

# WELCOME TO DATA TALKS

---

OLX Indonesia



Data Science  
Indonesia

We're not  
scientists,  
but we totally  
got space.  
giant leap from our  
tiny little planet to the  
vast, dark void of space.

# Join the Curious Squad!

## 3 Easy Steps to Join



Open browser



Go to slido.com



Enter event code

# Agenda for Today

Time	Duration	Activity
18.30-19.00	30'	Registration
19.00-19.05	5'	Opening by MC
19.05-19.10	5'	Opening Remarks from OLX Indonesia
19.10-19.15	5'	Opening Remarks from DSI
19.15-19.45	30'	Speaker 1: George Hodge & Anthony Mockler
19.45-19.50	5'	Break
19.50-20.20	30'	Speaker 2: Pim van Deurzen
20.20-20.50	30'	Speaker 3: Rahadian Rizki Prayoga
20.50-21.00	10'	Certificate Appreciation
21.00-21.30	30'	Networking Session



OLX GROUP



# Data Talks

OLX-DSI Initiative

Kamalesh Bathala

Jakarta/ 31st January 2019

# Quiz?

---

Owning a smartphone?

Used a cab/bike taxi in the last one week?

Transacted through online banking / card in last 1 month?

Bought anything online recently?

Social / Movies / Gaming online?

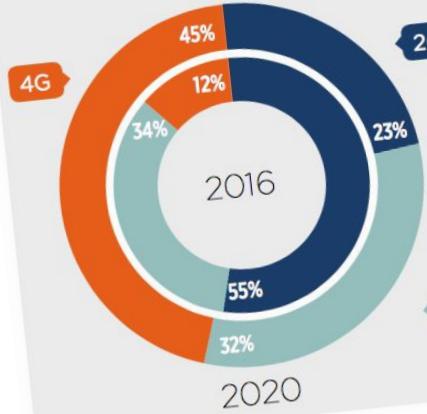
# WHAT IF I TOLD YOU

# YOU'RE NOT ALONE

imgflip.com

## Indonesia

### TECHNOLOGY MIX



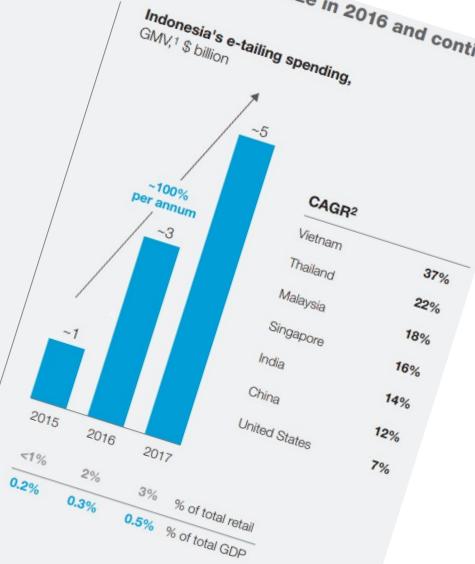
### SUBSCRIBER PENETRATION



**NEWS DESK**  
THE JAKARTA POST

Market  
total GDP  
Developing market  
cross merchandise value  
<sup>2</sup>Compound annual growth rate.  
Source: eMarketer; eshopworlds; Forrester; Statista; research partners: McKinsey analysis

e-tailing market, though still relatively small, doubled in size in 2016 and continues to grow rapidly, reaching \$5 billion in GMV<sup>1</sup> in 2017. This growth is driven by a combination of factors, including increasing smartphone adoption and a growing middle class.



