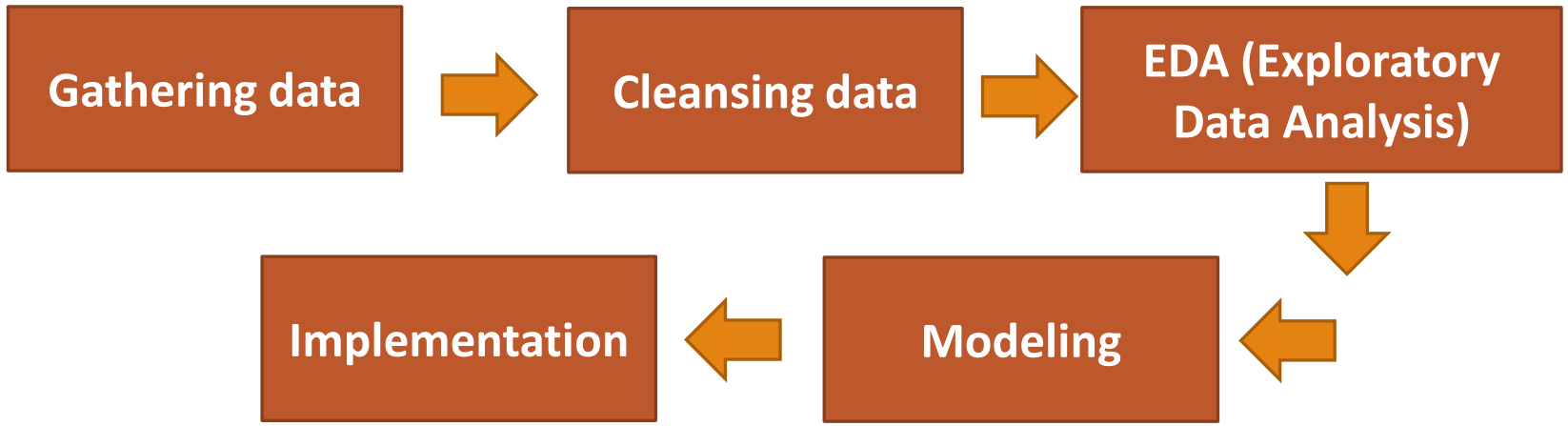
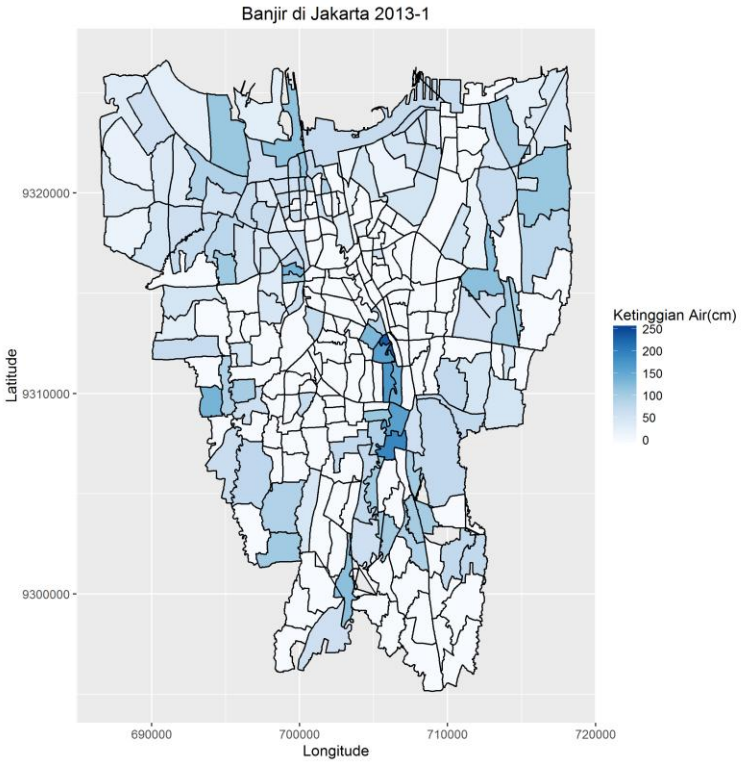
Flood in Jakarta

