

STATISTICS INDONESIA

Promising New Indicators for Tourism Statistics

By:

Setia Pramana

Associate Professor

Politeknik Statistika STIS

BPS Statistics Indonesia

Delivered on:
UNESCAP Stat Cafe

Time: 27 August, 2021

BIG DATA SOURCES

- *Mobile phone data*
- *Financial transactions*
- *Online search and access logs*
- *Citizen card*
- *Postal data*

Exhaust data

- *Satellite and UAV imagery*
- *Sensors in cities, transport and homes*
- *Sensors in nature, agriculture and water*
- *Wearable technology*
- *Biometric data*
- *Internet of Things (IoT)*

Sensing data

- *Social media data*
- *Web scraping*
- *Participatory sensing / crowdsourcing*
- *Health records*
- *Radio content*

Digital Content



What People Do

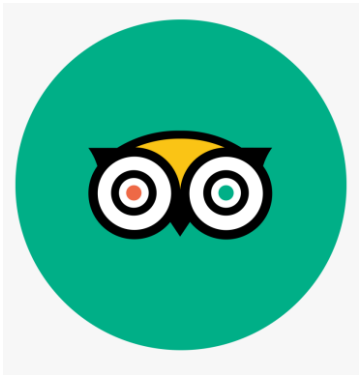


What People Say

BIG DATA SOURCES



Online Booking site: Booking.com and Agoda.com



Review Sites: Trip Advisor

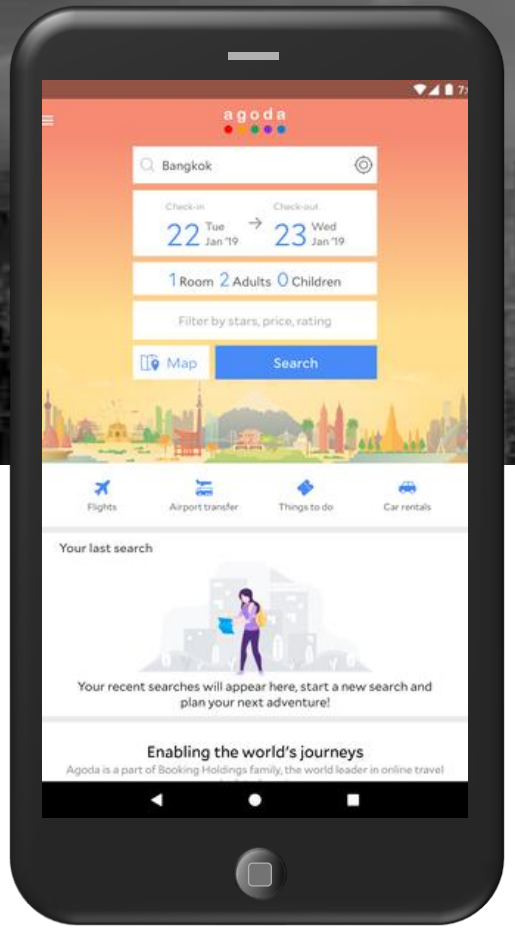


Google Trend



Google map

Online Booking Sites



Tourism statistics can be retrieved from various online accommodation booking service sites. These sites have quite diverse data such as the number of rooms available, rental prices, and other information.

Data Sources:
Website www.agoda.com, www.booking.com

Data Acquisition Technique
Web scraping

More detail : http://informationr.net/ir/25-4/paper885.html

Agoda: Collected Variables

01 Accommodation ID

02 Accommodation Name

03 Accommodation Type

04 Date of Collection

05 Location

06 Number of Rooms available

07 Total number of rooms

08 Room Prices

09 Star Rating

$$ROR = \frac{Total\ rooms - available\ room}{Total\ rooms}\ x\ 100$$

TripAdvisor: Collected Variables

01

Name

Name of the reviewer

03

Date

Date of the reviewer stay in that accommodation

02

Title

Feedback to the accommodation from the reviewer

04

Review_text

Explanation about the accommodation from the reviewer (can be positive feedback or negative feedback/improvement feedback)

Booking.com: Collected Variables

01 **url**
Link page of the review

02 **place**
Place that the reviewer visit

03 **reviewer_name**
Name of the reviewer

04 **reviewer_country**
Reviewer country of origin

05 **room**
Type of room that the reviewer used

06 **long_stay**
Duration of the stay from the reviewer

07 **time_stay**
Time of month when the reviewer visit

08 **review_time**
Time of the reviewer give the feedback/review of the accomodation

