Webinar UN-CEPAL, UNCTAD y FMI Medición de la economía digital y el comercio en América Latina y el Caribe

8-11 noviembre de 2022



Aprovechamiento de Big Data y la ciencia de datos para compilar estadísticas económicas

Margarida Martins
Departamento de Estadística, FMI

Contenido

- Panorama general de las fuentes de datos para la compilación de indicadores del PIB y de la actividad económica.
- Ejemplos de fuentes de datos no tradicionales para medir la actividad económica.
- Uso de la API de Google Places y los datos de Google Trends para desarrollar IFP.
- OECD Weekly Tracker of Economic Activity

Panorama general de las fuentes de datos para la compilación de indicadores del PIB y de la actividad económica

Motivación

- Los responsables de la formulación de políticas requieren datos oportunos y de alta frecuencia para evaluar las condiciones económicas en tiempo real:
 - ▶ La información anual oportuna es útil, pero no proporciona suficiente información para los responsables de la formulación de políticas sobre el ciclo económico;
 - ▶ Los datos trimestrales/mensuales publicados con un retraso de varios meses no permiten a los responsables políticos realizar ajustes en tiempo real.
- La pandemia de COVID-19 aumentó la necesidad de datos granulares, de alta frecuencia y oportunos para comprender mejor el impacto de la pandemia y el camino de la recuperación económica:
 - ► Aumentar la frecuencia de los indicadores económicos (por ejemplo, del PIB trimestral al mensual);
 - ► Indicadores económicos tradicionales Nowcast;
 - ▶ Análisis de cambios estructurales en evolución en tiempo real (por ejemplo, apertura de negocios, cierre);
 - ▶ Evaluar el impacto de la pandemia por tipo de negocio y ubicación.

Previsión

Previsión

Nowcast

Estimación temprana

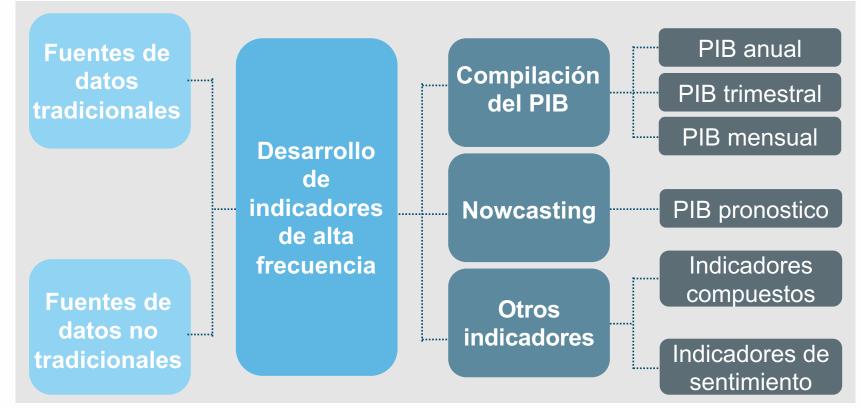
Estimación actual

exactitud

Resumen de fuentes de datos

- La compilación del PIB y otras estadísticas macroeconómicas tradicionalmente se basa en fuentes de datos que están disponibles con un desfase temporal.
 - ► Con la **digitalización** y la creciente **disponibilidad de datos administrativos**, está surgiendo una nueva gama de fuentes de datos no tradicionales que pueden utilizarse para mejorar las estimaciones del PIB, como insumo para ejercicios de predicción inmediata y para el desarrollo de indicadores económicos de alta frecuencia.

 Ejemplo de uso de indicadores de alta frecuencia



Resumen de fuentes de datos

- Las fuentes de datos no tradicionales presentan una mayor puntualidad, frecuencia y granularidad, pero es necesario abordar algunos desafíos:
 - ▶ Asegurar el acceso a los datos de forma regular con cobertura, variables disponibles y frecuencia consistentes;
 - Garantizar una cobertura suficiente de la actividad económica y series cronológicas que sean lo suficientemente largas;
 - ▶ Posible desalineación conceptual entre la fuente de Big Data y la estadística de destino que se está produciendo, ya que la fuente de Big Data en general no se desarrolló con el objetivo de producir estadísticas.