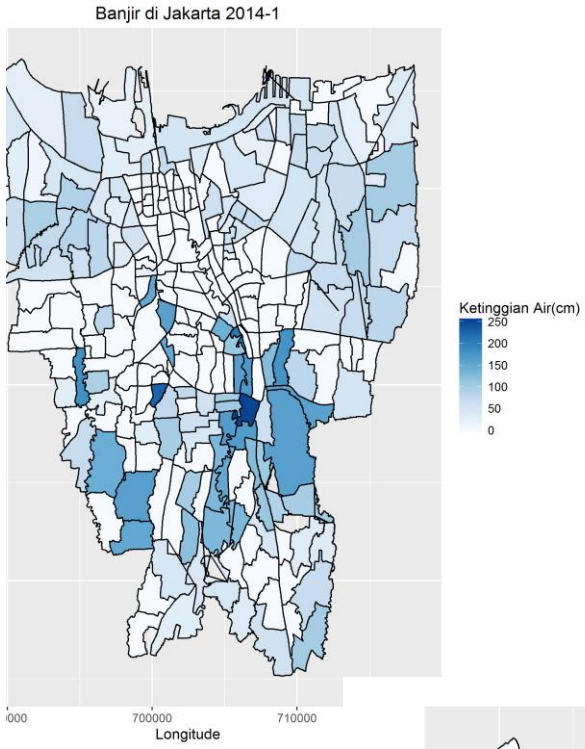
ADI WIJAYA SUCHIANA

Data Source & Workflow

- **Available data**
  - Hourly water level data (from http://bpbd.jakarta.go.id/waterlevel)
  - Flood data in Jakarta (2014-2016)
  - Average precipitation TRMM data(Januari 2014 – September 2016)
  - Flood reports from media social (Qlue, Twitter, Detik)
- **Workflow**