09 **review_reaction**
Reaction or comment feedback of the reviewer about the accommodation

10 **review_score**
Score given to the accommodation (1-10)

11 **like_comment**
Positive comment feedback from the reviewer about the accommodation

12 **dislike_comment**
Negative comment feedback/improvement feedback from the reviewer about the accommodation



About Google Trend

1. *Google Trends provides time-series data in the form of an index that shows the volume of queries entered by users in a region*
2. *Users can set their search data to start from the last hour or another range with a maximum for Google Trend data ranging from 2004 to the present. However, this maximum range is only available in the form of monthly data*
3. *It is a search ratio ranges from 100 to zero, where 100 indicates that the search term is relatively popular in the selected region.*
4. *According to Google, this normalization allows for smoother comparisons among search terms as search volume varies across different countries*
5. *Google Trend Index of different keywords are then combined by using weight resulted from Principle Component Analysis (PCA)*

Keywords

Bali	*Maldives*	*Phuket*
• Tourist Bali	• Tourist Maldives	• Tourist Phuket
• Hotel Bali	• Hotel Maldives	• Hotel Phuket
• Holiday in Bali	• Holiday in Maldives	• Holiday in Phuket
• Bali Tourism	• Maldives Tourism	• Phuket Tourism
• Bali Tourist Place	• Maldives Tourist Place	• Phuket Tourist Place
• Travel in Bali	• Travel in Maldives	• Travel in Phuket
• Bali	• Maldives	• Phuket

GOOGLE MAPS

Sources and Data Collection Methods.





Google API

- *The data collection process is carried out using the Python language with package **requests** and **urllib**.*
  1. *Collecting Hotel, tourist places, restaurant coordinates in the province using the Geocoding API*

Web Scraping

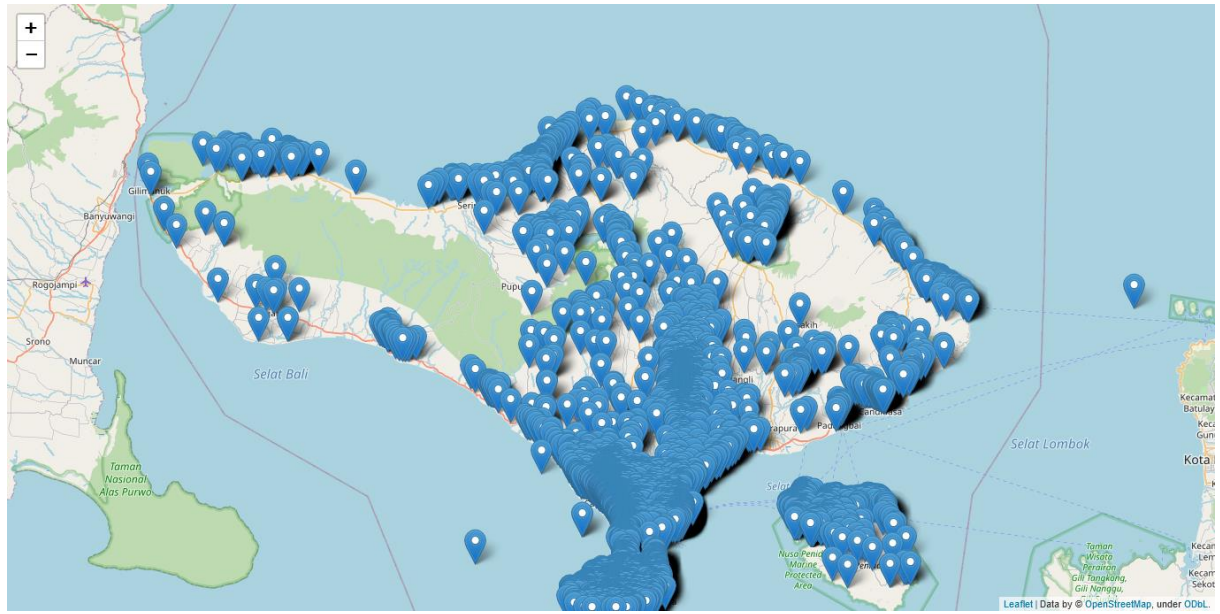
- *The data was collected using the web scraping package selenium method which was applied to the Python and Java programming languages.*
  1. *Hotel, tourism attractions, restaurant URL Collection*
  2. *Detailed location Data Collection*

