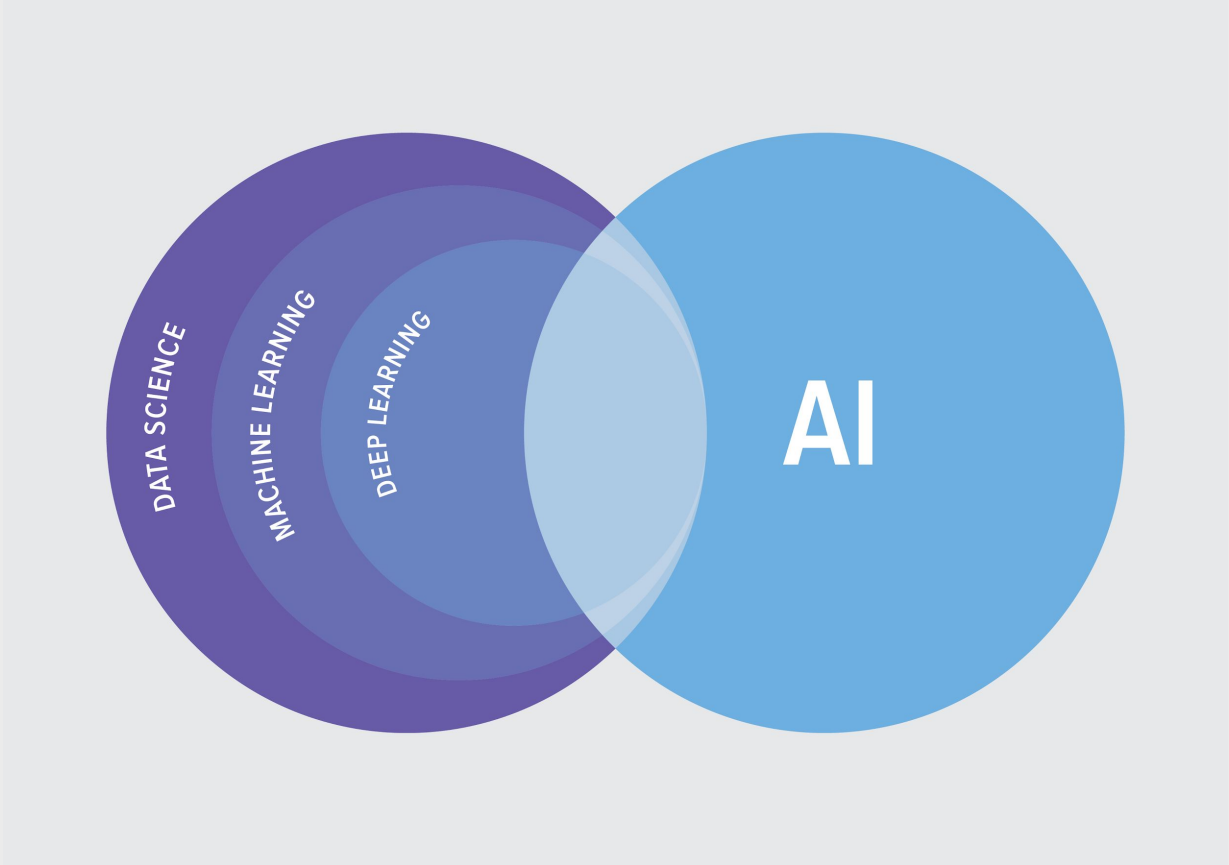
nodeflux

Extending Vision Beyond Imagination

Storytelling with Data

Before we start..



Outline

Introduction

Introduction

Deciding your objective(s), asking the right questions, metrics

Overview of the whole process

Data Gathering

How do we gather the data we need

The general way, and the “smart” way using AI

Data Visualisation & Analytics

Data Interpretation

How to understand the data that we have (basic and ML approach)

Applications

Business Implementation

Using the insights we get and leverage it in our business

Before anything else..

> *I have a bunch of data, but I don't know what to do with it,*
> *what can I get out of it?*

"Just in Case" ➡ "Collecting this to answer that"

"A good balance between intuition and data"

Reference: https://www.datapine.com/blog/data-driven-decision-making-in-businesses/

Objectives

- Knows what’s matters for your company
- Had very specific questions

Then you will know,

- What kind of data you’ll need to gather
- What technique/analytical method to understand it
- What informations you’d like to get, best way to visualise it

Reference: https://www.datapine.com/blog/data-driven-decision-making-in-businesses/

Metrics (that matter)

For example,

An e-commerce would care about their **conversion rate, customer lifetime value, retention rate.**

These metrics can then be divided into a more specific metric for a particular feature or business unit.

Usually we have a ***north star metric,*** the main thing that we would like to focus on.

e.g for Gojek: Completed transactions, Tokopedia: Gross Merchandise Volume

These “goal” can be used not only for your data analytics purpose, but to measure your team performance as well, e.g.:

Developer A creates a new method for a travel site’ recommendation & suggestion system.

To decide whether this new method works/not, we can see the **conversion rate**, from people seeing the pages to booking order.

Metrics (that matter)

Element	Function	Relevant metrics
Acquisition	Generate attention through a variety of means	Traffic, mentions, cost per click, search results, cost of acquisition, open rate
Activation	Turn the resulting drive-by visitors into users who are somehow enrolled	Enrollments, signups, completed onboarding process, used the service at least once, subscriptions
Retention	Convince users to come back repeatedly	Engagement, time since last visit, daily and monthly active use, churns
Revenue	Business outcomes (purchases, ad clicks, content creation, subscriptions, etc)	Customer lifetime value, conversion rate, shopping cart size, click-through revenue

Reference: Croll, A. and Yoskovitz, B., 2013. Lean analytics: Use data to build a better startup faster. " O'Reilly Media, Inc.".