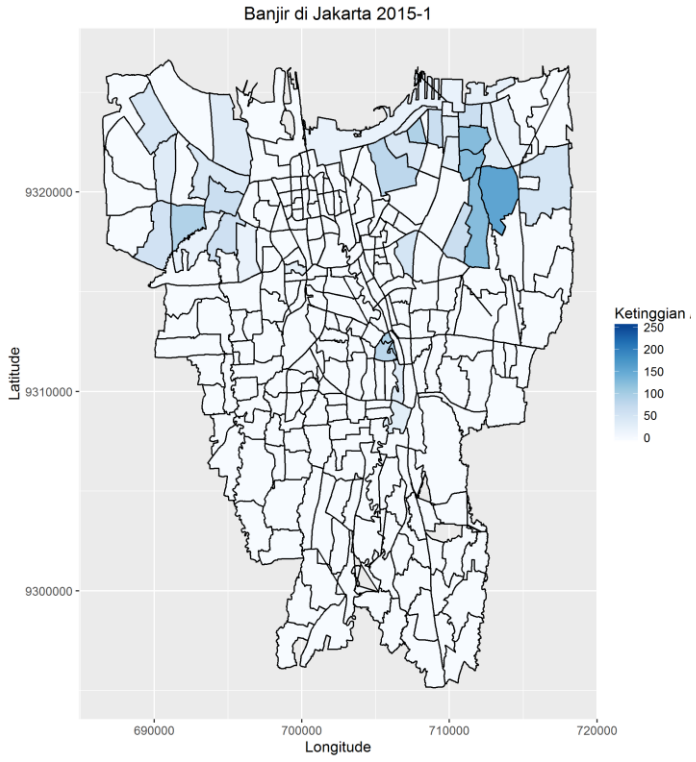


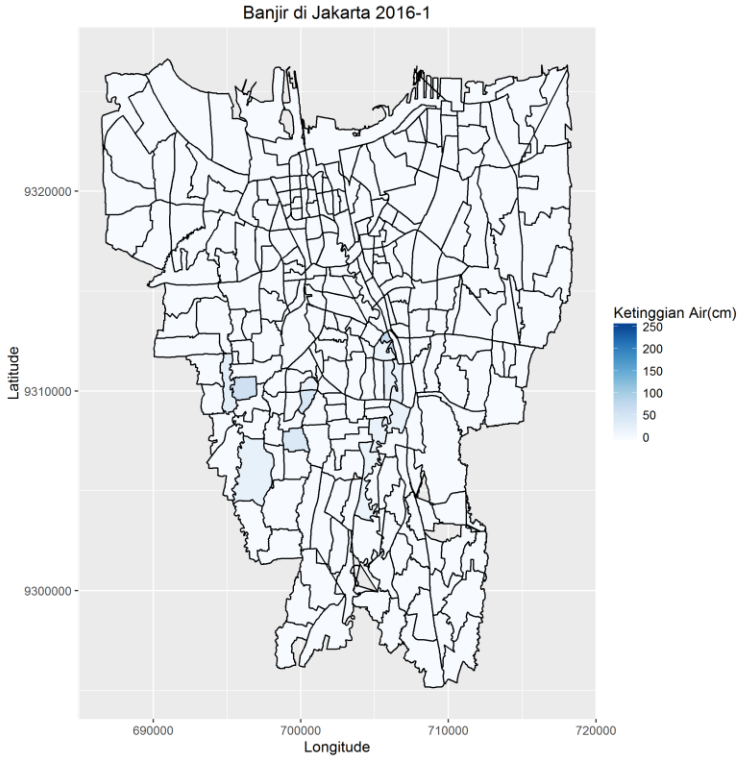
January

Exploratory Data Analysis.

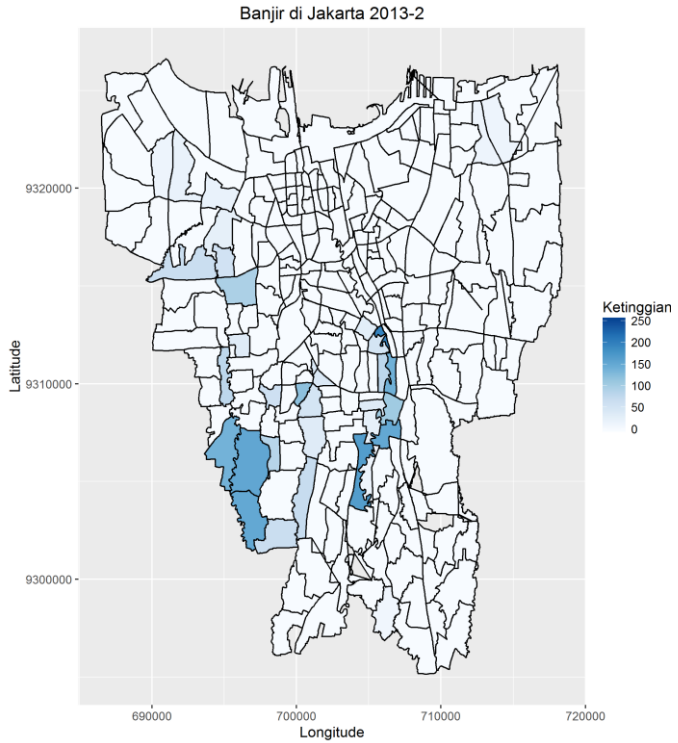
Jakarta flood map from 2014 – 2016 .

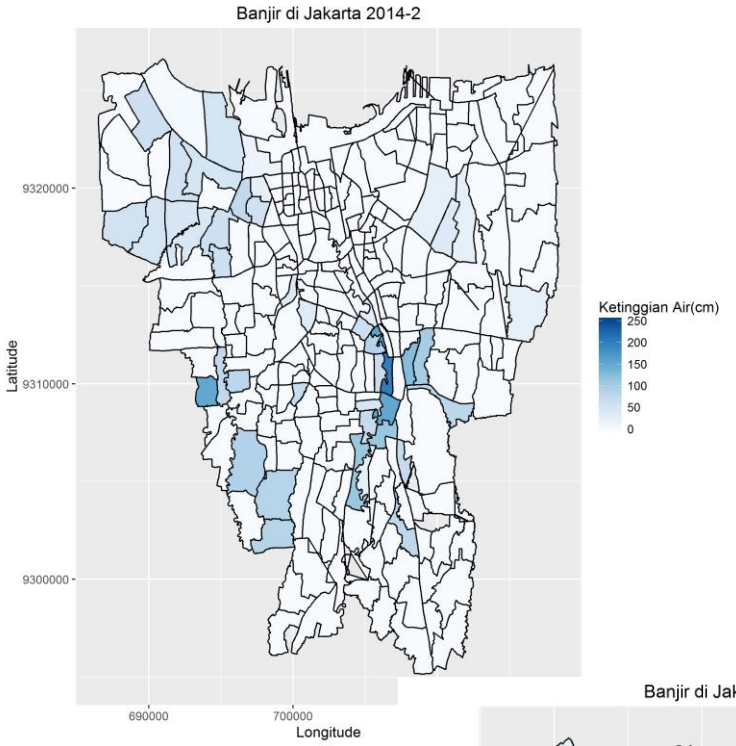
No	Kelurahan	Average.Ketinggian.Banjir(cm)
1	BIDARA CINA	160
2	KAMPUNG MELAYU	190
3	CIPINANG MUARA	125
4	CIPINANG BESAR SELATAN	100
5	CAWANG	225
...	...	...