GOOGLE MAPS

ATTRIBUTE	TYPE	EXPLANATION	SAMPLE DATA
Name	String	place name	Angkringan Rumah Cerita
Full_Address	String	place address	Jl. Tukad Citarum No.25, Renon, Kec. Denpasar Sel., Kota Denpasar, Bali 80234
Website	String	place website (may NULL)	business.site
Plus_Code	String	place plus code	869M+FR Renon, Denpasar City, Bali
Rating_all	float	rating in overall	4.3
Reviews_counts	integer	total number of reviews	67
Review	String	review text (may NULL)	(Diterjemahkan oleh Google) Tempat kecil (Asli) Small places
Rating	float	rating given each reviews	3
time	String	time the reviews was given	setahun yang lalu

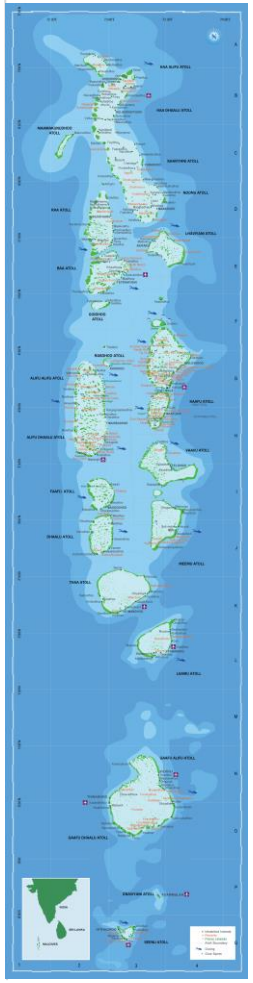
ACCOMMODATIONS LIST

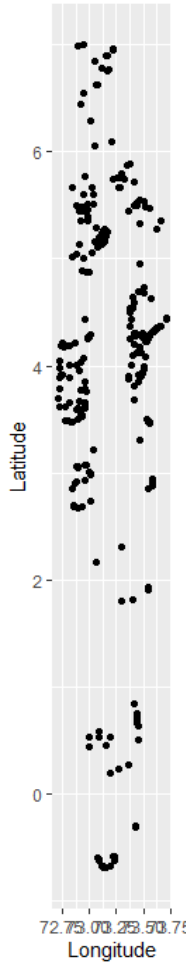
Bali

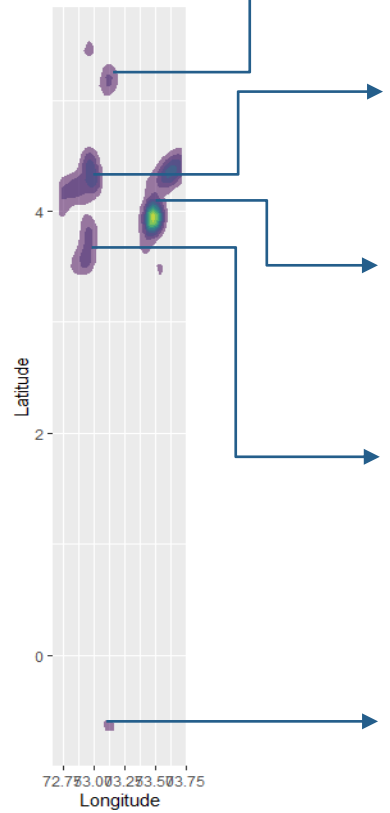


Type	Accommodation Type
Entire villa	1998
Guesthouse/bed and breakfast	1767
Hotel	1467
Resort villa	1186
Resort	1050
Villa	502
Homestay	407
Entire apartment	274
Entire bungalow	238
Hostel	190

