

Asia-Pacific Economic Statistics Week
Seminar Component
Bangkok, 17-21 June 2019

Name of author

Amanda Pratama Putra, Ignatius Aditya Setyadi, Siim Esko, Titi Kanti Lestari

Organization

*BPS – Statistics Indonesia
Positium LBS*

Contact address

*BPS - Statistics Indonesia
Jl. Dr. Sutomo No. 6-8 Jakarta 10710 Indonesia*

Contact phone

+62878 6396 9334 (Putra),

Email

amanda.putra@bps.go.id, adityasetyadi@bps.go.id, siim.esko@positium.com, titi@bps.go.id

Title of Paper

Measuring Commuting Statistics in Indonesia Using Mobile Positioning Data

Abstract

Commuting patterns refers to the movement of workers from their homes (origin) to work areas located in different cities. Commuter movements generally take place in metropolitan cities as centers of economic activity. Accurate and up-to-date data about commuting patterns needs to be obtained continuously as a basis for determining policies related to rapid urbanization and sustainable urban transportation.

The Population Census and surveys are the traditional methods to used to commuter statistics. These methods could provide an amount of information concerning commuting flow and the characteristics of workers grouped by their origin and destination. Indeed, BPS-Statistics Indonesia has conducted commuter surveys in several major cities in Indonesia since 2014. However, implementation of the survey could not be carried out simultaneously in all areas due to the limited costs. As a result, the commuter statistics could not be captured comprehensively. In addition, commuter statistics data from the survey cannot be updated regularly in a short period of time because it is expensive; and requires a long time to collect the data from the field.

In order to address the weaknesses of traditional methods, this paper utilized the availability of mobile positioning data as a complementary technique. From our experiment, the results show that commuting patterns in Bandung City measured using mobile positioning data provide the same commuting pattern as the survey results. Therefore, it seems mobile positioning data could be used as an alternative approach to measuring commuting statistics in all major cities in Indonesia simultaneously.

I. Contents

I. Contents	2
II. Introduction	3
III. Data and Methodology	4
Mobile positioning data	4
Home and work identification	5
Commuter identification	6
Household commuter survey data	7
Calibration with population projection	7
IV. Result	9
Result Comparison: MPD vs Traditional Survey	9
V. Conclusion	12
VI. References	13

II. Introduction

For citizens who live in major cities in Indonesia, the main places of activity, such as work or school that are administratively different from places of residence, are a common phenomenon today. More facilities offered in the city center the high price of land and building (including its tax) are the main attraction for residents living in the surrounding (suburbs) area. In addition, the high cost of housing in the city center makes residents prefer living in satellite areas that are close to the city center. They travel back and forth from home almost every day to work. This routine then forms a travel pattern, namely commuting patterns.

Commuting patterns refers to the daily movement of people especially workers from their homes (origin) to work area. The existence of commuters has a positive impact on the city center, which can reduce the density of the population living in the cities and the growing development of satellite areas as a place for commuters. However, the existence of commuters is a big challenge for big cities to improve public facilities, especially in the transportation sector that directly affect citizens daily lives. This relate to the SDGs Indicator 11 (Make Cities and Human Settlements Inclusive, Safe, Resilient and Sustainable). Target 11.1, by 2030, ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums. Target 11.2: By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons.

Commuter movements generally take place in metropolitan cities as the center of economic activity. Since 2014, BPS – Statistics Indonesia has conducted the commuter survey in some metropolitan areas in Indonesia to produce statistics and monitoring systems that are able to provide input for the formulation of population mobility policies and the development of origin and commuter destinations. The conventional methods like survey could provide the amount of commuting flow information and the characteristics of workers grouped by their origin and destination. However, one of the survey-based drawbacks is it takes a long time from the design of a survey to the release of its final result. Besides, the expensive survey costs make the implementation of commuter surveys in Indonesia unable to be carried out simultaneously in all metropolitan areas.

The accurate and up-to-date data about commuting patterns is very challenging to be obtained continuously as a basis for determining policies related to rapid urbanization and sustainable urban transportation. Consequently, different approaches and data sources are needed to measure the commuting statistics more intensively and comprehensively. Traditional census and administrative data are solid sources for long-term processes. For the

short term and everyday mobility, more flexible methods such as crowdsourcing, satellite-based methods, or modern sensing technologies are needed.