# The new Oil



# OLX Group

Horizontals



**17 Brands**

**40 Countries**

**5000+  
Employees**

# OLX Group



# OLX Indonesia



>2 Million Daily Users



>3 Million New Listing / month

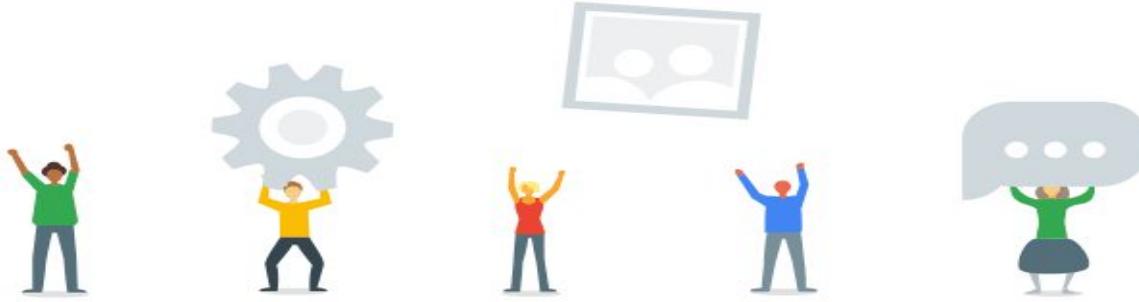


>3.5 Million monthly Buyers

# OLX Indonesia produces >XX petabytes data per day



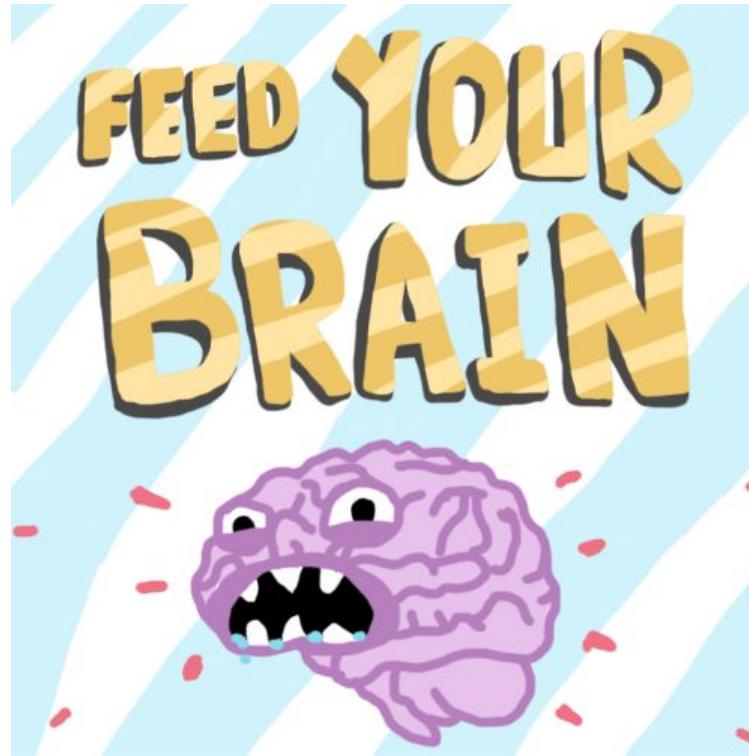
# OLX Indonesia team (Analytics & BI)



# OLX Indonesia team (Analytics & BI)



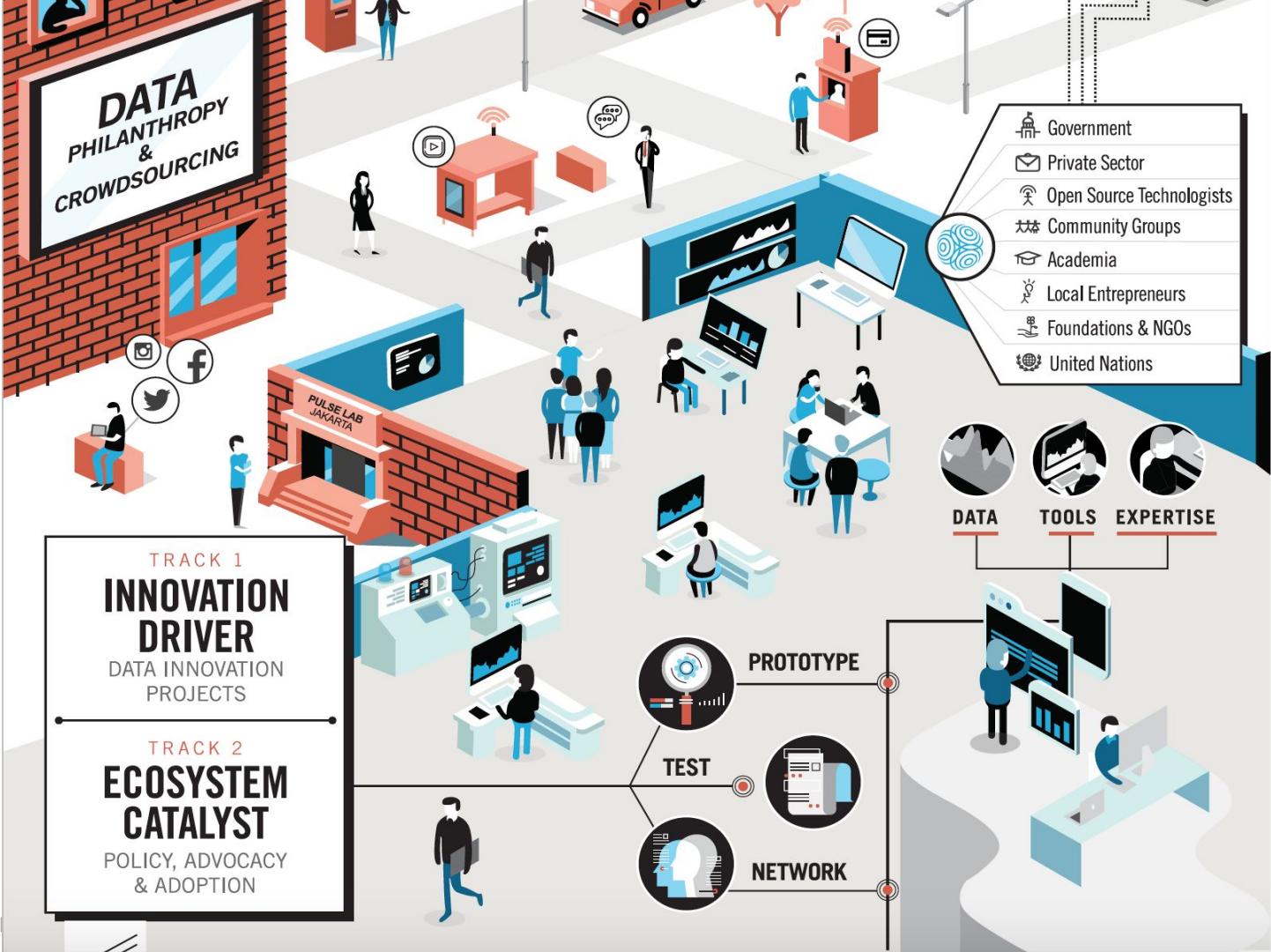
# Welcome to Data Talks! Happy Learning!!