Maldives

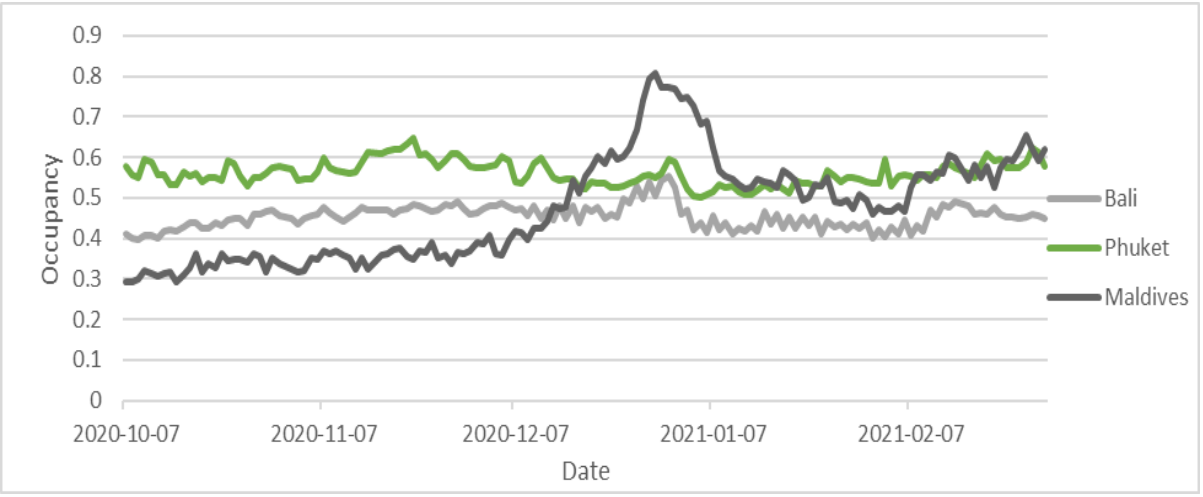




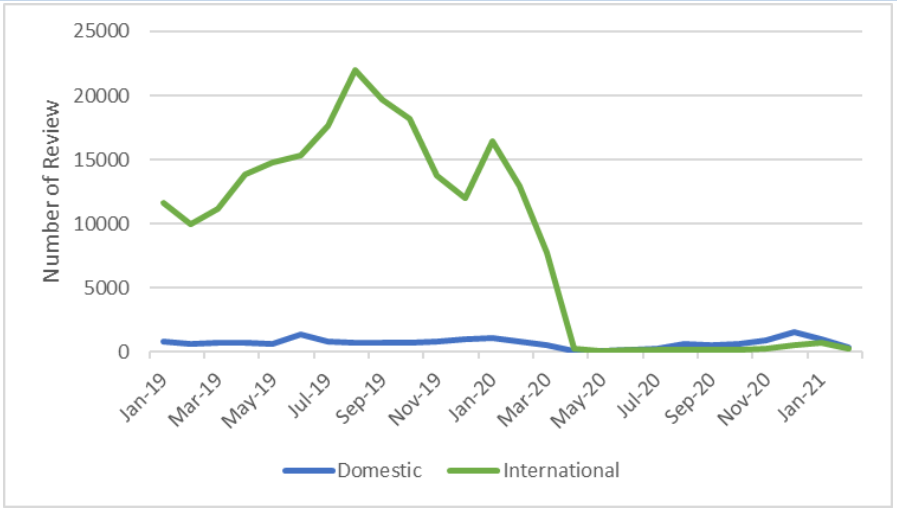


ROOM OCCUPANCY AND NUMBER OF REVIEWS

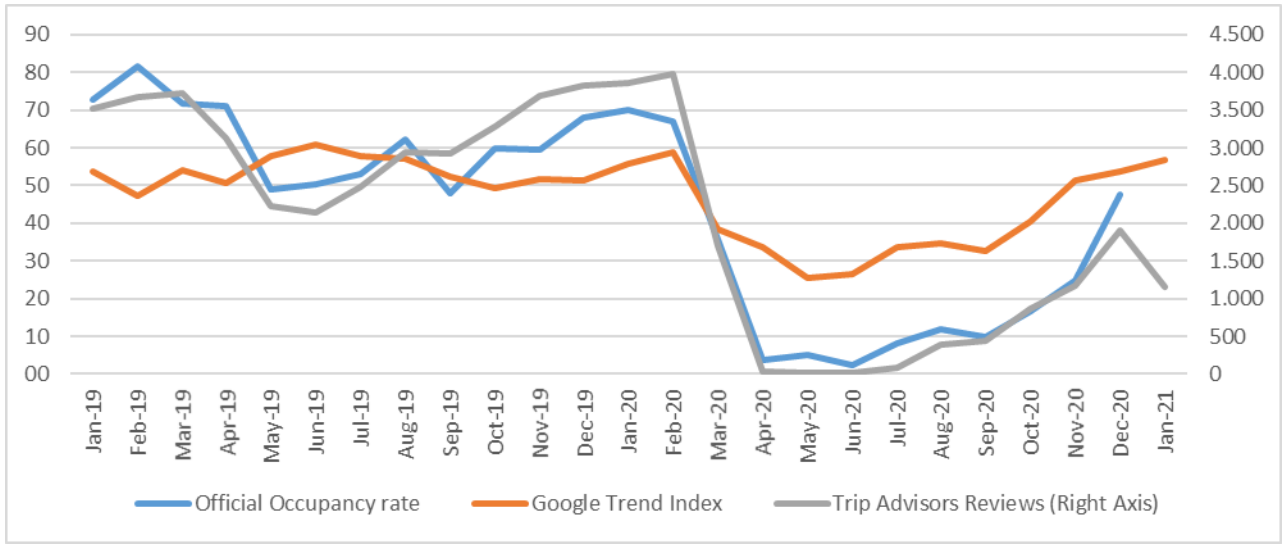
