STATISTICS

Developing High Frequency Economic Indicators using data from Google Platforms

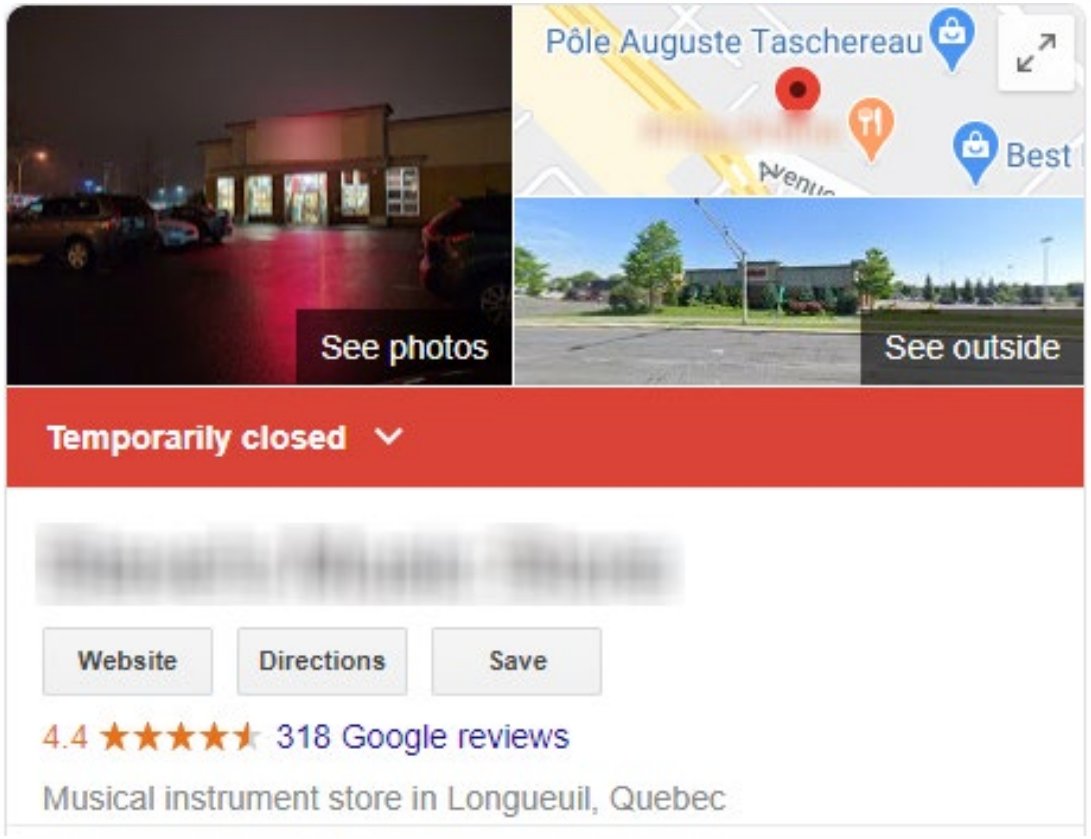
MAY 26, 2021

Marco Marini, Alberto Sanchez, Jim Tebrake,
Paul Austin, Chima Simpson-Bell
IMF Statistics Department

Disclaimer: The views expressed in this paper are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

User Demand in the Age of COVID: Finer, Faster and more Frequent

- Users are demanding more granular, high-frequency, and timely indicators to:
  - Nowcast of traditional economic indicators
  - Increase the frequency of traditional indicators (e.g., from quarterly to monthly GDP)
  - Examine evolving structural changes in real-time (e.g., business opening, closing)
  - Assess impact at a very granular level - by business types / geographic locations
- Could these indicators be developed using information from Google?



Is it possible to produce timely, high frequency economic indicators consistent with international classifications and statistical methods using data extracted from the Google Places API and Google Trends Platform.

(1) Data Sources – Google Places API

- ✓The Google Places API provides access to information that aligns closely in content and concept to the type of information typically recorded on a business register.
- ✓For example, the name, activity, comments could be used to assign an activity code and the “review”, “rating”, “price level” information can be used to provide an indication of size.
- ✓Shortly after the start of the COVID-19 pandemic Google added an additional field (business operating status) to the Google Places API.

Google Places Field	Google Places Value	Statistical Concept
Identifier	ChIJX8VNEYg2K4gRtgZVaeq3Zlc	Statistical Unit
Status	OPERATIONAL	Status
Addresss	2394 Bloor St W, Toronto, ON M6S 1P5, Canada	Geography
location.lat	43.6498786	Geography
location.lng	-79.4826734	Geography
northeast.lat	43.65115873	Geography
.northeast.lng	-79.48133312	Geography
southwest.lat	43.64845907	Geography
southwest.lng	-79.48403278	Geography
Country	Canada	Geography
City	Toronto	Geography
Name	Janchenko’s Bakery	Activity
	bakery	Activity
Rating	4.4	Size
Reviews	218	Size
price_level	NA	Size

