



COVID-19 Indonesia Social Media Monitoring

Team members: Laura Ralston (lead), Martha Dewi, Kiran Garimella, Alicia Mergenthaler, Louis Owen, Reksa Samudra, Dimas Setyonugroho (Social GP)

Contributions from Maham Khan, Nick Jones (GFDRR) and Kai Kaiser (Governance GP)

: Covid-19 Observatory in Indonesia – An Overview

Households

(phone based high frequency tracking)

Digital Merchants and Gig Workers

(surveys through digital platforms)

Social Media Monitoring

(harnessing big data from Facebook and Twitter)

This is an activity under the COVID-19 Observatory in partnership with the Poverty GP

We also draw from big data sources on mobility and remotely sensed information

WHY USE SOCIAL MEDIA?

- **Inform:** obtain policy relevant information for decision makers at national and local levels
- **Content:** citizens' understanding of the pandemic, the behaviors it triggers, sentiments towards government as well as rising concerns
- **Speed:** rapid monitoring, initial results available early May
- **Breadth:** 90% of Indonesians 170m internet users on SM
- **Value:** low cost relative to traditional methods: example per unit survey costs below 5 cents, big data collection free, costs are related to analysis time and computational resources

WHAT WE DID

- Monitoring conversations via social media and online news
- Crowdsourcing citizen information via online survey
- Comparing results across different data collection efforts
- Integrating with other big data analytics

WHAT WE KNOW SO FAR

- There is a high level of knowledge on COVID-19, with most respondents aware of common symptoms and preventative health measures. However, citizens face challenges in following these measures.
- Concerns related to COVID-19 have revolved around health care, food access, job loss and government handling. This matches own and other survey data.
- While the overall tone in the discussions on social media and in online news sources has been moderating, the government handling of the situation is widely discussed.

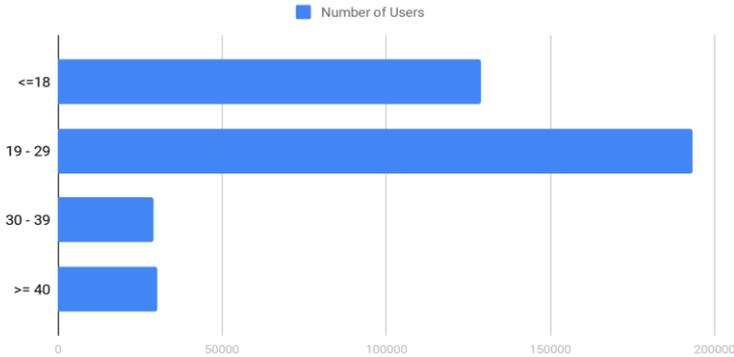
POLICY IMPLICATIONS

- Are there opportunities to make following preventative measures easier?
- Some anticipated concerns – food security, health access, job loss – are apparent. Need to monitor closely and unpack at subnational level to find appropriate entry points.
- Track if there are significant changes in tone of sentiment towards government. Are certain groups at greater risk of exclusion?

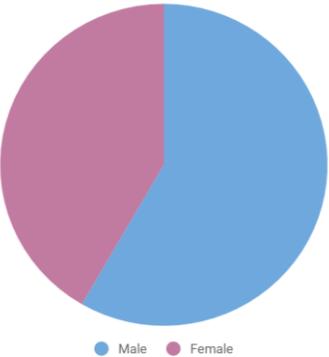
Methods Overview

Who? 440,000 Twitter users, 1.2M Instagram users, 220,000 online news articles, 3,300+ online survey respondents

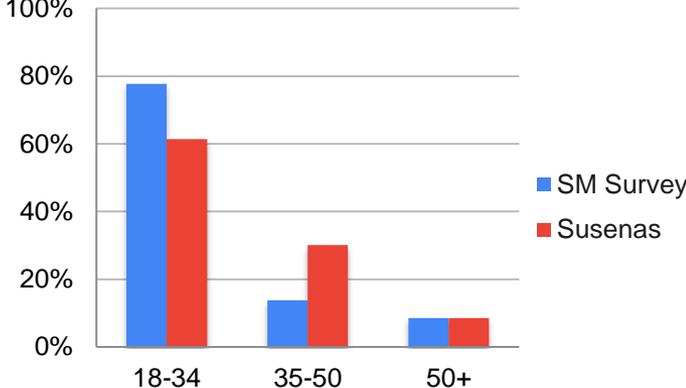
Twitter Users Age Distribution



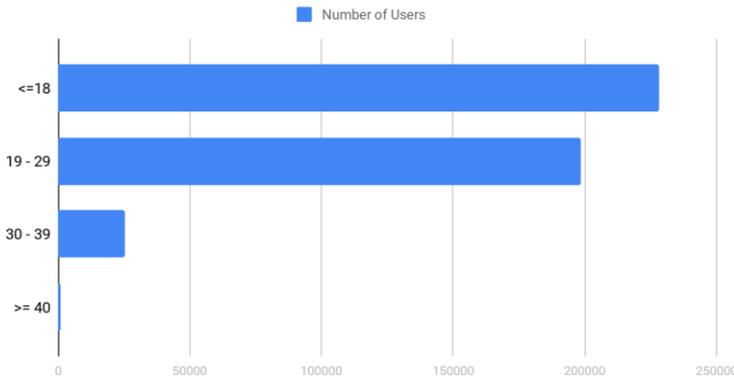
Twitter Users Gender Distribution



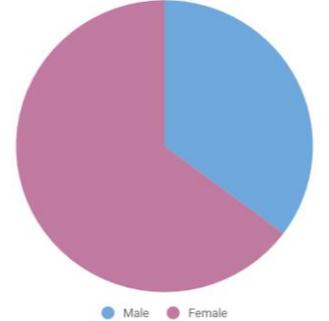
Online Survey vs Nat Pop



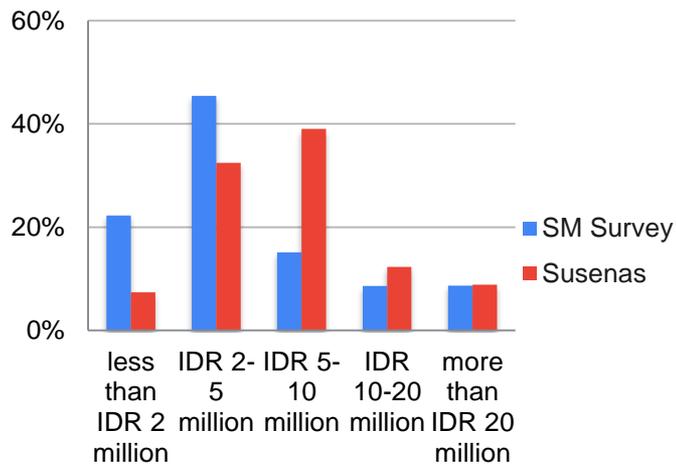
Instagram Users Age Distribution



Instagram Users Gender Distribution



Online Survey vs Nat Pop



How? Via comprehensive information collection and analysis approaches

Information collection

- 1 Collect posts with COVID related hashtags and geotag information since March 26th



>3mn
tweets



>5mn
posts

- 2 Extract COVID-19 related posts using GDELT repository of headlines



>200k
articles



~100
sources

- 3 

Compare online conversation with on-the-ground situation: online survey conducted from 20-Apr to 20-May

Content analysis & insights generation

- Content analysis using:
 - **Unsupervised Latent Dirichlet Allocation (LDA)** to review trending topics
 - **Word embedding methods** to identify the volume of conversation on pre-identified topics
- Apply **ex-post survey weights** to improve representativeness of sample

Robustness: cross-validation against other online survey data collection efforts

Institution	Timeline	Methodology	Sample	Modules
 PREMISE™	April 2, 2020 - now; Recurs every 3 days.	<ul style="list-style-type: none"> - Utilize global network via PREMISE crowdsourcing data collection phone app - Dynamic monitoring of changes on weekly basis 	2,749 Indonesian respondents (as of May 14, 2020)	Economics, social, health condition, concern about COVID-19
 J-PAL	March 29, 2020 - now; Recurs every week.	<ul style="list-style-type: none"> - Online using google survey, - Used Susenas 2019 to create ex-post sample weights - Results are pooled unless otherwise stated 	500 respondents every week, two groups: 1) 200 respondents aged 18+, 2) 300 respondents aged 35+	Employment, Food insecurity, Migration, Government assistance program, and Healthy lifestyle changes
 BADAN PUSAT STATISTIK	April 13-20, 2020	<ul style="list-style-type: none"> - Online survey using social media ads to re-direct to BPS official website (TBC) 	70,000+ (TBC)	Perception and personal behavior, Collective behavior, Wellbeing related to COVID, Economic impact, General (demographic)
 THE UNIVERSITY OF WARWICK Int. Academic Research Consortium	March 20, 2020 - now	<ul style="list-style-type: none"> - Online survey shared via various social networks - Website (covid19-survey.org) - Results are weighted based by country-level weights 	113,083 respondents worldwide; 1,618 respondents from Indonesia. (Dataset as of May 1, 2020)	Past and future behaviors, Personal attitudes about coronavirus measures, Perception about other's beliefs about coronavirus measures, Financial sanctioning of risky behaviors, Case predictions, Perceptions about government/public responses and efficacy; Worries battery; Depression questionnaire; Personality battery; Personal information
 WB Online Survey	April 20 – May 20, 2020	<ul style="list-style-type: none"> - Online survey using 5 sets of questionnaire, each 10 questions - Target audiences by age group, gender, location - Used Susenas 2019 to create ex-post sample weights 	3,375 (Java-Bali: 2,501, Jakarta: 659, Other regions: 215)	Knowledge, behaviors, concerns, vulnerability, sentiments, demographics

FINDINGS

AWARENESS AND BEHAVIOR CHANGE

Knowledge: In populations surveyed awareness is high, use of social media for information is common

Citizens already have high awareness about COVID-19



Have heard about COVID-19



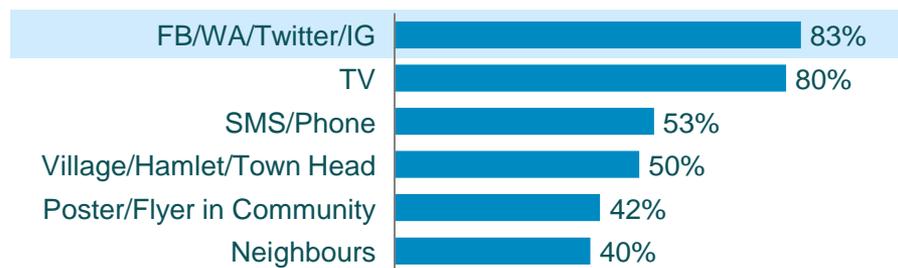
Have received info about social distancing & self-isolation



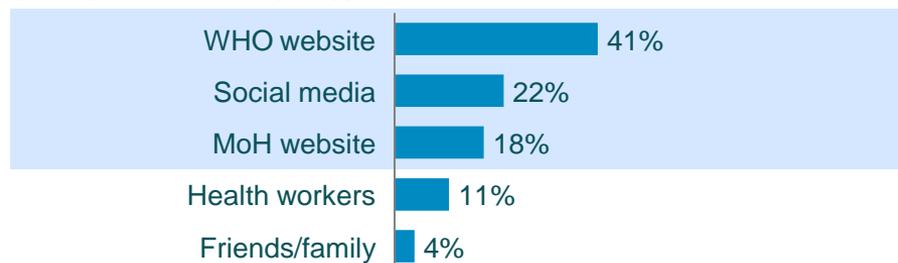
Know the three main symptoms

Social media is used as prominent source of information

>80% received info about social distancing on social media,
88% of them believe that it is reliable



Official website and SM are prominent source used to fact-check



Source: online survey; April 20th - May 20th, 2020; N=3,375 (Java-Bali: 2,501, Jakarta: 659, Other: 215)

Many popular SM posts focus on social distancing measures and staying at home



On Twitter: Most retweeted post comes from regular users, #stayathome is the most popular topic

Non-celebrities Twitter user

<p>Pribumi Kawe @jr_kw19 Dengerin pesan Dr Tirta</p> <p>Ini bukan saatnya kita menyalahkan pemerintah menyalahkan presiden menyalahkan menkes tp saatnya pemerintah dan dokter kita membantu dgn tinggal dirumah</p> <p>Jng sampai Indonesia kayak Italia krn warganya dablek susah diatur #DirumahAjaCekalCorona</p> <p><i>Advice to stay at home</i> Likes: 3.5K</p> <p>2,471 people are talking about this</p>	<p>Milk & Mocha @mikmochabear Someone is getting bored</p> <p>--- "Don't want to go outside" and "Forbidden to go outside" feel so different 😊 --- Feel free to mention Your loved ones, stay safe ❤️ #mikmochabear#WorkFromHome#dirumahaja</p> <p><i>Cartoons picture which encourage staying at home</i> Likes: 2.8K</p>	<p>Cookpad Indonesia @Cookpad_ID Pengumuman.</p> <p>Merespon situasi saat ini dan mendukung program #dirumahaja -- Cookpad memberikan akses gratis 30 hari ke layanan premium melalui Aplikasi Android/iOS.</p> <p><i>Cooking recipes for #Stayathome activities</i> Likes: 1.7K</p>	<p>dwi @DedicatedFanboy Dekor kamar biar aesthetic dan bikin nyaman #dirumahaja</p> <p>A Thread</p> <p><i>Room decoration advice for #Stayathome activities</i> Likes:3.5K</p>
--	---	---	---



On Instagram: Most liked post usually comes from celebrities with topics related to the public figure's life on quarantine or entertainment videos

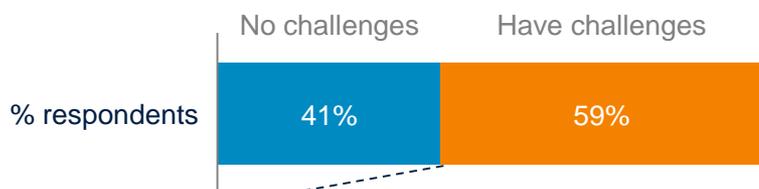
Celebrities Instagram User

<p>Indragiri @20666666</p> <p><i>Travel Parody video</i> Likes:85K</p>	<p>baimwang @21616666</p> <p><i>Online breakfasting</i> Likes: 96.2K</p>	<p>ashanty_ash @22016666</p> <p><i>Travel Parody video</i> Likes: 181K</p>	<p>nezerzahab @22016666</p> <p><i>Health workers parody video</i> Likes:182K</p>	<p>mesterorbuizer @61616666</p> <p><i>Home exercise</i> Likes:180K</p>
---	---	---	---	---

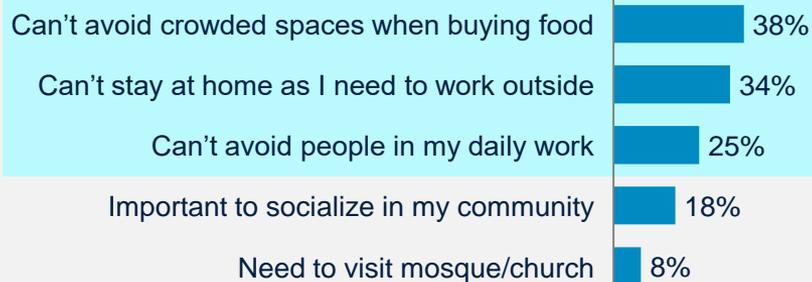
Source: Twitter and Instagram, 10-13 May 2020 (latest update). [Click here](#) to visit Most Liked Post on Twitter and [Click here](#) to visit Most Liked Post on Instagram

