



# USING BIG DATA ANALYTICS FOR IMPROVED PUBLIC TRANSPORT

**PARTNER: JAKARTA SMART CITY, JAKARTA PROVINCIAL GOVERNMENT**

**PROGRAMME AREA: REAL-TIME EVALUATION**

**TECHNOLOGY READINESS LEVEL: MINIMUM VIABLE PRODUCT (LEVEL 5)**

## SUMMARY

Pulse Lab Jakarta collaborated with Jakarta Smart City on a project to enhance transport planning and operational decision-making through real-time data analytics. Using data from TransJakarta – the city's rapid bus transit system – buses and passenger stations, the project mapped origin-destination trends and identified bottleneck locations, information which can be used to identify whether new routes are needed. The project also explored the possibility of using real-time data to determine passenger-waiting times in order to enhance the efficiency of the bus dispatching system. The collaboration contributed to improvements in TransJakarta's operations and enhanced capacity within Jakarta Smart City.

## BACKGROUND

Jakarta is well-known for its traffic jams, which, among other effects, prolong commutes and complicate the scheduling of public transportation. In an effort to improve public transportation services, the city government – through its Jakarta Smart City initiative – is looking to big data analytics for insights.

Established in December 2014, the Jakarta Smart City platform collects data through government, third parties and citizen crowdsourcing and provides information about the city to improve public service delivery. Applications such as Qlue, Waze and others deliver real-time information to the platform through citizen reporting.

The platform also includes data on the city's Bus Rapid Transit system, called TransJakarta, which was designed to provide Jakarta citizens with a fast public transportation system to help reduce commuting times and rush hour traffic. The buses use dedicated lanes and tickets are subsidized by the regional government. In 2016, 123.7 million passengers used the bus service (a 20% increase from 2015). Some of the challenges affecting TransJakarta include unpredictable service frequency, travel times and wait times.

In order to better understand these challenges, Pulse Lab Jakarta partnered with Jakarta Smart City on a project to analyse real-time TransJakarta data to help improve bus service planning and delivery. The project is part of efforts to support and advance the Government's data-driven policy and practice strategies.

## ANALYSING TRANSJAKARTA DATA TO IMPROVE SERVICE DELIVERY

The project focused on two aspects: using GPS bus data to map problematic locations, and using passenger tap-in data, or the moment a passenger starts a ride, to gain insights on passenger behaviour through origin-destination (OD) statistics and waiting times.

The TransJakarta fleet uses GPS devices that update bus information every five seconds, including geo-location, time and velocity data among other information. In order to identify problematic areas, the

project used the dataset to identify locations with bottlenecks and abnormal traffic behaviours and uncover where dedicated lanes are slowest.

The project also explored whether bus ticket tap-in data would reveal useful insights on passenger behaviour, which could improve TransJakarta planning and services. Based on this dataset, the project analysed origin-destination patterns, which included transaction and bus station ID, date and time of transaction, line number, card type and serial number. Although the tap-in system does not provide passenger details, the project used the existing data to determine average passenger waiting times.

## INSIGHTS AND OUTCOMES

The project used GPS data from April to June 2016 to analyse bus speeds, number of buses operating and bottleneck ratios in 12 dedicated corridors. This enabled Jakarta Smart City to map problematic lines and locations throughout the city in order to better detect anomalies and reduce traffic bottlenecks on the affected bus routes.

To gain insights on passenger bus use, the project explored tap-in data from May and June 2016 and selected a well-balanced one-week period to analyse weekly patterns. The analysis included 12 lines and 422,694 unique passengers, from which 202,933 passengers (48%) were included in the origin-destination pair analysis.

To infer origin-destination statistics, the project developed two different approaches. First, based on passenger transactions within a day, origin-destination was derived from two consecutive tap-ins: one at origin and the second at destination. Second, based on aggregated passenger transactions residence and workplace were derived from passenger behaviour indicated by their morning and evening tap-in locations.

Based on these methods, the project identified:

- Number of hours between consecutive trips (weekday, weekend);
- Three types of bus lines (peak time in morning/evening, peak time in morning, peak time in evening);
- Top destination lines; and
- Top origin-destination pairs at different levels: station, line, sub-district, city levels.

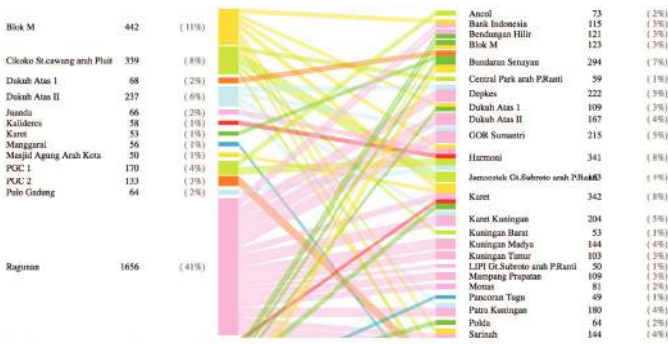


Image 1. Visualisation of TransJakarta's passengers points of origin and destination illustrating frequently congested locations.



Image 2. Mapping TransJakarta bus routes speed in problematic locations to improve services.

To determine which stations have the longest waiting times, the project used tap-in data and fleet GPS data based on number of buses (GPS) and number of passengers (tap-in). The methods and findings from the project were shared with TransJakarta to help improve route planning and regularity of service and inform plans to expand station facilities. The project highlighted key questions that TransJakarta could further research, such as exploring why some passengers only travel one-way instead of a round trip.

Based on feedback received from the project, the city administration has made a number of targeted changes and improvements to their operations, such as deploying more officers and barriers to secure dedicated lines in typically congested areas and adding buses on certain routes.

## CONCLUSIONS

The insights generated by this project have been used to improve transportation planning and demonstrate the potential for developing analytical models to enhance public service delivery.

The TransJakarta system generates a huge amount of data each day and these data sources can provide significant insights to inform improvements in services and increase overall ridership. The partnership with Jakarta Smart City provided an opportunity to develop internal government capacity while also demonstrating the benefits of data analytics for improved public services.

Development of the analytical model allowed Pulse Lab Jakarta to experiment with methods of inferring origin-destination statistics and passenger behaviour based on tap-in data. Jakarta Smart City found the OD pair analysis to be a useful way to understand passenger trip behaviour. As a result of the project, TransJakarta added the tap-out feature, which adds flexible fare pricing and enhances origin-destination analysis.

## IMPLICATIONS AND RECOMMENDATIONS

- In the future, similar projects would benefit from involving more partners in order to access expanded datasets. For example, when mapping problematic locations, GPS data from bus lanes could be combined with crowdsourced traffic data to compare regular traffic lanes with bus-restricted lanes.
- For more robust results, data analytics could be complemented by qualitative research. Though GPS data might pinpoint a problematic location, it can be challenging to identify possible causes and gain a complete picture of the issue without qualitative analysis.
- To validate findings, similar projects should consider incorporating passenger perspectives.
- Future models related to origin-destination statistical analysis could be strengthened by incorporating additional data, such as using CCTV information from video cameras, to assess bus and station capacity levels.

## HOW TO CITE THIS DOCUMENT:

UN Global Pulse, 'Using Big Data Analytics for Improved Public Transport,' Project Series, no. 25, 2017.