(1) Data Sources: Google Places API

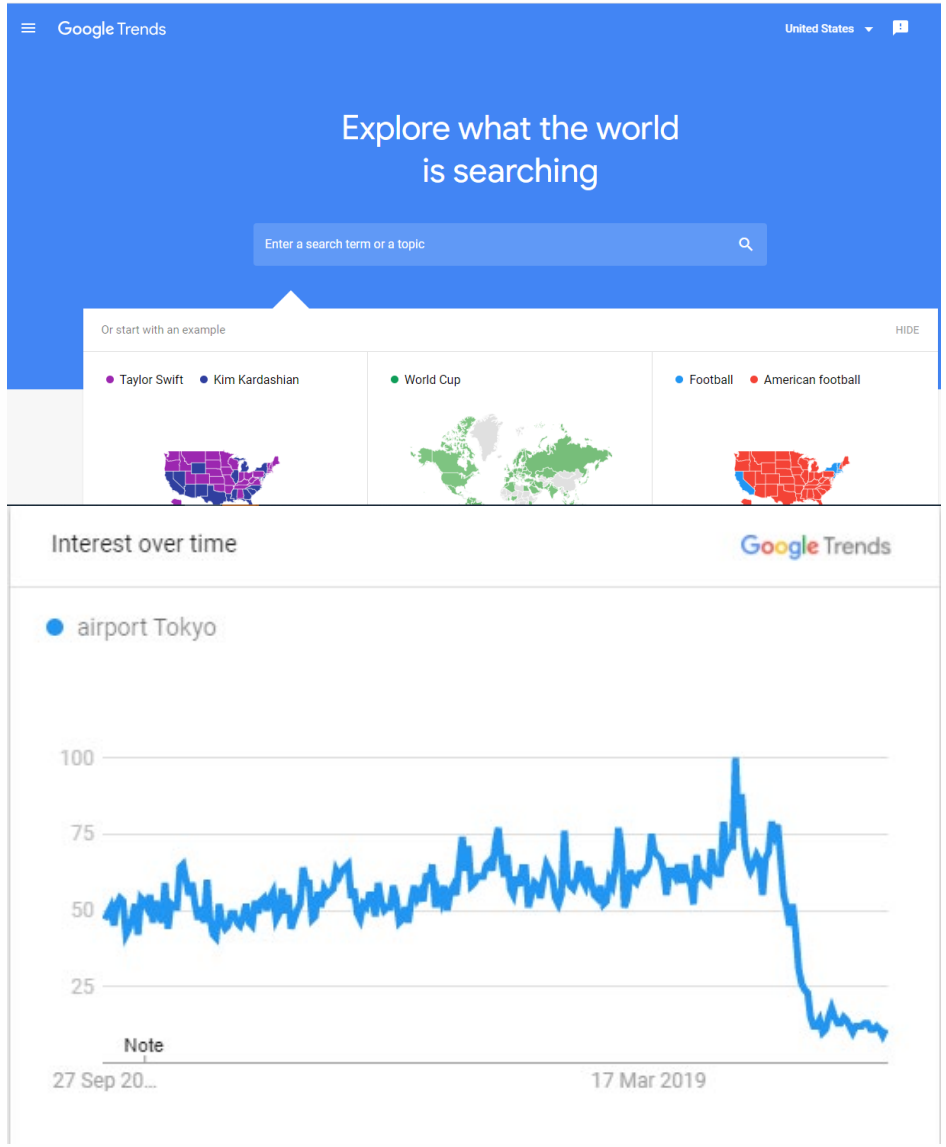
- ✓ Since the start of the COVID-19 Pandemic the IMF has been extracting data for a global sample of 90,000 “places”.
- ✓ The first extraction was on April 24th, 2020. The IMF Statistics Department has taken at least two extractions per month since that time.
- ✓ The database currently has over 3,000,000 observations and covers 21 of the largest cities in the world.

Type	Place_ID	Name	City
bar	ChIJYYa4bIRZwokRfF05m4Pogfw	BAR GOTO	New York

Period	business_status	Reviews
4/24/2020	CLOSED_TEMPORARILY	411
5/2/2020	CLOSED_TEMPORARILY	411
5/10/2020	CLOSED_TEMPORARILY	411
5/17/2020	CLOSED_TEMPORARILY	411
5/24/2020	CLOSED_TEMPORARILY	411
5/30/2020	CLOSED_TEMPORARILY	411
6/6/2020	CLOSED_TEMPORARILY	411
6/20/2020	OPERATIONAL	412
6/29/2020	OPERATIONAL	412
7/5/2020	OPERATIONAL	412
7/14/2020	OPERATIONAL	413
7/20/2020	OPERATIONAL	415
7/26/2020	OPERATIONAL	415
8/12/2020	OPERATIONAL	414
8/19/2020	OPERATIONAL	414
8/26/2020	OPERATIONAL	415
9/2/2020	OPERATIONAL	416
9/17/2020	OPERATIONAL	418
10/1/2020	OPERATIONAL	422

(2) Data Sources: Google Trends

- Google Trends is a measure of the interest in a topic relative to all topics over time.
- Google Trends provides access to anonymized, categorized and aggregated search requests.
- Google classifies search queries into 27 categories (***e.g. Transportation & Logistics***) at the top level and 241 categories at the second level using an automated classification engine.
- Recent OECD methodology uses Google Trends and machine learning to produce a weekly GDP tracker (Wolosko, 2020)



We are working with these data sources to deterime if ….

1. These data be used to develop a set of timely, high frequency economic indicators consistent with international statistical classifications and concepts.

2. These economic activity indicator be integrated with official statistics to improve their timeliness and frequency (e.g. be used to “nowcast” real GDP for a given country/region).