Ejemplos de fuentes de datos de tradicionales a no tradicionales

Tradicional No tradicional Encuestas a hogares Datos regulatorios Datos de Google Places y Encuestas empresariales Google API Datos aduaneros y fiscales Estadísticas de Finanzas Datos de observación de Sistema de seguridad social **Publicas** la Tierra Grandes corporaciones o Datos de comercio internacional Estadísticas de precios Datos de vuelo y buque asociaciones empresariales Datos de tarjeta de crédito financieras

PIB y Nowcasting

- Para la mayoría de las economías, existe un conjunto de información de alta frecuencia disponible
- Estos indicadores son útiles, pero individualmente no proporcionan una imagen agregada de todo el crecimiento económico.
 - ► Encontrar una manera de combinarlos resultaría en un indicador agregado para toda la economía.

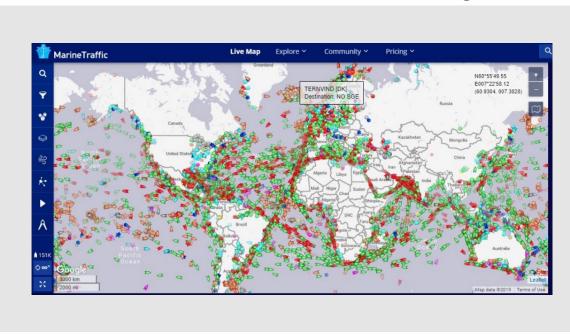
Ejemplo de tipos de indicadores por métodos de compilación y sincronía temporal

		MÉTODO DE COMPILACIÓN								
		Estadístico	Econométrico	Contabilidad						
ONÍA ORAL	Coincidente	Análisis del ciclo económico	Análisis del ciclo económico	Análisis descriptivo de la economía						
SINCRONÍA TEMPORAL	Principal	Alerta temprana de puntos de inflexión	Nowcasting/ Pronóstico							
	Ejemplos: Junta de Conferencias, Indicadores principales compuestos (OCDE).		PCA, modelos factoriales, ADL, ARIMA, VAR.	PIB mensual de Canadá.						

Ejemplos de fuentes de datos no tradicionales para medir la actividad económica

- Datos del sistema de identificación automática (AIS)
- La información incluye la ubicación, la velocidad y el estado de los buques (por ejemplo, escalas semanales en el puerto y volumen de comercio).
- Posible uso del comercio de nowcast.

Una instantánea de los datos AIS basados en el tráfico global de buques



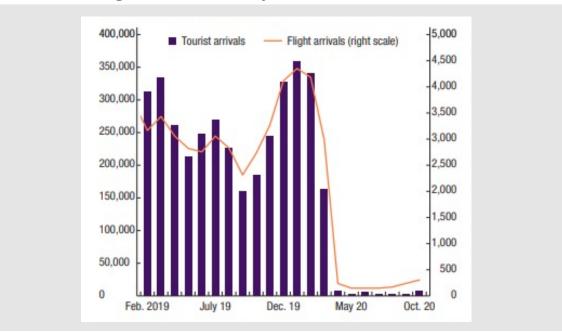
Fuente: https://www.FMI.org/~/media/Files/Publications/WP/2019/wpiea2019275-print-pdf.ashx, basado en datos de MarineTraffic.

Nota: Diferentes tipos de recipientes se muestran en diferentes colores.

Datos de vuelo

 Posible uso en estimaciones de viajes aéreos, estimaciones para la actividad turística.

