Egyptopia Places Dataset Feature Descriptions

Feature Name

- **place_id**: Unique identifier assigned to each place in the dataset.

- **Name**: The official name of the place.

- **Category**: The general category or type of place (e.g., museum, park, monument).

- **Google Maps Link**: URL linking to the place's location on Google Maps.

- **City**: The city in Egypt where the place is located.

- **Rate**: Average user rating of the place based on online reviews.

- **Total Rates**: Total number of ratings submitted for the place.

- **Description**: Brief textual description of the place, highlighting notable features.

- **Tourism Type**: The type of tourism associated with the place (e.g., cultural, natural).

- **city_id**: Unique identifier for the associated city.

- **Bayesian Average**: Adjusted rating score calculated using the Bayesian Average method to address data sparsity.

- **Latitude**: Geographic latitude coordinate of the place, extracted from Google Maps.

- **Longitude**: Geographic longitude coordinate of the place, extracted from Google Maps.

- **Popularity**: Feature indicating the relative popularity of the place, used to diversify recommendations between popular and less popular destinations.

Dataset Versions

The dataset evolved in four stages: adding Bayesian scores for cold-start recommendations, extracting geographic coordinates for location-based suggestions, and introducing a popularity feature to balance recommendations between well-known and lesser-known places.

egyptopia_places_v1

place_id	Name	Category
City	Rate	Total Rates
Tourism Type	city_id	Description
	Google Maps Link	

egyptopia_places_v2

place_id	Name	Category	Google Maps Link
City	Rate	Total Rates	Description
Tourism Type	city_id	Bayesian Average	

egyptopia_places_v3

place_id	Name	Category	Google Maps Link
City	Rate	Total Rates	Description
Tourism Type	city_id	Bayesian Average	Latitude
Longitude			

egyptopia_places_v4

place_id	Name	Category	Google Maps Link
City	Rate	Total Rates	Description
Tourism Type	city_id	Bayesian Average	Latitude
Longitude	Popularity		