(1) Supply Indicators – Business Status

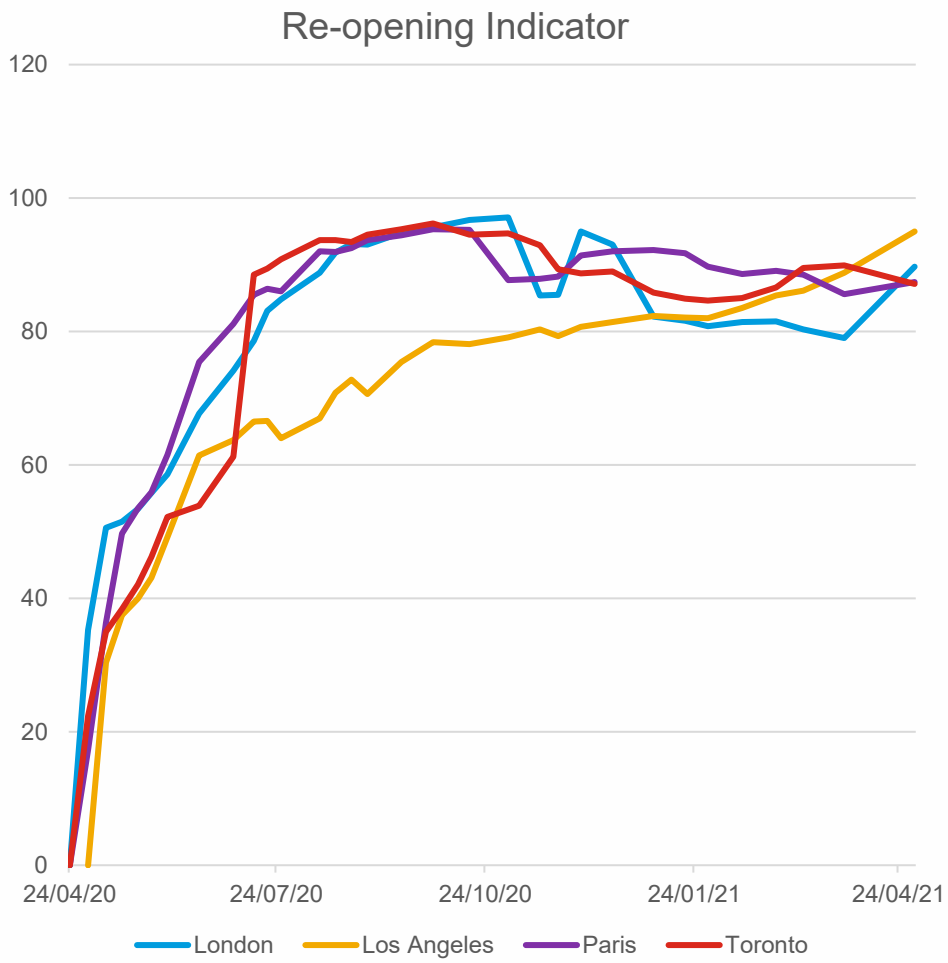
The **reopening indicator** measures the percentage of businesses temporarily closed on a baseline date that have since re-opened.

	April 24	May 1	May 8	May 15	May 22	May 29
Est 1	C	C	C	C	C	C
Est 2	C	C	C	O	O	C
Est 3	C	C	C	C	O	O
Est 4	C	C	C	C	C	C
Est 5	C	O	O	O	O	O
Reopening Indicator	0	20	20	40	60	40

- ✓ Start with an initial sample of businesses that temporarily closed due to COVID-19.
- ✓ Track their "status" at each collection (bi-weekly)

(1) Supply Indicators – Business Status

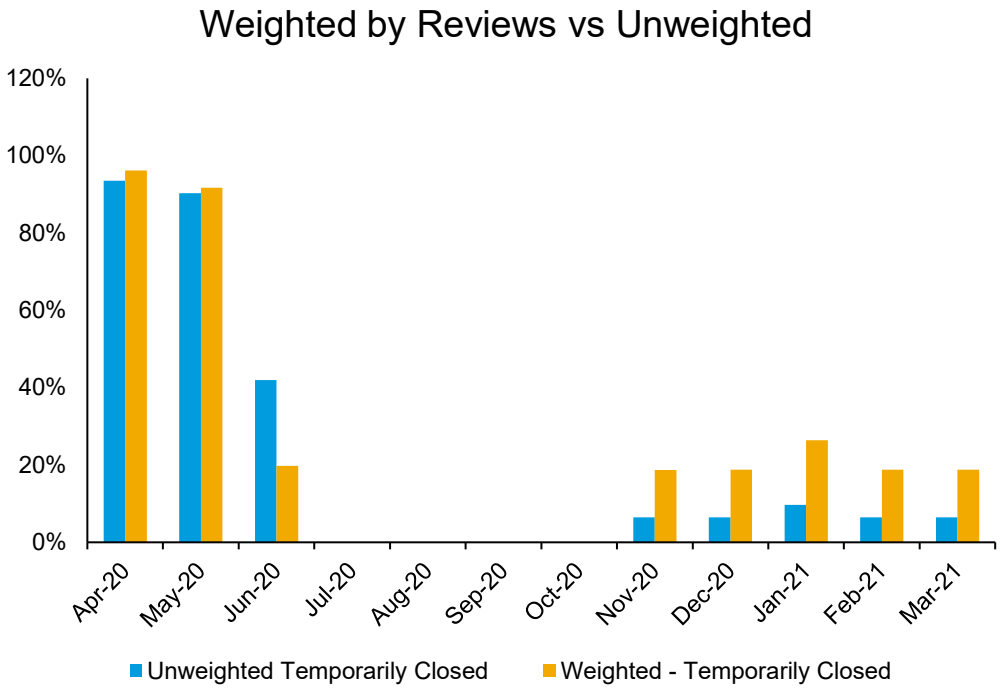
grouping_var	base_value	period_baseline	2020-04-24	2020-05-02	2021-01-06	2021-01-20	2021-01-30	2021-02-14
Atlanta	503	2020-05-02		0.0	97.4	96.6	96.8	97.0
Bogota	339	2020-05-02		0.0	89.7	90.0	90.0	90.6
Casablanca	209	2020-05-17			53.6	53.6	54.5	55.5
Istanbul	566	2020-04-24	0.0	14.8	84.6	84.6	83.9	85.2
Johannesburg	71	2020-05-17			95.8	95.8	97.2	97.2
Lagos	180	2020-04-24	0.0	12.2	66.1	66.7	58.3	66.7
London	842	2020-04-24	0.0	35.3	82.2	81.6	80.8	81.4
Los Angeles	1001	2020-05-02		0.0	82.3	82.1	82.0	83.5
Madrid	1437	2020-04-24	0.0	19.3	92.9	92.4	92.9	91.7
Manila	2750	2020-04-24	0.0	13.9	87.9	87.1	88.1	88.4
Milan	936	2020-05-24			81.7	80.9	81.9	82.9
Mumbai	2939	2020-04-24	0.0	17.3	92.4	92.8	92.8	93.4
New York	1278	2020-04-24	0.0	25.7	92.6	92.3	92.3	93.1
Paris	1645	2020-04-24	0.0	17.3	92.2	91.7	89.7	88.6
Rome	1343	2020-05-17			87.6	87.6	88.3	89.3



Share of “places” temporarily closed on April 24, 2020 that have since re-opened.