---

# DATA SCIENCE FOR PUBLIC POLICY





# Pulse Lab Jakarta - A brief history



**2008**

During the global financial crisis, world leaders - including former President Yudhoyono - were concerned about the impact it would have on communities in their countries, and expressed the need for more timely information to enable response.



**2009**

Global Pulse was created by the UN Secretary General, based on a recognition that digital data offers the opportunity to track the human impacts of crisis as they unfold, and to get real-time feedback on how well policy responses are working.



**2015**

In 2015, PLJ received new funding from the Government of Australia.



**Now**



**Australian Government**

## Pulse Lab Jakarta - Data Partners



# BIG DATA FOR SUSTAINABLE DEVELOPMENT





# NOWCASTING AIR QUALITY

We investigated a method to nowcast air quality level using deep learning by fusing conventional data (Meteorological and Satellite imagery) with photos shared on social media. Our approach:

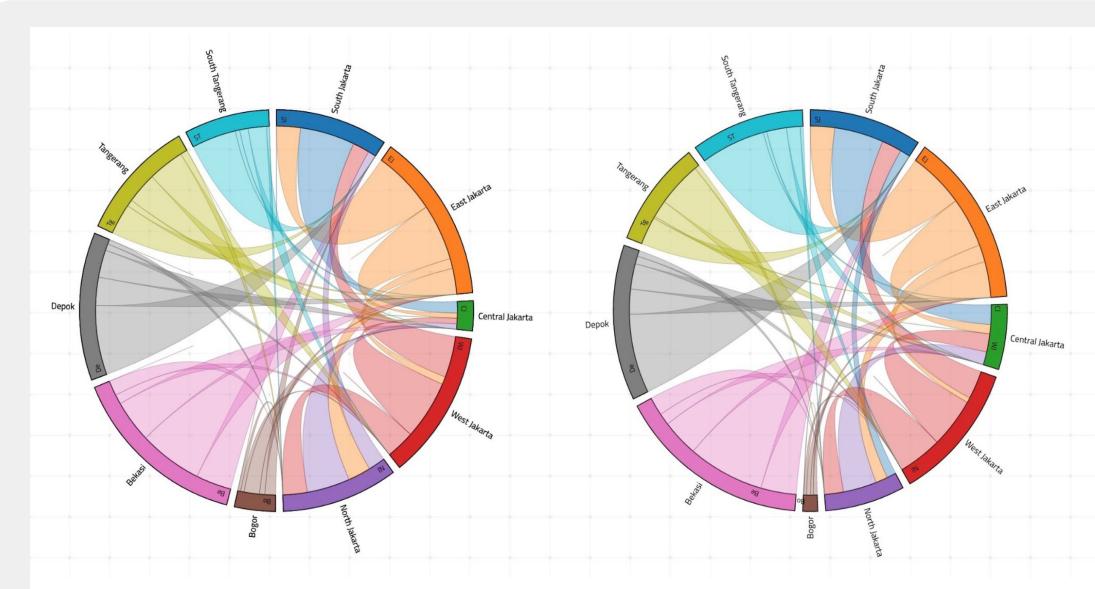
- **Classifying Outdoor Photos:** preprocessing step, classify 2 types of images; outdoor and non-outdoor images
- **Inferring visibility levels from outdoor photos:** Haze/visibility levels are defined as the difference between original and dehazed images
- **Producing near real-time air quality:** Model patterns of AQI level with features extracted from AQI, social media photos, meteorological, and satellite data

This method has potential in nowcasting air quality indicators at the city level.

# COMMUTING STATISTICS FROM SOCIAL MEDIA

Pulse Lab Jakarta worked with *Sekolah Tinggi Ilmu Statistik* (STIS) to test whether location information from social media on mobile devices could reveal commuting patterns.

The chord diagrams show that this could be a promising source of data for inferring commuting statistics in Greater Jakarta.