However, challenges remain in following public health, social distancing, and self-isolation measures

~60% indicated to have challenges in following prevention measures¹



For those with challenges, top challenges are typically due to daily work and buying food



Only the top 5 options are shown

Some of them said it has impacted their income and jobs



Needed to return to village / hometown since they have lost their jobs



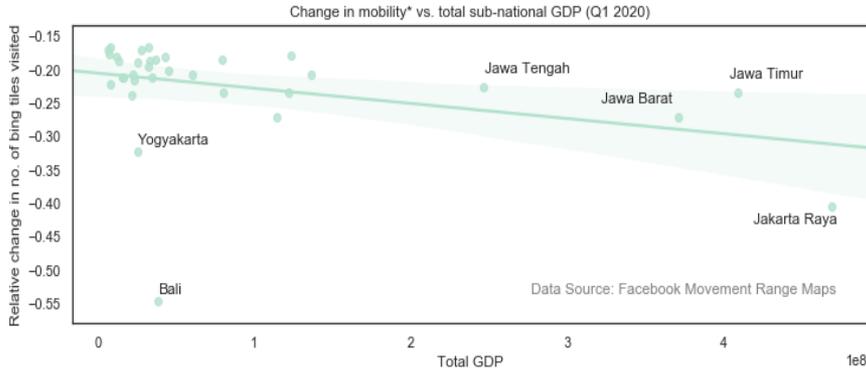
Mostly an issue for males 35-50 years old



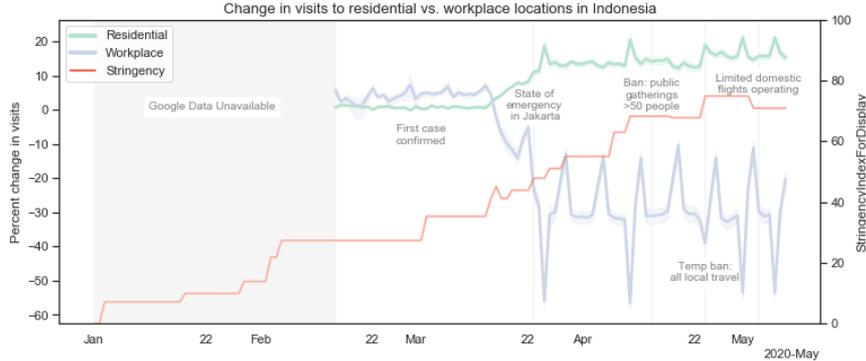
All challenges are more severe for **lower income group**

Our survey result is also confirmed with mobility insights from big data analytics

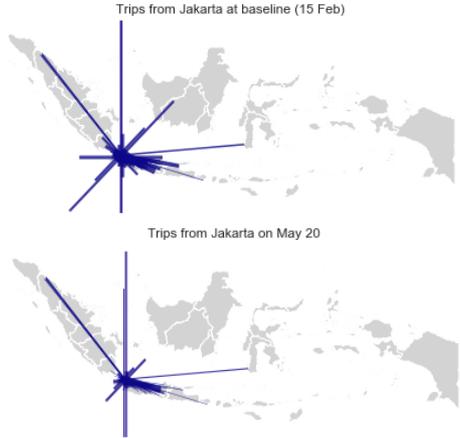
LSRR and travel ban policy have reduced the mobility of people in hotspot locations..



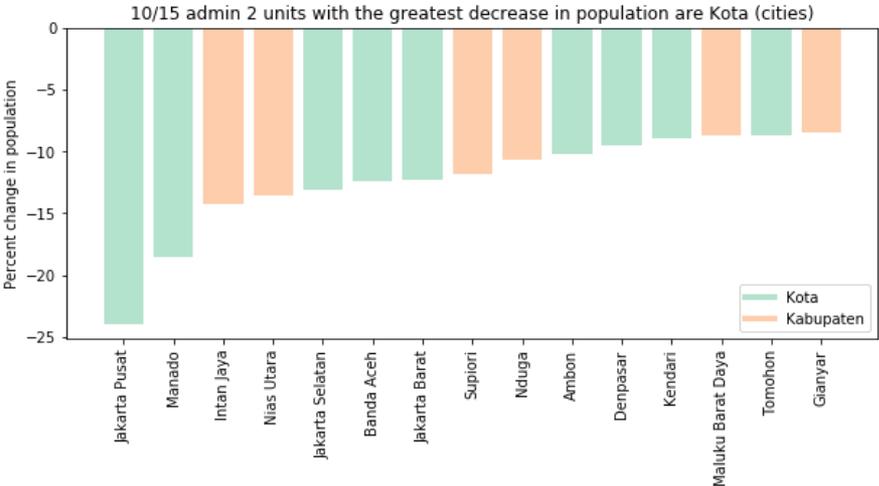
.. and cause most people to spend more time at home...



..with fewer trips in and out of **Jakarta**



Despite that, there is still evidence of **urban flight**..



Comparison of findings across surveys



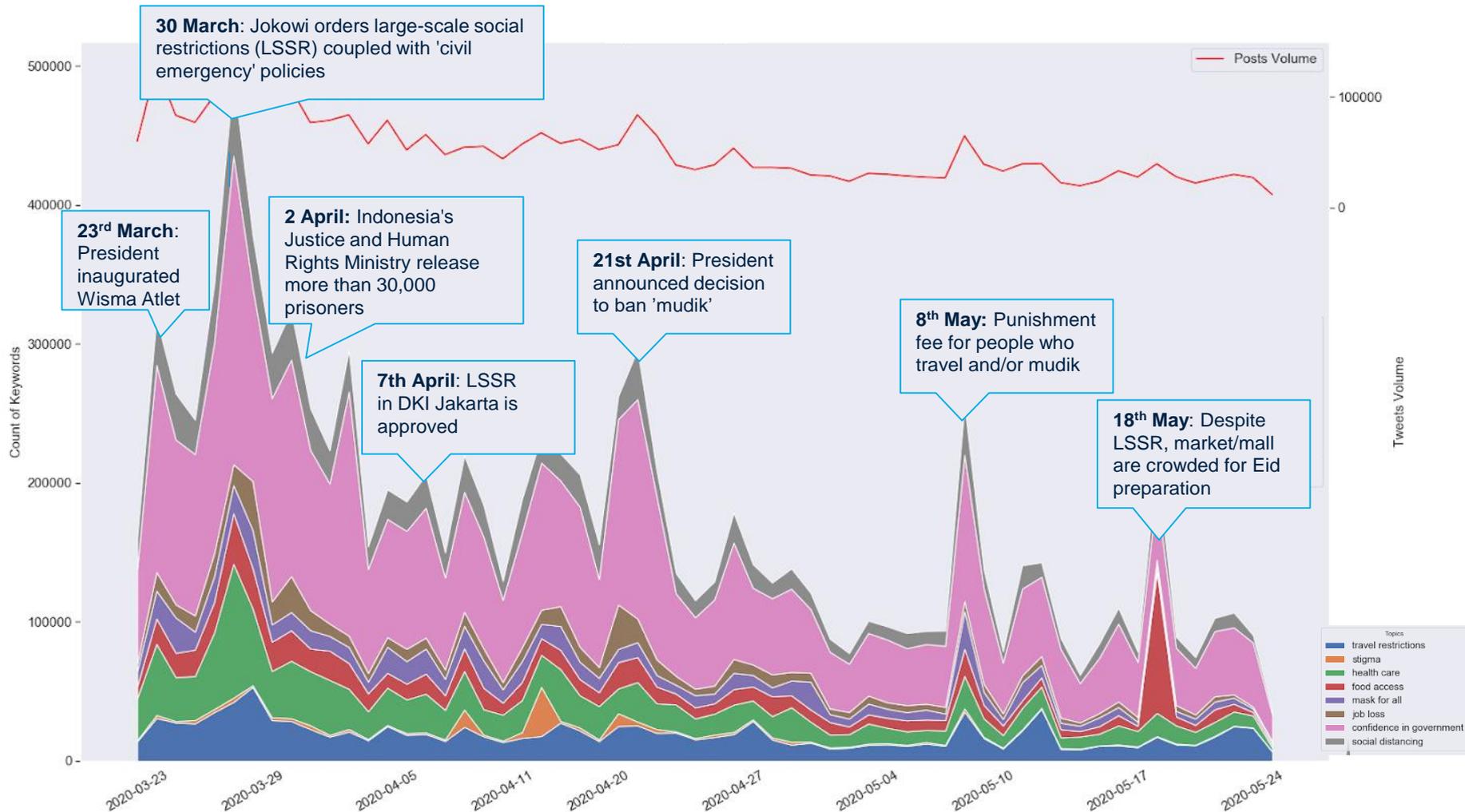
Topic	WB Online Survey	Premise	JPAL	BPS	Warwick
Knowledge	High level of awareness	✓	N/A	High level of awareness on LSSR	N/A
Health behavior change (sanitation and hygiene)	High level of change in hygiene behavior	✓	✓	✓	✓
Challenges with social distancing	Moderate (only 41% have no problem implementing)	Moderate (29% can't practice social distancing at work, 20% can't when using public transport)	Strong (only 21% made changes to improve social distancing)	N/A	N/A
Migration	Low	N/A	High – especially within same kabupaten/city	N/A	N/A

Legend: ✓ = similar result to our survey

CONCERNS AND VULNERABILITIES



Spikes in Tweet volume respond to the recent policy or announcement enacted by the government

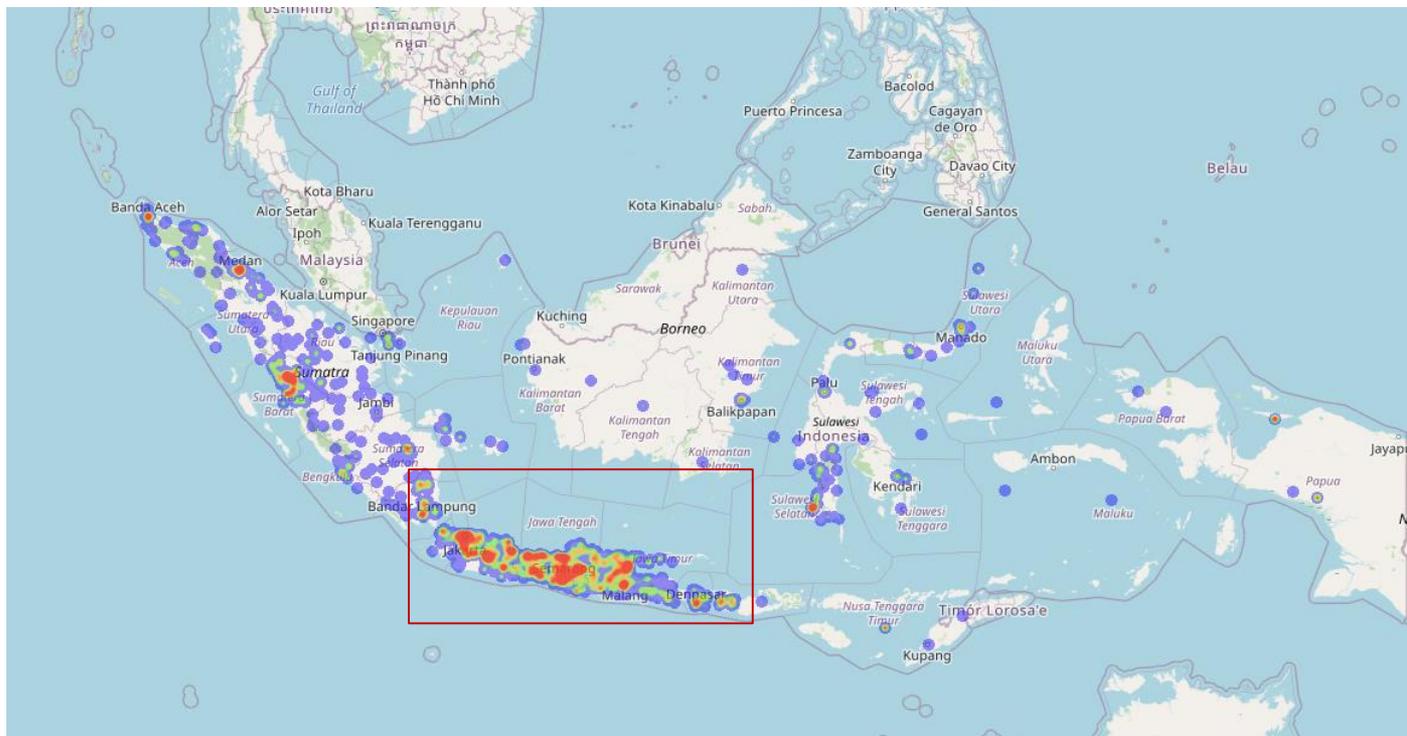


Source: Policy Issue Identification Word Embedding

Almost all provinces in Indonesia discussed COVID-19 in SM

But there was more discussion in areas with high case loads

Heatmap of Fraction of Tweets Mentioning Corona or COVID-19



[Click here for more information](#)

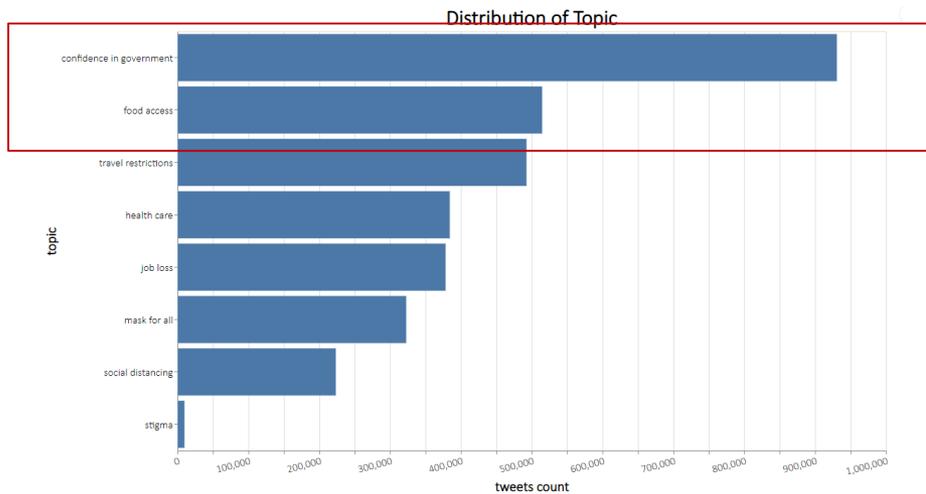
Insights:

- COVID19 is mostly discussed in Java-Bali region where 66%¹ cases COVID19 are located
- Low volume of conversation on COVID19 in Eastern part of Indonesia

Government handling of situation, food access and health care are salient topics discussed in both Twitter and Instagram



Distribution of topic discussed on Twitter



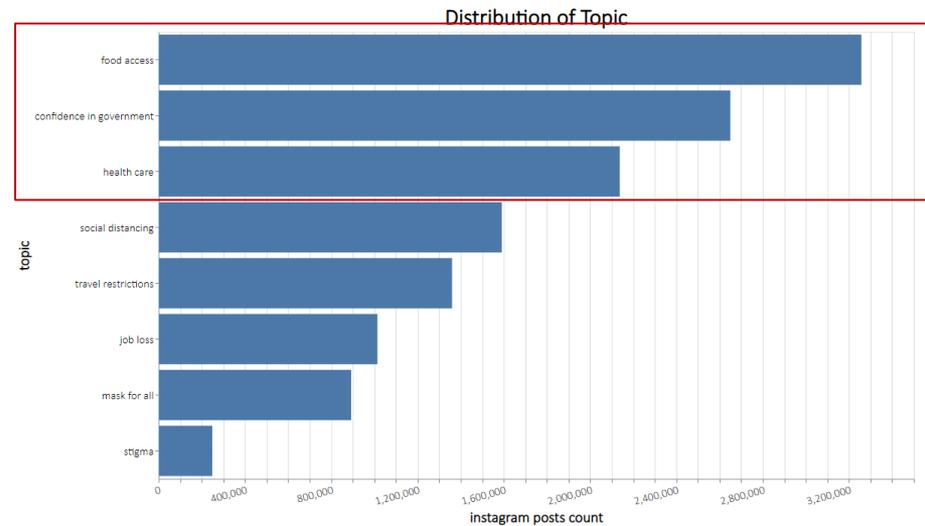
[Click Here](#) to See the Interactive Visualization