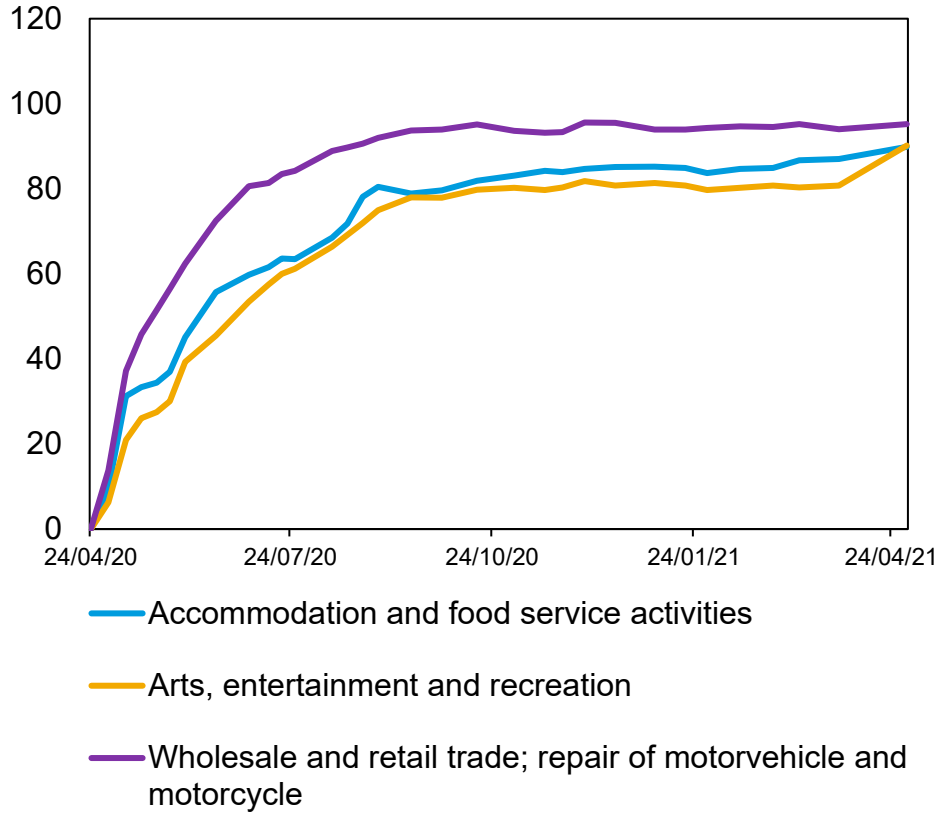
(1) Supply Indicators – Business Status

Toronto: Hair Care Temporarily Closed



- ✓ Using the reviews as weights provides a different perspective on the potential severity of the closures.

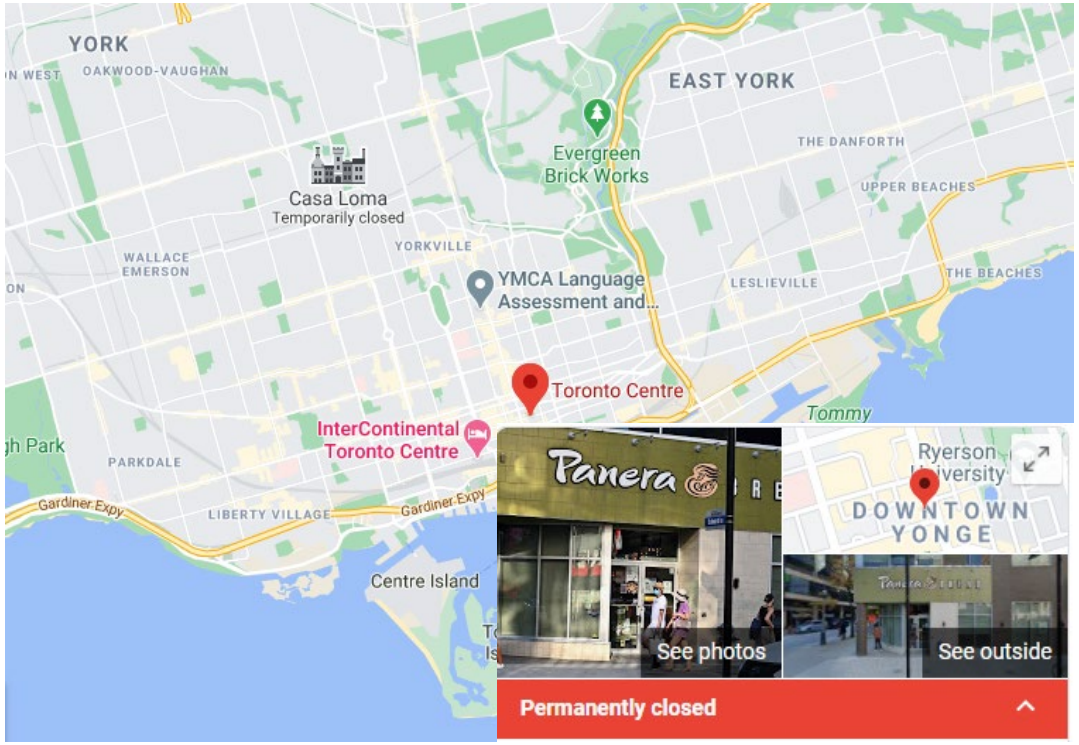
Re-opening Indicator



- ✓ Using the name and Google Activity code it is possible to assign an ISIC class.

(1) Supply Indicators – Exits

Toronto Center



Toronto: Exits

Date	Name	Type	Weighted Ratiings
D20200424	Family Video	movie_rental	2.4
D20200424	Seaton Video Rental	NA	0
D20200424	Family Video	movie_rental	3.04
D20200424	Videoflicks	NA	54.52
D20200424	Blockbuster	movie_rental	0
D20200524	Pier 1	home_goods_store	42
D20200524	Pier 1	home_goods_store	63
D20200524	Canvas	home_goods_store	3
D20200819	Gap	clothing_store	71
D20200819	Rexall Centre	stadium	851
D20200917	Panera Bread	meal_takeaway	830
D20201017	Prop Haus Inc.	movie_rental	-
D20201219	Grand Hotel And Suites	lodging	1,279
D20210331	Liberty Arts Gallery and Shop	art_gallery	2

✓It is also possible to measure entries and exits through the business status indicator. Using review information will provide an indication of the size of the exit.