To reduce the time lags and cost in producing the commuting statistics, there are a lot of attempts related to experimental work to use other data sources. Novak et al. [1] in 2013 explore the potential use of mobile phone location data to map the commuting patterns in Estonia. They found that mobile phone location data represent a significant alternative source in cases where traditional data are missing or are not appropriate or the method of collection is not cost effective. It implies that the commuting behaviors from big data (mobile positioning data) can allow the public sector to produce more frequent and more detail statistics.

As part of innovation and exploiting the new sources of data, we are currently working with alternative data sources for commuting statistics using mobile positioning data (MPD). MPD is the pinpoint location of mobile network subscribers which extracted from the location of exposed base transceiver station (BTS). The data is obtained from Telkomsel, one of the largest mobile network operator (MNO) in Indonesia. Since 2017 BPS has begun to collaborate with Telkomsel to explore MPD as an alternative and complementary data source for some official statistics production. The data obtained is based on contract for a year. While for the methodology for commuting we got help from Positium LBS. So, the methodology that we use is similar as in Novak et. al paper. This paper aims to describe the result of the experience using MPD to detect the commuting pattern, that will be used as the basis for measuring commuting statistics. Finally, we evaluate the result with official commuting statistics based on the survey result.

The paper is organized as follows. Section 3, describes the data and methodology. Section 4, provides the results. The last section is the conclusion.

III. Data and Methodology

Mobile positioning data

Mobile positioning data in this paper refers to the large-scale location data of subscribers of MNOs that are processed and stored in operator systems (Altin et al. 2015). The location of the mobile devices can be obtained in real-time or historically. The data is highly sensitive but also very valuable. It could be used anonymously and aggregated in the high-level aggregation to ensure customer privacy. MNOs store huge amounts of data for their business purposes including “location events” (e.g. network infrastructure quality monitoring (signaling quality), operational activity data, billing data, data for marketing and CRM, handover, Location Area updates, active positioning requests, etc.) (Tiru, 2014). In Indonesia, call detail records (CDR) and signaling data are used since Telkomsel is collecting, storing, and commercializing it. The data period used for measuring the commuting pattern are during

2018. Both CDR and signaling data have time and location information (a detailed latitude and longitude coordinate for cell towers) for every detected event. Telkomsel processes the CDR and signaling data into staypoint based on Mobility Genome Framework. Staypoint is data points that have been grouped to the constraint of average distance and the total duration of stay. according to Dang, et. al. (2017), “a stay point is constructed if the average distance from its centroid to all other points is less than Δd and total duration spent is more than Δt ”.

Owing to the privacy policy, there are several restrictions in the cooperation agreement with Telkomsel including data access rights. Telkomsel applies a sandbox mechanism approach in the process of developing algorithms and accessing data to protect data confidentiality. The data provided is staypoint (processed by Telkomsel) due to access to raw data is restricted. Then the algorithm is set to work on staypoint data. Because of the limitation, we were unable to estimate in the low administrative level (Desa/Kelurahan or even Kecamatan), so we decided to use the Kabupaten level which this level also adequate to compare with official statistics.