Official commuting  
statistics 2014

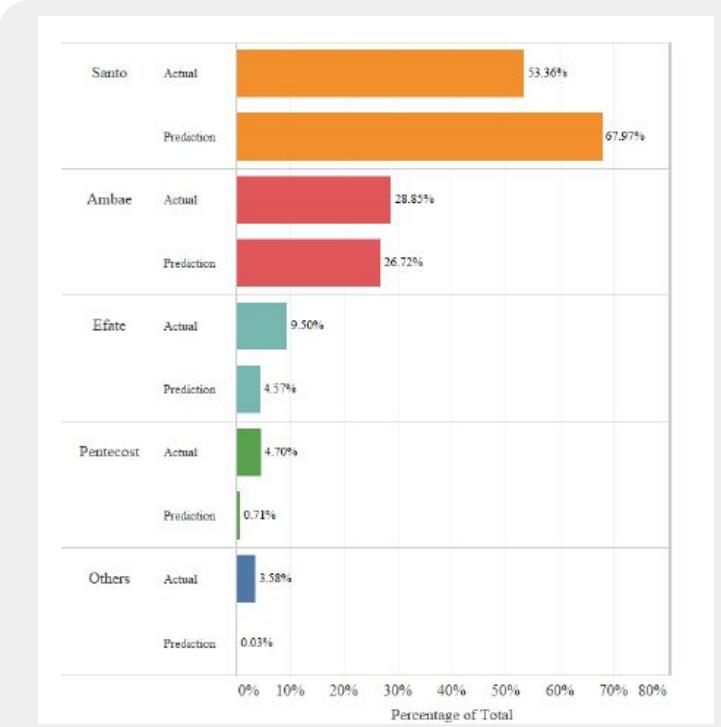
Twitter data  
2014

# ANALYSING CITIZEN DISPLACEMENT FROM MOBILE NETWORK DATA

This research aimed to inform preparation and planning by developing a method to predict evacuation destinations before a disaster happens by applying machine learning approaches to mobile network data.

The eruption of Mount Monaro (Vanuatu) in 2017 was used to test the prediction performance of the model in a real disaster scenario. We explored 273 features, extracted from over one-hundred-million anonymised mobile network records, to describe (a) basic phone usage, (b) active user behavior, (c) spatial behavior, (d) regularity, and (e) diversity.

The results show that supervised machine learning methods can produce promising results in predicting evacuation destinations.