(2) Demand Indicators – Google Interest by Industry

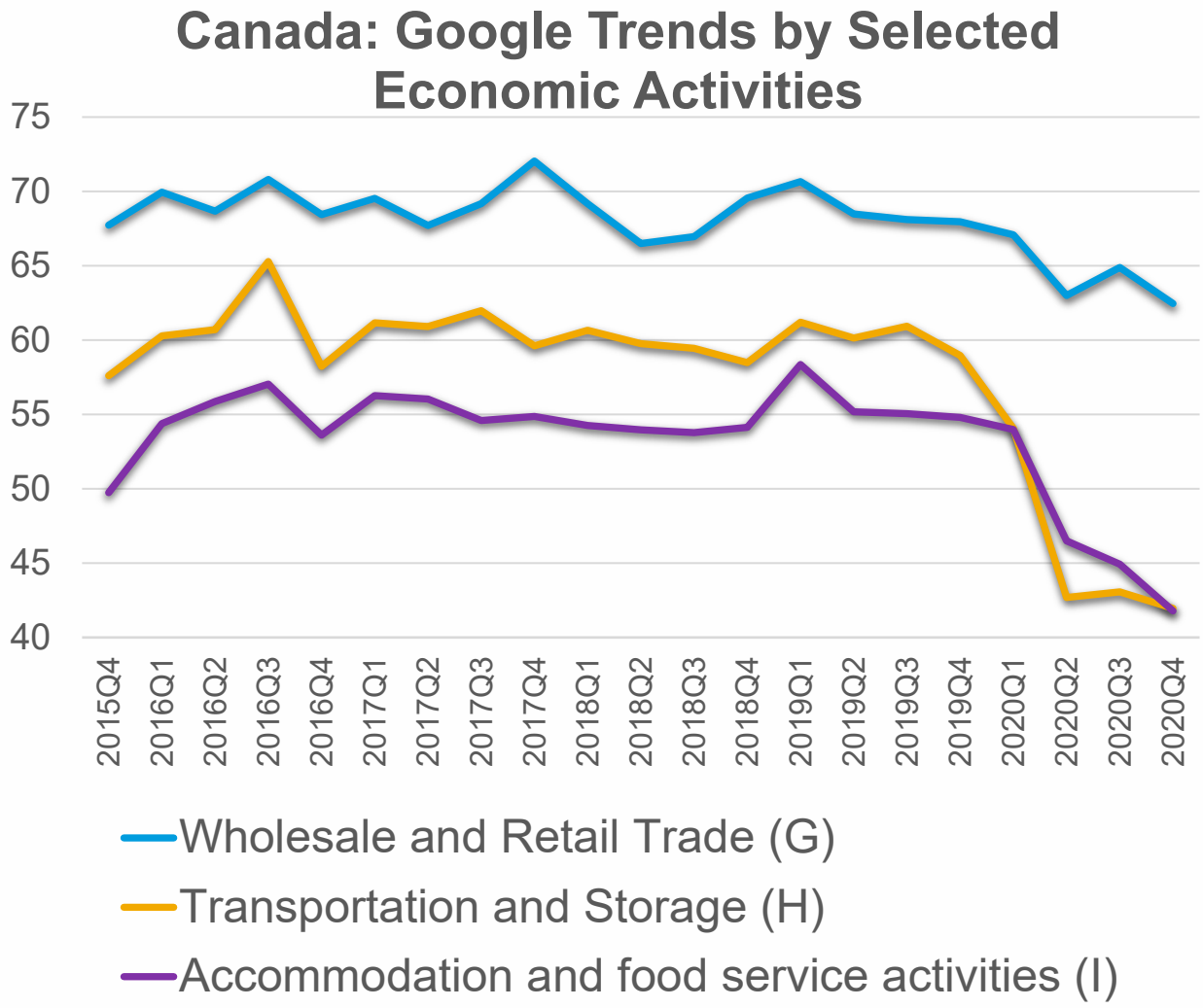
- ✓ Given the infinite number of possible search terms - Google has developed an algorithm to aggregate searches into 1000+ “trend” categories.
- ✓ For example, the category “Consumer Electronics” is an aggregation of search topics as indicated below (this is an example for Australia).

Category	ID
Educational Resources	374
T-Shirts	428
Children's Clothing	985
Clothing Accessories	124
Watches	987
Consumer Electronics	78
Televisions	305
Flowers	323
Sporting Goods	263
Ticket Sales	614
Adventure Travel	707
Mountain & Ski Resorts	1119
Vacation Offers	1019

Related Topics		
Radar	Apple	Ultra-high-definition television
Bureau of Meteorology	Television	Kmart Pharmacy
The Good Guys	Canon	Kmart
Xbox	Canon	Rain
Xbox One	JB Hi-Fi	Meaning
Camera	Apple	Weather radar
Headphones	Battery charger	Soundbar
Australia	Sony	Bunnings Warehouse
Xbox	Price	Fitbit
PlayStation 4	Fortnite	Smart TV
Loudspeaker	PlayStation 4 Pro	JBL
Television set	Microsoft Xbox One X	Watch
Garmin Ltd.	Nintendo Switch	Nintendo
Samsung Electronics	New South Wales Education Standards Authority	Netflix
Samsung	AirPods	Garmin Forerunner 235
Samsung Group	Oppo	reddit
4K resolution	Reddit	