Home and work identification

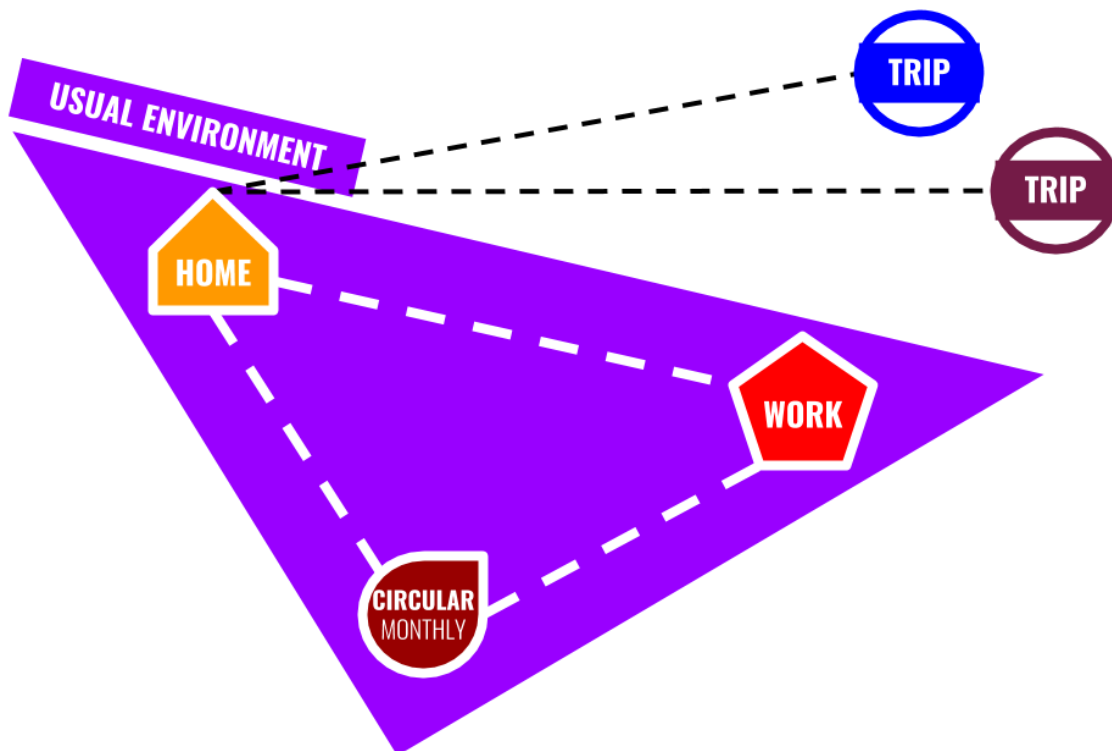


Figure 1. Usual Environment (Saluveer, E)

Implementing home and work to mobile positioning data is quite a tricky part. Therefore, BPS-Statistics Indonesia asked about 63 employees in all over Indonesia if they like to be volunteers. BPS uses the MPD of volunteers to develop an algorithm of usual

environment identification, especially home and work and then validate the result for each individual with their actual home and work location at the municipal level. At first, because the home and work location for each volunteer is estimated based on one past month data only, the results become inaccurate (only 58,92% accurate). Then, home and work location are estimated based on past 6 months data which showed more accurate results (up to 90% accurate) compared to home and work locations that were estimated based on 1-month data.

The long-term home and work location decided for the past 6 months and for every month select highest ranking home and work (based on duration). Then sum durations of home and work candidates by grouping them by Kabupaten (municipality) and Kabupaten with the highest duration will be the actual location for home/work.

Commuter identification

In this paper, commuters are defined as subscribers who live and work in different municipal and make a trip from home to work and return home within 24 hours routinely. Currently, it is determined that for the observation month, a subscriber will be classified as a commuter when he has at least 4 days during that month he has made a trip that follows the aforementioned rule. These 4 days also have to be registered on 4 different weeks, meaning that if a subscriber travels between home and work municipal on daily basis for the first two weeks and then the next two weeks are on a vacation, they will not be counted as a commuter.

Commuters with very distant locations (no direct transportation connections) are removed (outliers). For example, if we have just couple of subscribers commuting from Jakarta to distant kabupaten X, then those commuters will be removed. Commuters with non-standard work patterns, such as night workers might not be identified by this approach. Additionally, commuters who are on holiday, ill or otherwise absent from work in some specific month may not be identified because we process the algorithm monthly. Figure 2 describe the Commuters algorithm proposed by BPS and implemented by Telkomsel.

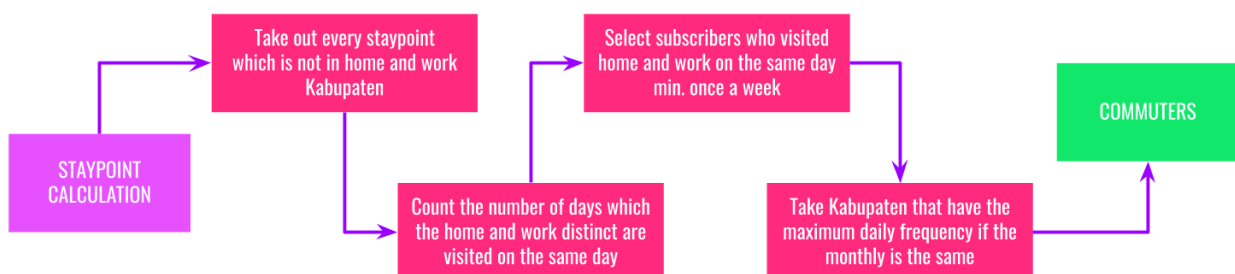


Figure 2. Commuters algorithm

Household commuter survey data

The latest commuter survey conducted by BPS-Statistics Indonesia was held in 2017, with domain estimation covering two different areas, the (1) Bandung Raya and (2) Gerbangkertosusila.