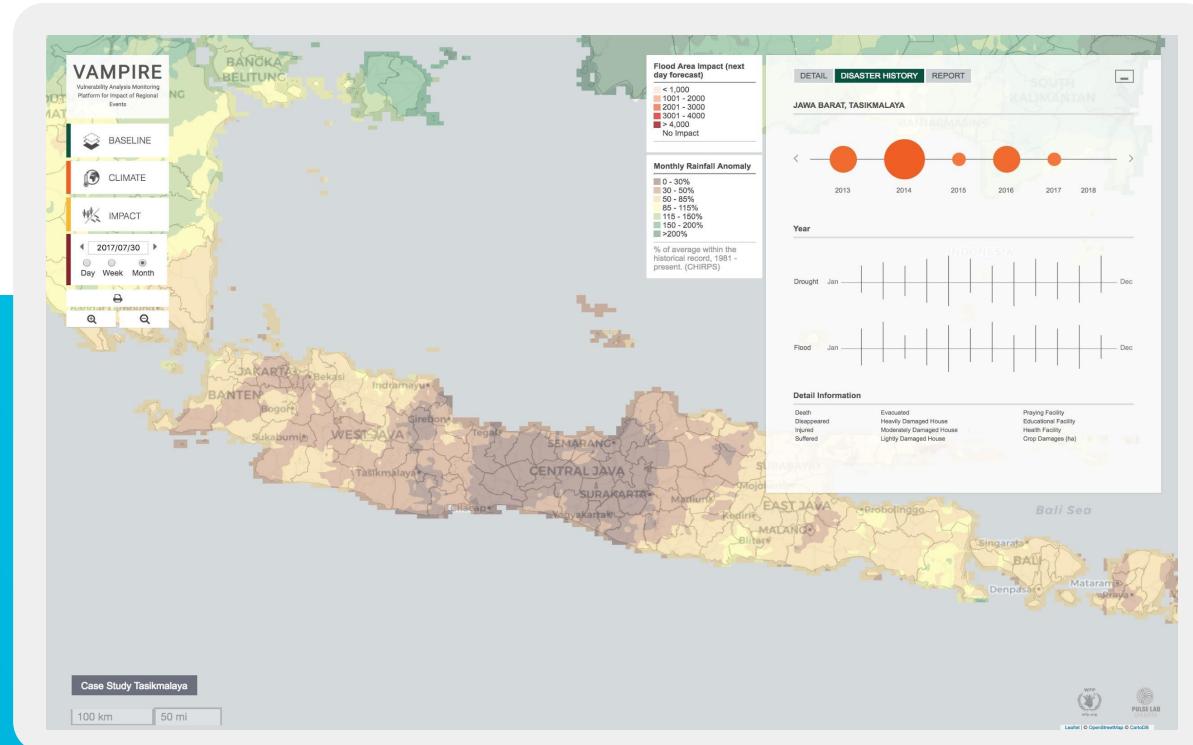
# ANAL DISPL MOBI



# DATA TOOLS

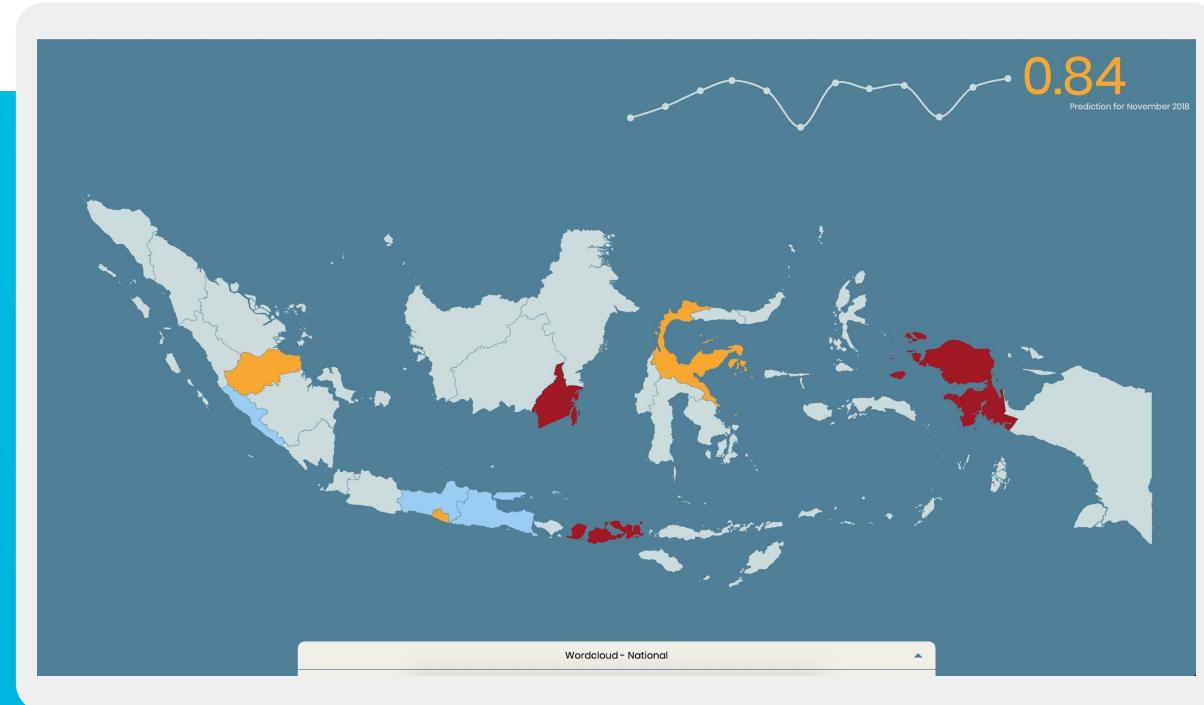
Some of the more successful proofs-of-concept that our lab has been working on have been turned into working prototypes. These include a tool developed together with World Food Programme that can help gain insights on the extent of drought-affected areas and affected populations throughout Indonesia.

The data tool has been adopted as part of an integrated early warning system in the Executive Office of the President of Indonesia. These platforms also help show other potential government partners what is possible.



# EARLY WARNING SYSTEM FOR INFLATION

This prototype platform integrates several sources of data to estimate the level of inflation down to the province level. This includes incorporating price data of major commodities that affect inflation, social media data, and mass media to create an early warning platform to monitor levels of inflation and flag the need for anticipative action. PLJ is working to improve the accuracy and usability of the model with OLX data.



# RESEARCH DIVES FOR DEVELOPMENT

Research dives are three-day research sprints that bring together academia, government and development practitioners to explore the potential of particular datasets for insights into key development issues.

These have been proven successful in also building a community of practice around advanced data analytics, as well as connecting the lab to key domain experts. This year's themes include urban planning, artificial intelligence to estimate poverty, and health.





**PULSE  
LAB JAKARTA**

Harnessing data for  
development.  
Translating insights  
for social innovation.



plj@un.or.id  
@PulseLaabJakarta



@PulseLabJakarta



@PulseLabJakarta



@PulseLabJakarta



Pulselabjakarta.org

# Ready with your questions?

## 3 Easy Steps to Join



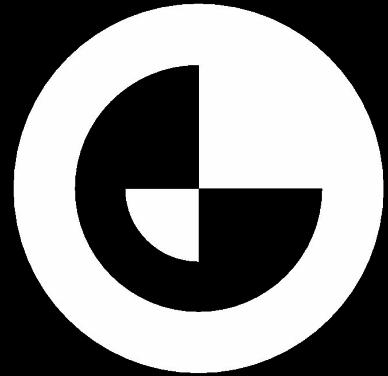
Open browser



Go to slido.com



Enter event code



OLX GROUP



OLX GROUP

AutoTrader.ro Avito AUTOVIT.RO

Domofond dubizzle Fixly

imovirtual letgo OLX

otodom. OTOMOTO property24

shedd STANDVIRTUAL STRADIA

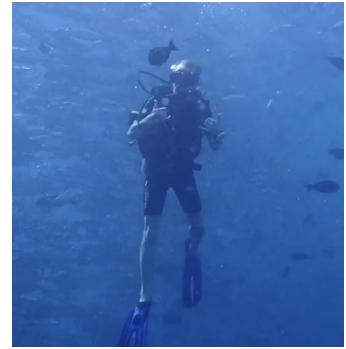
storia TRADUS

# Building strong measurement & data tracking capabilities to drive decisions

Pim van Deurzen

Jakarta, 31 January 2019

# Senang bertemu dengan kalian!



# On OLX we have 3 types of listings

Depan > Semua iklan > Mobil > Mobil Bekas

Platinum



4 Okt  
Mini Cooper S 2012 Multimedia  
Mobil Bekas » Mini Cooper Jakarta Pusat

PLATINUM



26 Des

All New Xtrail 2.5XT At 2015 Low Km ANTI...

