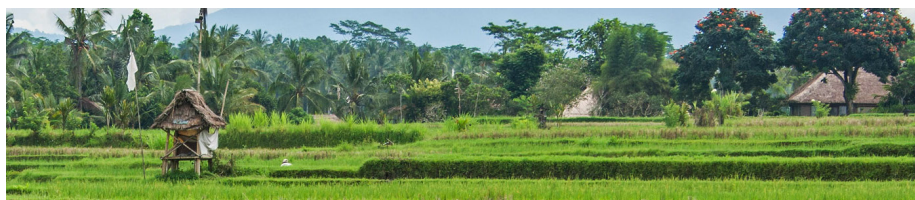


Bali Airbnb April 2018

Una mirada al listado de propiedades de Airbnb en Bali, Indonesia



CC BY SA-3.0 — *A view of Bali*

Autor

David Moriano <dmoriano@uoc.edu>

Contexto

Airbnb afirma ser parte de la “economía colaborativa” y de romper la industria hotelera. Analizando los datos en una ciudad o región podremos descubrir por nosotros mismos el impacto real de “compartir el hogar”.

Contenido

Este conjunto de datos contiene más de 12000 listados de propiedades Airbnb en Bali, Indonesia, recogidos en Abril del 2018. No pretende ser una colección completa de todas las propiedades disponibles en Bali pero con suerte puede ser un conjunto representativo.

Los datos han sido recogidos de Airbnb.com mediante un scraper creado con [Scrapy](#). El scraper itera sobre los sectores de Bali que proporciona Airbnb y sobre el rango de precios (en incrementos de 10 USD). Esta metodología no garantiza la recopilación de todas las propiedades listadas en una ciudad o territorio, pero permite profundizar más allá de las apenas 300 entradas que Airbnb ofrece a través de su web.

El scraper debería ser lo bastante genérico como para funcionar en otras ciudades. No obstante, no se dan garantías al respecto.

Los datos incluyen los siguientes campos:

Columna	Descripción	Tipo
<code>_id</code>	Listing identifier in Airbnb	Numeric

Columna	Descripción	Tipo
name	Name of the listing	String
user_id	User identifier of the host	Numeric
user_name	First name of the host	String
city	City name	String
neighborhood	Neighborhood	String
latitude	Latitude of the neighborhood	Numeric
longitude	Longitude of the neighborhood	Numeric
section_offset	Page number on which it was listed	Numeric
bathrooms	Number of bathrooms	Numeric
bedrooms	Number of bedrooms	Numeric
beds	Number of beds	Numeric
business_travel_ready	Can accommodate business travelers	Boolean
host_languages	Host languages	Array
new_listing	Is it a new listing?	Boolean
person_capacity	Guests capacity	Numeric
property_type	Type of property	Numeric
refundable	Is it refundable	Boolean
reviews_count	Number of review	Numeric
room_type	Room type	Numeric
scrim_color	Scrim color	String
star_rating	Star rating given by the guests	Numeric
superhost	Is the host a superhost	Boolean
tier_id	Tier identifier	Numeric
verified	Is it a verified host	Boolean
price_rate	Price rate in USD	Numeric
price_rate_type	Price rate type	Numeric
weekly_price_factor	Weekly stay discount factor	Numeric
monthly_price_factor	Monthly stay discount factor	Numeric

Inspiración

Claramente, este conjunto de datos se puede utilizar para encontrar el lugar ideal para las próximas vacaciones en Bali (y potencialmente en otro destino). O también se puede utilizar para determinar el precio ideal de tu propiedad si la quieres listar en Airbnb. Estas cuestiones prácticas son sin duda de mucho interés, pero también se puede mirar a este conjunto de datos desde un contexto diferente.

El interés en los datos de Airbnb surge de la investigación realizada por Murray Cox ([InsideAirbnb.com](https://insideairbnb.com)) sobre el impacto de Airbnb en la [gentrificación](#) de New York. Murray Cox intenta responder a algunas preguntas relevantes dado el tamaño de las actividades de Airbnb: ¿Cómo se está usando Airbnb? ¿Cómo está afectando a los diferentes barrios de New York? Su investigación ha sido

expandida recientemente por el grupo de investigación Urban Planning and Governance de la universidad McGill en Canada en un [informe](#) que responde a estas preguntas:

1. Dónde se localiza en New York la actividad de Airbnb, y cómo está cambiando.
2. Quién hace dinero con Airbnb en New York.
3. Cuantas viviendas se han perdido en New York en Airbnb.
4. Es Airbnb responsable de gentrificación en New York.

La revista [CityLab](#) ha publicado un interesante [artículo](#) sobre esta investigación.

Está claro que nos podemos hacer las mismas preguntas sobre otras ciudades.

Bali es el centro cultural de Indonesia y uno de los lugares más turísticos del mundo; el turismo supone el 80% de la economía de esta isla. La actividad de Airbnb en Bali es muy significativa, es natural preguntarse cual es su impacto.

Licencia

Los datos se publican con licencia:

[Creative Commons Attribution-NonCommercial 4.0 International License](#).

Esto significa que los datos se pueden:

- Compartir - copiar y redistribuir en cualquier medio o formato
- Adaptar - reorganizar, transformar y construir sobre los datos

Con las siguientes condiciones:

- Atribución - Se debe dar crédito, proporcionar un enlace a la licencia, e indicar si se han hecho cambios, de cualquier modo que sea razonable, pero no de forma que se sugiera que el licenciante te respalda a ti o al uso que hagas de los datos.
- No-comercial - No se pueden usar los datos para propósitos comerciales.

Se pueden consultar los detalles completos de la [licencia](#) en este enlace.

La razón de esta licencia es que este conjunto de datos se ha construido en un contexto académico con propósito educativo o de investigación. La idea es mantener esa intención.

Agradecimientos

Gracias a Airbnb por permitir el scraping en su website. Y gracias a los desarrolladores de [Scrapy](#) por esta magnifica herramienta.

Código

El código del scraper está disponible en el repositorio github.com/daben/m2851 en el directorio [airbnb_scraper](#).

El código se publica bajo [licencia MIT](#).

Para ejecutar el scraper se ejecuta:

```
$ scrapy crawl airbnb -a query="Bali, Indonesia" -o bali_listings.csv
```

Files

File	Description
.gitignore	Specifies intentionally untracked files to ignore in the repository
LICENSE	Data license
Pipfile	See https://github.com/pypa/pipenv
Pipfile.lock	ditto
README.md	This document
airbnb_scraper/	Scraper project folder
../LICENSE	Code License
../scrapy.cfg	Scrapy configuration file
../airbnb_scraper/	Scraper code
../../_init_.py	Module initialization
../../items.py	AirbnbListing item
../../middlewares.py	Middleware (not used)
../..pipelines.py	Pipelines: remove duplicates
../..settings.py	Scraper settings
../..spiders/	Spiders folder
../..../_init_.py	Spiders module initialization
../..../airbnb.py	Airbnb spider
assets/banner.jpg	Banner image

Dataset

- [listings_bali_201804.csv](#) (~3MiB)