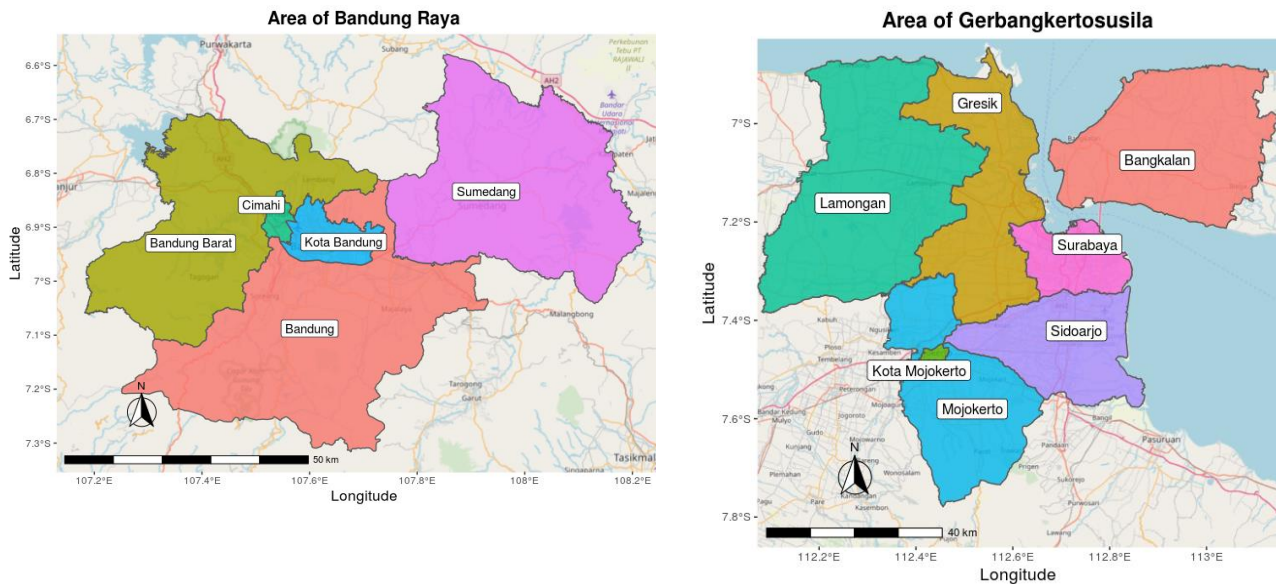


Figure 3. Area of Bandung Raya and Gerbangkertosusila

This survey aims to understand the behavior of commuters in 5 cities or regencies in Bandung Raya (Kabupaten Bandung, Kabupaten Sumedang, Kabupaten Bandung Barat, Kota Bandung, and Kota Cimahi) and 7 cities or regencies in Gerbangkertosusila (Kabupaten Gresik, Kabupaten Bangkalan, Kabupaten Mojokerto, Kota Mojokerto, Kota Surabaya, Kabupaten Sidoarjo, and Kabupaten Lamongan). The commuter survey is a household survey covering 10.600 households from 1.060 enumeration areas. The time reference used in this survey is June 2017. The survey provides the estimation of flows in the areas and also provides the demographic characteristics of the commuters as well.

Calibration with population projection

BPS has the duty to produce the latest population data and information, and its estimates in the future as a basis for determining development policies. Population projection is a scientific calculation based on the assumptions of the components of the population growth rate, namely birth, death, and displacement (migration). These three components will determine the number and age structure of the population in the future. The population includes all residents in the geographical area of the Republic of Indonesia who has stayed

for six months or longer and those who intended to stay even though their length of stay was less than six months.

To validate our usual environment algorithm, especially home identification in aggregated level, we compare the number of subscribers in a municipal area with the population projection for all 514 regions. MPD's home was only obtained from one MNO (only Telkomsel), then we have to calibrate those home with the population projection data in each municipality, so that can reasonably be compared with the statistics result from the survey. The result is shown in Table 1. For example, the calibration ratio for Kota Bandung is 1.44, Kota Surabaya is 1.22, and Kabupaten Sumedang with 1.04.