Step Towards Data-Driven

Data Gathering

Analytics

Story

Visualisation

Insights

Data Gathering

Information can come from range of sources:

- Interviews
- Questionnaires and surveys
- Documents and records
- etc,

In modern way,

- API and Web Services
- Using AI tech

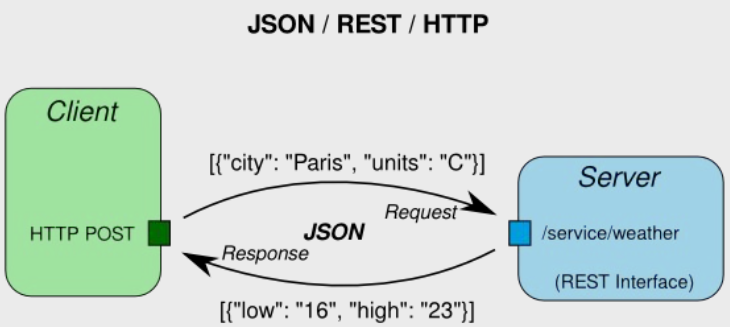


Image credit: Safety Net

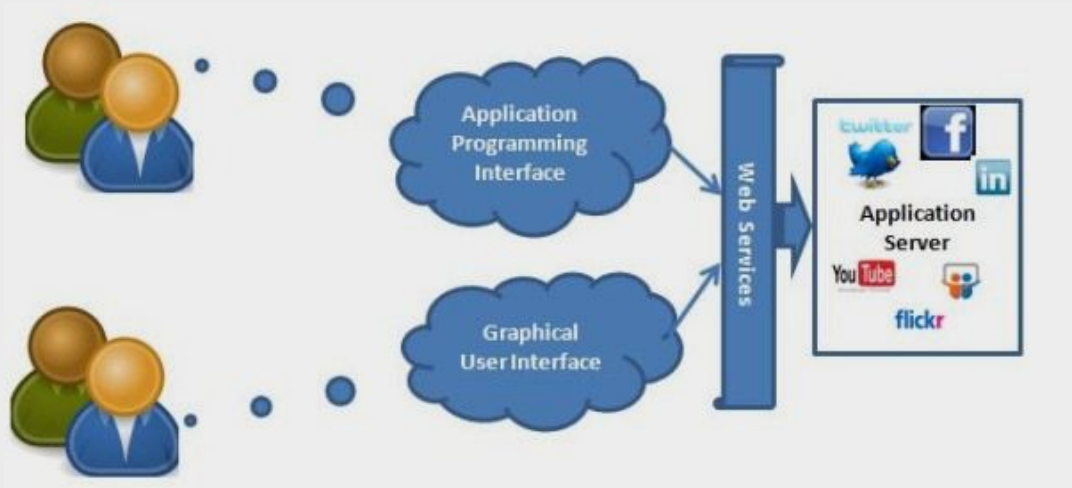


Image credit: sociolytics.files.wordpress.com

Nodeflux - Gojek

Problem

Inability to gain insights on **competitors dissemination** on the road

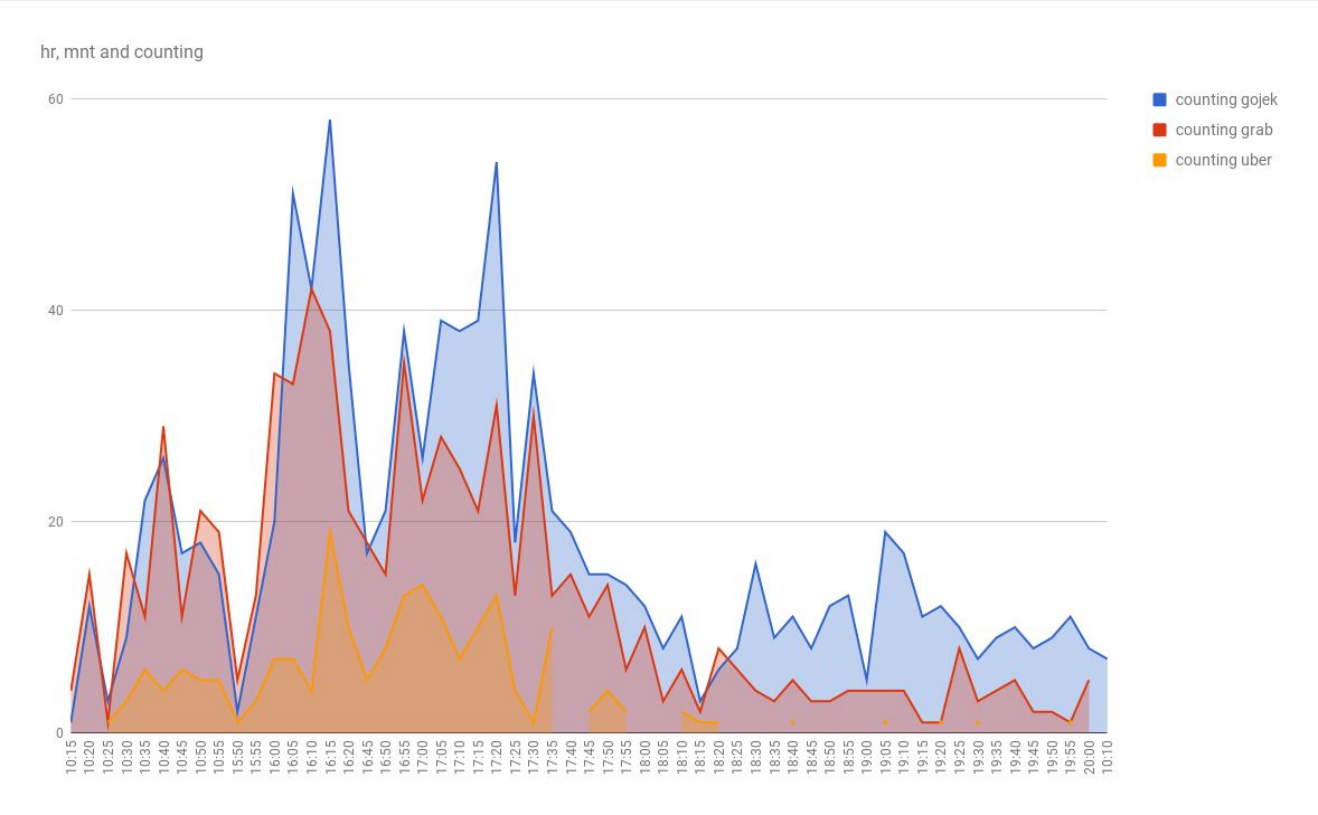
Some things to consider

- CCTV angles
- Spot

We use data (labelled video) to train our model, than to generate new data (new real-time vehicle detection analytic). Our new data can then be used for another insights.