Costa Rica: Impacto de la pandemia en el sector turístico – Llegadas de turistas y vuelos internacionales

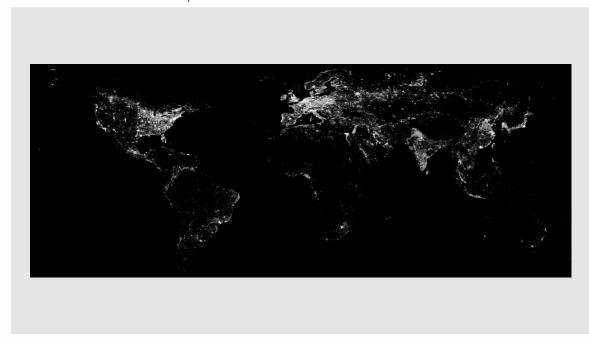


Fuente: https://www.FMI.org/en/Publications/Departmental-Papers-Policy-Papers/Issues/2021/02/19/Tourism-in-the-Post-Pandemic-World-Economic-Challenges-and-Opportunities-for-Asia-Pacific-49915, basado en Banco Central de Costa Rica; FlightRadar24; Instituto Costarricense del Turismo; e estimaciones del personal técnico del FMI.

Datos de observación de la Tierra

- ▶ Uso de datos de imágenes satelitales para mejorar las estadísticas oficiales sobre una amplia gama de temas, incluidos la agricultura, el clima, la actividad empresarial y el transporte.
- ▶ Los proyectos piloto incluyen densidad de cultivos, estadísticas agrícolas, estadísticas de cobertura y uso de la tierra, sistemas urbano-rurales, datos climáticos e inventario de petróleo crudo.

Mapa de luces nocturnas en 2010



Fuente: https://www.FMI.org/en/Publications/WP/Issues/2019/04/09/Illuminating-Economic-Growth-46670.

Brasil y México: Ejemplos de iniciativas que utilizan datos satelitales

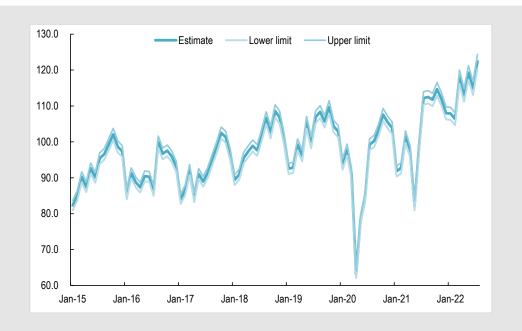


Fuente: http://brazildatacube.org/en/about-brazil-data-cube-2/ y https://www.inegi.org.mx/investigacion/geomediana/#Metadatos

Datos de consumo de electricidad

Posible uso para estimar la actividad de sectores económicos como la industria manufacturera, o como insumo para el PIB nowcast.

Colombia: Indicador de estimación temprana para la industria manufacturera e intervalos de confianza del 95 por ciento

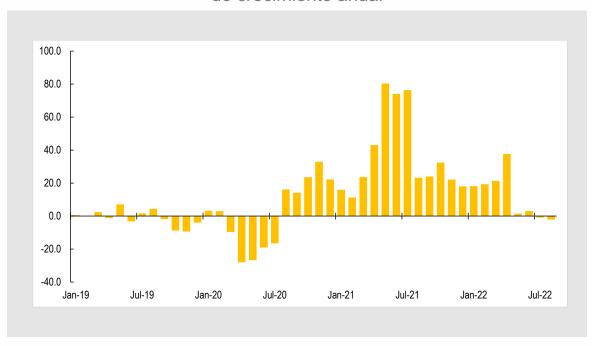


Fuente: FMI, basado en datos del DANE (al 12 de septiembre de 2022), disponible en Estadísticas experimentales (dane.gov.co).

Datos fiscales

- ► El Impuesto sobre el Valor Añadido (IVA) y otros datos fiscales capturan tanto el volumen de negocios como el gasto.
- Posible uso para rastrear el comercio, hoteles y restaurantes, y otras actividades.

Chile: Índice de ventas minoristas diarias, promedio mensual, tasa de crecimiento anual

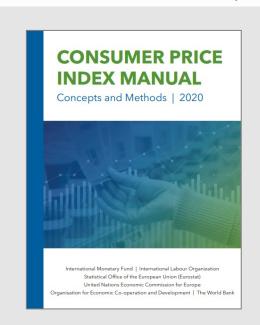


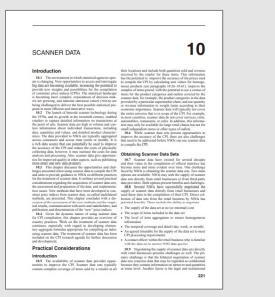
Fuente: FMI, a partir de datos del Banco Central Chile (al 13 de septiembre de 2022), disponible en Estadísticas Experimentales - Banco Central de Chile (bcentral.cl).