Data collected from March 23th – May 24th

#Tweets: ~ 3.3 million



Distribution of topic discussed on Instagram



[Click Here](#) to See the Interactive Visualization

Data collected from March 23th – May 24th

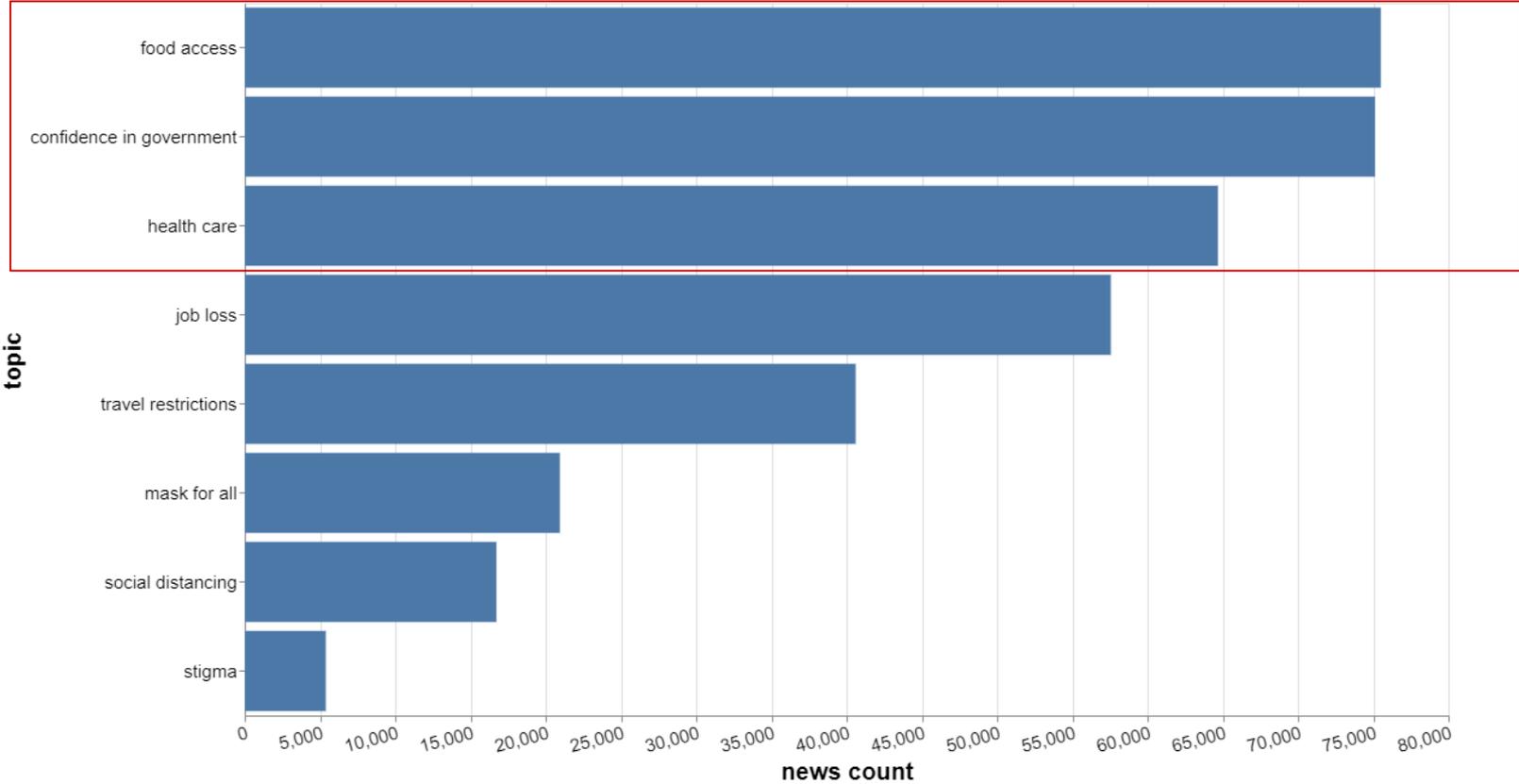
#Posts: ~ 5.7 million

Note: There is strong degree of overlap of topics discussed – see annex slides.

Source: Policy Issue Identification Word Embedding

News media resonates with SM: Food access, confidence in government and health care are salient topics

Distribution of topic discussed on GDELT News Data



Source: Policy Issue Identification Word Embedding

Data collected from March 23th – May 17th
#News: ~ 220k



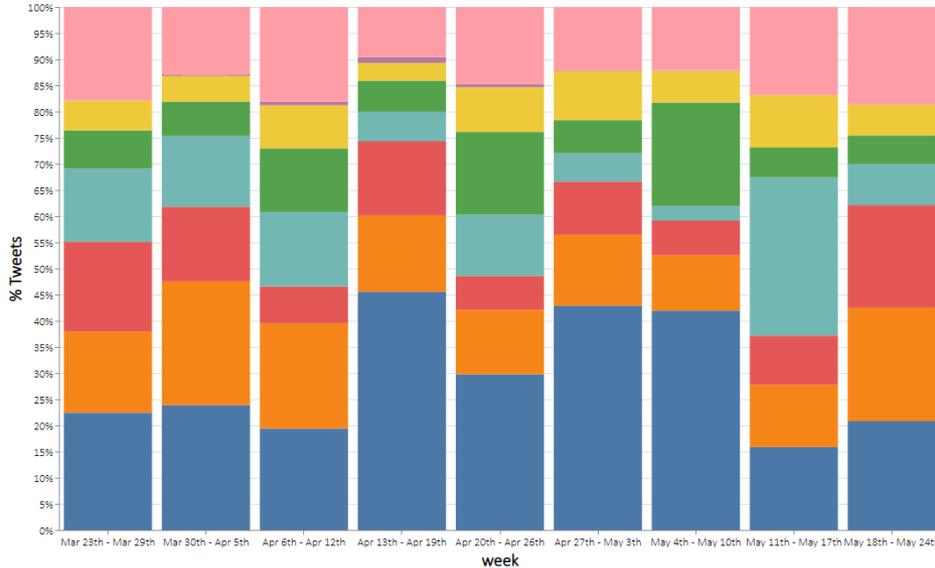
Note: There is strong degree of overlap of topics discussed – see annex slides.

Discussions have shifted over time in both Twitter and Instagram



Attention to job loss was significant in mid-May; food access and health care are more current concerns

Distribution of Topic Weekly



Date: March 23th – May 24th

#Tweets: ~3.3 million

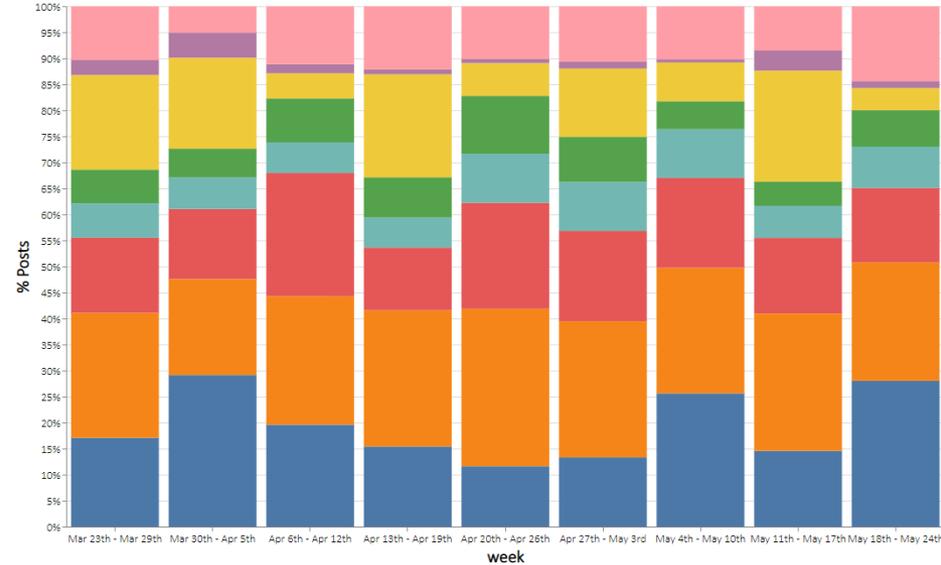
[Click Here](#) to See the Interactive Visualization

Source: Policy Issue Identification Word Embedding



Concerns on food access remains high throughout April; discussion on government rose in early May

Distribution of Topic Weekly



Date: March 23th – May 24th

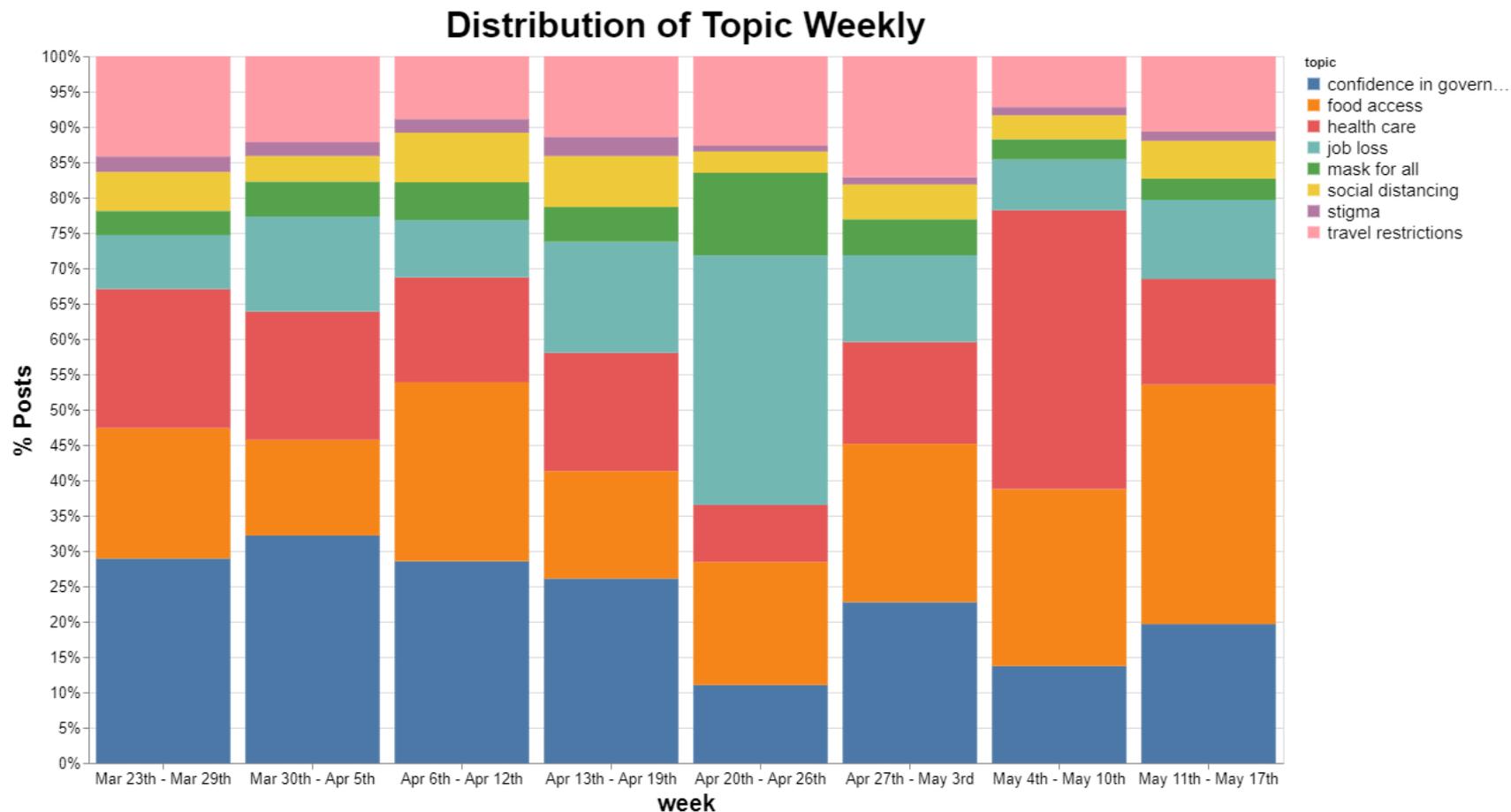
#Posts: ~5.7 million

[Click Here](#) to See the Interactive Visualization

- confidence in government
- food access
- health care
- job loss
- mask for all
- social distancing
- stigma
- travel restrictions

Shifts in discussion also happened in news media:

Confidence in government receives much attention on earlier (late March until mid-April); discussion on economic impacts of COVID-19 peaked recently (job loss peaked on late April and food access peaked on mid-May); health care was a prominent concern in early May



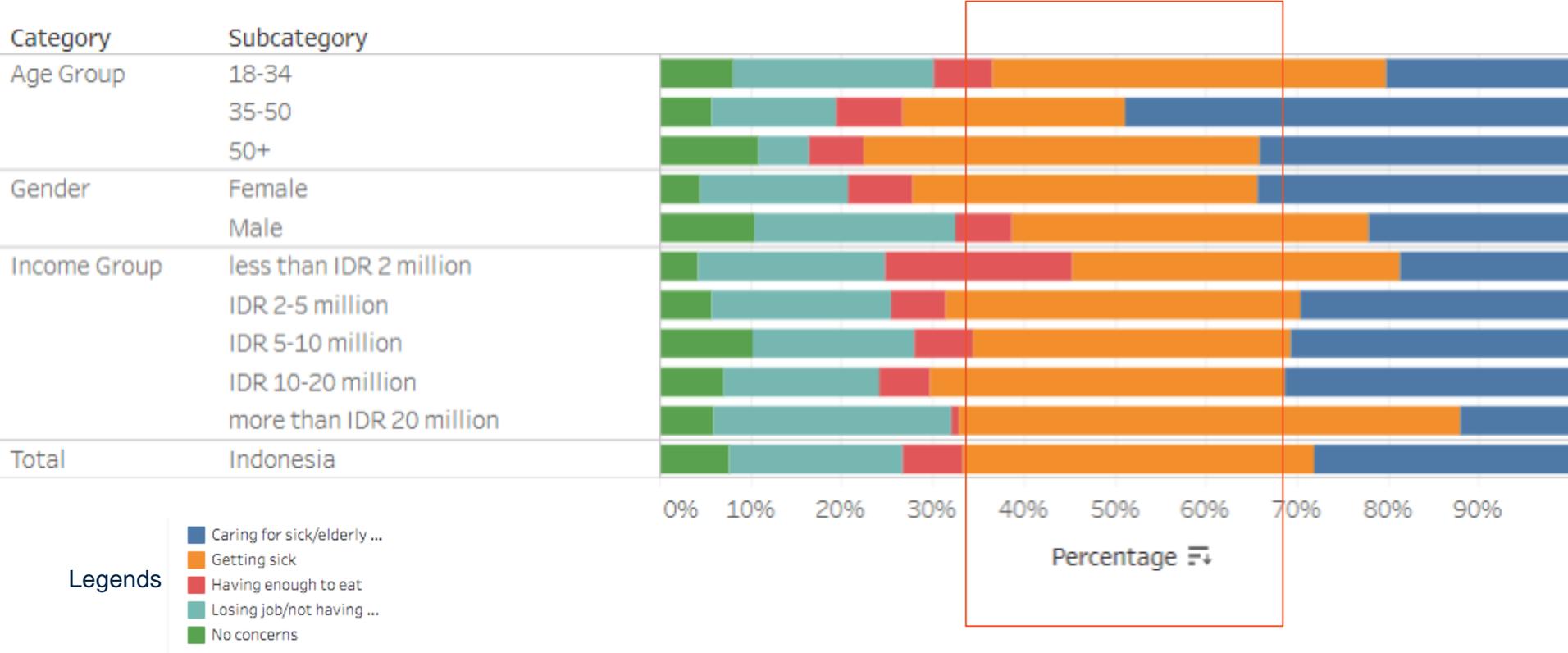
Source: Policy Issue Identification Word Embedding

Concerns raised on SM also found in online survey:

In general everyone is most worried about **getting sick**, then **having enough to eat** and **losing job** are important concerns

Concerns are more acute among low income groups

When asked about healthcare access 77% indicated they could access the care they needed



Legends

- Caring for sick/elderly ...
- Getting sick
- Having enough to eat
- Losing job/not having ...
- No concerns

Health and Economic Concerns are confirmed by other surveys..