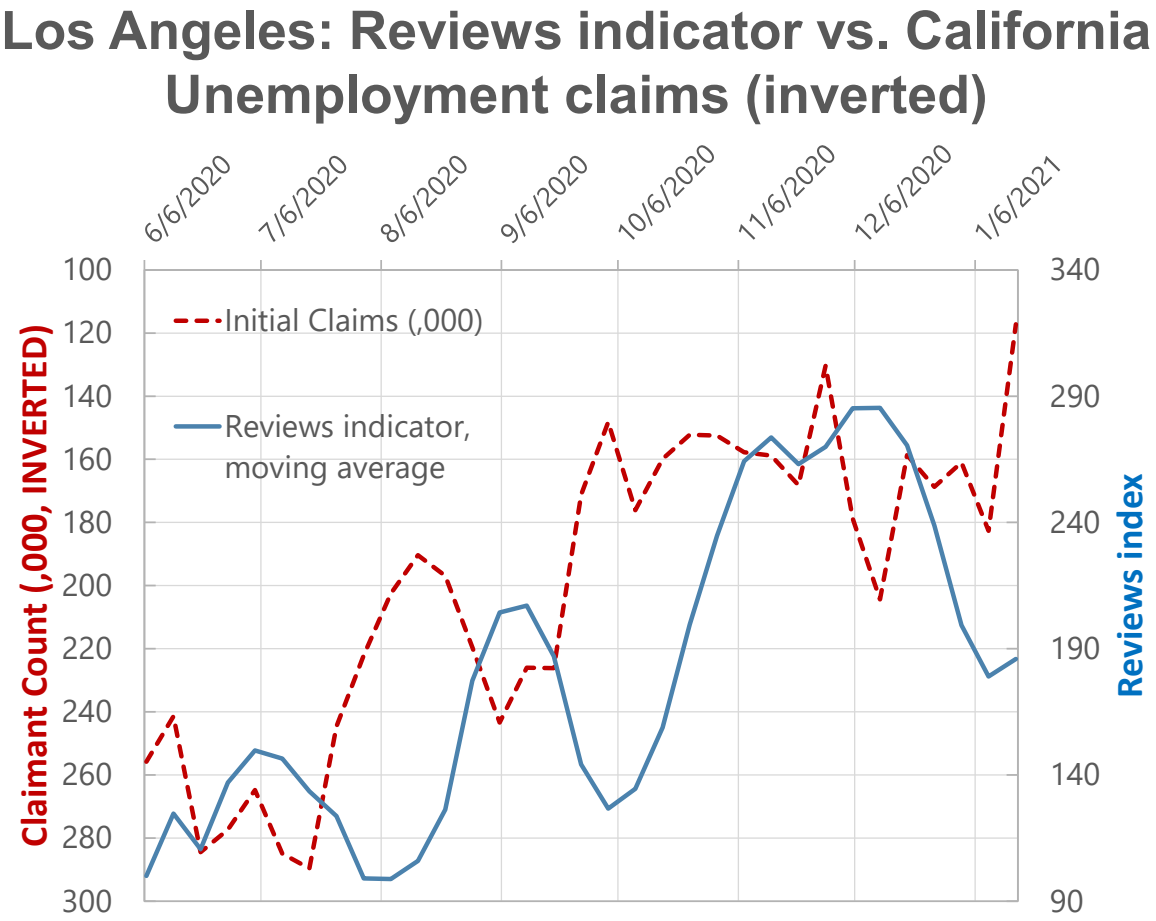
(2) Demand Indicators - Google Interest by Industry

- ✓ We map the Google Trends categories to the *International Standard Industrial Classification of All Economic Activities* (ISIC) using Natural Language Processing methods (word2vec) that score the Google Trends category description against the ISIC description (at the 4-digit level).
- ✓ We are currently testing various methods to aggregate the ISIC assigned Google Trends categories to derive an estimate of Google Trends by ISIC for a given country.



(2) Demand Indicator – Google Reviews

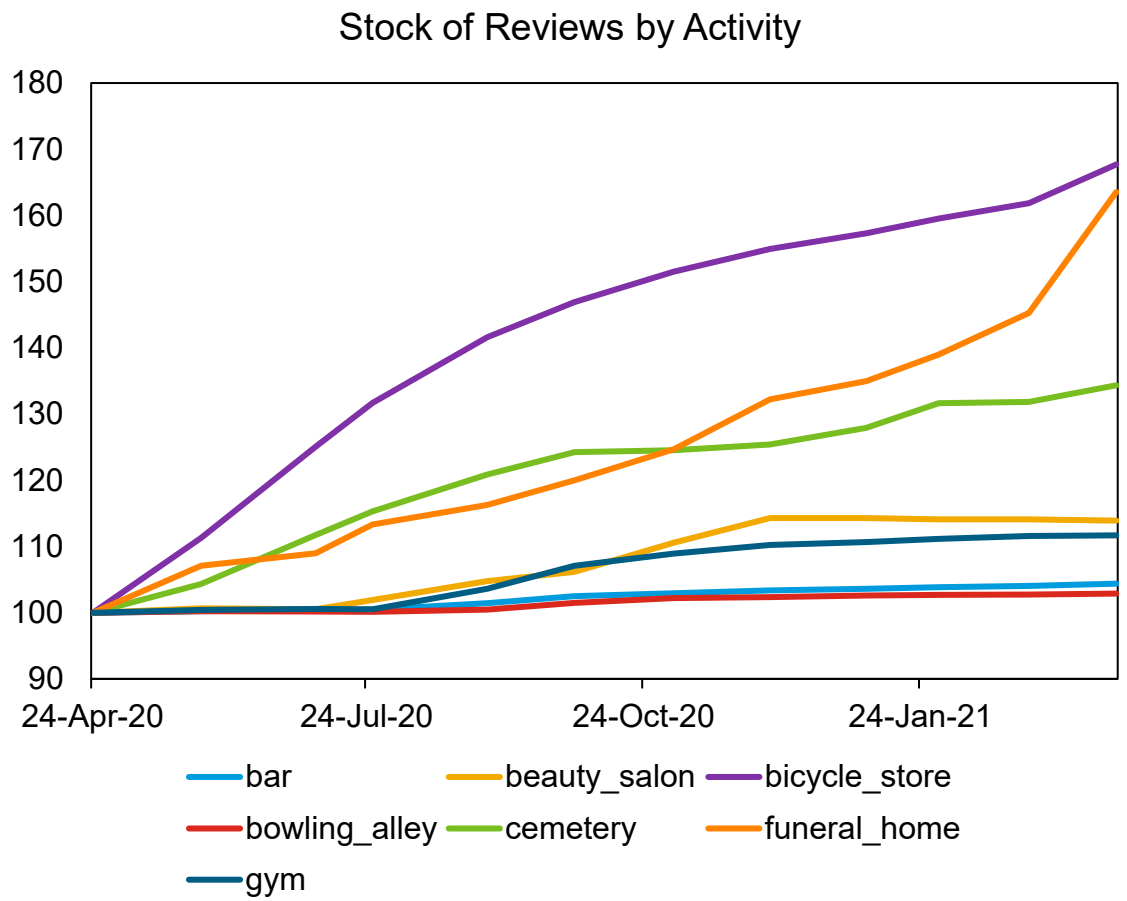
- ✓The reviews indicator is an index based on the number of new customer reviews received by a fixed group of businesses, relative to a baseline date
- ✓Since customer reviews follow engagement with a business, the reviews indicator provides timely information on recent sales activity
- ✓The new reviews can also be weighted by their average rating (from 1 to 5) to give less weight to negative reviews



