The Use of Mobile Phone Data for Tourism and Commuting Statistics

2nd International Seminar on Big Data

Titi Kanti Lestari
Mobile Phone Data can be used as

Complement for other data sources
✓ Complement for immigration data (administrative data)
✓ Inbound tourism, when there is no immigration checkpoint, destination analysis
✓ Outbound tourism, to obtain country of destination and length of stay in each country

Replacement for other data sources
✓ Domestic tourism, replacement of household survey
✓ Commuting, replacement of household survey
✓ Inbound tourism, replacement of shuttle trade (cross border) survey
✓ Event analysis, replacement of survey or ticket sales
Use Cases in Tourism and Commuting Statistics

✓ Inbound Tourists (number of tourists, length of stay, place visited)
✓ Outbound Tourists (number of tourists, length of stay, country visited)
✓ Domestic Tourists (number of tourists, length of stay, place visited, O-D matrix)
✓ Event Analysis (number of visitors, venue visited)
✓ Commuting (number of commutes, O-D matrix)
How do we implement the tourism concept to mobile phone data

CLASSIFICATION OF INBOUND TRAVELLERS

- INBOUND TRAVELLERS
  - TOURIST (Overnight Visitors)
  - SAME-DAY VISITORS (Excursionists)

- INBOUND VISITORS
  - Nationals Residing Abroad
  - Other Non-Residents (Foreigners)

- OTHER INBOUND TRAVELLERS
  - Border Worker
  - Seasonal Workers
  - Other Short-term Workers
  - Long-term Workers
  - Nomads & Refugees
  - Transit Passengers Not entering The Economic & Legal Territory
  - Crews on Public Modes of Transport
  - Person Entering The Country to Establish There Their Country of Residence
  - Long-term Students & Patients and Their Family Joining Them
  - Other Travellers Deemed Not to Enter The Economic Territory:
    - Diplomats, Consular staff, Military Personnel & their Dependents
    - Armed Forces on Maneuver

International Recommendation for Tourism Statistics, UNWTO, 2008
What data is used

Signalling (probe)

☑ Capture more data (very big, especially for domestic tourism)
☑ Good for tourism statistics and commuting
☑ Add noises (statistical and non statistical)

Call Detail Record (CDR)

☑ Less data
☑ Possible under coverage, especially for inbound and outbound
Signalling vs CDR

Signalling contributes most in hard-to-reach areas

1. Islands
2. Border to less developed countries

Number of trips from CDR data

Number of trips from signalling data (multiplier ratio)

- More trips
- Even more trips from signalling data
- Signalling contributes more
- Signalling contributes less

- Malaka on the Timor Leste border
- Sangihe islands
- Boven Digoel near Papua
- Kupang on the Timor Leste border
- Talaud islands
- Anambas islands
- Natuna island

1. Islands
2. Border to less developed countries
Statistical and Non Statistical Noises

- Fast fliers
- Seamen
- Accidental Roamers
- Other transit

Methodology is important

- Filtering method
MPD for Official Tourism Statistics

Press Release

https://www.bps.go.id/pressrelease.html
MPD for Official Tourism Statistics

Publications

STATISTIK KUNJUNGAN WISATAWAN MANCANEGARA
TAHUN 2019
INTERNATIONAL VISITOR ARRIVALS STATISTICS 2019

LAPORAN SURVEI WISATAWAN NASIONAL
(OUTBOUND)
TAHUN 2019
For Official Statistics

1. Quality Assurance
   - In line with UN-QAF, Big Data QAF, NSO’s QAF

2. Sound Methodology
   - Various methodologies
   - Choose that reflect reality

3. Privacy-Preserving Processing
   - Privacy protected
   - Aggregate data
Quality Assurance

- In-line with BPS QAF Handbook (for Census, Survey and Administrative data)
- In-line with UN QAF and Unece QAF for Big Data
- Quality check (Input, Throughput, Output)

Input Quality Checking (First gate)
- Data gaps
- Missing data
- Incorrect timestamps
- Duplicate record

Throughput Quality Checking (Second gate)
- Errors in data processing
- Overwrites

Output Quality Checking (Third gate)
- Anomalies checking
- Coherence with other data
- New phenomena can be explained
- Passed Calibration/Comparison with other data
Some Quality Assurance Results

Natural hourly rhythm

 Logical daily present

Steady data flow

 Logical daily present
Usual Environment

- Outside usual environment, tourist
- Home-work, commute
- Changing home, over a year, internal migration
Privacy Protected through Pseudonymization and k-Anonymity

The subscribers is masked with hash, when data scientists processed

The data produce is aggregate data (tables)
Data Architecture of Sandboxing

For Mobile Positioning Data
MPD vs Conventional Survey Result (at Jakarta Greater Area)

MPD (2019)

Commuter Survey (2019)
Commuting Statistics

Commuters Before and After the Covid-19 Pandemic

Percentage of Commuters in MA - Jabodetabek

Percentage of Commuters in MA - Mebidang
Central City

Threshold 5 percent

Threshold 1.5 percent

Presidential Regulation No. 4 of 2018

MPD for MA Delineation on Cekungan Bandung
Contribution of Tourism in Indonesia, 2016 - 2019

The chart above illustrates the contribution of tourism in Indonesia from 2016 to 2019 for GVATI, TDGVA, and TDGDP. Here are the values for each year:

**GVATI**
- 2016: 7.10
- 2017: 7.08
- 2018: 7.00
- 2019: 7.15

**TDGVA**
- 2016: 4.63
- 2017: 4.67
- 2018: 4.90
- 2019: 4.97

**TDGDP**
- 2016: 4.65
- 2017: 4.67
- 2018: 4.91
- 2019: 4.97
MPD for Tourism Satellite Account

- SDG Goal 8 (Indicator 8.9.1)
- MPD give better coverage than household survey, better match with supply side
THANK YOU