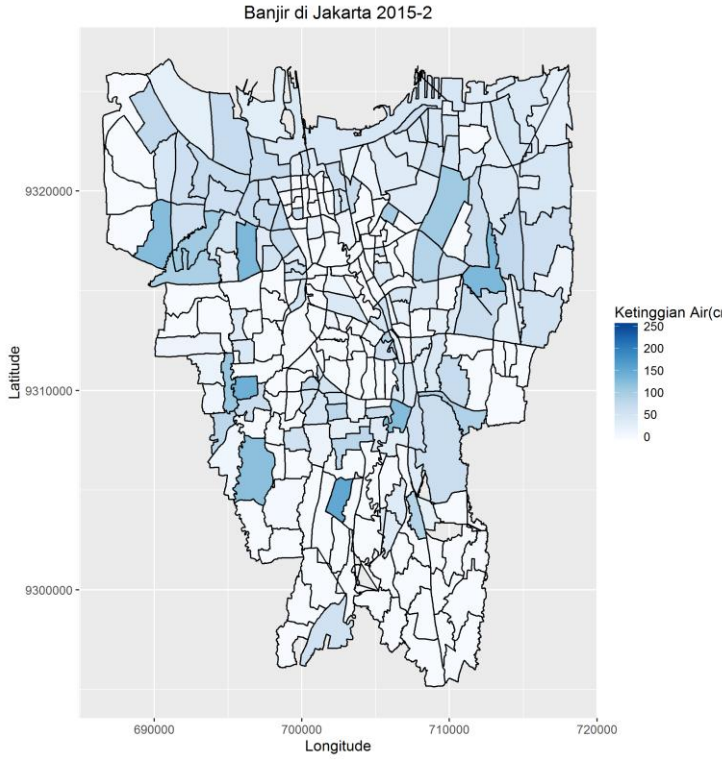


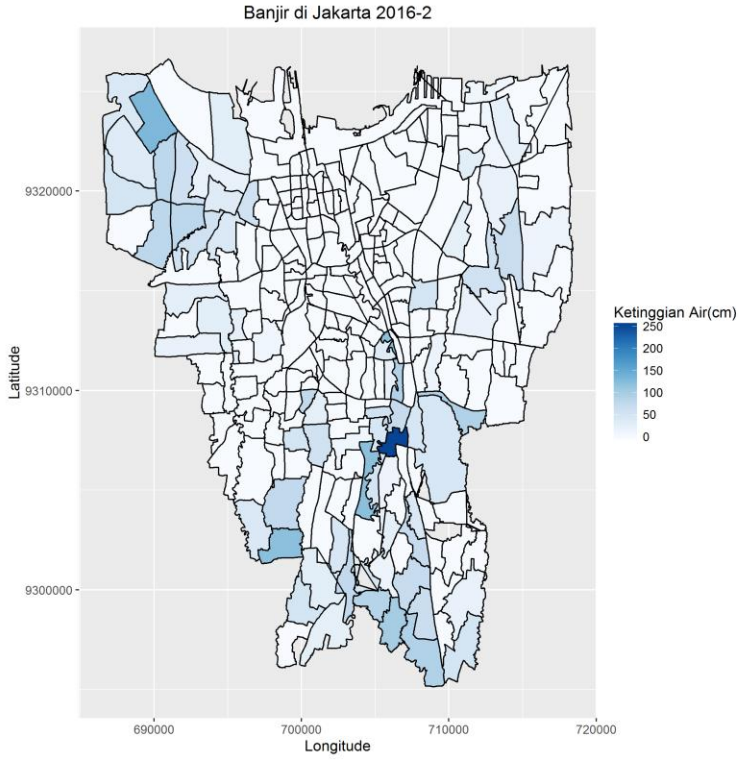
February





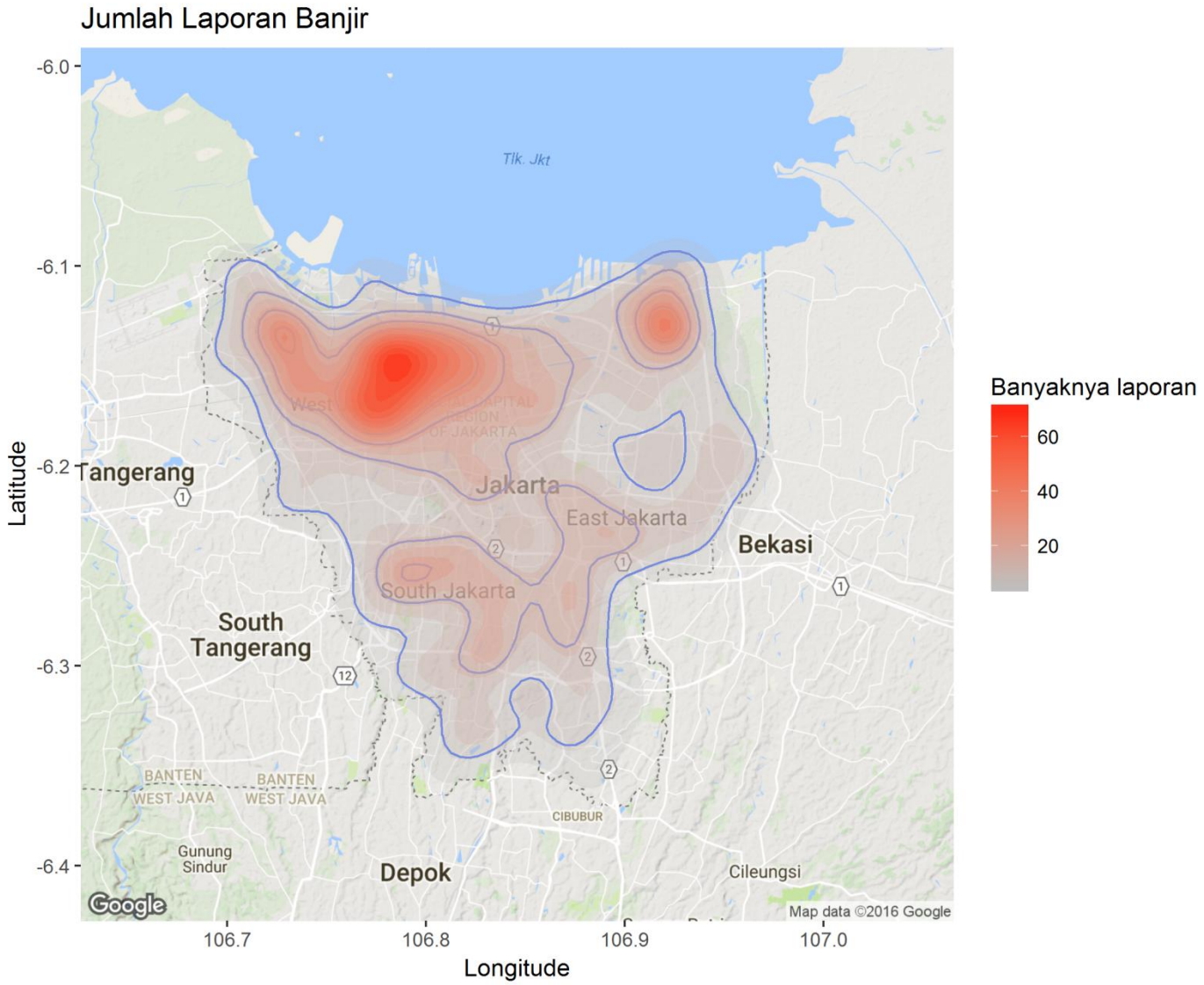
Exploratory Data Analysis.
Flood map shows that in 2013-2014 heavy flood occurred in January, while in 2015-2016 it was shifting to February.





Exploratory Data Analysis

We collected flood report data from social media (Qlue, Twitter, and Detik) to monitor flood area in 2016 as shown in heat map ‘Jumlah Laporan Banjir’. North and West Jakarta are the area with higher number of flood reports.