Mobil Bekas » Nissan Yogyakarta Kota

Yogyakarta Kota, Yogyakarta D.I.

25.000 - 30.000km 2015

Rp 255.000.000

Nego

GoGo Mobilindo

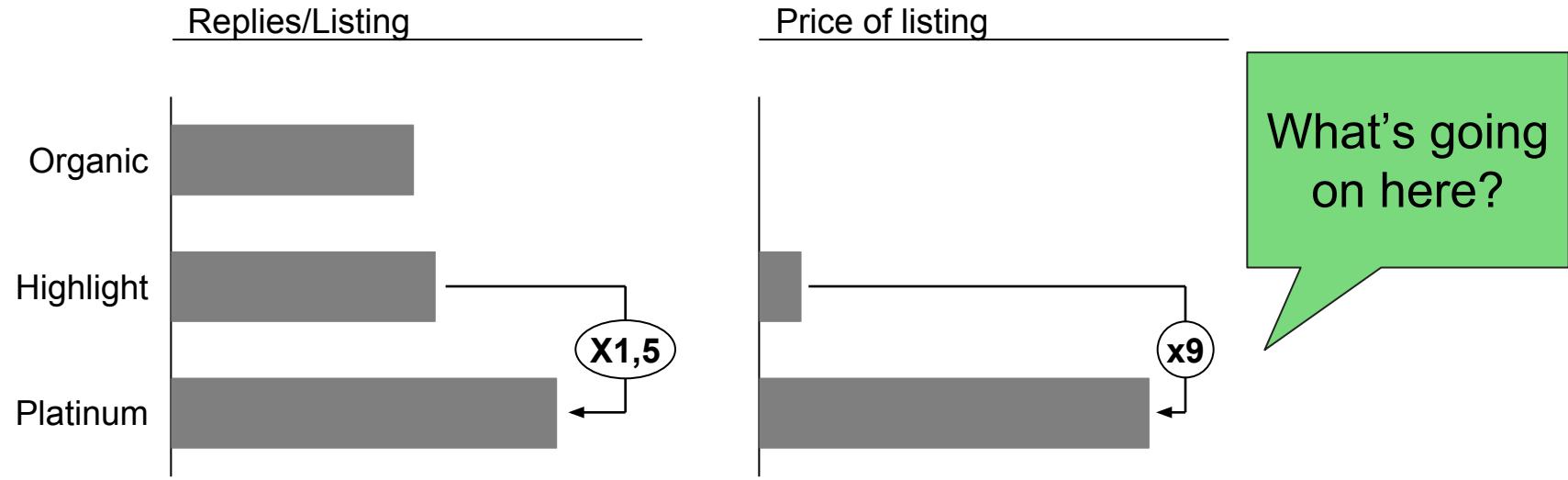
Tampilkan nomor telepon



Introduced in August to help our sellers get more replies



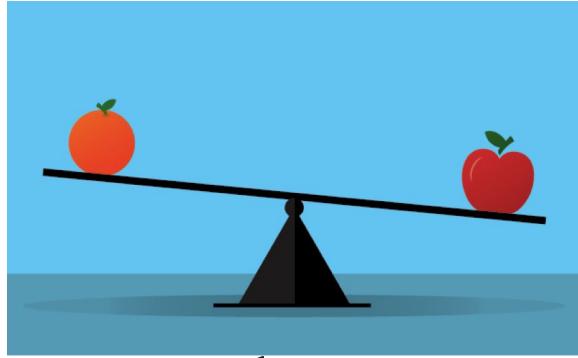
However, our first analysis showed that we only increased replies by 1,5x while the price was much higher!



# So, we started an investigation



Were we looking at  
the right data?



Were we comparing  
apples to apples?



Were our ads  
simply not  
working?

# And we found the answer!

PLATINUM

26 Des

All New Xtrail 2.5XT At 2015 Low Km ANTI...

Mobil Bekas » Nissan Yogyakarta Kota

Yogyakarta Kota, Yogyakarta D.I.