Just like story yang bersambung

Nodeflux - Gojek



Nodeflux - Retail

Problem

Inability to gain insights for **consumers acquisition** on the offline store

Visitor Queue

Visitor Heatmap

Visitor Pathflow

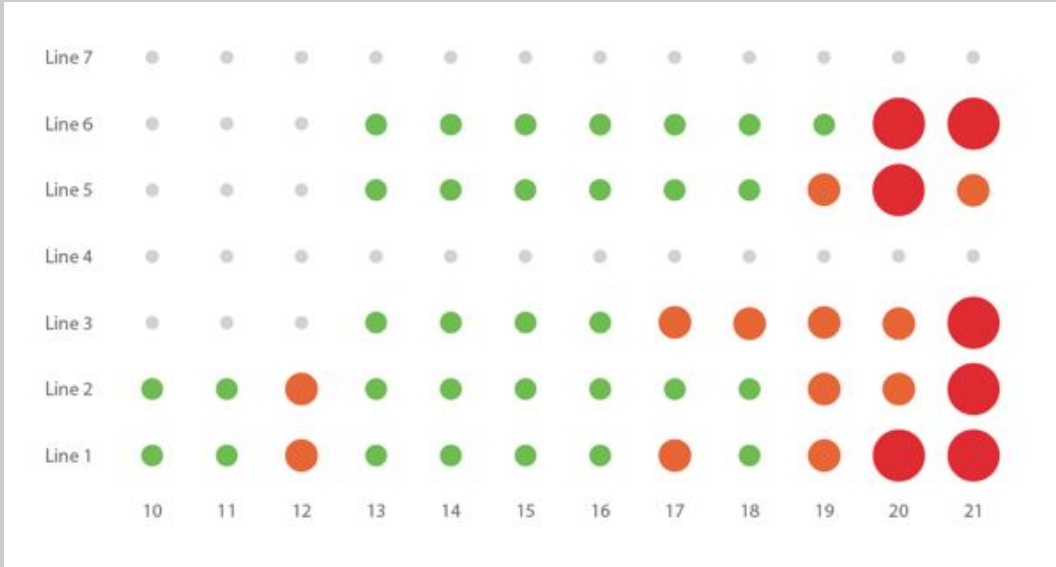
Visitor Counting

Visitor Demographic

Product Ranking

Visitor Queue

Queue estimation will give you more insight about your queue. You can aligning queue data with cashier line availability and slice it within certain time period.



Visitor Heat Map

Visualizing overall activity level within an area, to indicate hot spots, according to where customers stop at often or for longer periods of time.

Visitor Path Flow

Tracks the visitor path within an area. This helps you to identify the dominant paths taken by customers within a designated timeframe.

Visitor Counting

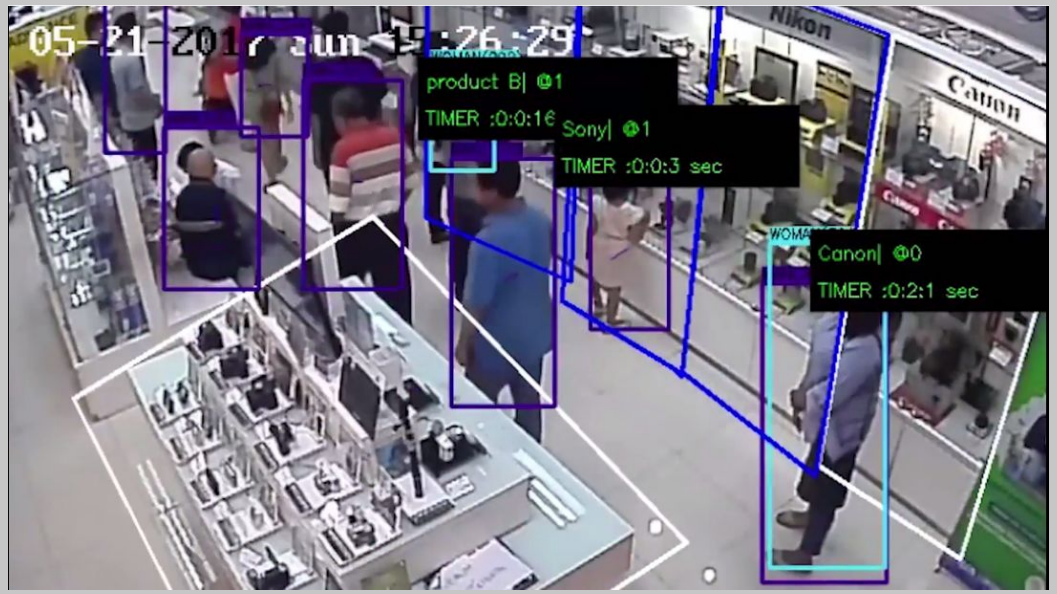
Visitor counting provides data that enables end users to better understand their operation and maximize potential opportunities by analyzing closing ratio and identify opportunity for improvement.