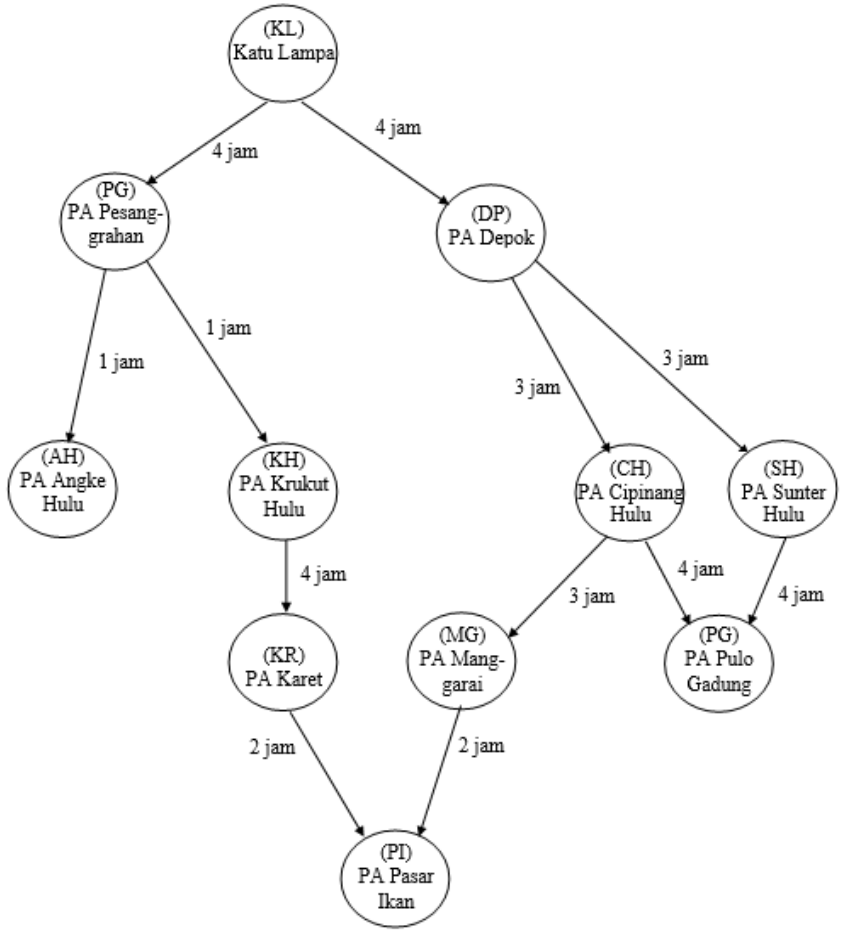


Water Level Model Jakarta

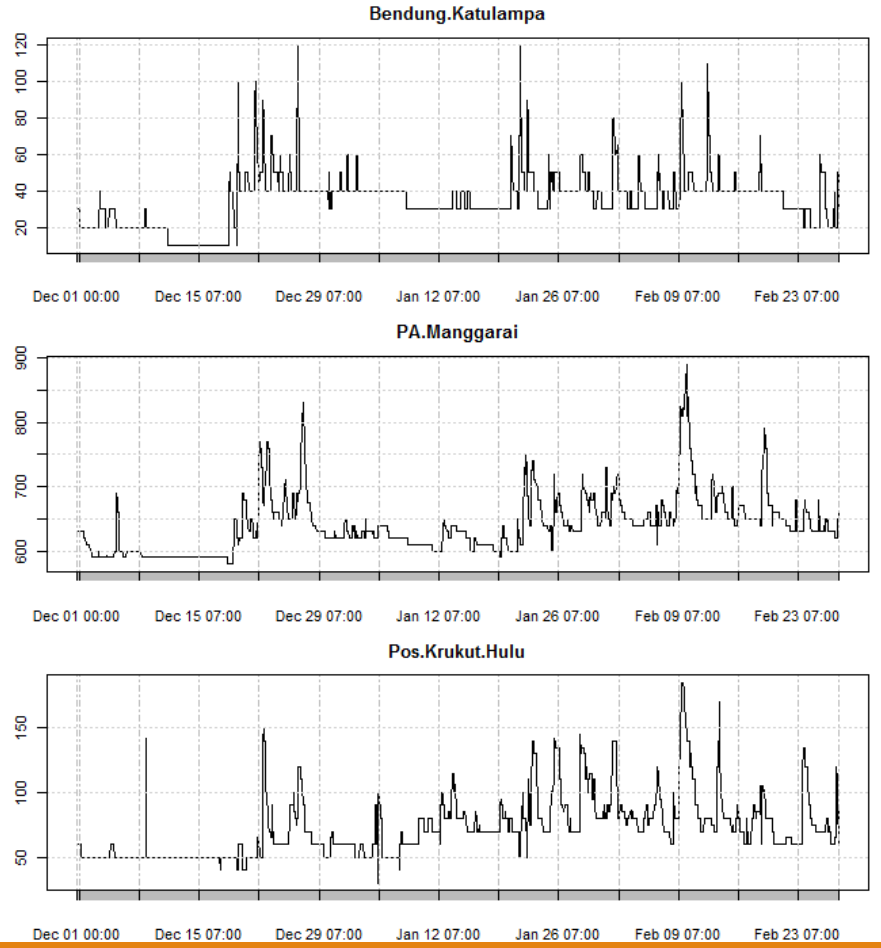
To predict flood in Jakarta, we are using indirect approach. Instead of directly creating flood model, we created model to predict water level in each river gauge across Jakarta (11 river gauges)