25.000 - 30.000km 2015

Rp 255.000.000

Nego

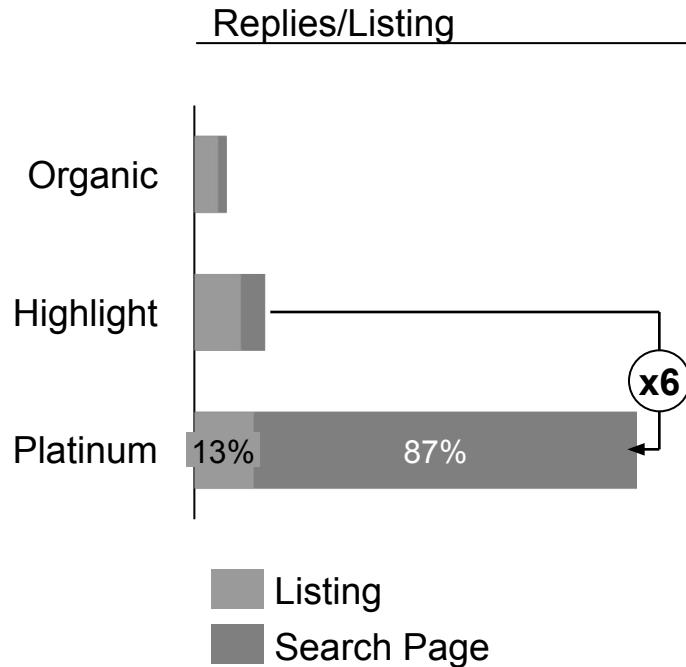
GoGo Mobilindo

Tampilkan nomor telepon

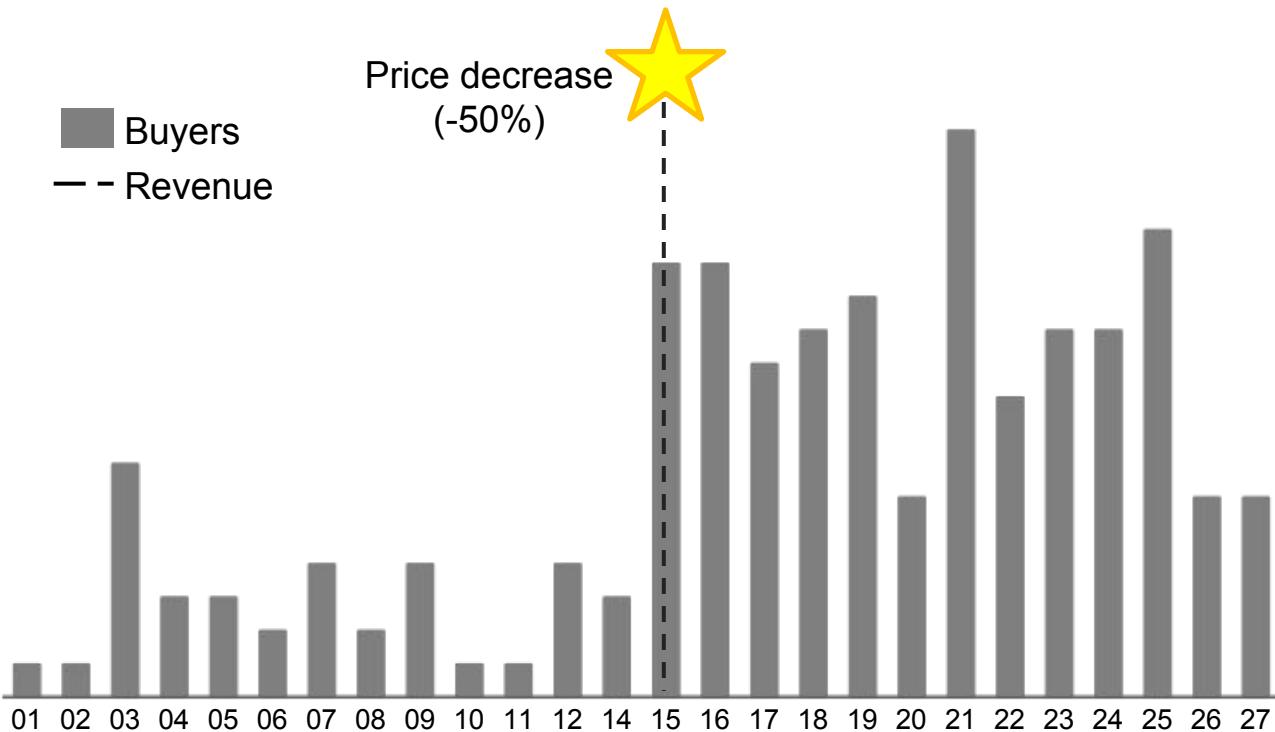
VIP

It turns out we were  
not tracking the clicks  
on this button properly

# As it turned out, our Platinum listing was delivering extra value!



# The price decrease, drastically improved the adoption rate and revenue!



# In the end it was all worth it!

---

From this painstaking process we learned that:

- We need to make sure we are tracking all the new product features
- Keep digging if results don't make sense (95% of the times, there is something wrong)
- We can always improve, analysis should always lead to actions