Result of health concern in our survey..



Respondents indicated **getting sick** as their **primary concern**¹

PREMISE™ ..is confirmed by Premise with higher magnitude²



are concern that they will **personally contract COVID-19**



are concern about the spread of COVID-19 **in their community**



are concern about **access to adequate health care**

Note: 1. Question: What is your biggest concern in the next week? 2. % of respondents answering 'concern' and 'very concern' 3. Question: Do you work less hours, fewer days? Source: World Bank Online Survey, Premise Indonesia Dashboard, Online Survey on Economic Impact of COVID-19 in Indonesia Results by JPAL from Week 6



Job loss is prominent economic concern in our survey..



Respondents are **working less** – fewer days and hours³



..as also indicated by JPAL survey - **66%** respondents are **not working**



56% of men and **58% of women** were **working before**, but are no longer working

...as are food security and food price concerns



14% of our respondents are prone to food insecurity...¹



35%

indicated that it's due to **price surge**

35%

have **needed help** from friends and family

10%

have **borrowed money**

...which aligns with JPAL and PREMISE study

J-PAL Food insecurity has **higher magnitude** in JPAL study

35%

of households report **having to eat less** than they should often

23%

households ate as usual in the last week

PREMISE™

44%

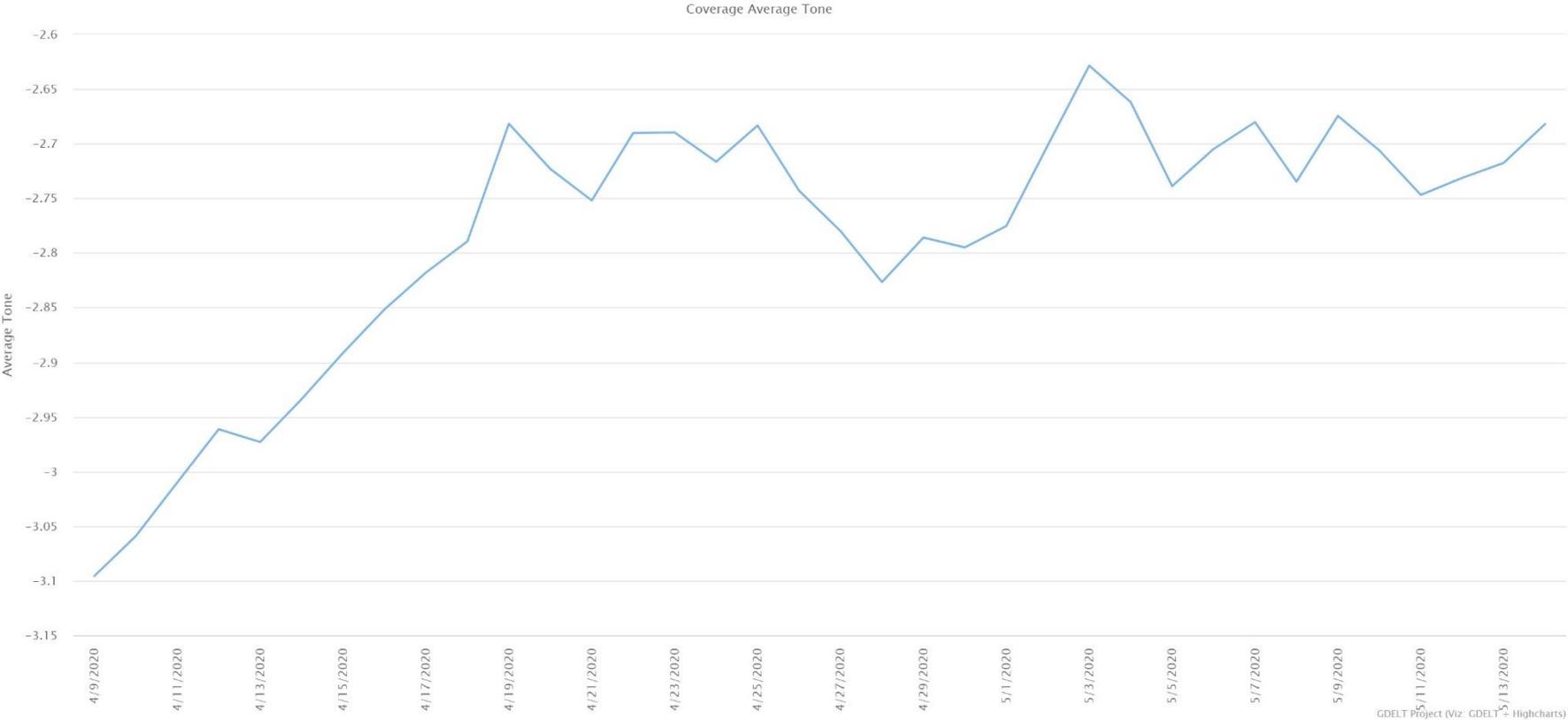
people are affected by **sudden price surge**²

Note: 1. Question: Did you or any household member eat less, substitute foods, skip meals, or not eat, even when hungry?? 2. % of respondents answering having 'moderate effect' and 'major effect'. Source: World Bank Online Survey, Premise Indonesia Dashboard, Online Survey on Economic Impact of COVID-19 in Indonesia Results by JPAL from Week 6



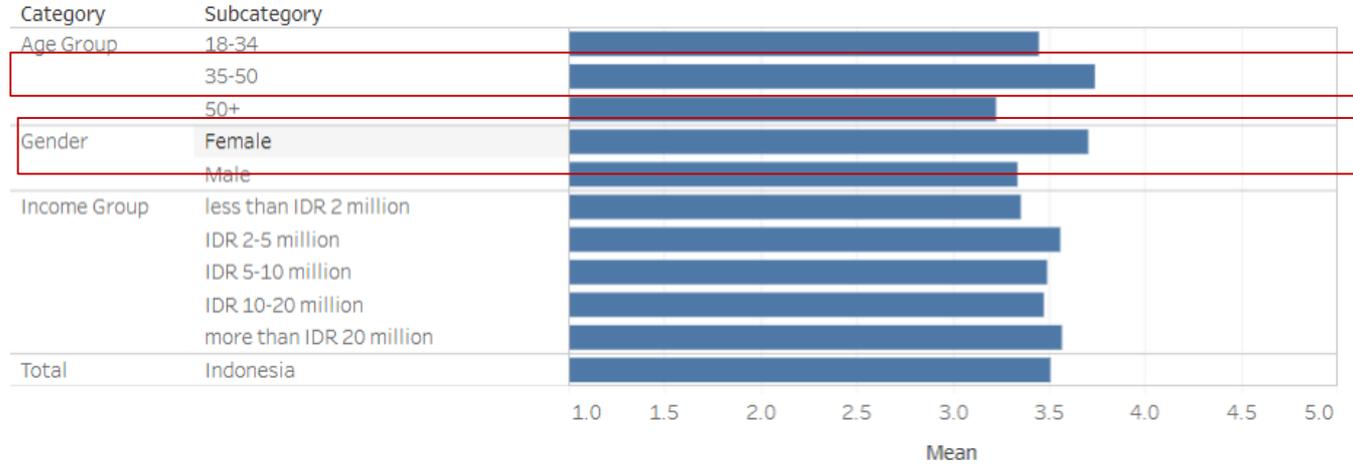
SENTIMENT ANALYSIS

The average tone of news media has been negative but is moderating

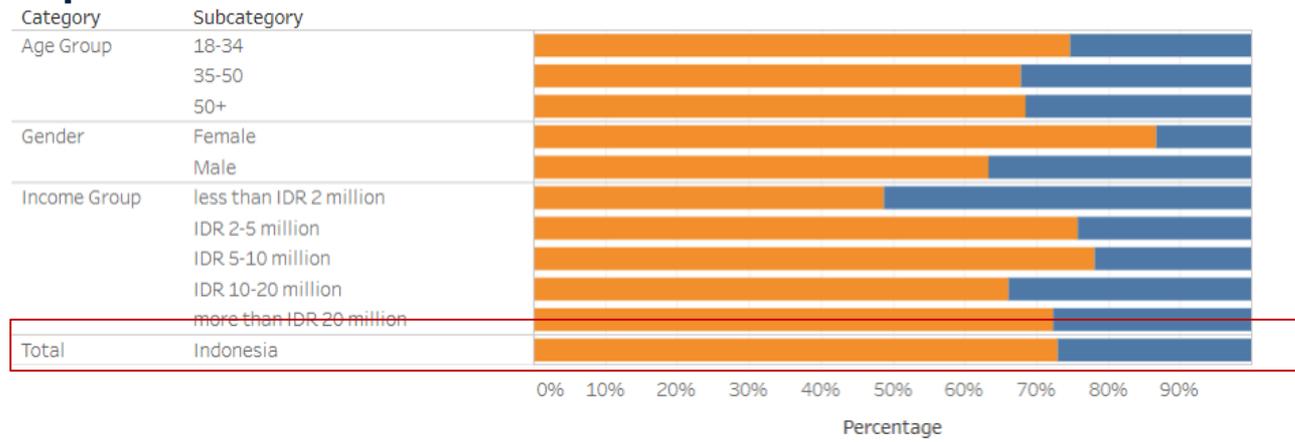


However, many survey respondents indicated they were **moderately to significantly worried** about the situation, with **73% concerned about unrest**

People aged 35-50 and women are more worried regarding the current situation



73% of respondents are concerned about social unrest



Sentiment to government:

People show more confidence in provincial than central government

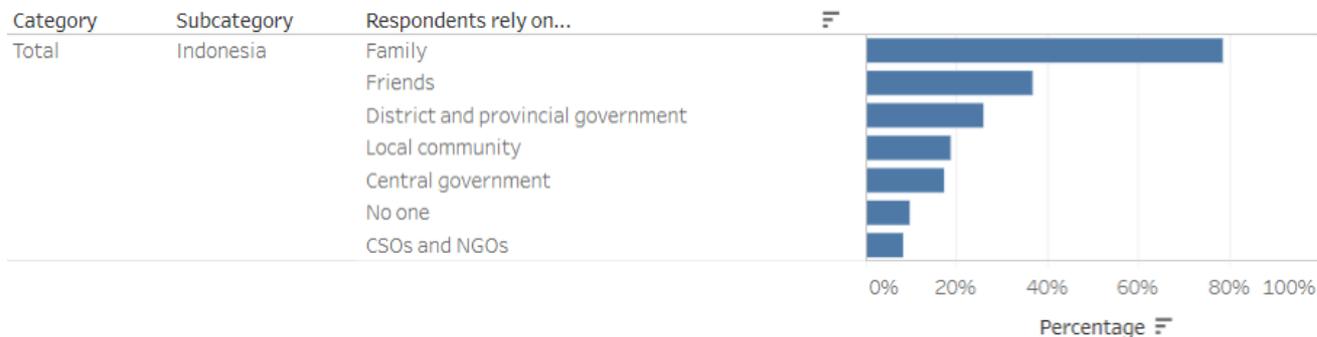
Do you think the reaction by the central government to the current coronavirus outbreak is appropriate, too extreme or not sufficient?



Do you think the reaction by the provincial government to the current coronavirus outbreak is appropriate, too extreme or not sufficient?



Who do you count on most to support you through this period?

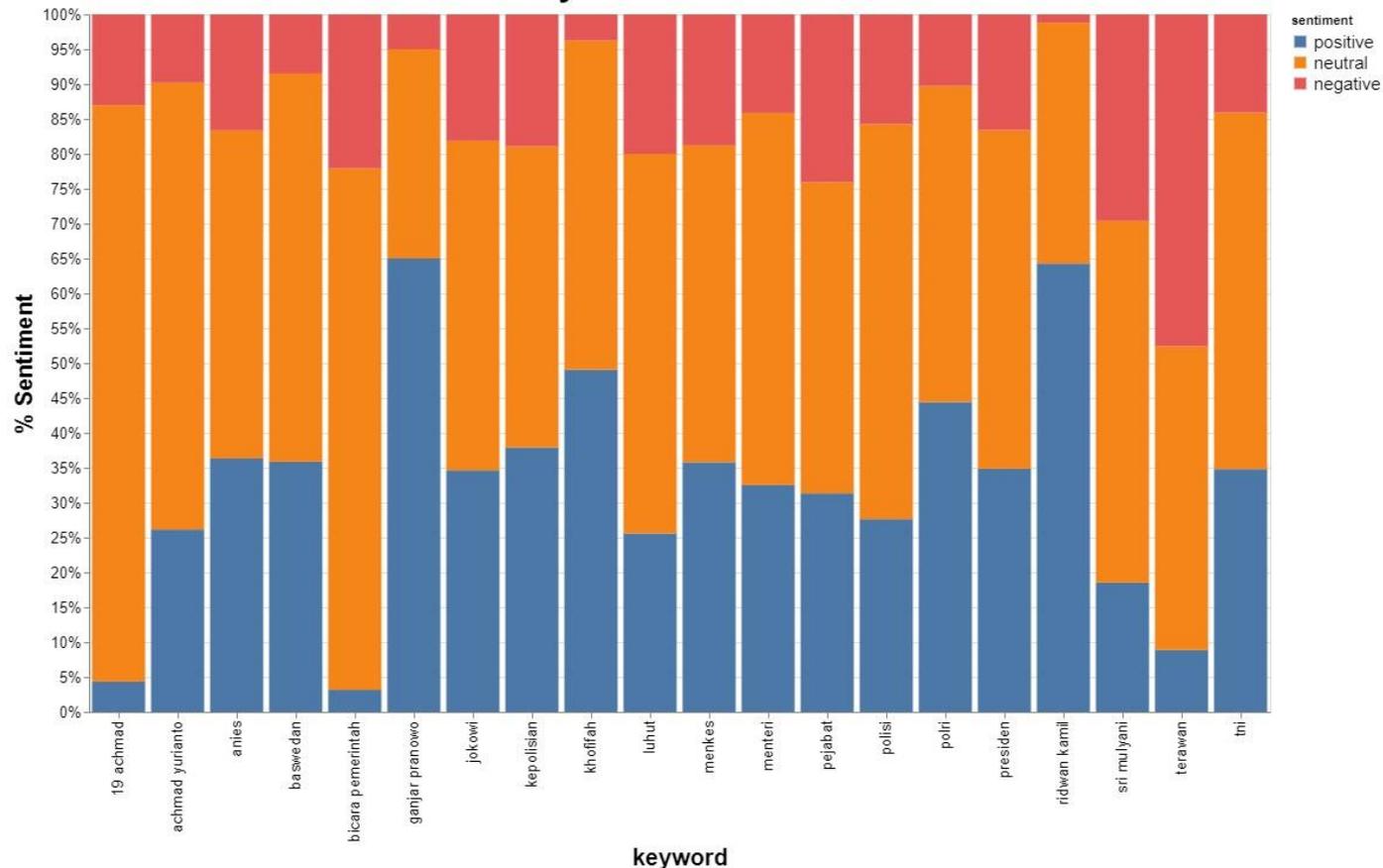




Twitter Topic Keyword Sentiment: Confidence in Government

Twitter users seems to have more favorable sentiment toward local government compared to national government