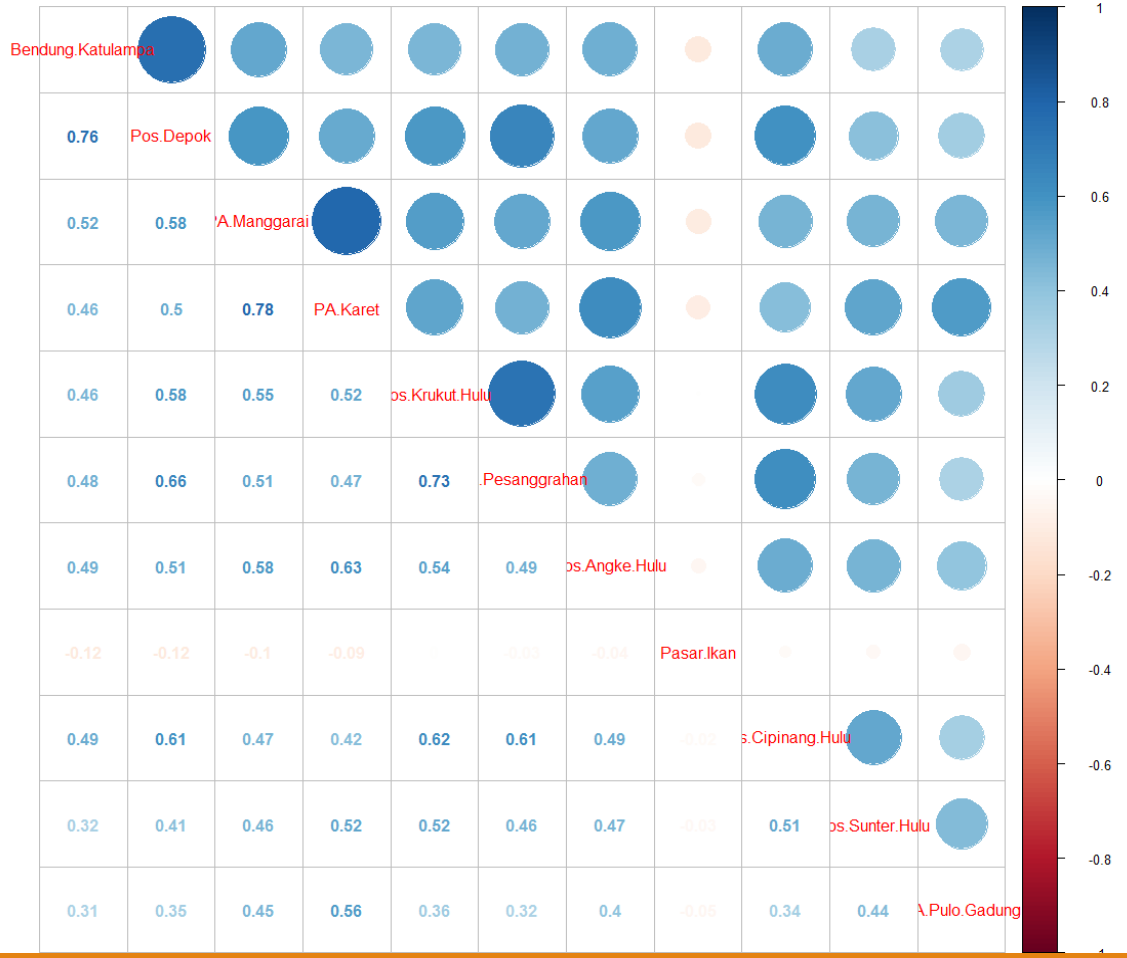
We can rely on this model because based on experiences. If we know the water level, we can determine flood area.