Table 1. Calibration rate between Home from MPD and Population Projection

No	BPS Population Projection	BPS Population Projection	Home for Distinct Telkomsel Number	Ratio Population Projection and Home
(1)	(2)	(3)	(4)	(5)
1	Bandung	3,697,346	1,510,948	2.44
2	Sumedang	1,148,766	657,616	1.74
3	Bandung Barat	1,677,985	518,348	3.24
4	Kota Bandung	2,501,823	1,736,766	1.44
5	Kota Cimahi	605,575	372,089	1.62
6	Gresik	1,294,343	652,830	1.98
7	Bangkalan	976,223	329,405	2.96
8	Mojokerto	1,105,643	323,267	3.42
9	Kota Mojokerto	127,948	57,635	2.22
10	Kota Surabaya	2,881,942	2,350,916	1.22
11	Sidoarjo	2,205,717	1,213,393	1.82
12	Lamongan	1,188,773	437,846	2.72

IV. Result

Result Comparison: MPD vs Traditional Survey

For this study, we only use the number of commuters who have a home location and work location within the area that we just mentioned. So, the number of commuters used in this study (100%) does not take the number of commuters who have work locations outside the area mentioned. Furthermore, in table 1 and table 2, the number of commuters is estimated according to the location of the home and work respectively. In the table, two estimation numbers are obtained from both, MPD and traditional surveys. the numbers in brackets and thickened are estimated numbers from MPD. The commuting patterns in Bandung Raya visually presented in the figure 4.

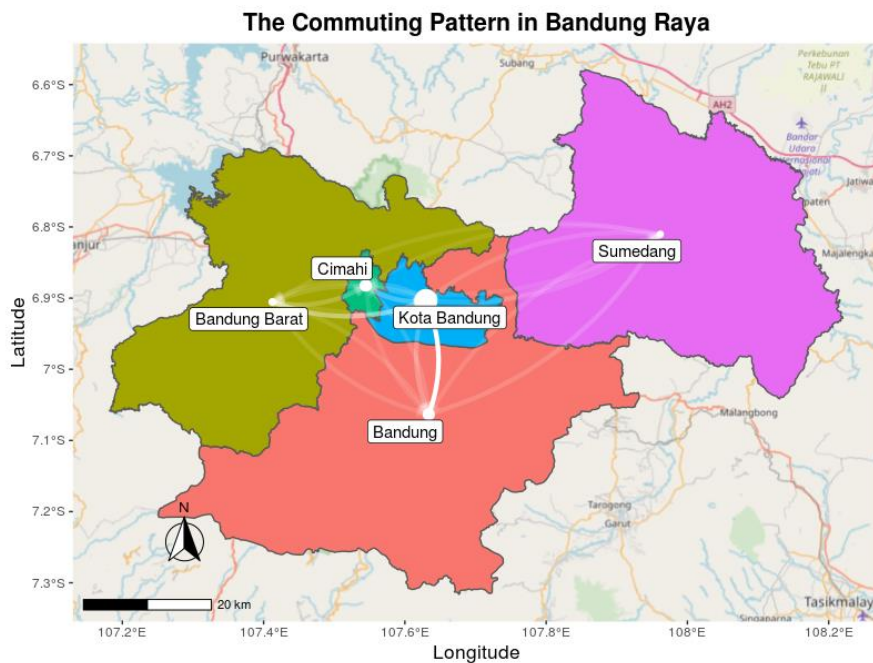


Figure 4. The Commuting Pattern in Bandung Raya

Table 2 is the origin destination matrix (ODM) for commuters in Bandung Raya to shows the result comparison of MPD and traditional survey. For instance, in the total estimated number of commuters in Bandung Raya, the traditional survey estimate 36.05% of people commute from Kabupaten Bandung to Kota Bandung, while MPD estimate for 43.22%. The result also shows that Kota Bandung is the most favorite destination for commuters in every Bandung Raya origin, except commuters from Kabupaten Sumedang that only 2.88% commute to Kota Bandung.