Hotel Room Occupancy Rate



Number of Reviews on Bali



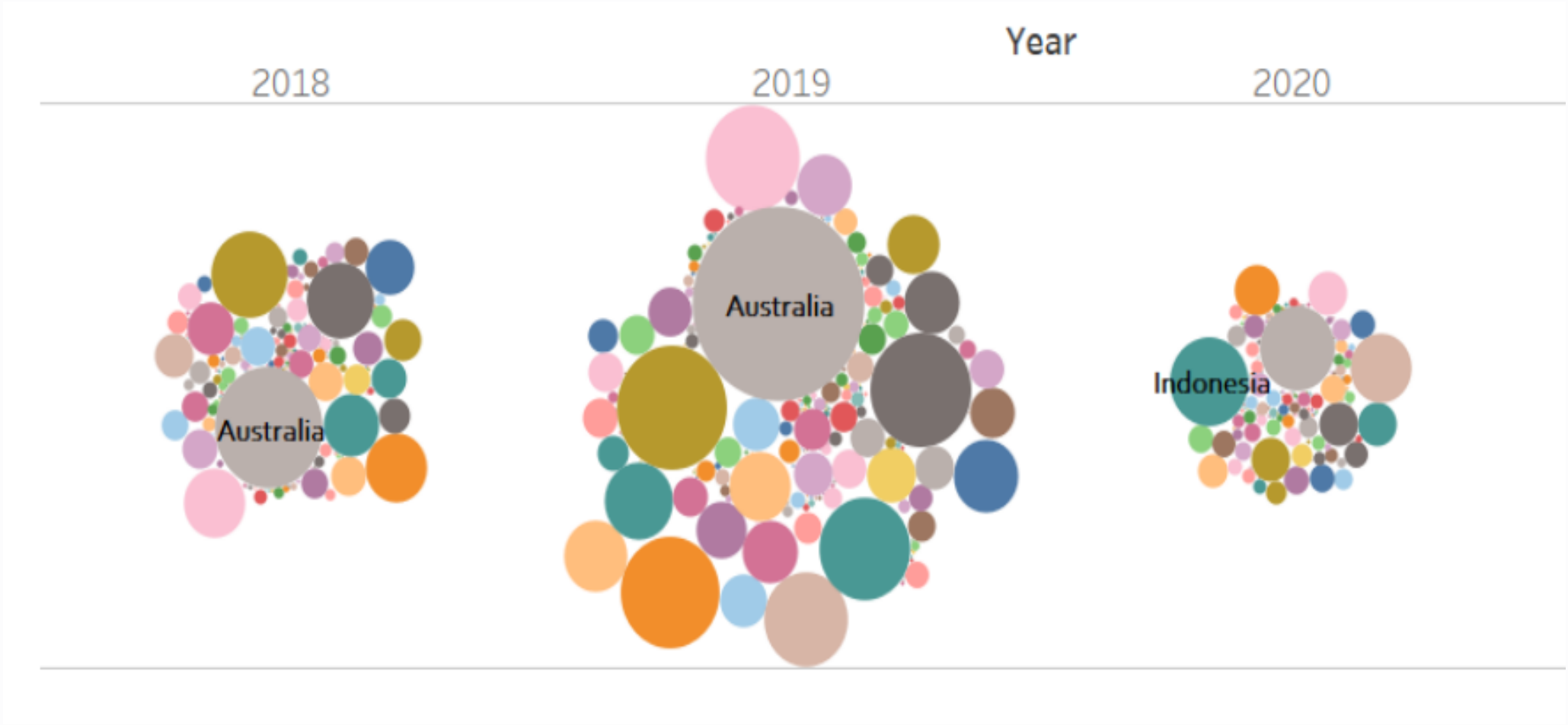
Hotel Room Occupancy Rate from Agoda, BPS and Google Trend in Bali