Data: Times series data recorded hourly from January 2014 – September 2016 with relationship shown in the diagram.



Water Level Data (Januari 2014 – September 2016)





Time series from 3 river gauges

Cross Correlation between river gauge time series

Implementation

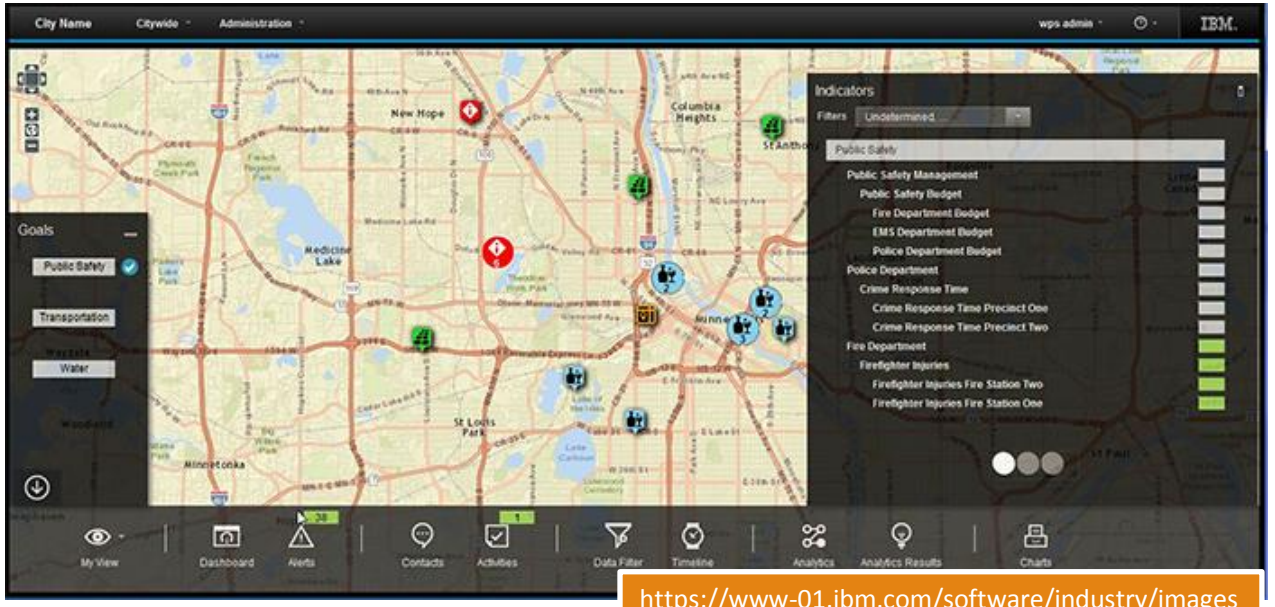
There are 33 models representing 11 river gauges (3 model for each).
The models are:

1. Auto-regressive with adjustment (adding delta variable based on river gauge's relationship)
2. Auto-Arima (Hyndman, 2008)*
3. LSTM Recurrent Neural Network (Keras)

Notes:

LSTM RNN has the strongest performance but we use model 1(Auto-Regressive) due to its simplicity in the implementation

*Hyndman, R. J. and Y. Khandakar. 2008. Automatic Time Series Forecasting: The forecast
Package for R. Journal of Statistical Software. Vol 27.



https://www-01.ibm.com/software/industry/images

The model is implemented inside IOC-IBM for early warning system. Unfortunately, the system is not run automatically to predict flood in the future, still need input from user.

Conclusions and Future Works

- Conclusions
  - Heavy flood in Jakarta is shifting to February
  - North and West Jakarta are the area with higher number of flood reports
  - In production phase, simple model is more reliable due to its simplicity

- Future works
  - Create direct flood model using hourly water level and also rainfall data
  - Create real-time system that run automatically to predict water level and flood area