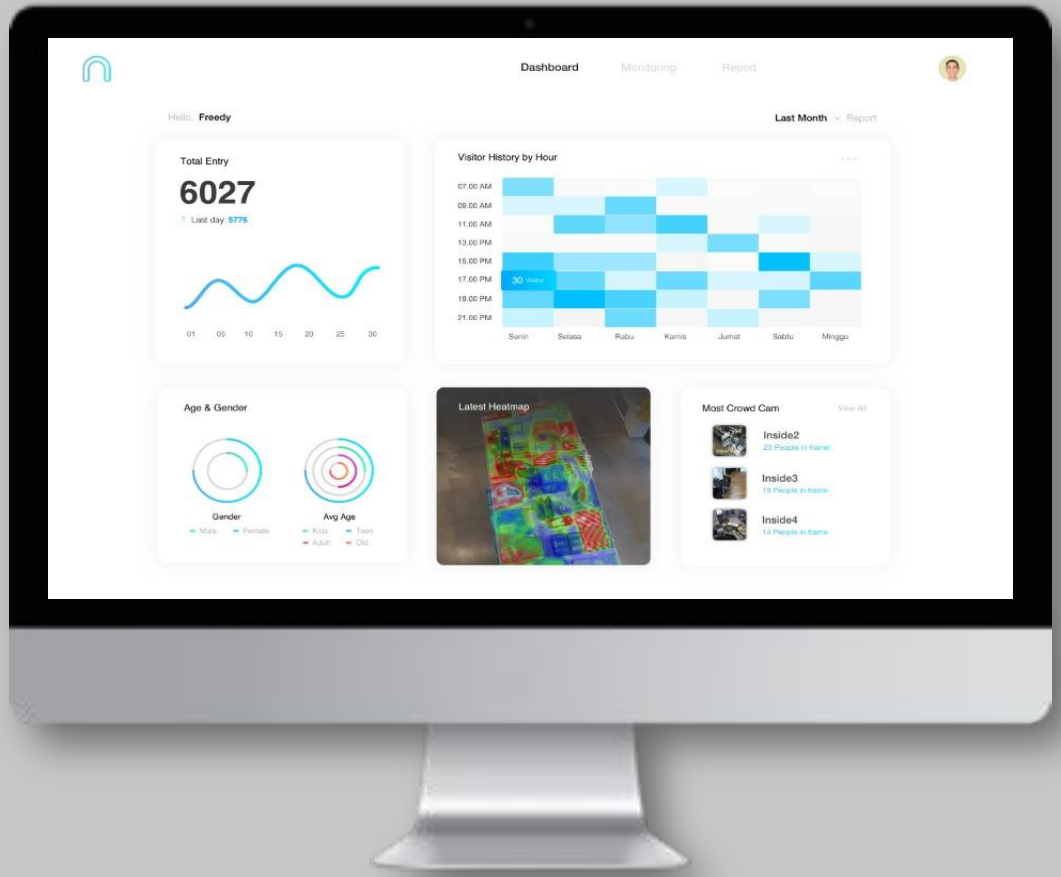
Visitor Demographic

Understand your visitor better and get their demography as a reference data to optimize your business activity and better targeting your market.

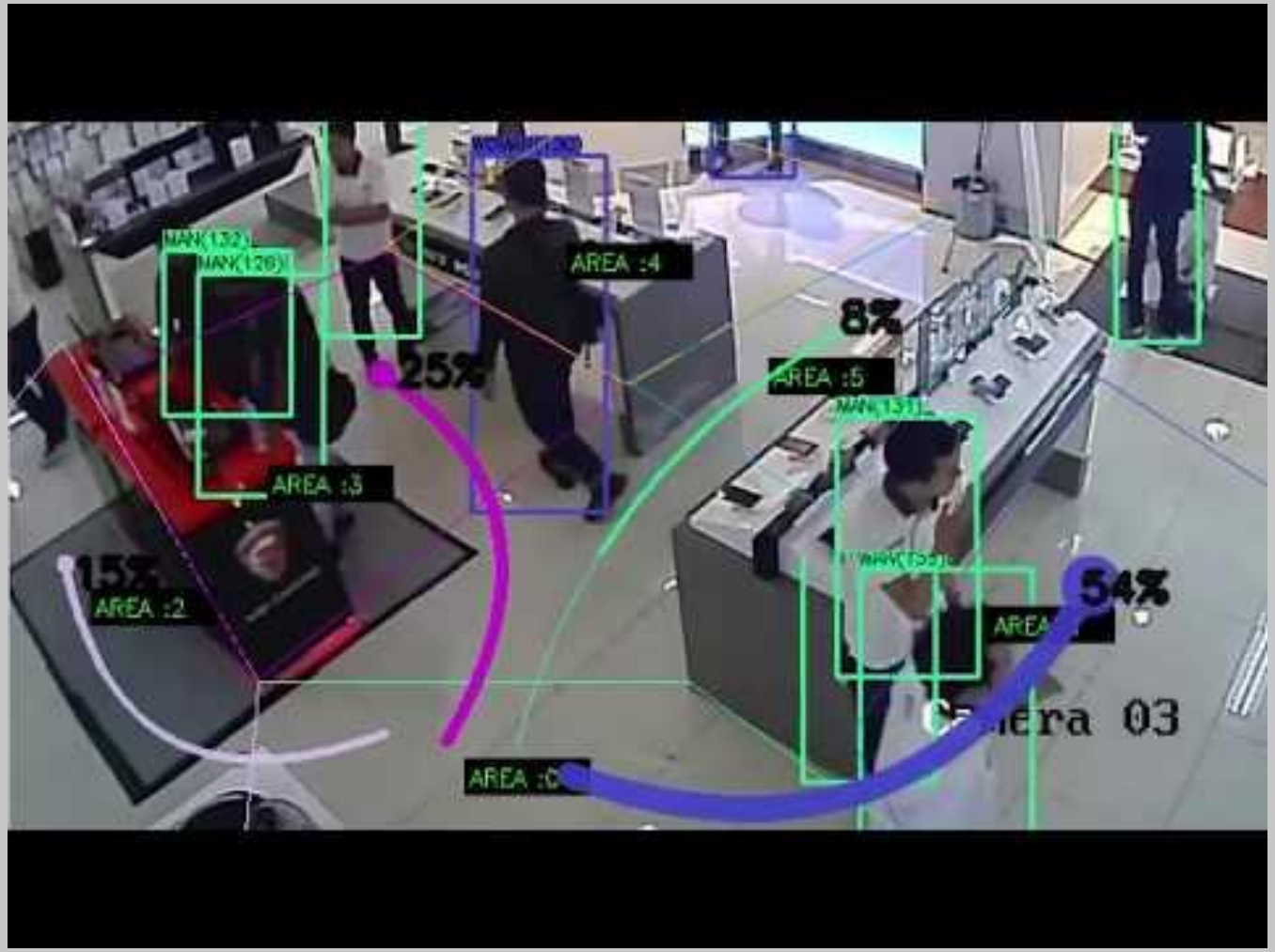
Product Ranking

Use visitor dwelling spot and time to obtain additional information such as product ranking. Analyze which product that can attract visitor more. Visitor favorite spot and interaction time could be valuable information in order to optimize your selling operation





Retail Store
Analytics Dashboard



Store Analytics Video


COUNTER MULTIPLE LINE
LANE 0: CAR: 0 | TRUCK: 0 | MOTORBIKE: 8 | BUS: 0 |
LANE 1: CAR: 17 | TRUCK: 3 | MOTORBIKE: 45 | BUS: 2 |
03-11-2017 17:09:55
LANE 0
LANE 1

Analytics

Sometimes it doesn't need to be "sophisticated" method such as using machine learning.