Datos del escáner

Uso de datos de escáner de cadenas de supermercados y otros minoristas, así como precios en línea obtenidos de web scraping, para compilar índices de precios.

Manual del IPC 2020: Capítulo sobre datos del escáner



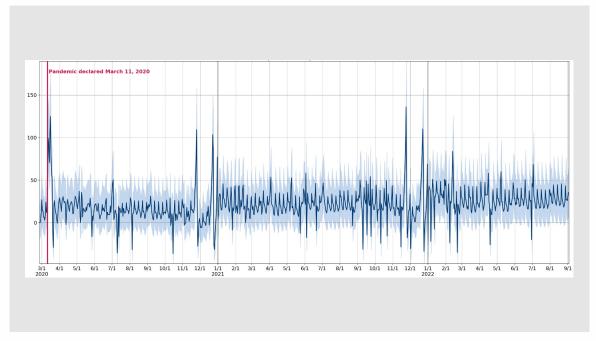


Fuente: https://www.FMI.org/-/media/Files/Data/CPI/cpi-manual-concepts-and-methods.ashx.

Datos de tarjeta de crédito

Posible uso para comprender los cambios en los patrones de consumo y estimar el gasto de consumo.

Estados Unidos: Gasto en tiendas de alimentos y bebidas: un estudio de eventos basado en transacciones con tarjeta de pago



Fuente: Oficina de Análisis Económico de los Estados Unidos, disponible en https://www.bea.gov/recovery/estimates-from-payment-card-transactions (a partir del 13 de septiembre de 2022).

Nota: El gráfico muestra la diferencia con respecto al nivel típico de gasto sin cambios en la economía relacionados con COVID-19. El nivel típico corresponde a un valor de 0. El área sombreada representa bandas de intervalo de confianza del 95 por ciento.

Uso de Google Places API y Google Trends datos para desarrollar Indicadores de Alta Frecuencia

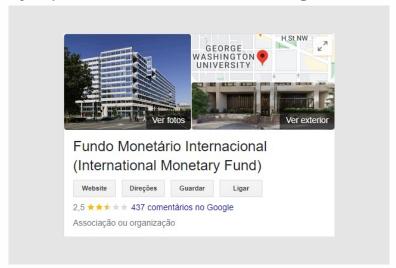
API de Google Places

- ► La API de Google Places permite a los usuarios extraer información sobre lugares de la plataforma Google Maps.
- ▶ Los usuarios pueden extraer 23 campos de información para cada lugar, incluido el nombre, la dirección, la identificación del lugar, el estado comercial, la latitud/longitud, las revisiones, el nivel de precios y las horas de operación.

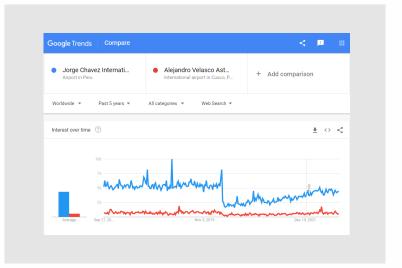
Datos de Google Trends

- ▶ Google Trends es una medida de interés en un tema en relación con todos los demás temas a lo largo del tiempo. Un tema puede ser cualquier cosa, desde una persona o evento hasta un negocio o producto específico.
- Los números representan el interés de búsqueda en relación con el punto más alto en el gráfico para la región y hora dadas. Un valor de 100 es la popularidad máxima del término. Un valor de 50 significa que el término es la mitad de popular. Una puntuación de 0 significa que no hubo suficientes datos para este término.
- Cuando los temas se relacionan con un negocio, industria o producto, la tendencia podría ser indicativa, al menos hasta cierto punto, de la actividad económica.