Table 2. Origin Destination Matrix (ODM) of Commuters in Bandung Raya

Home (Origin)	Work (Destination)				
	Bandung	Bandung Barat	Kota Bandung	Kota Cimahi	Sumedang
Bandung	-	0.89% (1.21%)	36.05% (43.22%)	3.31% (1.99%)	5.72% (3.41%)
Bandung Barat	1.12% (1.53%)	-	10.00% (15.92%)	8.98% (6.20%)	0.00% (0.23%)
Kota Bandung	6.53% (6.55%)	1.78% (1.74%)	-	4.48% (2.56%)	0.56% (0.87%)
Kota Cimahi	0.89% (0.80%)	1.88% (1.66%)	10.61% (8.31%)	-	0.12% (0.06%)
Sumedang	4.11% (1.83%)	0.00% (0.09%)	2.88% (1.78%)	0.09% (0.04%)	-

Table 3 is the ODM of commuters in Gerbangkertosusila. Following the result, both MPD and traditional survey estimate Kota Surabaya as the top destination for commuters. On the other hand, Kota Mojokerto that also labeled as a city like Kota Surabaya is only favorite for the commuters from Kabupaten Mojokerto.

Table 3. Origin Destination Matrix (ODM) of Commuters in Gerbangkertosusila

Home (Origin)	Work (Destination)						
	Gresik	Bangkalan	Mojokerto	Kota Mojokerto	Kota Surabaya	Sidoarjo	Lamongan
Gresik	-	0.00% (0.20%)	0.83% (0.31%)	0.00% (0.02%)	10.89% (13.28%)	2.17% (1.35%)	2.27% (2.05%)
Bangkalan	0.00% (0.30%)	-	0.00% (0.00%)	0.00% (0.00%)	1.23% (2.58%)	0.00% (0.08%)	0.00% (0.00%)
Mojokerto	1.62% (1.12%)	0.10% (0.02%)	-	8.65% (3.34%)	2.34% (3.01%)	4.99% (3.25%)	0.29% (0.54%)
Kota Mojokerto	0.05% (0.07%)	0.00% (0.00%)	2.09% (1.04%)	-	0.42% (0.40%)	0.76% (0.35%)	0.00% (0.01%)
Kota Surabaya	5.07% (4.41%)	0.32% (0.52%)	0.37% (0.39%)	0.00% (0.08%)	-	9.86% (8.79%)	0.00% (0.21%)
Sidoarjo	5.25% (2.43%)	0.00% (0.12%)	3.89% (1.44%)	0.65% (0.27%)	29.00% (42.27%)	-	0.00% (0.11%)

Lamongan	4.77% (3.48%)	0.00% (0.01%)	0.00% (0.25%)	0.00% (0.02%)	2.03% (1.73%)	0.09% (0.15%)	-
-----------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	---

