

Comments on the Future of High-Frequency Data

Alberto Cavallo
Harvard Business School

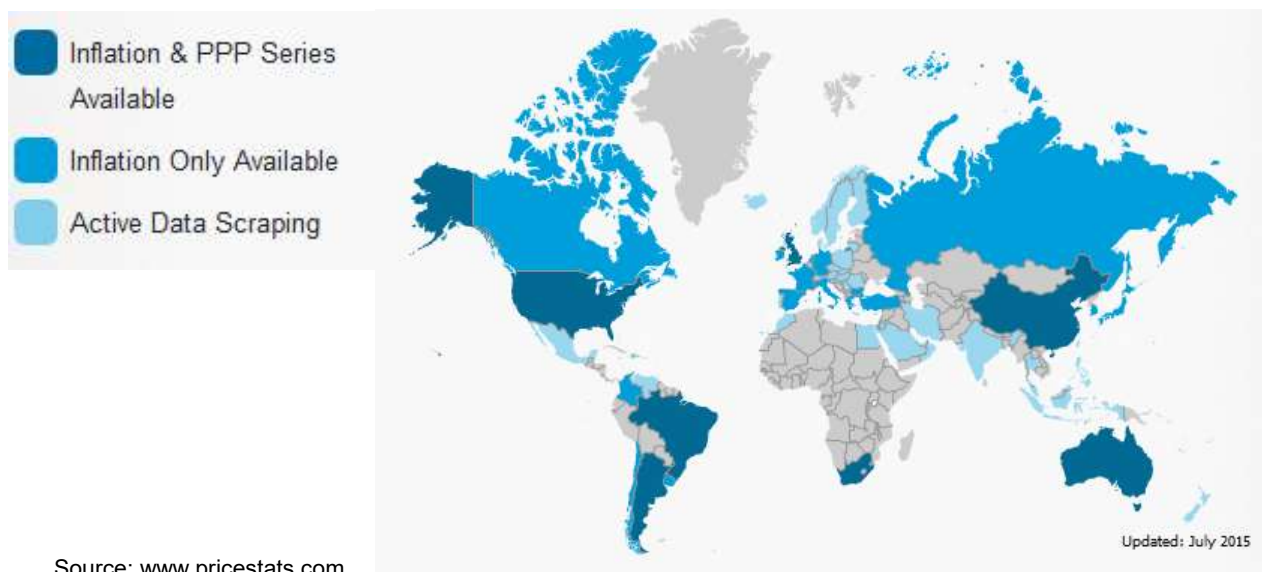
ECB Workshop on High Frequency Data - October 16th, 2020

Big Data and Economic Statistics

- New data sources
 - Gov administrative data (eg. tax & other records)
 - Scanner data (eg. Nielsen)
 - Search/Social data (eg. Google, Indeed, Facebook)
 - Satellite data (eg. lights, parking lots, tanker and crop heights)
 - Sensor data (smart phones, IOT devices)
 - Crowd-sourced data (web, mobile phones)
 - Online data (eg. Billion Prices Project, PriceStats)

Online Inflation Measurement : The Billion Prices Project

- 2008 → Daily online index for Argentina (www.inflacionverdadera.com)
- 2010 → Daily index for the US on the Billion Prices Project website
- 2011 → PriceStats collects the micro data and publishes daily inflation in 22 countries in real-time.



VENEZUELA: CROWDSOURCING INFLATION



InflacionVerdadera.com



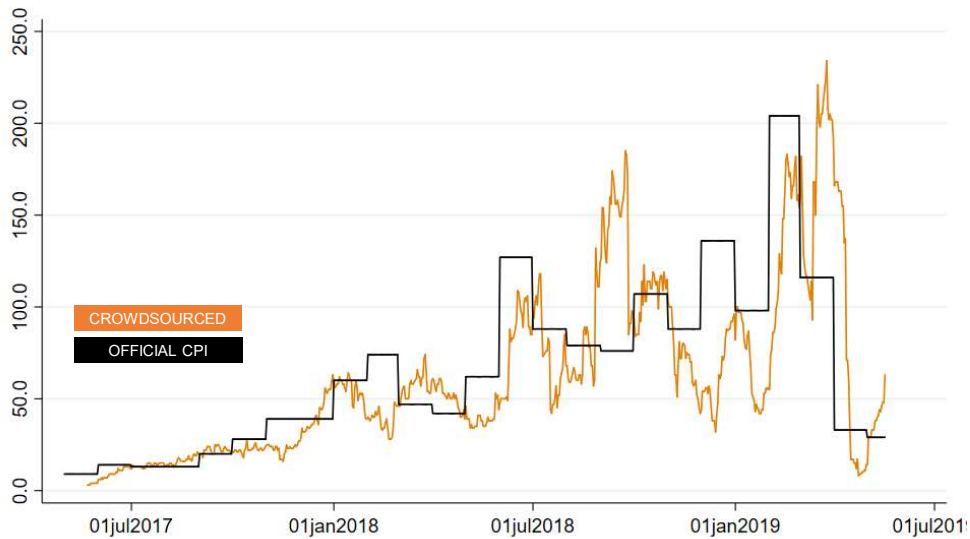
Cual es el precio que se ve en la imagen?
El total a pagar, incluyendo el IVA (PMVP + IVA)