Ejemplo de datos de la API de Google Places



Ejemplo de datos de Google Trends



- Método: (A) Indicadores de estado de apertura
 - ▶ Los lugares reciben el estado de "Abierto", "Cerrado temporalmente" o "Cerrado permanentemente". Esta información se puede utilizar para producir varios indicadores dinámicos de negocio útiles.

Indicador operacional

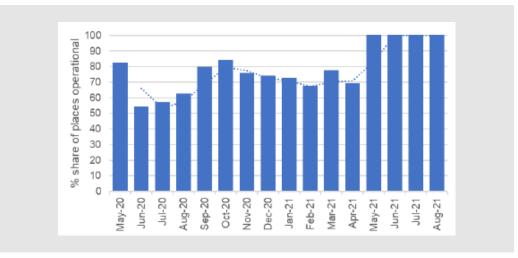
- Representa la proporción de lugares en cada región geográfica que están operativos en un momento dado, ponderada por el número de revisiones.
- La ponderación por revisiones tiene como objetivo capturar el impacto del tamaño del negocio, en el que las empresas con más revisiones tendrán un mayor impacto en el movimiento en el indicador.

Ejemplo de estado de negocios por semana

	Wee	k 1 - Initial Status:	Places A, B, C, and D oper	rational				
Business	Reviews	Share of Reviews	Business Status	Operational Indicate				
Α	1000	40	Operational					
В	500	20	Operational					
С	500	20	Operational					
D	100	4	Operational					
E	400	16	Temporarily Closed					
	2500	100		84				
			e F is temporarily closed					
Business	ess Reviews Share of Revie		Business Status	Operational Indicato				
F	1000	37	Temporarily Closed					
G	700	26	Operational					
Н	500	19	Operational					
1	100	4	Operational					
J	400	14	Operational					
	2700	100		63				
Week 3 – All places are operational								
Business	Reviews	Share of Reviews	Business Status	Operational Indicato				
F	1000	33	Operational					
K	700	24	Operational					
L	500	17	Operational					
M	400	13	Operational					
J	400	13	Operational					
	3000	100		100				

Fuente: www.elibrary.FMI.org/view/journals/001/2021/295/001.2021.issue-295-en.xml.

Ejemplo de indicador operativo: São Paulo - bares



Fuente: www.elibrary.FMI.org/view/journals/001/2021/295/001.2021.issue-295-en.xml

- Método: (A) Indicadores de estado de apertura
 - ▶ Los lugares reciben el estado de "Abierto", "Cerrado temporalmente" o "Cerrado permanentemente". Esta información se puede utilizar para producir varios indicadores dinámicos de negocio útiles.

Indicador de reapertura de negocios para centros urbanos seleccionados

Indicador de reapertura de negocios

Se utiliza para rastrear la ruta y el ritmo al que reabren las empresas que están cerradas temporalmente en una región, siendo de particular interés para evaluar el impacto de las regulaciones gubernamentales durante la pandemia de COVID-19.

- ▶ El indicador comienza con la selección de una cohorte de referencia de lugares que están temporalmente cerrados. Cada semana se examina el estado de cada uno de estos lugares para ver si han abierto o permanecen temporalmente cerrados.
- ► El indicador representa la proporción de lugares en cada región geográfica que están operativos en un momento dado ponderado por el número de revisiones.