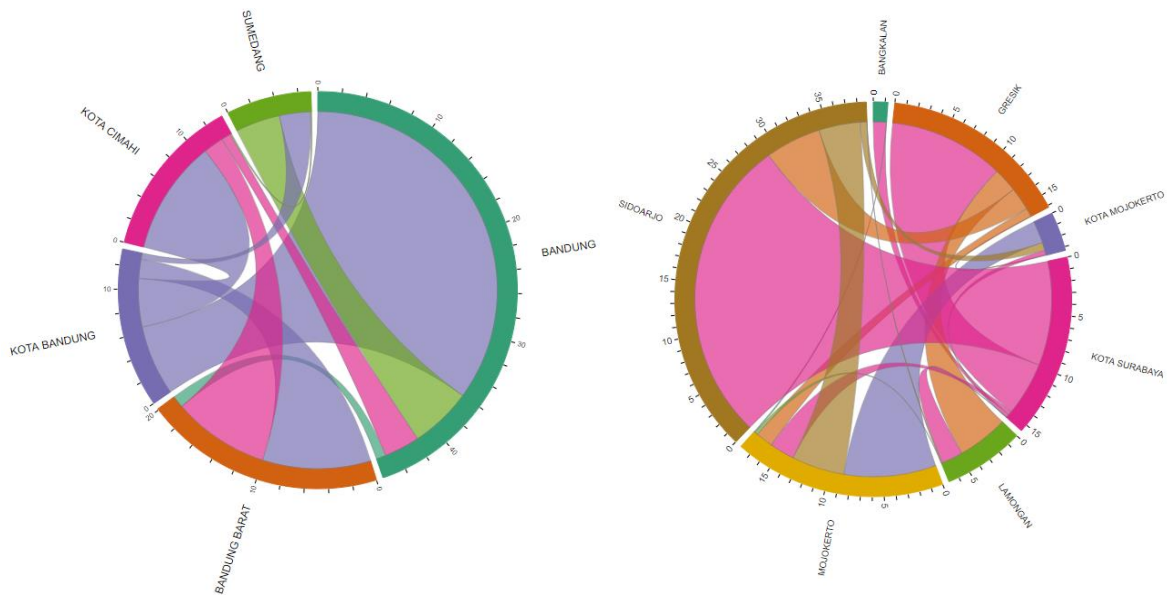


Figure 5 Commuting Pattern from Traditional Survey

The cities with more facilities offered like Kota Bandung and Kota Surabaya are the main places for commuters to work and make a living. It makes sense because the two cities are the center of provincial government and they are located right in the middle of their area. Figure 5 shows the commuting pattern from the traditional survey and figure 6 shows the commuting pattern from MPD. In the diagrams, each flow visualizes a specific flow between origin to destination. It describes clearly that these two different methods relatively produce the same flow.

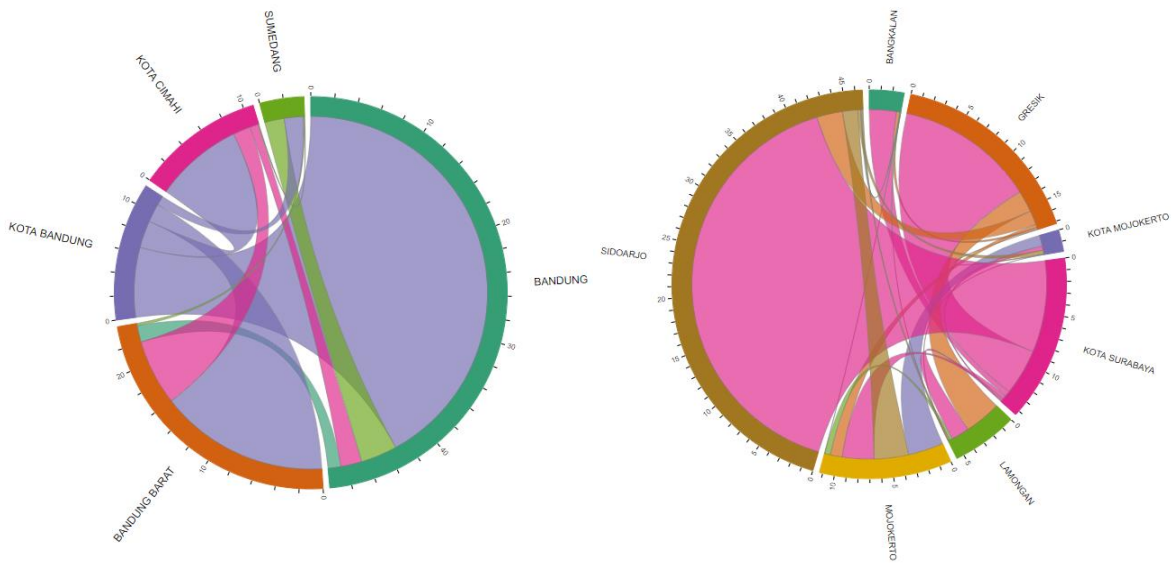


Figure 6 Commuting Pattern from MPD

Mathematically, we provide the calculation of Pearson Correlation Coefficient (PCC) indicates that there is a high positive association between the result from official statistics and MPD. The PCC for the estimation in Bandung Raya is 0.98 (strong correlation) and for Gerbangkertosusila is 0.96 (strong correlation).