# More questions, please!

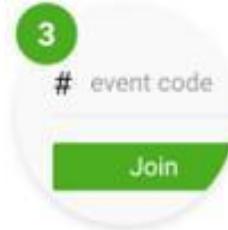
## 3 Easy Steps to Join



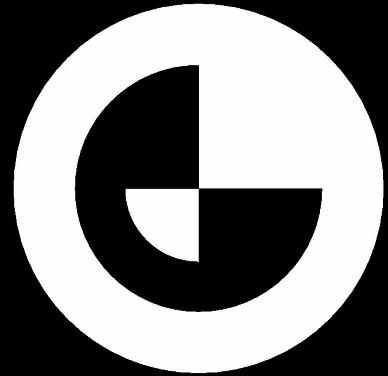
Open browser



Go to [slido.com](https://slido.com)



Enter event code



OLX GROUP

# **BUILDING ANALYTICAL INFRASTRUCTURE USING SIMPLE ETL TOOL**

**Rahadian Rizki Prayoga**

# Profile

## **Education :**

- SHS 1 Jember
- Sekolah Tinggi Ilmu Statistik
  - Economic Statistics

## **Experience :**

- FX Portal (Data Scientist)
- Cekindo Bisnis Group (Market Research & Business Reporting)
- iLotte (Business Intelligence Analyst)
- Akseleran (Senior Data Scientist)



# **How to start your first analytics project ?**

- Gather Requirements
- Build templates and formats
- Build infrastructure
- Develop metrics
- Present and evaluate

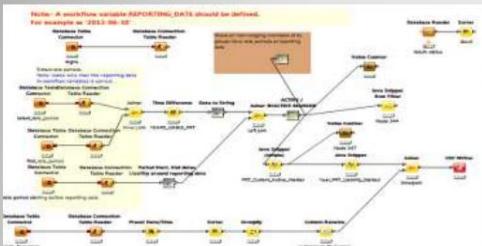
# Problems

- First analytics project in the company / organization
- Current scheme is manual reporting and analytics
- Dependency to IT guy (Backend / Database Engineer) is too high
- Many requests from users
- User requests are made regularly (daily, weekly, monthly, etc)
- Data needs to be integrated from more than one database source
- Type of results wanted are both raw data or analysed data

# **KNIME**

- Stands for **Konstanz Information Miner**
- Free and Open-Source
- Friendly graphical UI
- Fast and easy learning
- Modular data-pipelining
- Enrich with coding via coding snippets (R, Python, Java)
- Tableau Connector
- Usage :
  - Extract, Transform & Load (ETL)
  - Data Integration
  - Machine Learning
  - Data Mining

- Canvas programming



```

if ($THRESHOLD_ALLOCATION_METHODS.equals("M")) {x = "M_Missing";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("Non Grouped") && $IS_WAGES_OSRs == 0 && !$CLAIMING_THRESHOLD_INDs.equals("N")) {x = "0_N_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("Non Grouped") && $IS_WAGES_OSRs == 0 && $CLAIMING_THRESHOLD_INDs.equals("N")) {x = "1_NG_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("Non Grouped") && $IS_WAGES_OSRs > 0 && !$CLAIMING_THRESHOLD_INDs.equals("N")) {x = "2_NG_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("Non Grouped") && $IS_WAGES_OSRs > 0 && $CLAIMING_THRESHOLD_INDs.equals("N")) {x = "3_NG_WithThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("NTC") && $IS_WAGES_OSRs == 0) {x = "4_NTC_NSW_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("NTC") && $IS_WAGES_OSRs > 0) {x = "5_NTC_IS_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("DGE") && $IS_WAGES_OSRs == 0 && !$CLAIMING_THRESHOLD_INDs.equals("N")) {x = "6_DGE_NSW_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("DGE") && $IS_WAGES_OSRs == 0 && $CLAIMING_THRESHOLD_INDs.equals("N")) {x = "7_DGE_NSW_WithThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("DGE") && $IS_WAGES_OSRs > 0 && !$CLAIMING_THRESHOLD_INDs.equals("N")) {x = "8_DGE_IS_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("DGE") && $IS_WAGES_OSRs > 0 && $CLAIMING_THRESHOLD_INDs.equals("N")) {x = "9_DGE_IS_WithThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("SL") && $IS_WAGES_OSRs == 0) {x = "8_SL_NSW_ThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("NL") && $IS_WAGES_OSRs == 0) {x = "9_NL_NSW_NoThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("SL") && $IS_WAGES_OSRs > 0) {x = "8_SL_IS_ThresholdClaim";}

else if ($THRESHOLD_ALLOCATION_METHODS.equals("NL") && $IS_WAGES_OSRs > 0) {x = "9_NL_IS_NoThresholdClaim";}

else if ((($THRESHOLD_ALLOCATION_METHODS.equals("PRC") || $THRESHOLD_ALLOCATION_METHODS.equals("PER")) || $THRESHOLD_ALLOCATION_METHODS.equals("BHK") && $IS_WAGES_OSRs == 0), x = "10_OTHER_NOWL");

```

- Coding

## Java Snippet (simple)



# Questions...?

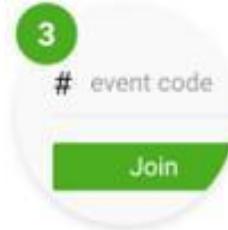
## 3 Easy Steps to Join



Open browser



Go to slido.com



Enter event code

In God we trust.  
All others must bring Data.

*-William Edwards Deming*