city	sam ple size	baseline	24- Apr -20	24- May -20	26- Jul- 20	26- Aug -20	3- Nov -20	30- Jan -21	31- Mar -21	1- May -21	1- Jul- 21	10- Aug -21
Atlanta	503	2-May-20		59	82.9	96	97	96.8	97.6	98.6	98.8	99
Bogota	339	2-May-20		37.2	76.4	76.7	87.6	90	92	100	97.9	94.1
Casablanca	209	17-May-20		9.1	32.5	40.7	47.4	54.5	60.3	93.8	65.1	67.9
Istanbul	566	24-Apr-20	0	41.7	64.7	78.3	83.6	83.9	88.5	94.9	88.7	91.5
Lagos	180	24-Apr-20	0	25	38.3	38.9	53.9	58.3	62.8	98.3	76.7	76.1
London	842	24-Apr-20	0	53.4	84.8	93.1	97.1	80.8	79	89.7	97.9	98.5
Los Angeles	1,001	2-May-20		40	64	72.8	79.1	82	88.8	95	95	96.6
Madrid	1,437	24-Apr-20	0	44.6	75.9	87.3	92.3	92.9	94.7	98.3	95.8	95.9
Manila	2,750	24-Apr-20	0	41	70.1	79.6	84.6	88.1	89.8	96.6	92.1	92.8
Milan	936	24-May-20		0	59.1	77.1	84	81.9	86.3	95.8	92	90.5
Mumbai	2,939	24-Apr-20	0	45.6	66	72.8	85.7	92.8	94.2	97.2	93.4	93.9
New York	1,278	24-Apr-20	0	47.4	75.7	84.7	92.4	92.3	94.3	96.9	97.2	97
Paris	1,645	24-Apr-20	0	53.5	86	92.5	87.7	89.7	85.6	87.4	96.2	96.9
Rome	1,343	17-May-20		25.4	64.6	82.7	88.2	88.3	89.4	97.2	93.8	94.6

Source: www.elibrary.FMI.org/view/journals/001/2021/295/001.2021.issue-295-en.xml

Método: (B) Indicadores de actividad empresarial

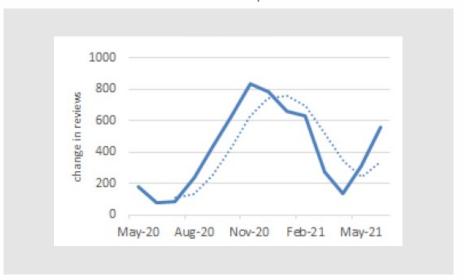
▶ Las Tendencias y Reseñas de Google se utilizan como un proxy de la actividad comercial, suponiendo que existe una relación entre los cambios en el interés en un(os) tema(s) y los cambios en la actividad comercial.

Google Reviews como indicador de la actividad empresarial

La API de Google Places permite a los usuarios extraer el número de reseñas publicadas para un lugar determinado y la calificación promedio (las calificaciones varían de 1 - pobre a 5 - excelente). Se supone que un mayor cambio en las calificaciones se correlaciona con una mayor actividad económica.

► La calificación se utilizó para ajustar el número de reseñas de tal manera que un Lugar con 100 reseñas mal calificadas tendría un peso menor que un Lugar con 100 reseñas altamente calificadas.

Ejemplo de indicador de actividad de opiniones: São Paulo – Cambio en la opinión del restaurante



Source: www.elibrary.FMl.org/view/journals/001/2021/295/001.2021.issue-295-en.xml Nota: Las líneas punteadas en los gráficos son promedios móviles de tres meses.

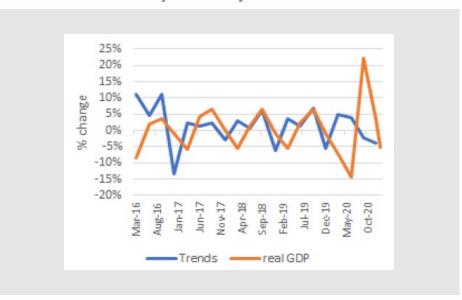
Método: (B) Indicadores de actividad empresarial

▶ Las Tendencias y Reseñas de Google se utilizan como un proxy de la actividad comercial, suponiendo que existe una relación entre los cambios en el interés en un(os) tema(s) y los cambios en la actividad comercial.

Google Trends como indicador de actividad

Dado el número infinito de términos de búsqueda posibles, Google ha desarrollado un algoritmo para agregar búsquedas en categorías de "tendencia" (por ejemplo, categoría de "electrónica de consumo").

► Las categorías de Google Trends se asignan a la Clasificación Industrial Internacional Uniforme de Todas las Actividades Económicas (CIIU) para crear indicadores de actividad empresarial. Ejemplo de indicador de actividad empresarial: Brasil -Comercio al por mayor y al por menor; Reparación de vehículos automotores y motocicletas; Transporte y almacenamiento; Actividades de alojamiento y servicio de alimentos.