Sometimes a simple Excel formulas might do.

- How many customers do we have as of today?
- How many customers that unfortunately decide to leave us this month?
- How many customers that recently join us this month?
- Average deposit in their accounts?
- etc.

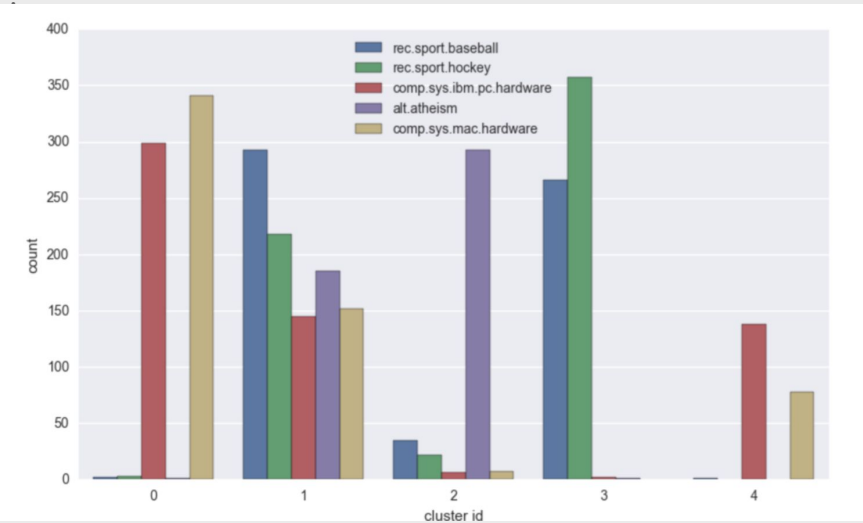
Questions like those above can be answered by using aggregation technique to your past transactional data.

Analytics

Say you have a list of customers that has left your company, and you would like to know their characteristics, that associate all of them so you could improve your service accordingly.

There is an algorithm called ***apriori algorithm,*** or you could leverage ***clustering method***.

Basically, you don’t know the similarities between each data point, and you’d want the machine to do that for you



Visualisation

> *"What is the point? What is the message?, What is the story here? or the concise, "So what?"*

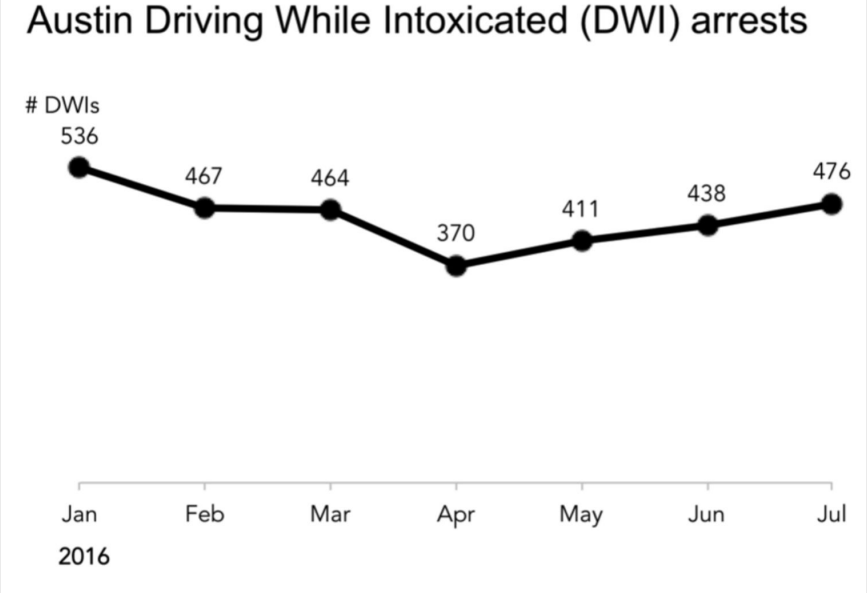
Everyone wants to "tell a story with data." But very often, when we use this phrase, we don't really mean story. We mean what I mentioned above—the point, the key takeaway, the so what?

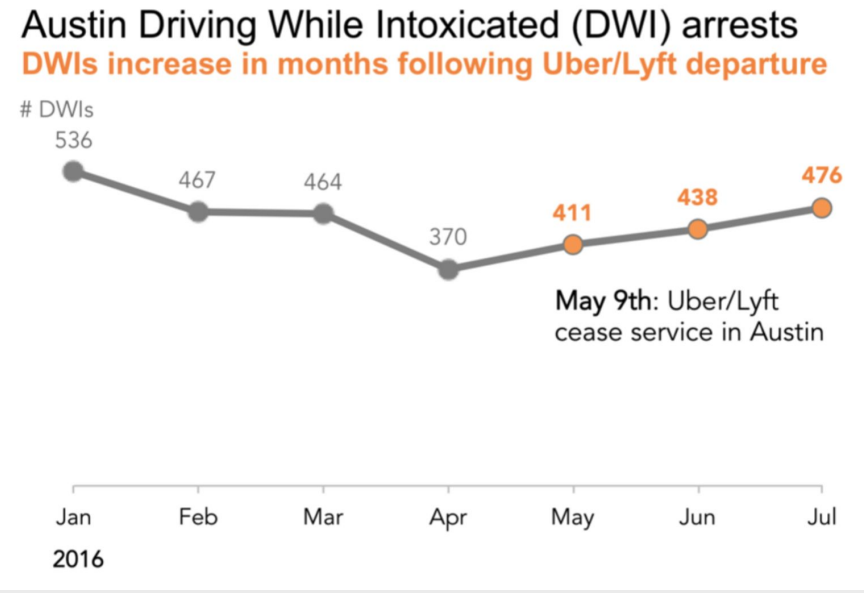
A Story has key critical components—there is a structure, a shape—it has a plot, a rising action, a point of climax where tensions reach their highest, a falling action, and a resolution.

The "story" for when you are communicating for explanatory purposes with data, is not really story at all, but rather the point—the so what? For every graph you show, for every slide you show: make the point clear to your audience.