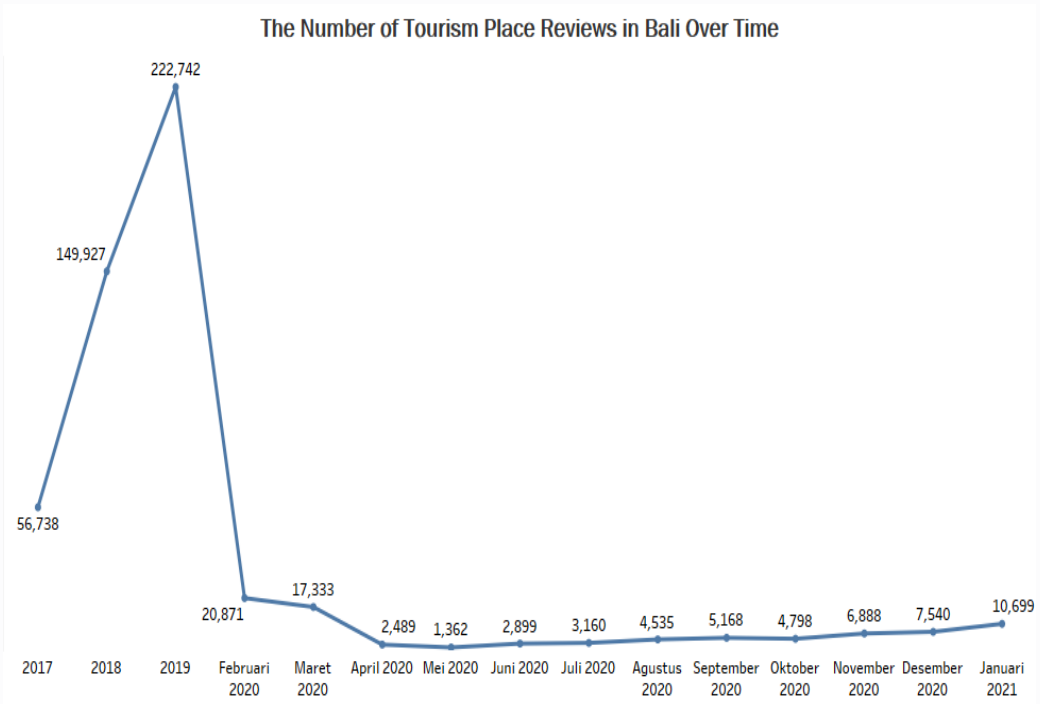
NUMBER OF REVIEWERS BY COUNTRY

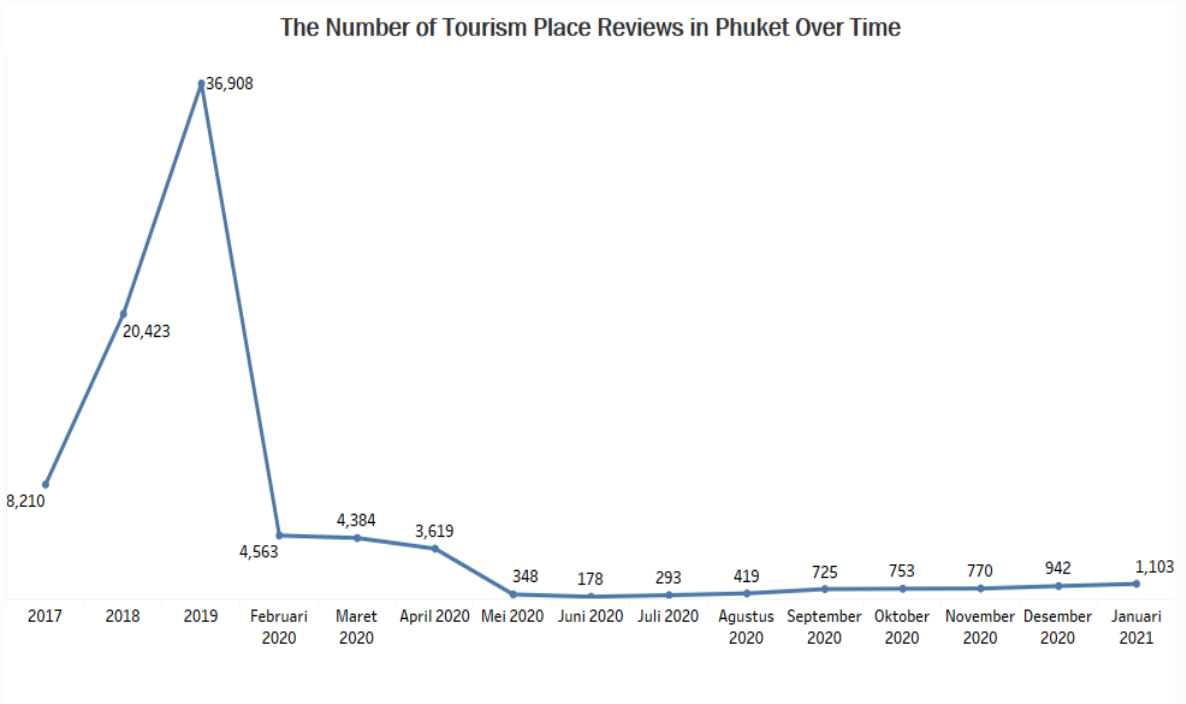
Reviewer Country (by Year)



RESULTS: GOOGLE MAPS

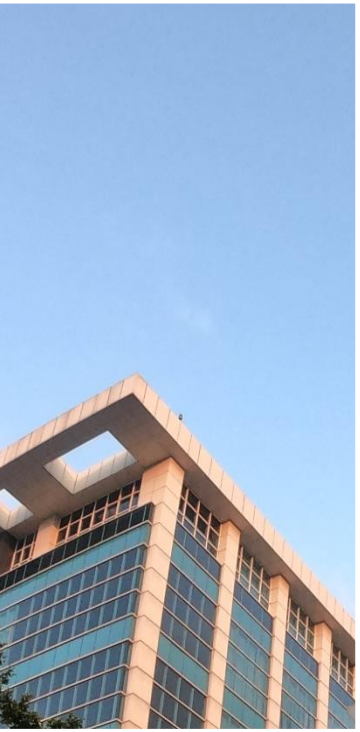
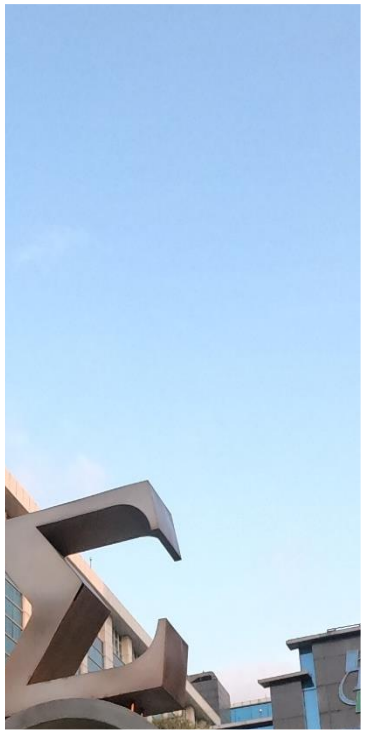
Number of Reviews of Tourist Places





Take Home Messages

- *Various Data Sources can be used as alternative for Tourism Statistics*
- *The variables, information, and data availability are different across the platforms and sources.*
- *Need data cleaning, aggregation and integration to combine them into a system of official statistics*
- *These alternatives data sources can be used as complement of the existing tourism statistics*





Thank You

Setia Pramana
Assoc. Professor Politeknik Statistika STIS,
Coordinator for Statistical Model Development
BPS-Statistics Indonesia
setia.pramana@stis.ac.id