Fuente: www.elibrary.FMI.org/view/journals/001/2021/295/001.2021.issue-295en.xml.

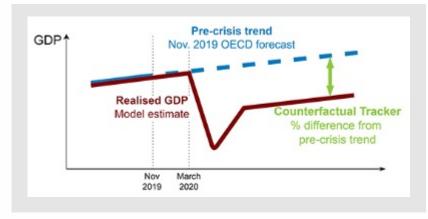
Nota: El indicador se refiere a la CIIU Rev. 4, secciones G, H e I.

OECD Weekly Tracker of Economic Activity (Registro Semanal de Actividades Económicas de la OCDE)

OECD Weekly Tracker of Economic Activity

- ► El Rastreador semanal de crecimiento del PIB de la OCDE proporciona un indicador de alta frecuencia en tiempo real de la actividad económica utilizando el aprendizaje automático y los datos de Google Trends.
 - ▶ Disponible para 46 economías de la OCDE y el G20.
 - ► Aplica un modelo de aprendizaje automático a un panel de datos de Google Trends y agrega información sobre el comportamiento de búsqueda relacionado con el consumo, los mercados laborales, la vivienda, el comercio, la actividad industrial y la incertidumbre económica..

The OECD Contrafactual Tracker



Source: https://www.oecd.org/economy/weekly-tracker-of-gdp-growth/

- ► Hay tres series del Weekly Tracker:
 - ► Tracker (año a año): estimaciones del PIB semanal en relación con la misma semana del año anterior. Cubre el período comprendido entre principios de 2020 y hoy.
 - ► Tracker (año a 2años): estimaciones del PIB semanal en relación con la misma semana en los dos años anteriores (la diferencia de 104 semanas). Cubre el período comprendido entre principios de 2020 y hoy.
 - ► Tracker (contrafactual): estimación de la diferencia porcentual entre el PIB semanal y la tendencia del PIB anterior a la crisis (como lo representan las previsiones de la OCDE realizadas en noviembre de 2019). Disponible hasta finales de 2021.

Cada serie tiene sus propios intervalos de confianza del 95% (bandas inferior y más alta).

OECD Weekly Tracker of Economic Activity

Ejemplos del monitoreo semanal de la OCDE: PIB semanal en relación con el año anterior



Fuente: https://www.oecd.org/economy/weekly-tracker-of-gdp-growth/.

Nota: El Weekly Tracker proporciona una estimación del PIB semanal basada en los datos de búsqueda de Google Trends y el aprendizaje automático...

¡Muchas gracias!

Referencias

- ► Arslanalp, S., M. Marini. y P. Tumbarello, 2019, *Big Data on Vessel Traffic: Nowcasting Trade Flows in Real Time*, Documento de trabajo del FMI No. 19/275 (Washington: Fondo Monetario Internacional). Disponible en: www.FMI.org/~/media/Files/Publications/WP/2019/wpiea2019275-print-pdf.ashx
- ► Austin, P., M. Marini, A. Sanchez, C. Simpson-Bell, y J. Tebrake, *Using the Google Places API and Google Trends Data to Develop High Frequency Indicators of Economic Activity*, Documento de trabajo del FMI No. 21/295 (Washington: International Monetary Fund). Disponible en: www.elibrary.FMI.org/view/journals/001/2021/295/001.2021.issue-295-en.xml
- ► Hu, Y. y J. Yao, 2019, *Illuminating Economic Growth*, FMI Working Paper No. 19/77 (Washington: Fondo Monetario Internacional). Disponible en: www.FMI.org/en/Publications/WP/Issues/2019/04/09/Illuminating-Economic-Growth-46670
- ▶ Woloszko, N., 2020, *Tracking activity in real time with Google Trends*, Documento de trabajo del Departamento de Economía de la OCDE No. 1634 (Paris, OCDE). Disponible en: www.oecd-ilibrary.org/economics/tracking-activity-in-real-time-with-google-trends_6b9c7518-en.