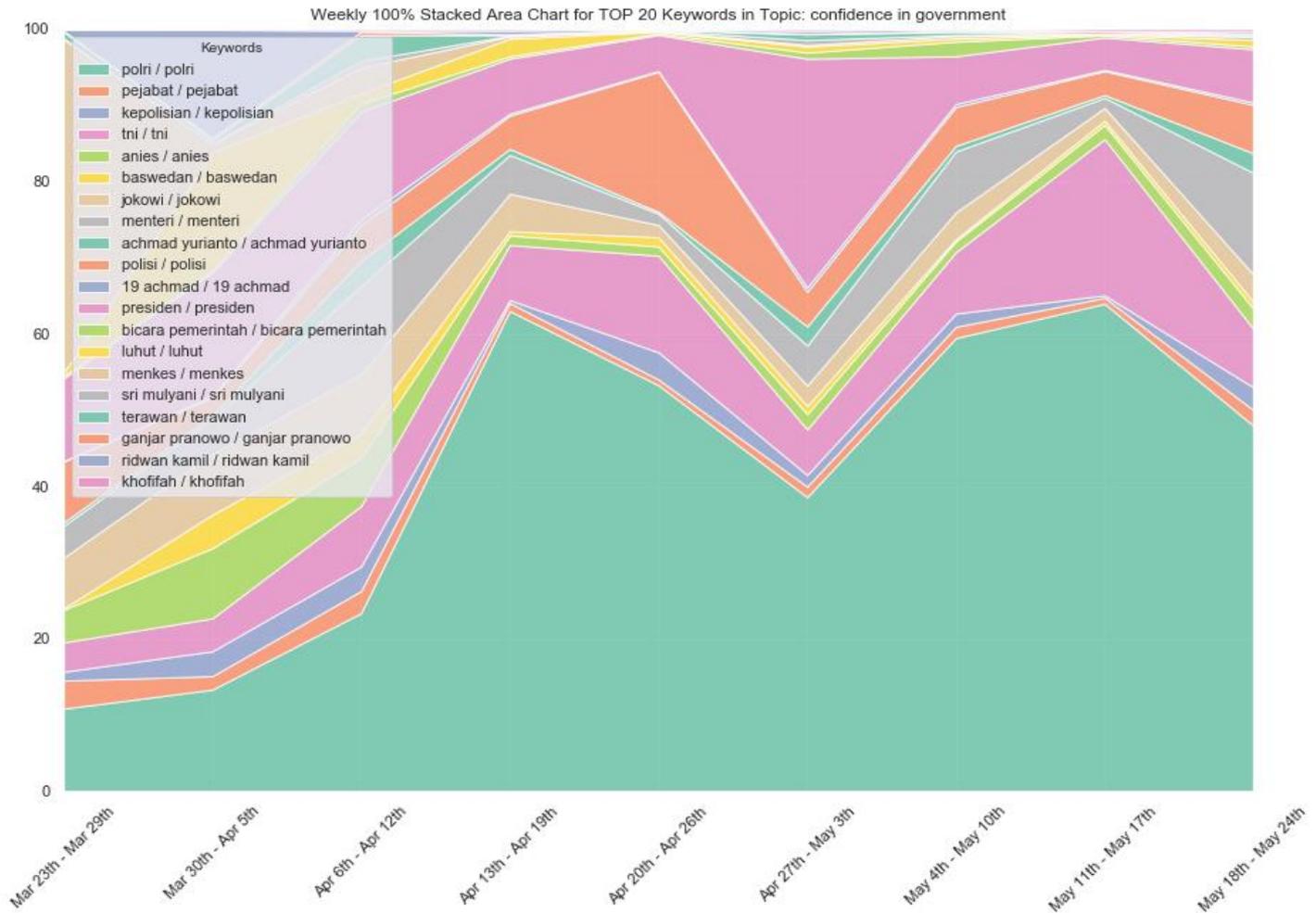
Sentiment of TOP 20 Keywords for Confidence in Government



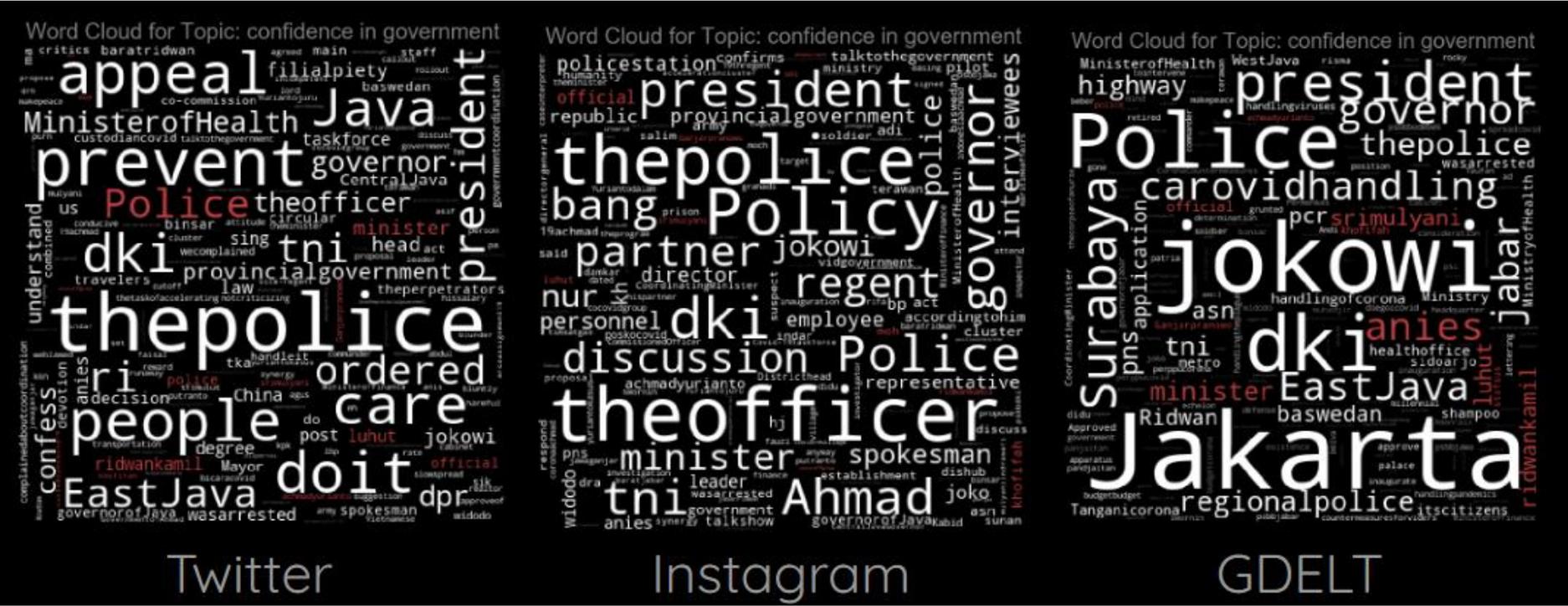


Twitter Topic Keyword Evolution: Confidence in Government

National police remains the most widely-discussed topic on Twitter as they are the main actor to ensure that LSSR is implemented



Comparison of discussion on government across platforms



- Police, national government, and regional government consistently becoming main topic across all platforms
- In mainstream news media, Jakarta and Surabaya are the two prominent cities covered due to the number of cases

Source: Policy Issue Identification Word Embedding

Comparison of results in sentiments: Our results more moderate than Warwick survey but stronger than with PREMISE survey

Public and Government Response Perception (in percentage)

	WB Survey	Warwick Indonesia ¹	Warwick Global ¹	PREMISE ²
Government response is too extreme	Central Government:2 Local Government: 3	0.4	8.3	11
Government response is insufficient	Central Government:68 Local Government: 57	83.8	49.2	22
Government response is enough	Central Government: 30 Local Government: 40	15.8	42.5	65

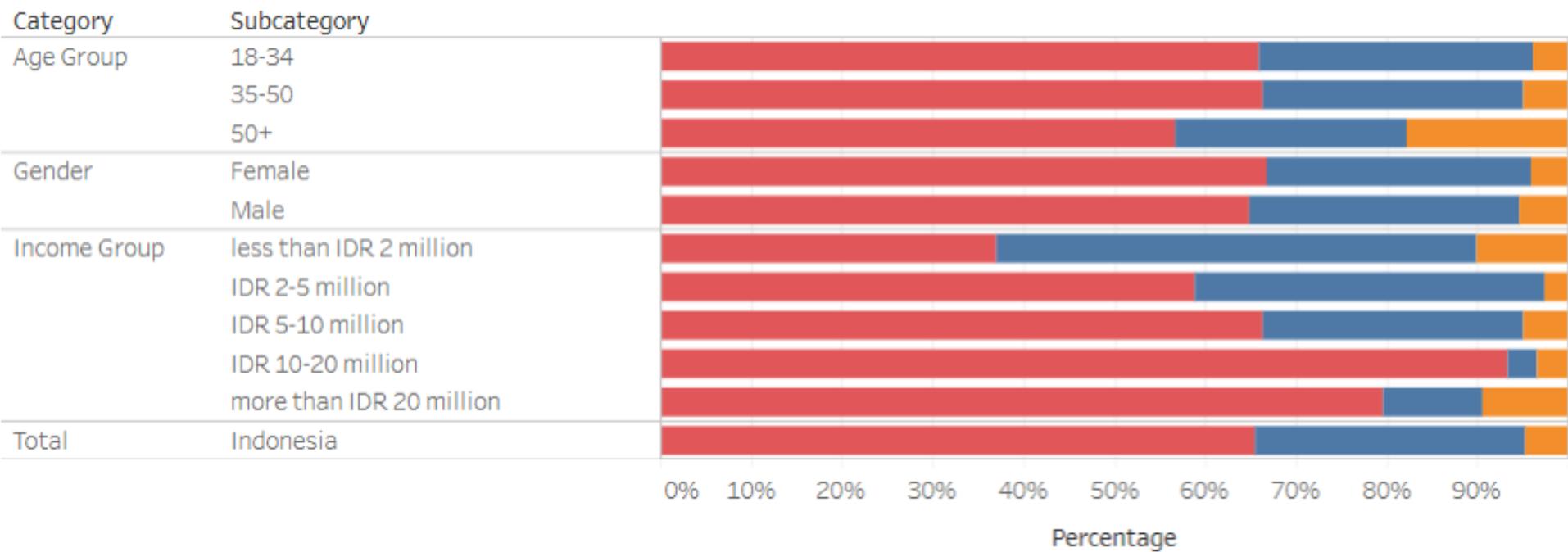
Source: Warwick International Survey on Coronavirus (2020), World Bank staff calculation on Indonesia numbers.

Note: N for Indonesia = 1,618 respondents, N for World = 113,083. Dataset analyzed is May 1, 2020 version and PREMISE Indonesia Dashboard

- **Warwick survey confirms our result that most Indonesians think that government's response is insufficient**
- However, in PREMISE, up to **65%** respondents believe that government's response is sufficient.
- Based on PREMISE survey, **48%** trust their national government more (only **28%** trust their local government more)

Most respondents believe that it will take >2 months for things to get back to normal

Female and people with higher income group tend to believe that it will take longer for the country to recover



- Legends
- 1-2 weeks
 - 1-2 months
 - >2 months

Source: Online Survey. Based on data collected between April 20th - May 20th, 2020; N=3,375 (Java-Bali: 2,501, Jakarta: 659, Other regions: 215)

NEXT STEPS

Future activities

- Continued monitoring of social media and online news
- New rounds of online survey data:
 - In which sectors are job losses occurring?
 - Which food prices are increasing?
 - Awareness and support of government response programs
 - Knowledge of COVID-19 cases
 - Target questions to SM users based on content they are sharing
- Google trends analysis: what does this show us about evolving concerns?
- Credibility and independence of news coverage
- Misinformation assessment

Looking ahead: Policy Implications



Likely **reduction in policy stringency**, especially Java/Jakarta



Early warnings for localized "second wave"