Reference: http://www.storytellingwithdata.com/blog/2017/3/22/so-what

Visualisation



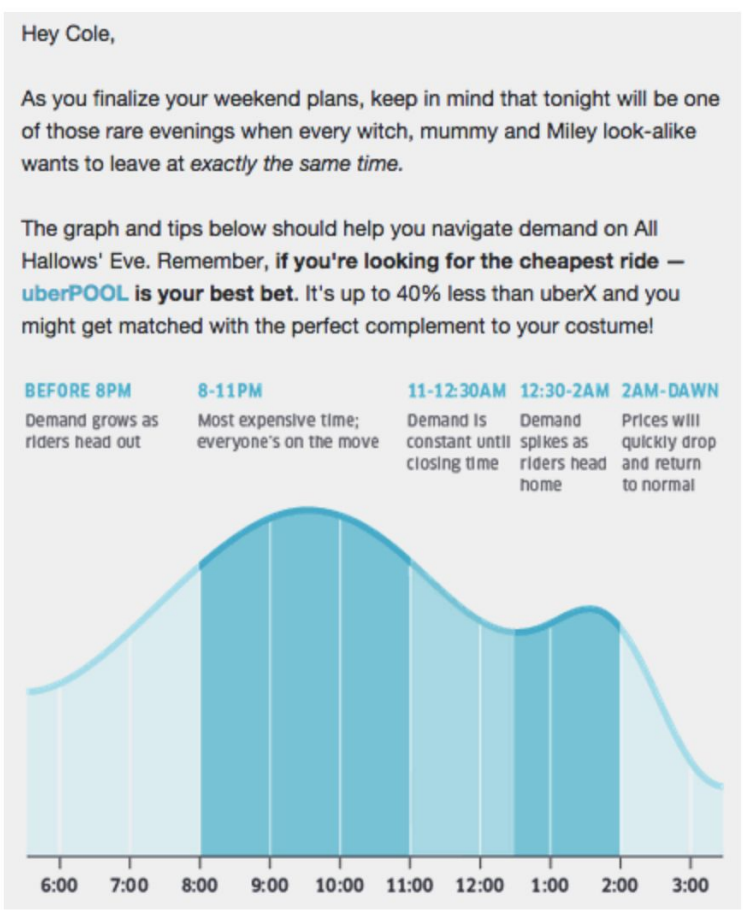


We can clearly see the trends, like it was decreasing from Jan to April, and start to show positive trend afterwards, but *so what?*

I put the main point into words via the subtitle. I did my audience the added bonus of tying these words visually to the relevant data points through consistent use of color.

Reference: http://www.storytellingwithdata.com/blog/2017/3/22/so-what

Visualisation



Another good example of effective data visualisation, where Uber inform their customers in the email' body about the traffic, with the use of annotated line graph

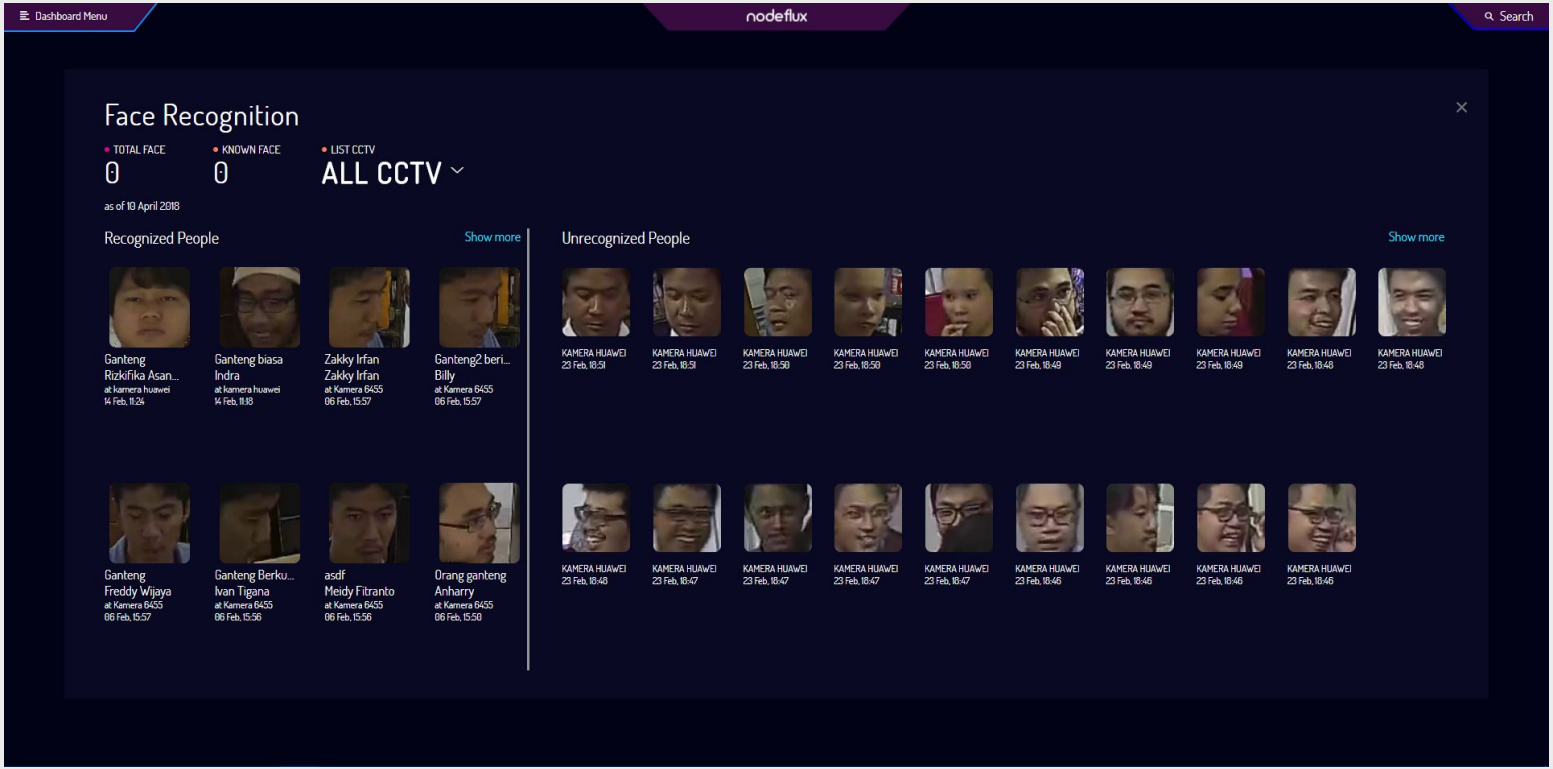
What not to do in data visualisation:

