preferences	Dogs sizes preferred by sitter
happy_to_meet_in_person	Is true if the sitter is happy to meet up in person
happy_to_video_call	Is true if the sitter is happy to video call
interested_in_remote_working	Is true if the sitter is interested in remote working
interested_in_dogs	True if the sitter is interested in looking after dogs for potential homeowners
interested_in_cats	True if the sitter is interested in looking after cats for potential homeowners
interested_in_reptiles	True if the sitter is interested in looking after reptiles for potential homeowners
interested_in_horses	True if the sitter is interested in looking after horses for potential homeowners
interested_in_fish	True if the sitter is interested in looking after fish for potential homeowners
interested_in_poultry	True if the sitter is interested in looking after poultry for potential homeowners
interested_in_farm_animals	True if the sitter is interested in looking after farm animals for potential homeowners
interested_in_birds	True if the sitter is interested in looking after birds for potential homeowners

interested_in_small_pets	True if the sitter is interested in looking after small pets for potential homeowners
--------------------------	---

Table: listings

Description: Each record in this table represents an approved listing that has been created. It brings together dimensions and metrics associated with each listing relating to the home details, photos, assignments and applications. Listings -> Users is a one-to-one relationship. There shouldn't be users with multiple listings.

Field	Field Description
id	The listing id
user_id	The user id of the owner who created the listing
approved_dt	The date when the listing was approved by trustedhousesitters
modified_dt	The most recent date that the listing was modified
latitude	latitude of the listing
longitude	longitude of the listing
pct_complete	The percentage of the listing that is complete
home_type	The type of home, e.g. house or apartment. Null if incomplete
car_required	Boolean field to flag whether or not a car is required to carry out the sit
car_included	Boolean field to flag whether or not a car will be available to the sitter
disabled_access	Boolean field to flag whether or not there is disabled access to the property
wifi_available	Boolean field to flag whether wifi is available in the property
family_friendly	Boolean field to flag whether the property is family friendly or not
local_attractions	List of attractions such as beach, countryside, city, mountain
nb_of_photos	The total number of photos attached to the listing and the related pet profiles
photo_tags	A list of the tags attached to each of the related listing and pet photos
nb_of_pets	A count of the total number of pets for the listing
nb_distinct_pets	A count of the unique pet types like cats, dogs
pets	A comma separated list of the unique pet types e.g. cats, dogs
other_animals	The copy added to the listing to describe any other animals not otherwise listed
nb_assignments_published	The total number of assignments published for the listing
nb_assignments_filled	The total number of assignments with a confirmed sitter
nb_unique_sitters	The number of unique sitters who have been confirmed on an assignment for the listing
nb_repeat_sitters	The number of sitters who have completed more than one assignment for the listing
nb_domestic_sitters	The number of times a domestic sitter has been confirmed on an

	assignment
avg_nb_apps_per_assg	The average number of applications received per assignment
nb_invites	The number of invites that the owner has sent in relation to their listing
accessible_by_public_transport	A boolean field to show whether the property is accessible by public transport
completed_dt	The date when the listing was first completed
children_preferences	A list of the age ranges that are suitable for children to be at the listing
family_preferences	A list of the suitable family situations for the listing
minutes_pet_can_be_left_alone	The time in minutes that the pets can be left at the property for
pets_welcome	A boolean field to flag whether the sitters pets are welcome at the property
suitable_for_home_working	A boolean field to flag whether the property is suitable for a home worker

Table: pets

Description: Each row in the pets model represents an individual pet associated with a listing. The table contains information on the pets, including their name, size, etc, as well as their needs during the stay.

Field Name	Field Description
id	Unique identifier for each individual pet
listing_id	The unique ID of the listing associated with the pet
pet_name	The name of the pet
description	A description of the pet
animal	The type of animal the pet belongs to
breed	The breed of the pet
year_of_birth	The year the pet was born in
daily_activity_minutes	The number of minutes the pet needs walking
requires_regular_medication	Whether the pet requires regular medication or not (true/false)
size	The size of the pet, in XS/S/M/L/XL notation
extra_personality	Selected adjectives the owner has chosen to describe the personality of their pet

Table: users

Description: Each row in the users model represents a distinct user (includes both partials and members), detailing attributes such as registration date and method, preferences, and membership status.

Field	Field Description
id	Unique identifier for each user
registered_dt	Date the user registered
registration_method	The method used to register: facebook, google, apple or email. The logic is based on whether a facebook/google/apple user ID exists.
marketing_comms_accepted	Indicates whether the user has accepted marketing comms
alerts	The user's settings for alerts: daily, real, none
can_contact_phone	Indicates whether the user wants to be contacted on the phone
nb_raf_emails_sent	The number of RAF (refer a friend) emails sent by the user
billing_city	The city of the billing address provided
billing_state	The state of the billing address provided
billing_country	The country of the billing address provided

Table: feedback

Description: Each record in feedback represents feedback left by the housesitter for the homeowner and their sitting.

Field	Field Description
id	Unique identifier for the feedback
assignment_id	The ID of the sitting receiving the feedback
feedback_ts	Date and time the feedback was submitted
listing_id	The ID of the listing receiving the feedback
owner_user_id	The user ID of the owner who is receiving the feedback
profile_id	The ID of the profile associated with the sitter who is giving the feedback
sitter_user_id	The user ID of the sitter who is giving the feedback
recommended	The overall rating given to the sit out of 5
reply_sent_ts	Date and time the reply was sent
requested	Is true if the feedback was requested by the owner

Table: reviews

Description: Each record in reviews represents a review left by the homeowner for the housesitter and their stay at their place.

Field	Field Description
id	Unique identifier for each review
assignment_id	Unique identifier for assignments
review_ts	The date the review was submitted
listing_id	Unique identifier for each listing
owner_user_id	Unique identifier for each homeowner
profile_id	Unique identifier for each housesitter profile
sitter_user_id	Unique identifier for each housesitter user
organised	Score out of five on how organised a sitter was
reliable	Score out of five on how reliable a sitter was
self_sufficient	Score out of five on how self sufficient a sitter was
tidy	Score out of five on how tidy a sitter was
pet_care	Score out of five on how good care did the sitter take of the pet(s)
avg_score	Average score out of the five scores the sitter received: organised, reliable, self sufficient, tidy and pet care
overall_score	Overall score given to the sitter
reply_ts	The date a reply to the comment was left
requested	Whether the review was requested or not

Table: applications

Description: Each row in the applications model represents an application to a sit (or assignment) created by the sitter, or invitations created by the owner. This model includes metrics and dimensions such as created/modified dates, success metrics, and sitter, owner and sitting details. Both sitters and owners (or listings) can make/receive multiple applications.

Field	Field Description
id	Unique identifier for each application/invite
created_dt	Date the application was created
modified_dt	Date the application was last modified
assignment_id	ID of the assignment/sit the application is for
assignment_dates_changed	Indicates whether the assignment/sit dates have changed
listing_id	ID of the listing
owner_user_id	ID of the owner receiving the application or sending the invite
profile_id	ID of the sitter's profile
sitter_user_id	ID of the sitter sending the application or receiving the invite
req_type	Whether the application was a sitter application or an owner invite
acknowledged	Indicates whether the application/invite has been acknowledged by both parties
short_listed	Indicates whether the application/invite has been short listed
owner_confirmed	Is true if the owner confirms the application
owner_declined	Is true if the owner declines the application
owner_cancelled	Is true if the owner cancels the application
sitter_confirmed	Is true if the sitter confirms the application
sitter_declined	Is true if the sitter declines the application
sitter_cancelled	Is true if the sitter cancels the application