Shift from response to recovery
more nuanced set of issues

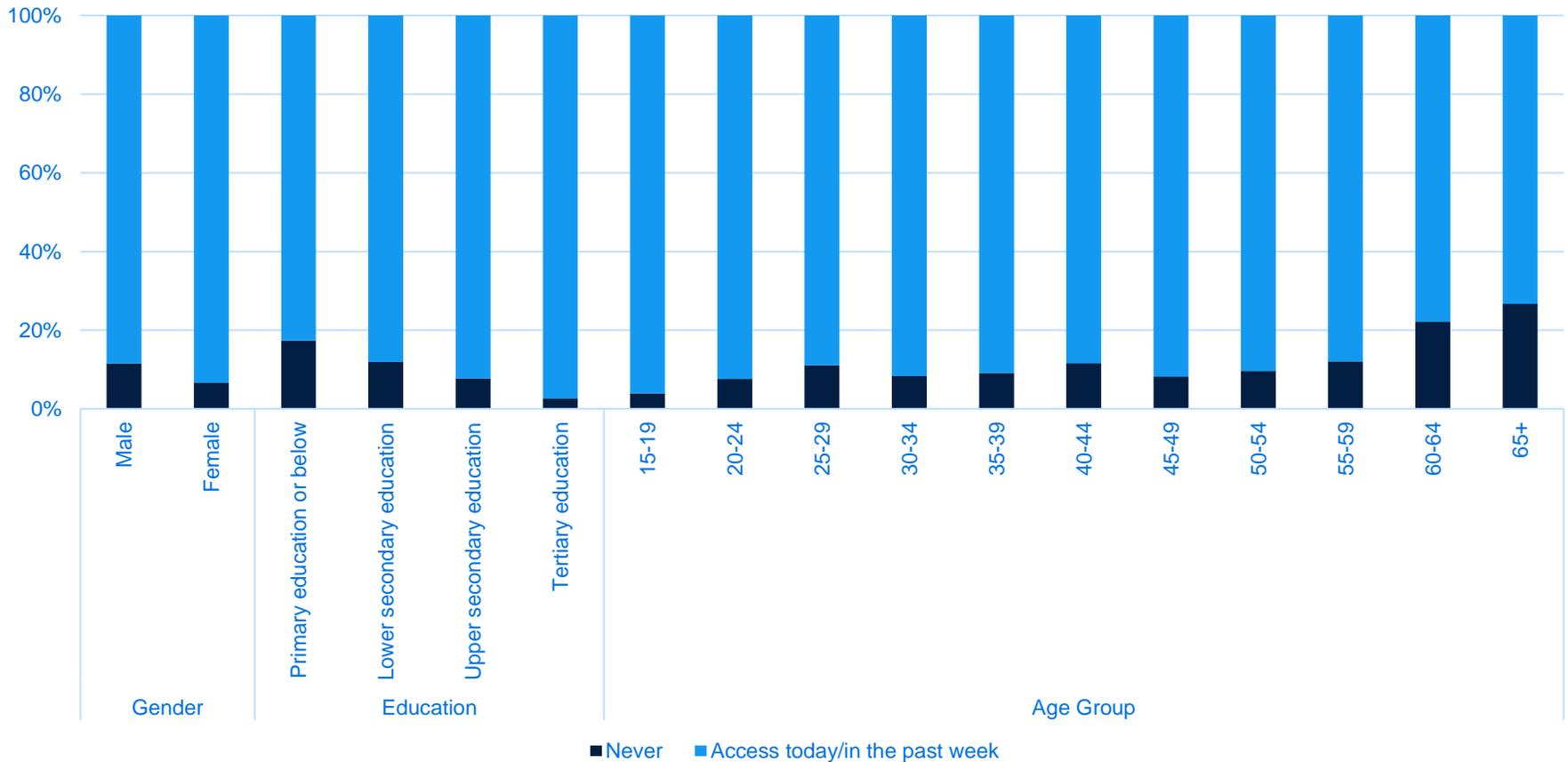


Important to **keep tracking numbers**

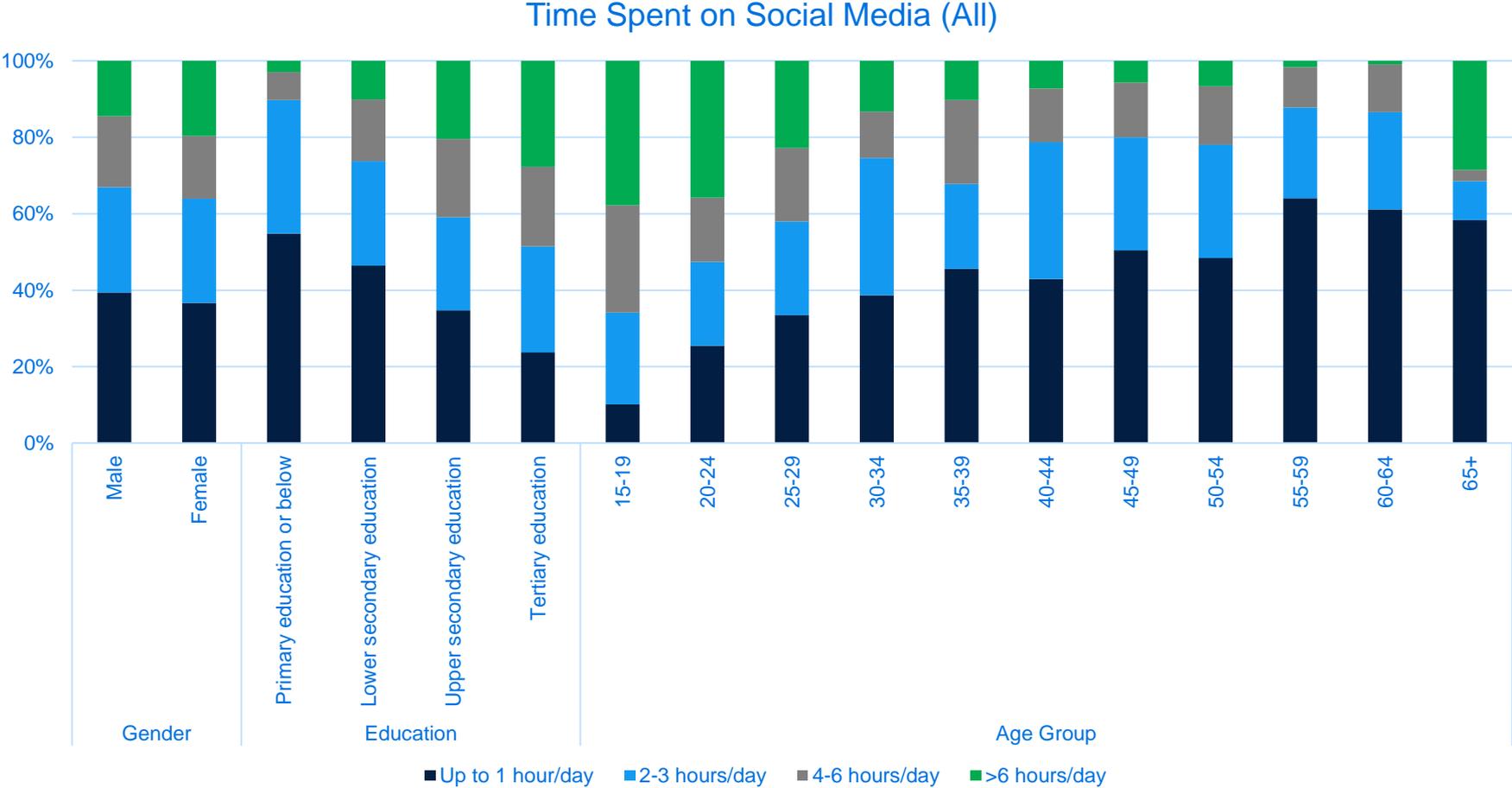
Extra Slides

Social media use is prevalent in Indonesia – over 90% penetration of SM

Accessed Any Social Media Platform



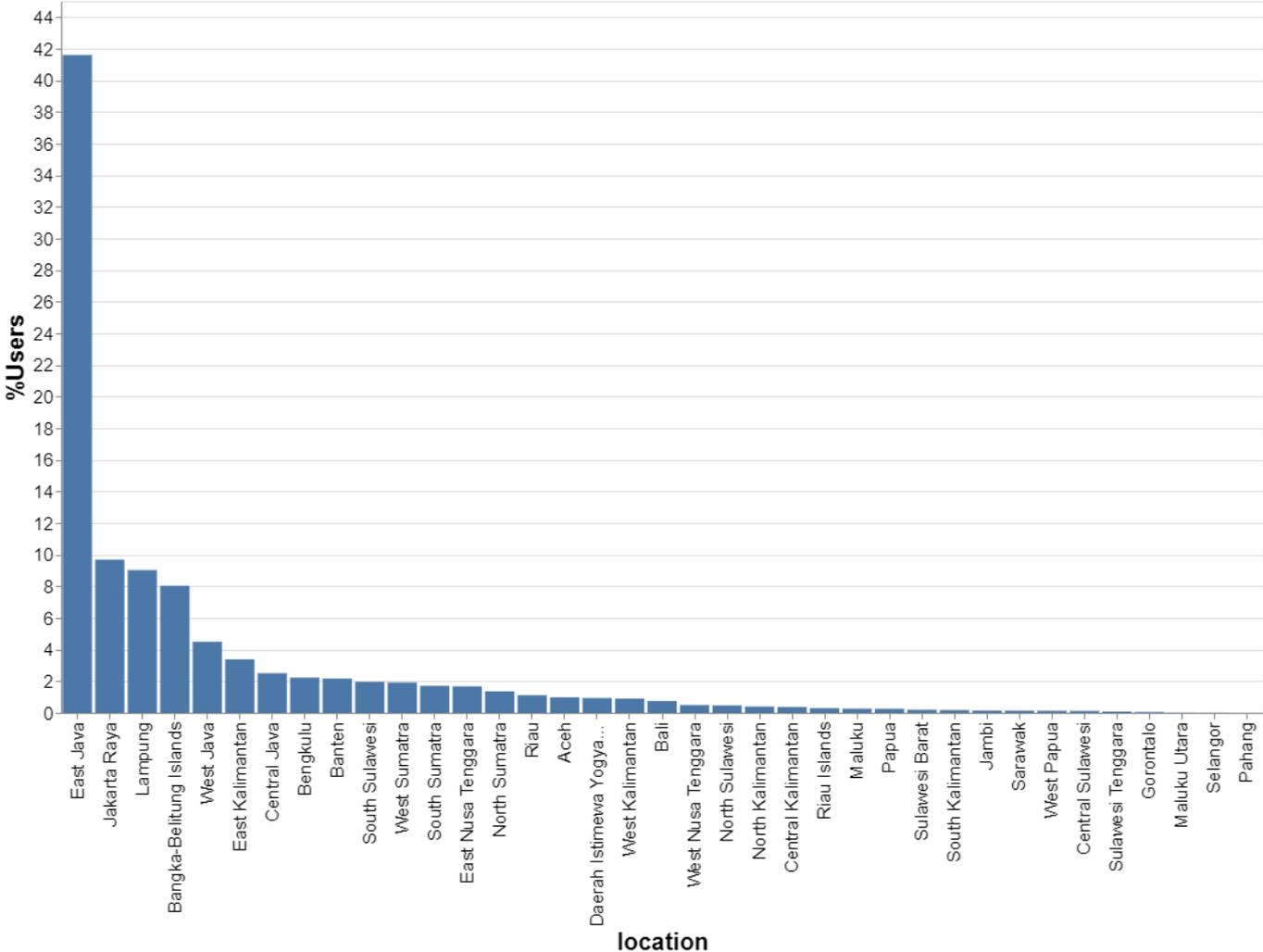
...and intensive – over half of respondents spend more than 2 hours per day on SM



Who we monitor: Twitter Users Demographic Across Region

*12% of twitter users self-report location information

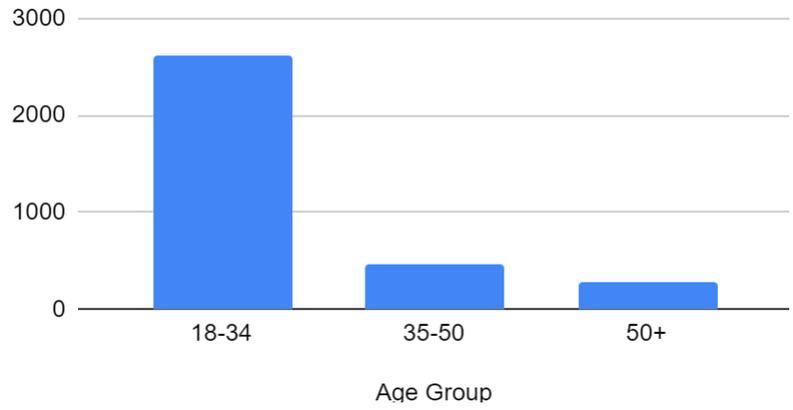
Percentage of Twitter Users Across Region



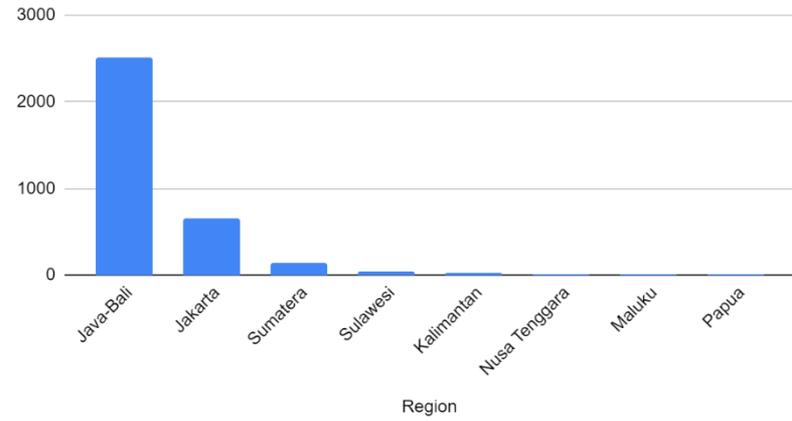
Date: March 23th – May 24th

As per May 20th 2020, we managed to have **3,375** respondents with varied characteristics

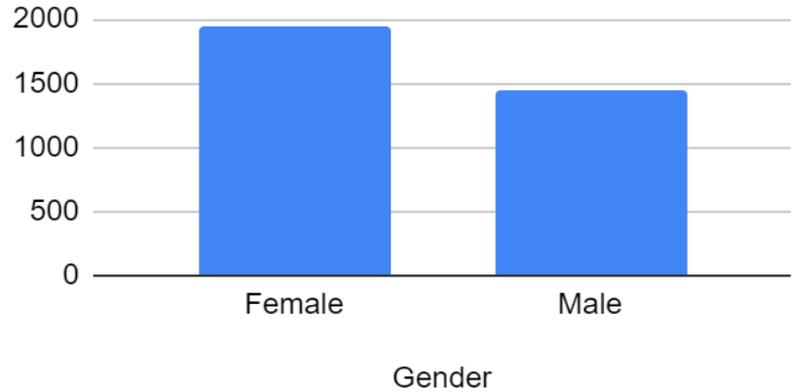
By Age Group



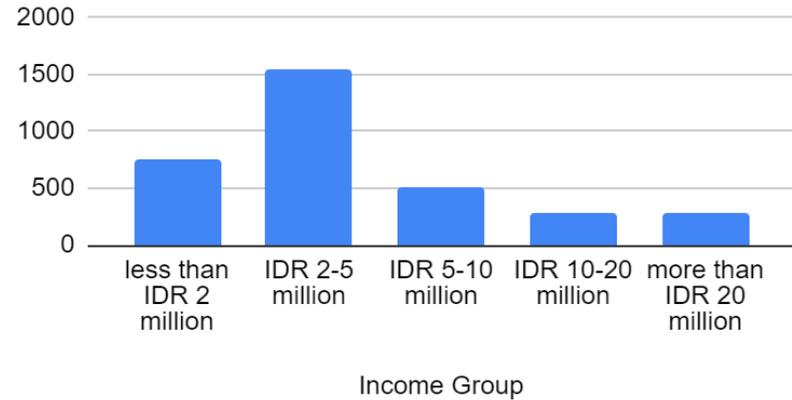
By Region



By Gender



By Income Group

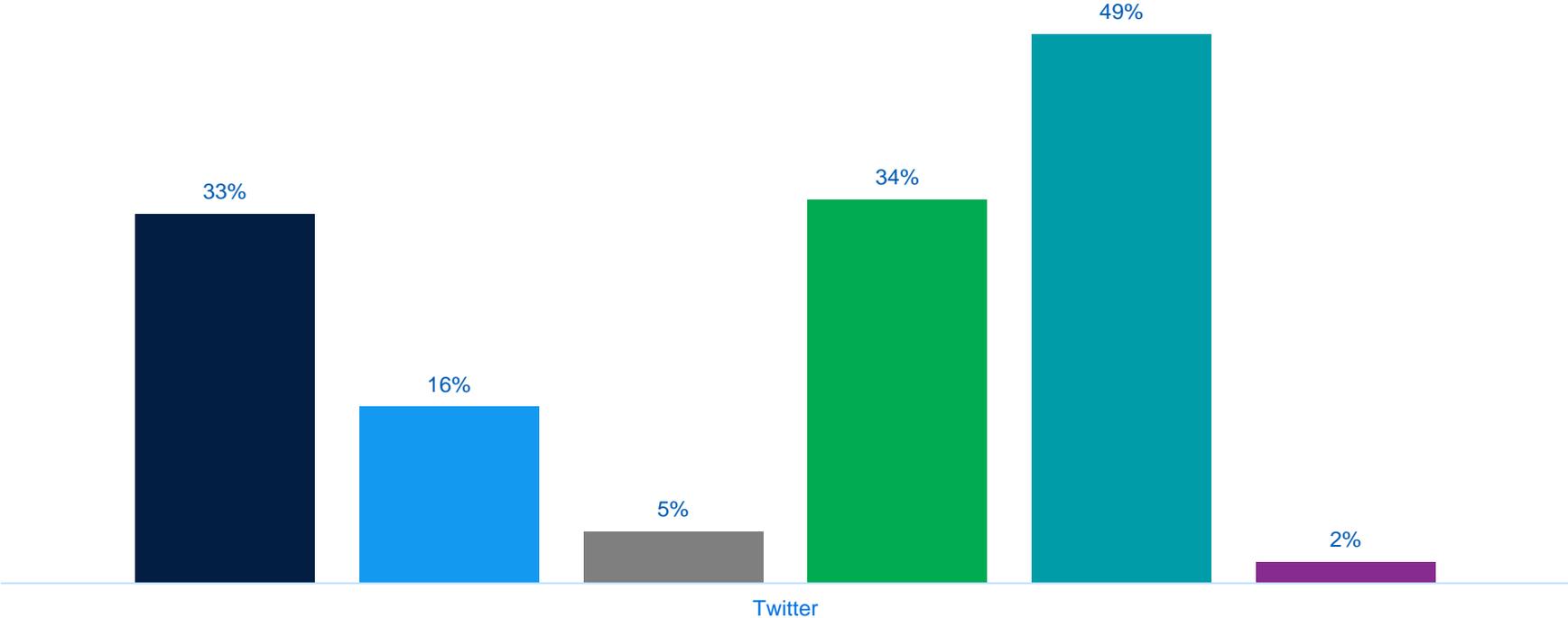


Source: Online Survey. Based on data collected between April 20th - May 20th, 2020; N=3,375(Java-Bali: 2,501, Jakarta: 659, Other regions: 215)