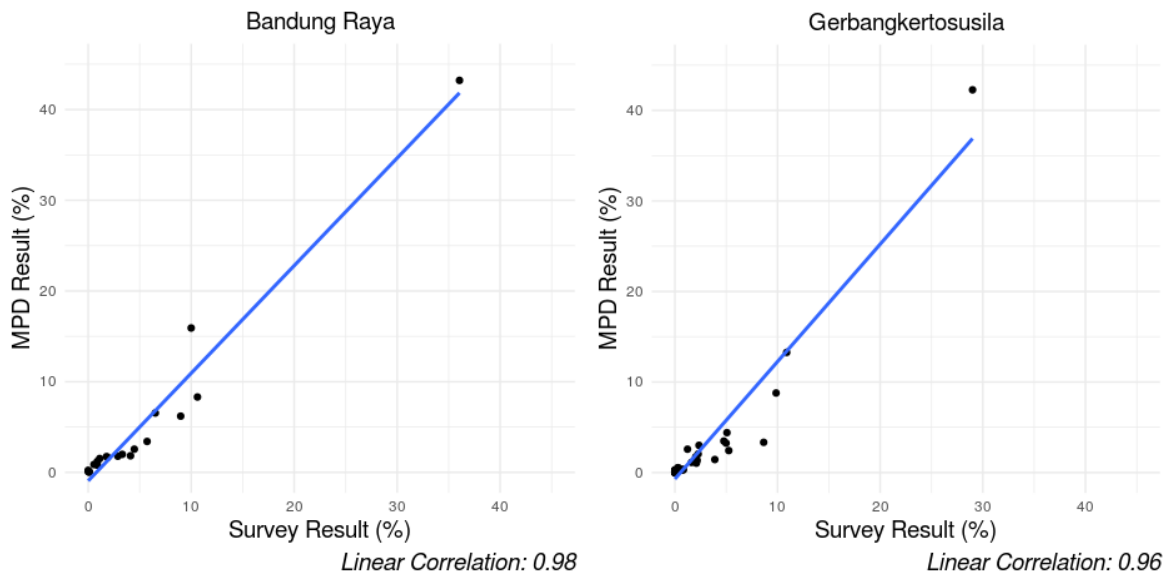


Figure 7 Linear correlation

V. Conclusion

The use of mobile positioning data changes the process of design, build, data collection, data processing, and dissemination. Data collection will be more efficient in terms of cost and time. Near real-time dissemination can be achieved by using mobile positioning data. Big data is a part of the data revolution and needs to develop as a validated source for statistics.

The results showed that mobile positioning data has the potential to increase the level of disaggregation, timeliness, and usefulness for commuting analysis. The method for calculating commuters is using mobile positioning data and calibrating it with population projection data and also comparing with Commuter Household Survey for benchmarking. The commuter household survey can only be conducted every two years and only covered some municipalities. So, the data frequency is quite low and the data quality of household survey is depending on respondent's answer, while Mobile Positioning Data can provide commuter data from all the municipalities (514 Kabupaten in Indonesia), more frequent data and also it depends on methodology that can be improved over time and data recalculated as the

methodology improving. According to the limitation of our current methodology for MPD that only get data source from single mobile networks and not be able to detect commuters with non-standard work patterns, we believe that our approach is still under-coverage even after we calibrate the result with the population projection.

VI. References

- [1]. Ahas, Rein. Siiri Silm, Erki Saluveer, and Olle Jaarv, (2009). Modelling Home and Work Locations of Population Using Passive Mobile Positioning Data, <https://www.researchgate.net/publication/227086074>.
- [2] Ahas, Rein. Siiri Silm, Olle Jaarv, Erki Saluveer, and Margus Tiru (2013). Using Mobile Positioning Data to Model Locations Meaningful to User of Mobile Phones, *Journal of Urban Technology*, 17:1, 3-27.
- [3] Altin, Laura. Tiru, Margus. Saluveer, Erki. Puura, Anniki. (2015). Using Passive Mobile Positioning Data in Tourism and Population Statistics, the New Techniques and Technologies for Statistics 2015 proceedings. Brussel: Eurostat.
- [4] Anh Dang, The & Deepak, Jayakumaran & Wang, Jingxuan & Luo, Shixin & Jin, Yunye & Ng, Yibin & Lim, Aloysius & Li, Ying. (2017). Mobility Genome™- A Framework for Mobility Intelligence from Large-Scale Spatio-Temporal Data. 449-458. 10.1109/DSAA.2017.20.
- [5] Novak, Jakub. Ahas, Rein. Aasa, Anto. and Silm, Siiri. (2013). Application of mobile phone location data in mapping of commuting patterns and functional regionalization: a pilot study of Estonia. *Journal of Maps*, 9:1, 10:15, DOI: 10.1080/17445647.2012.762331.
- [6] Raun, Janika and Rein Ahas. (2016). Defining usual environment with mobile tracking data , *the 14th Global Forum on Tourism Statistics, Venice, Italy*.
- [7] Tiru, Margus. (2014). Overview of the Sources and Challenges of Mobile Positioning Data for Statistics, presented in the International Conference on Big Data for Official Statistics. Beijing: UN Trade Statistics, Positium, and University of Tartu.