(2) Demand Indicator – Google Reviews by activity

- ✓Reviews can be summarized by activity, weighted by rating.
- ✓Indicators tracking the stock of reviews or the number of new reviews can be developed and used as a short-term indicator of change in demand.
- ✓These indicators are highly correlated with the supply side indicators.

Toronto: Stock of Reviews and Change in Reviews by activity



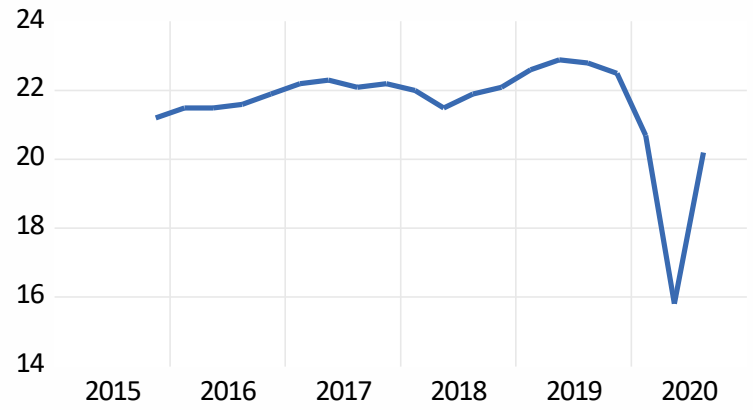
Application to Selected Countries: Nowcasting Quarterly GDP during COVID-19 using Google Based High Frequency Indicators by Economic Activity

Google Data Allow to Track in Real-Time the Impact of the Pandemic

- ✓ Google Trends by Economic Activity and the Reopening Indicator track well the fall and recovery of those sectors particularly hit by the pandemic (e.g., transportation and storage in France)

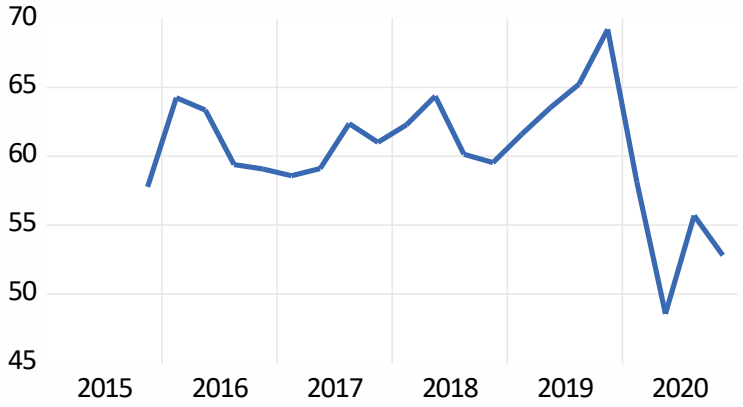
GDP Value Added - Transportation and Storage (H)

France



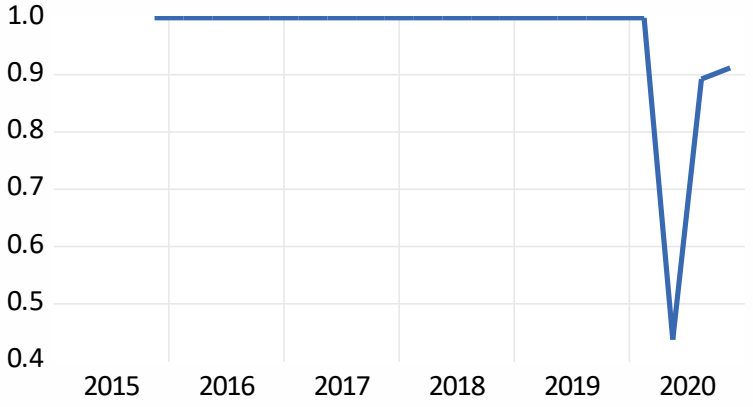
Google Trends - Transportation and Storage (H)