Attention to government activities may reflect the nature of conversations on Twitter

Topics discussed on Twitter

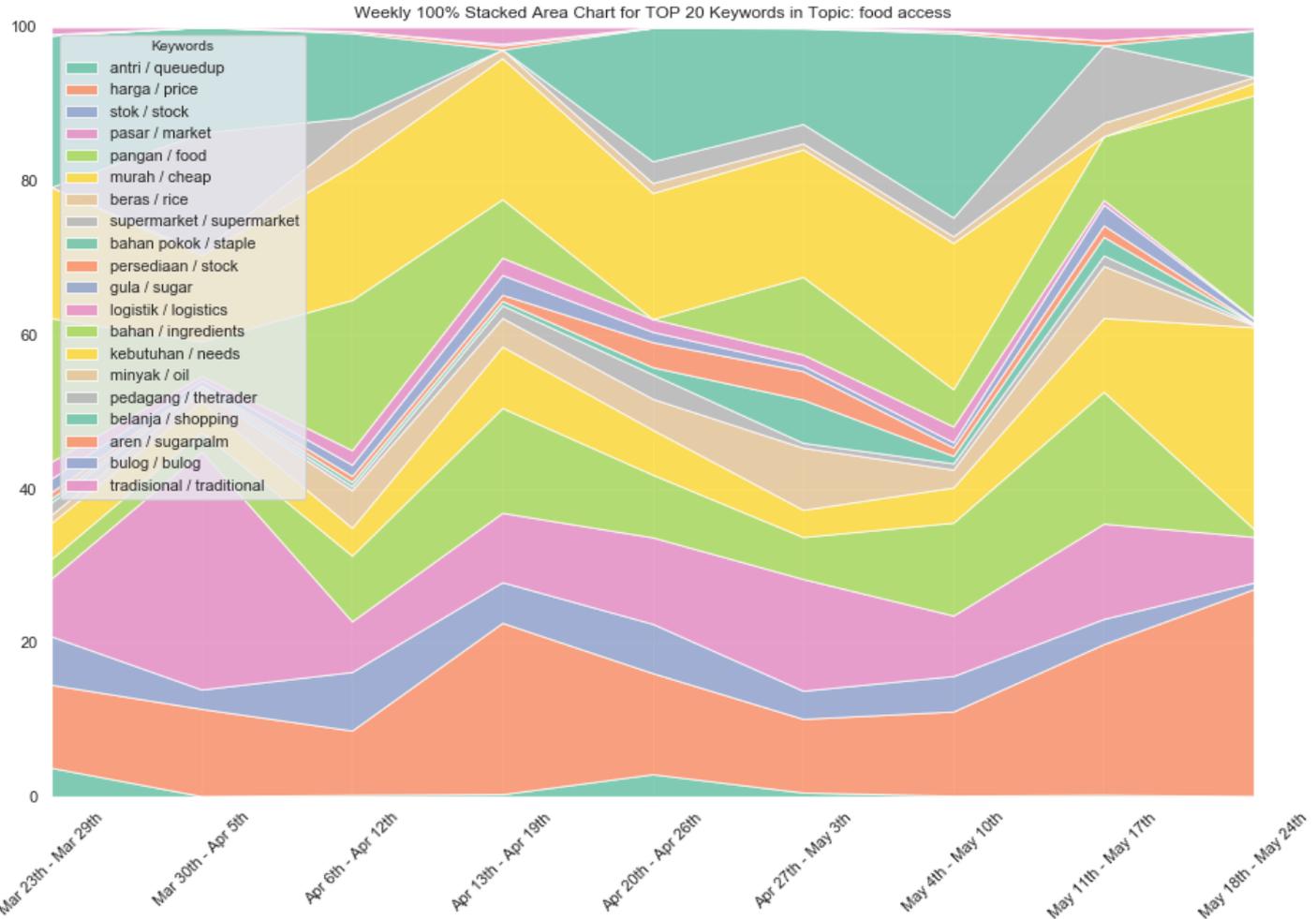
■ Politics ■ Religious matters ■ Gossip ■ Academic/public policy issues ■ Hobby/lifestyle ■ Work/business





Twitter Topic Keyword Evolution: Food Access

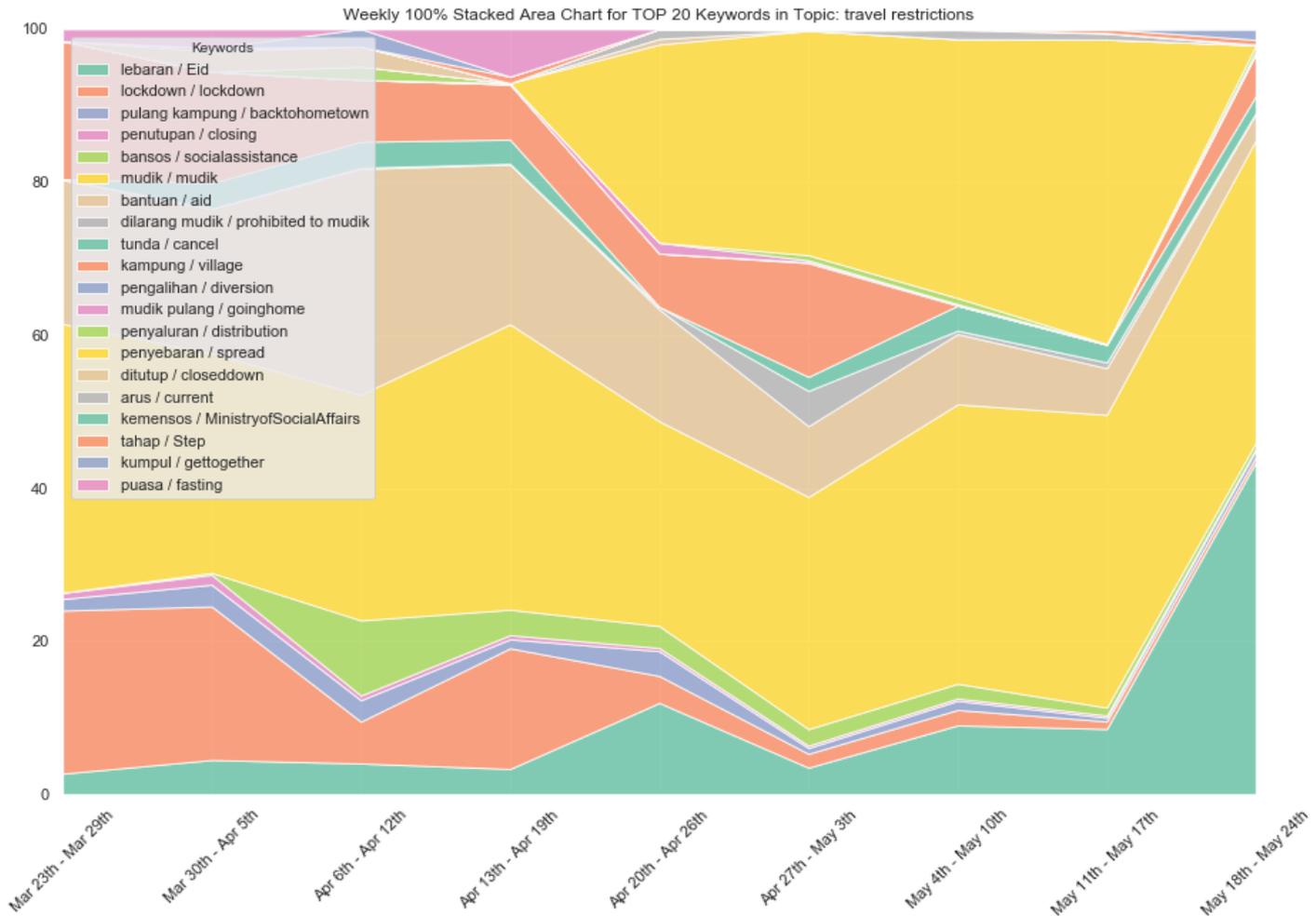
In late March wet markets were a major topic as they changed opening hours; "belanja recently becomes trending because people need to buy necessities for Eid preparation





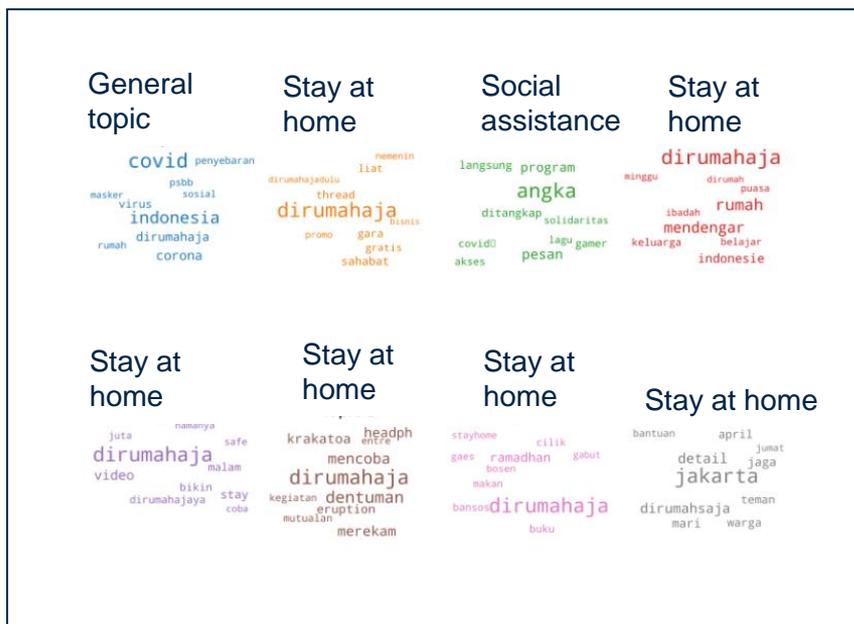
Twitter Topic Keyword Evolution: Travel Restrictions

‘Mudik’ has been a trending topic as many people are still discussing whether to go home despite the government’s travel advisories.



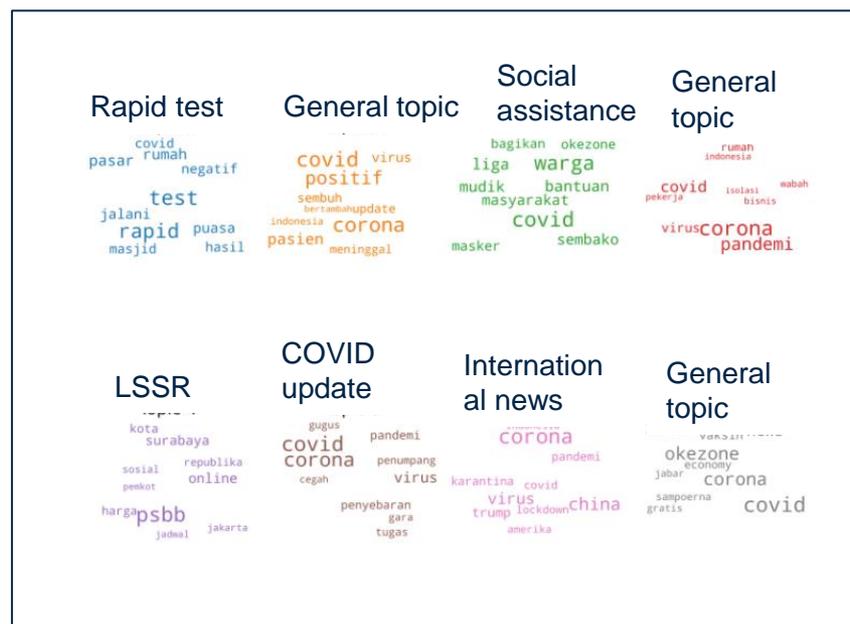
These concerns are also reflected through mainstream news media and unsupervised Twitter topic analysis

Twitter Hashtag Data Topics (COVID-19 Tagged Tweets)



In 6 out of 8 topics, **#dirumahaja** is a top keyword showing social distancing is one of the most discussed topic in Social Media

GDELT Bahasa COVID News Topics



The news cover more topics and focus on **highlighting the government response** in COVID-19 prevention (LSSR, social assistance, rapid test) and also COVID-19 general news and update. GDELT data covers over 200,000 published articles on COVID-19 in Indonesia since 3/29/2020; 98 different news sources covered.