viz.wtf

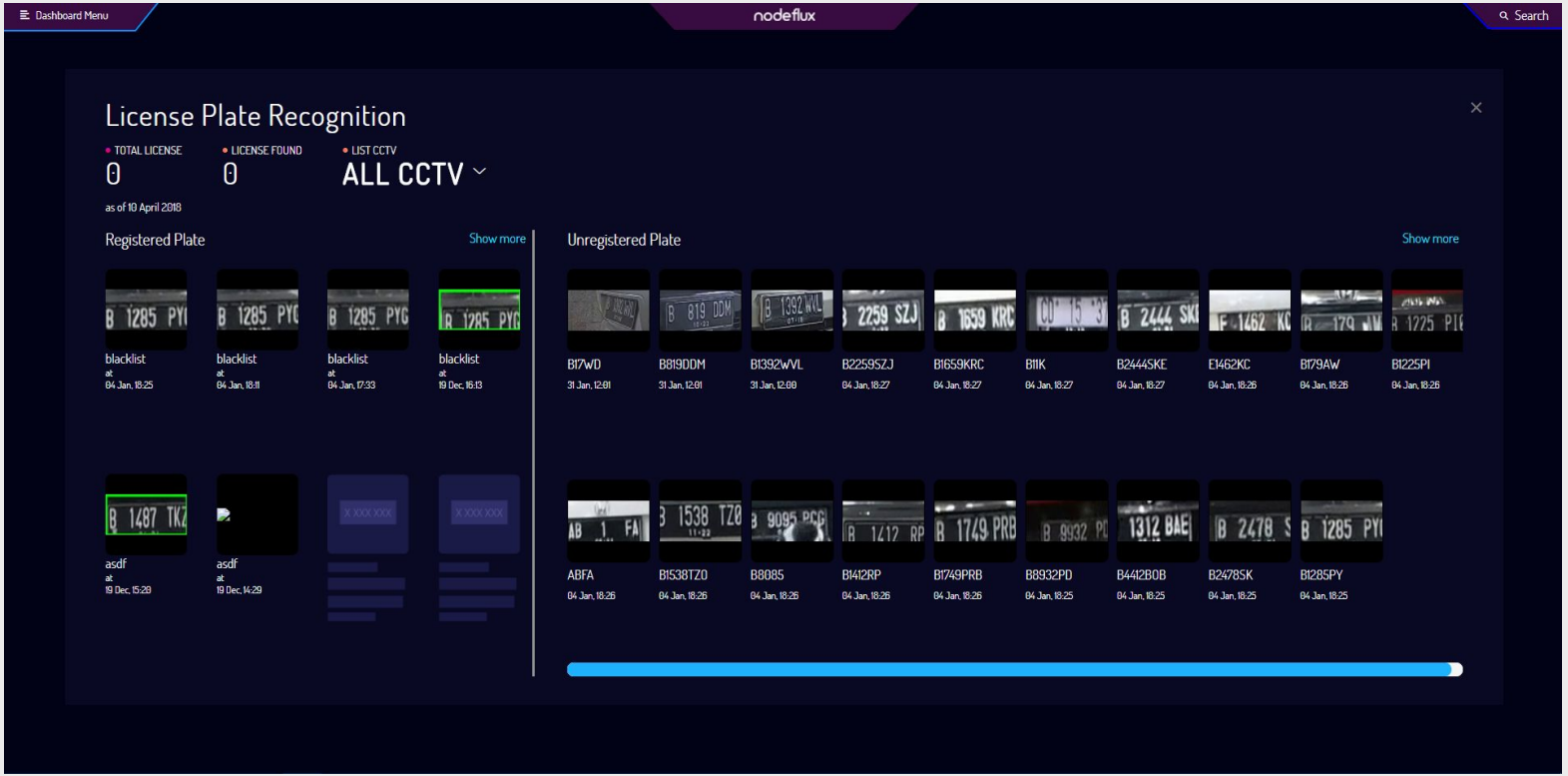
Reference:

http://www.storytellingwithdata.com/blog/2014/10/annotated-line-graph-from-uber

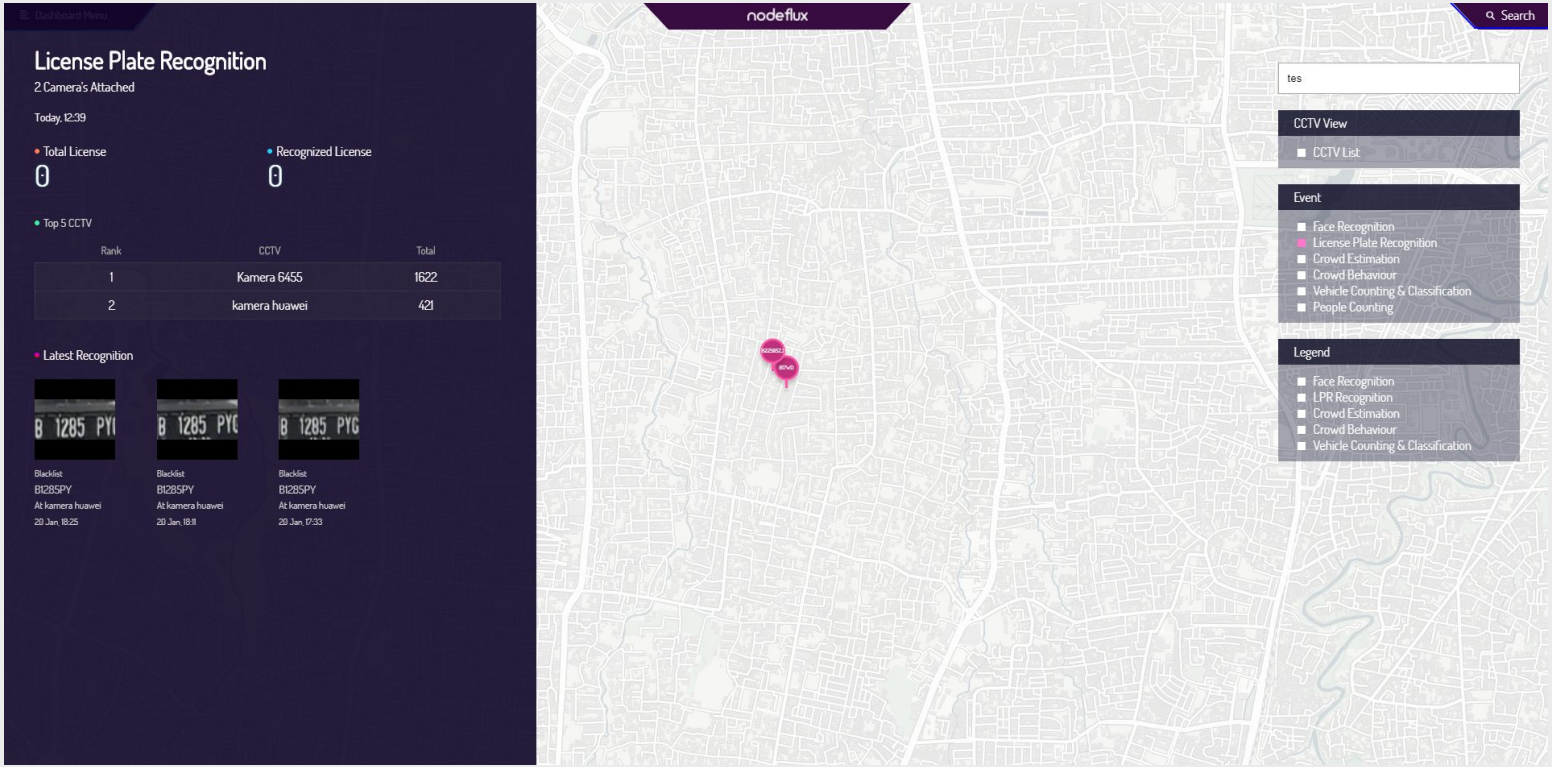
Visualisation



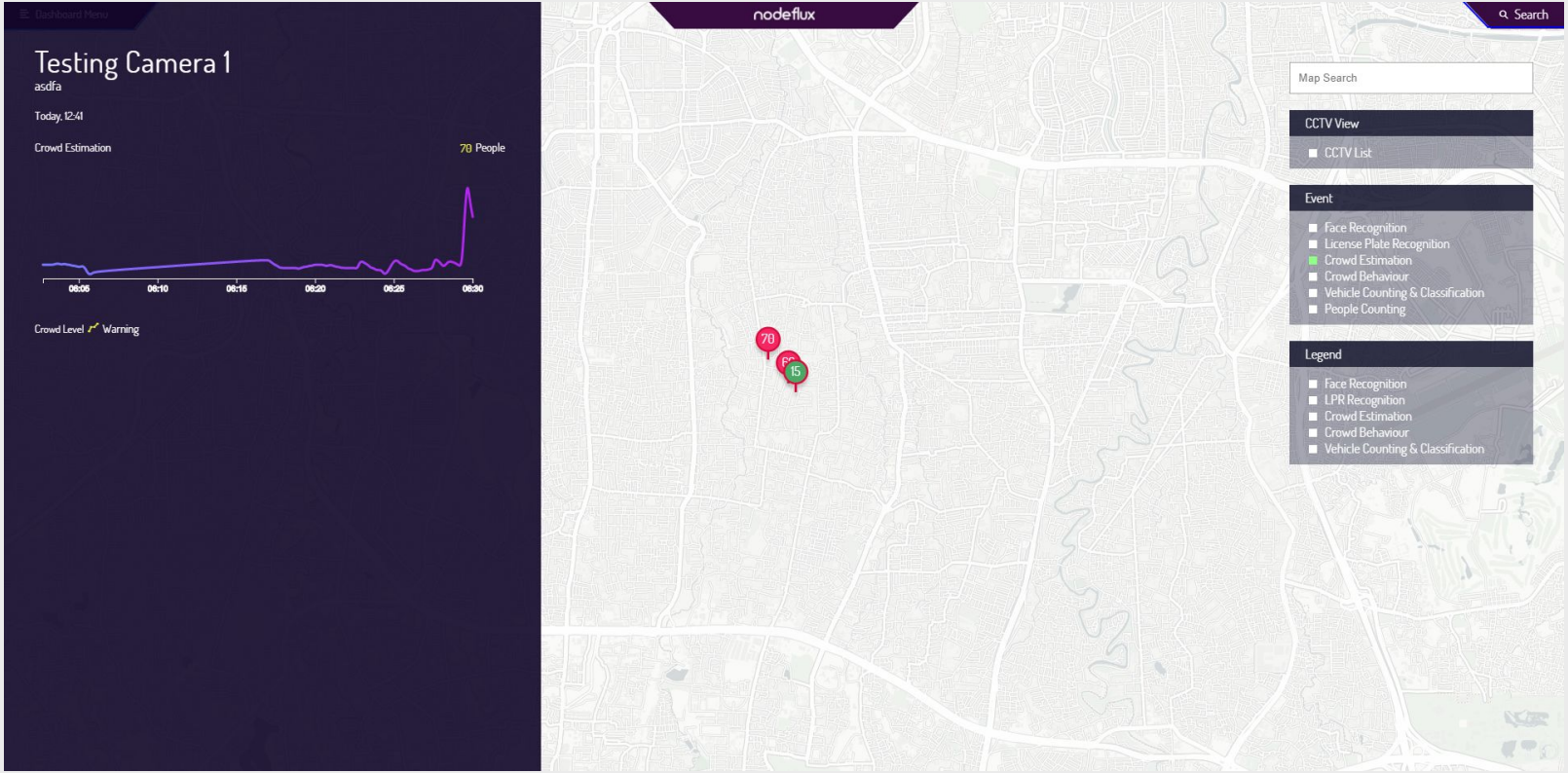
Visualisation



Visualisation



Visualisation



Intermezzo

- A developer gathered our data through Facebook app, for “research purposes”
- Turns out, it got to the wrong hands, e.g. CA, who use the data to somehow manipulate public opinion about a certain candidate in an election, and develop fake news.

Imagine,
You know Person A doesn’t like Pak Jokowi.
You know Person A also doesn’t like littering.
You also know when Person A usually access FB.
You could easily show a “fake” news of Pak Jokowi littering, when Person A most likely will see it.

richard@nodeflux.io