France



Google Places - Reopening Indicator

France



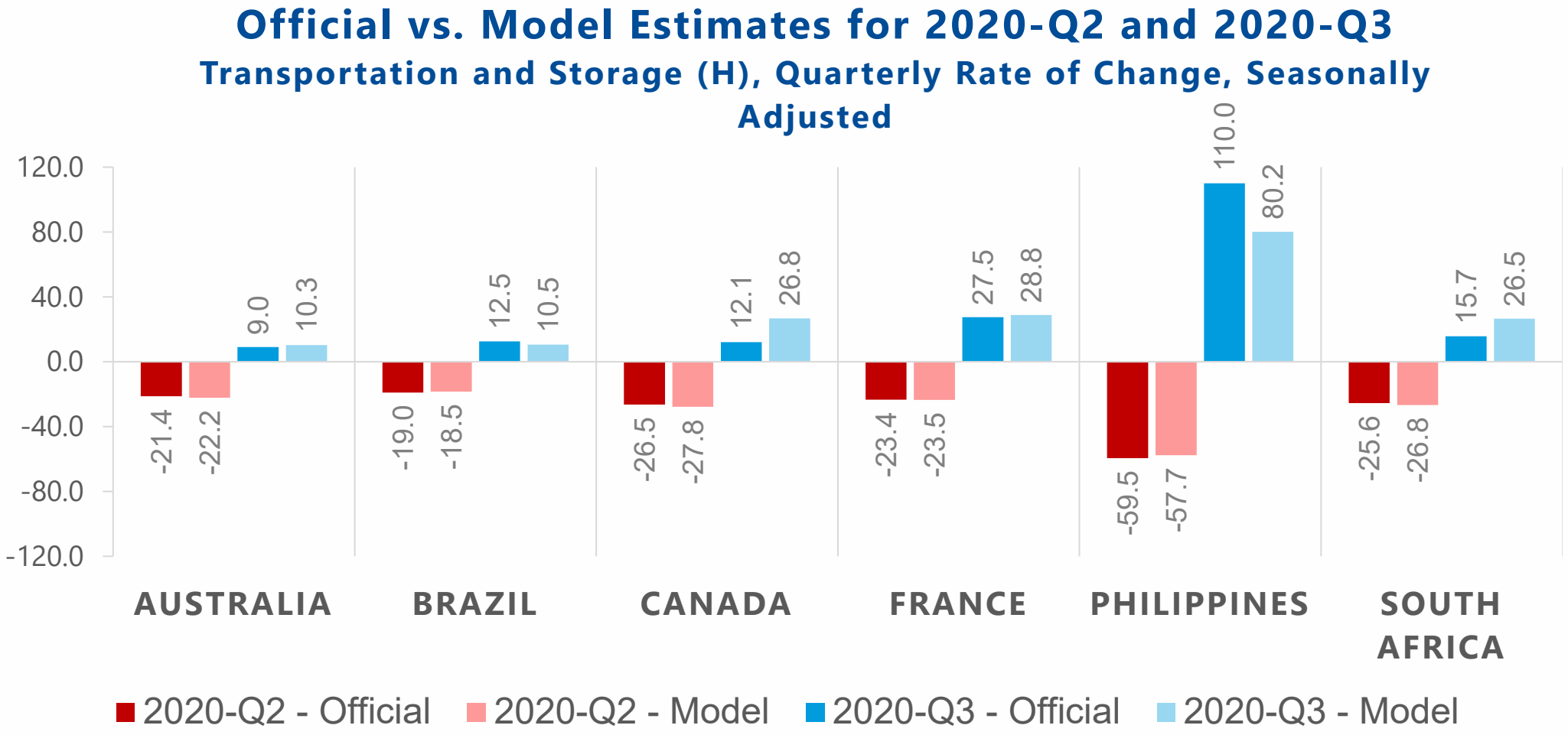
Indicators Developed in this Study Prove to Be Powerful Predictors for Some Economic Activities…

✓ Regression model for Quarterly Value Added of **Transportation and Storage** (Section H of *ISIC rev. 4*) in six countries

		Google Trends by Economic Activity		Reopening Indicator	
	R^2	Coeff	t-stat	Coeff	t-stat
Australia	0.83	0.43	3.22**	0.19	2.43**
Brazil	0.91	0.10	2.21**	0.30	13.20**
Canada	0.88	0.89	5.57**	0.38	6.44**
France	0.92	0.25	2.15**	0.48	7.28**
Philippines	0.95	0.45	6.40**	1.28	15.96**
South Africa	0.92	0.25	2.47**	0.34	8.90**

Estimation Period: 2015q4-2020q3
Model specification in logs, with constant. All data seasonally adjusted

… Producing Accurate (and Timely) Nowcasts of the Second and Third Quarter of 2020



Conclusions and Next Steps

Objective: Develop a repeatable and highly accessible methodology that can be used to produce high frequency indicators by economic activity based on Google data for real-time monitoring and nowcasting

Next Steps:

- We are refining our methodology to
  - ✓ Improve weighting scheme for aggregation of categories (currently, a simple average is used)
  - ✓ Develop activity indicators for selected businesses/industries by combining information from Google Places and Google Trends
- We are in the process of documenting the methodology and results in an IMF working paper and intend to include all the R scripts a compiler would require to stand-up these indicators for a given geographic region.

Google support to the Global Statistical System

1. Support for National Statistical Systems

- Many National Statistical organizations are unable to pay the Google Places Access fee. Is there any way the credits provided by Google through the not-for-profit COVID-19 response efforts could be extended to National Statistical organizations?

2/ Historical "Start-up" Data

- For National Statistical Organizations interested in using the Google Places data as an input into their business register and business dynamics program is Google able to provide a "Start-up" time-series of 2-3 years?

Thank you!

Marco Marini (mmarini@imf.org)
Alberto Sanchez (asanchez3@imf.org)
Jim Tebrake (jtebrake@imf.org)
Paul Austin (paustin@imf.org)
Chima Simpson-Bell (csimpson-bell@imf.org)

IMF Statistics Department

Annex – Mapping Trends Categories to ISIC

- ✓ Google search categories are selected automatically through text comparison with description of 4-digit level of the ISIC rev. 4
- ✓ For example, the following categories are selected for Transportation and Storage (all within "Transportation & Logistics" Trends category)

Level 1	Level 2	Level 3
Business & Industrial: 12	Transportation & Logistics: 50	Aviation: 662
Business & Industrial: 12	Transportation & Logistics: 50	Distribution & Logistics: 664
Business & Industrial: 12	Transportation & Logistics: 50	Freight & Trucking: 289
Business & Industrial: 12	Transportation & Logistics: 50	Mail & Package Delivery: 1150
Business & Industrial: 12	Transportation & Logistics: 50	Maritime Transport: 665
Business & Industrial: 12	Transportation & Logistics: 50	Moving & Relocation: 291
Business & Industrial: 12	Transportation & Logistics: 50	Packaging: 290
Business & Industrial: 12	Transportation & Logistics: 50	Parking: 1306
Business & Industrial: 12	Transportation & Logistics: 50	Public Storage: 1347
Business & Industrial: 12	Transportation & Logistics: 50	Rail Transport: 666
Business & Industrial: 12	Transportation & Logistics: 50	Urban Transport: 667