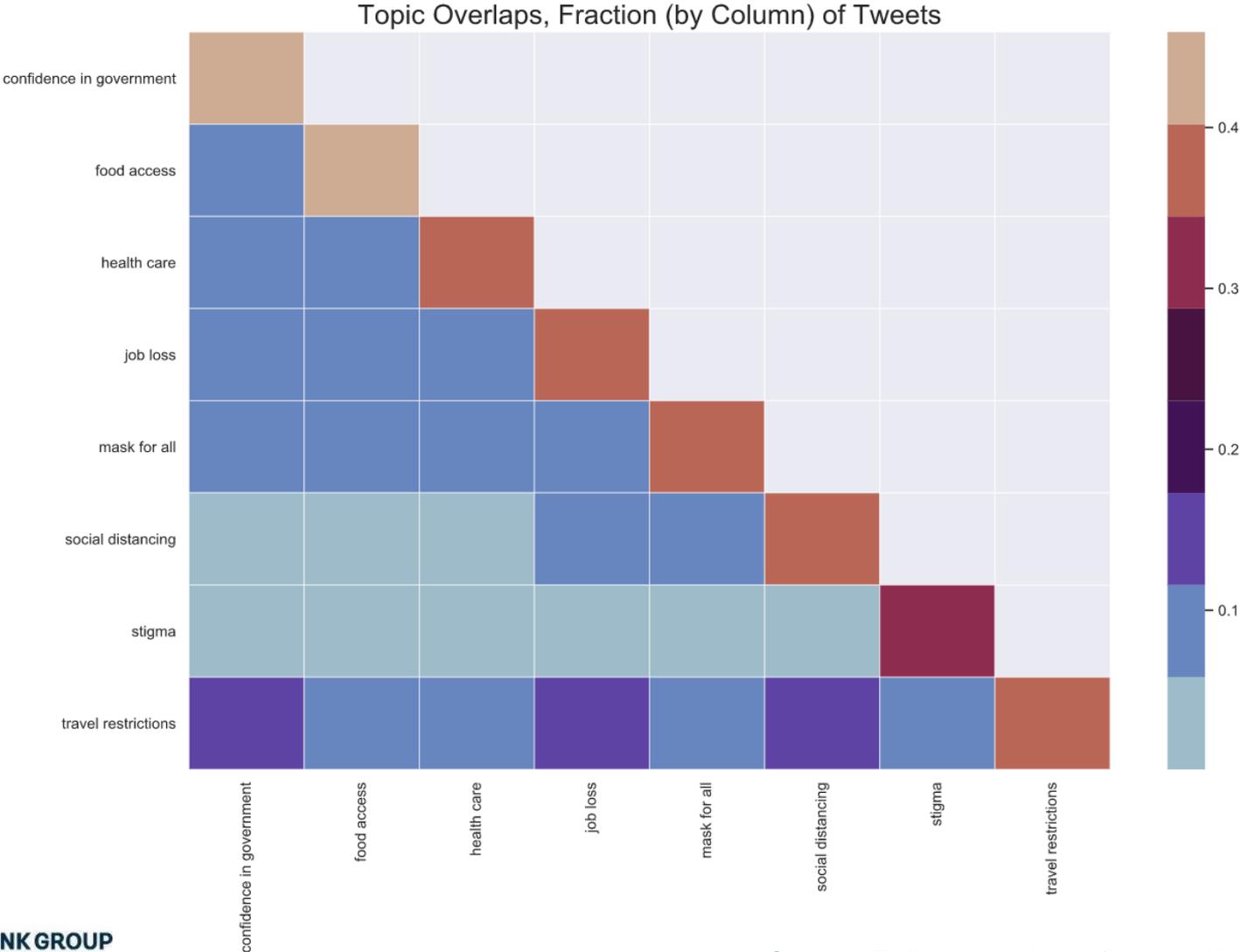
Source: GDELT and Twitter LDA Analysis

Summary: Comparison of findings across surveys

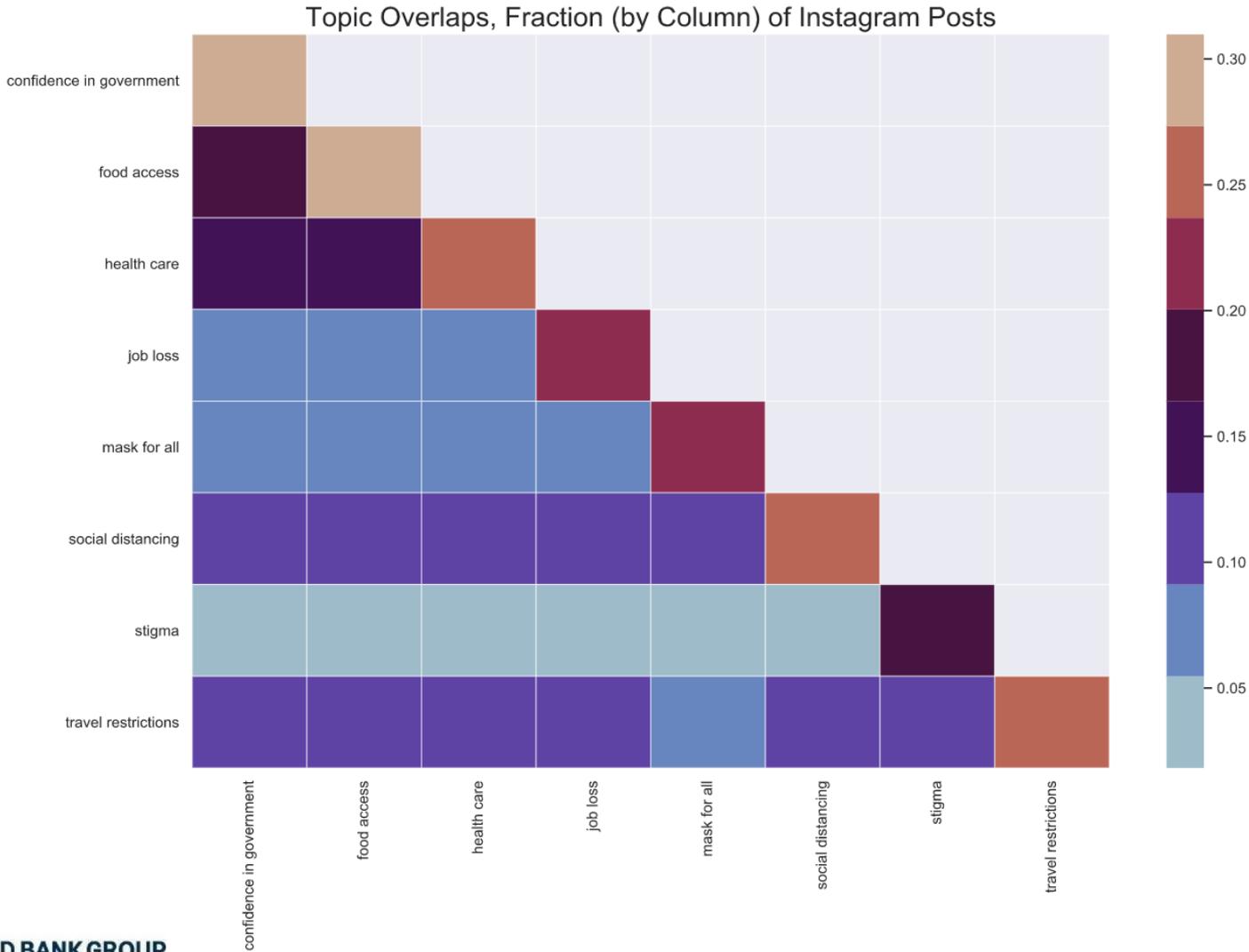
Topic	WB Online Survey	Premise	JPAL	BPS	Warwick
Knowledge	High level of awareness	✓	N/A	High level of awareness on LSSR	N/A
Health behavior change (sanitation and hygiene)	High level of change in hygiene behavior	✓	✓	✓	✓
Challenges with social distancing	Moderate (only 41% have no problem implementing)	N/A	Strong (only 21% made changes to improve social distancing)	N/A	N/A
Migration	Low	N/A	High – especially within same kabupaten/city	N/A	N/A
Food insecurity	Low (14%)	N/A	Moderate (35%)	N/A	N/A
Employment	High impact (54% working less)	N/A	High impact (57% used to work now not working)	N/A	N/A
Greatest concerns	Getting sick, losing job, having enough to eat	Getting infected and economic impacts	N/A	N/A	N/A
Intensity of concerns	Moderate (3.5 out of 5)	High on short-term outlook	N/A	N/A	N/A
Sentiment towards government	Moderate dissatisfaction	Moderate satisfaction	N/A	N/A	Strong dissatisfaction (but earlier data)

ANNEXES

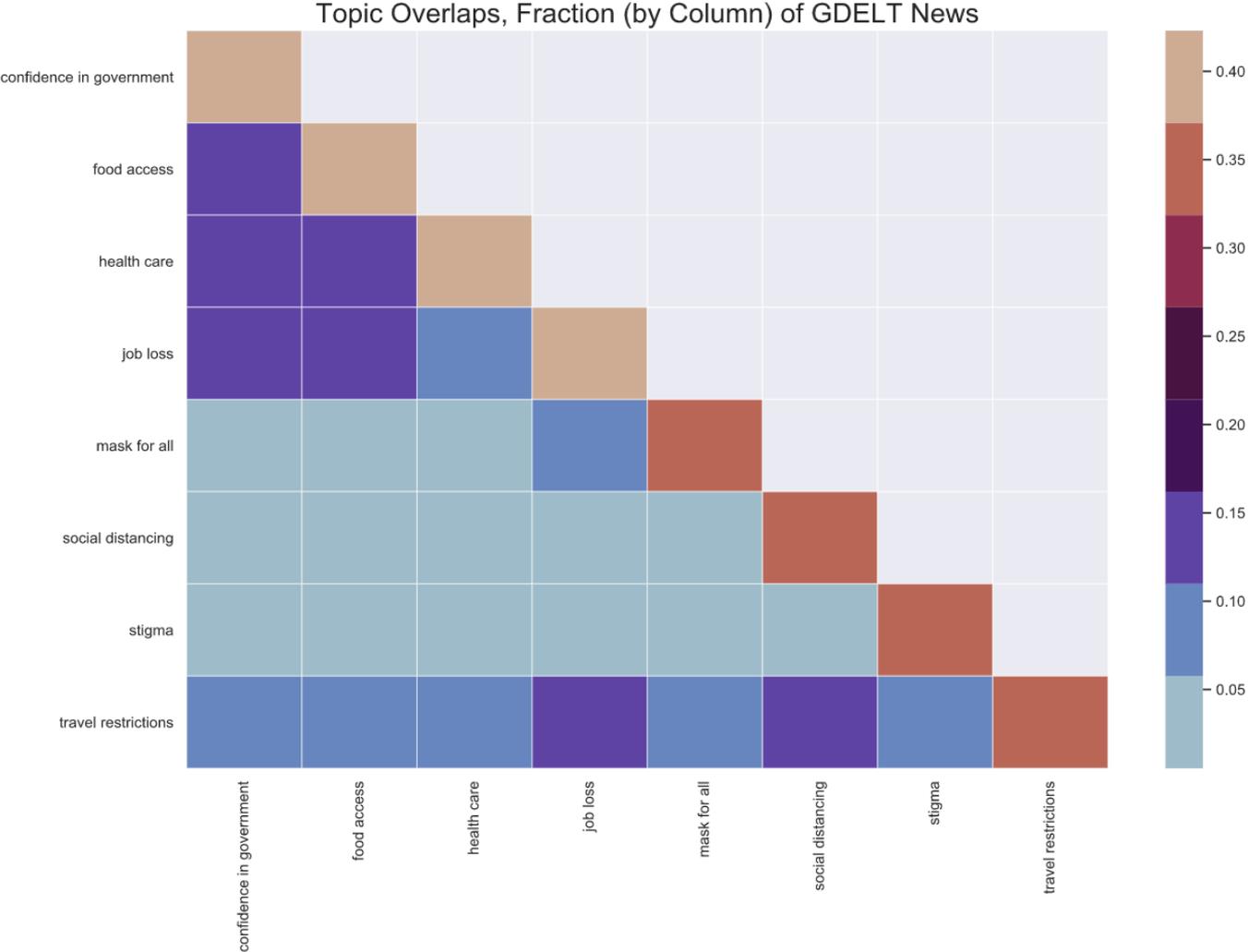
Many Twitter users who discuss about government also interested in travel restrictions



Many Instagram users who discuss about food access also interested in health care, social distancing and travel restrictions



Topic Overlaps on GDELT News



Weighting Method for SM COVID-19 Survey (1/2)

- This weighting framework uses 18+ years old individual with access to social media in 2019 in Indonesia as the reference population. According to the Susenas 2019, the number of this reference population is approximately 82 million individuals.
- After choosing our reference population, we look at the distribution of reference population across multiple “boxes”. “Boxes” here refer to the finite combination of groups formed by **a** category of region, **b** category of age group, **c** category of gender, and **d** category of income group (to proxy for income in Susenas, we will use the expenditure data instead). For each of the box, we define **x** as the total number of population in a single, unique combination of **a**, **b**, **c**, and **d**.
- After obtaining the value of **x** for each box, we distribute **x** uniformly within the box by dividing **x** with the total number of observation that available in each of the box. We define the result of the division between **x** and total number of surveyed observation (**n**) on each box as **y**. This can be written as:

(please continue to the next page)

Weighting Method for SM COVID-19 Survey (2/2)

$$y_{a,b,c,d,i} = \frac{x_{a,b,c,d}}{n_{a,b,c,d}}$$

, where:

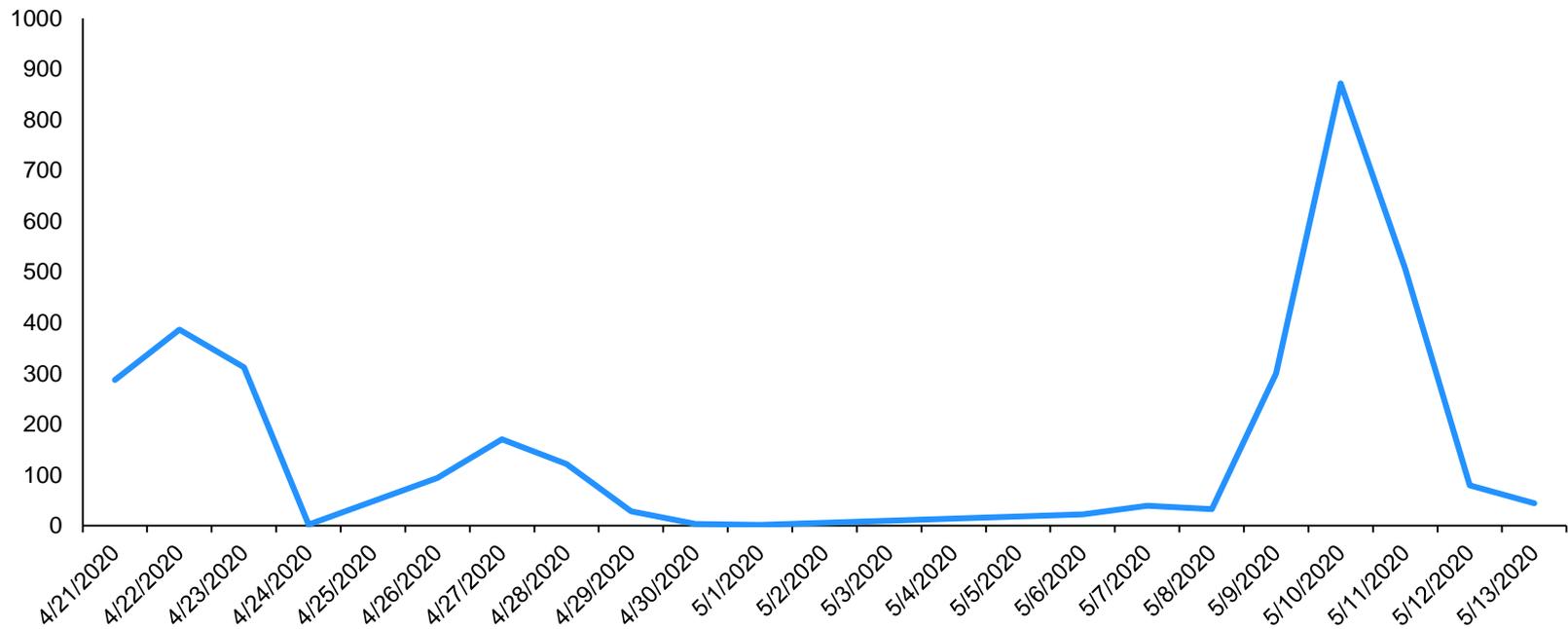
$y_{a,b,c,d,i}$ = weight of observation i of region a , age group b , gender c , and income group d .

$x_{a,b,c,d}$ = total population of region a , age group b , gender c , and income group d (derived from Susenas).

$n_{a,b,c,d}$ = total surveyed observation of region a , age group b , gender c , and income group d .

In the end, we assign y as the weight for each i observation in the SM COVID-19 Survey data.

Number of FB Survey's sample per day



Misinformation Prediction Methodology

- TL;DR: extract keywords from misinformation manually -> expand keywords using word embedding (only for verbs) -> keyword matching between misinformation news corpus and tweets corpus. Misinformation news is clustered manually

Keyword Matching Methodology