- El precio es:
- No se ve el precio / No se ve la imagen

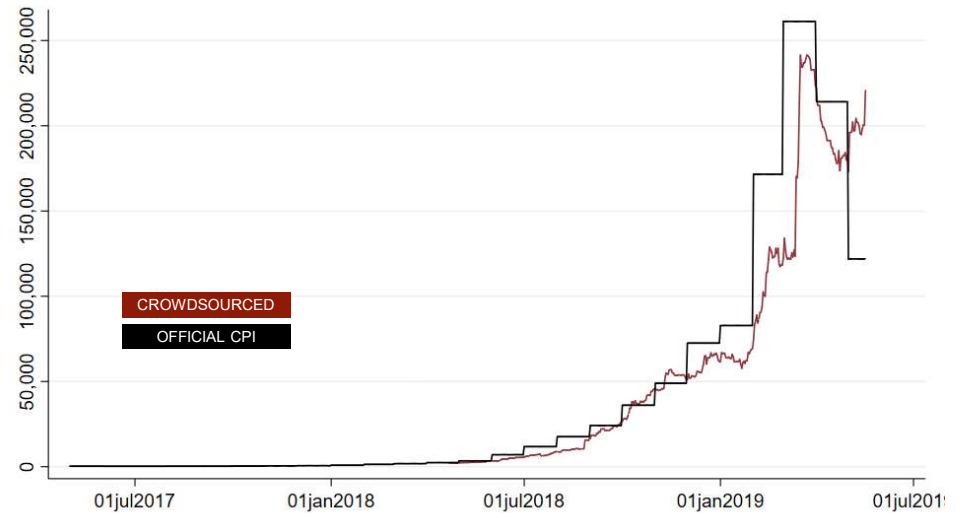
<http://www.inflacionverdadera.com/venezuela/>

Venezuela's official data was finally published on May 29th, 2019

Monthly Inflation Rate (Food and Beverages)

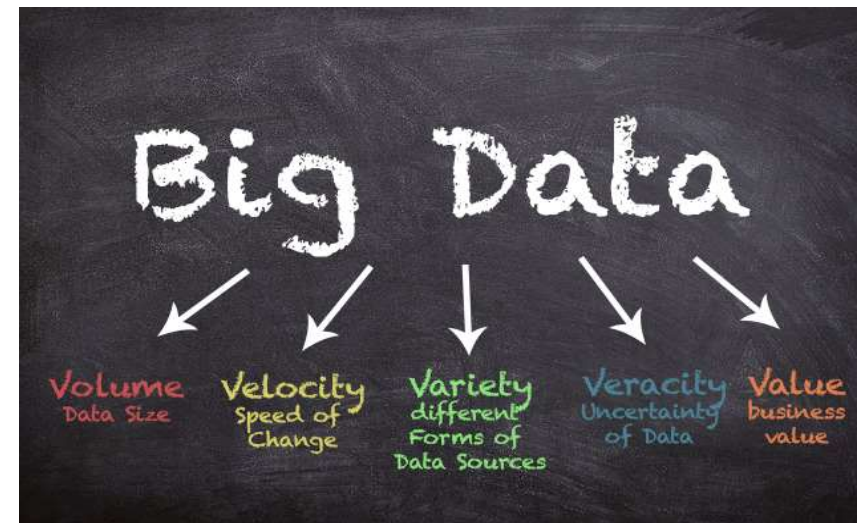


Annual Inflation Rate (Food and Beverages)



Big Data and Economic Statistics

- New data sources
 - Gov administrative data (eg. tax & other records)
 - Scanner data (eg. Nielsen)
 - Search/Social data (eg. Google, Indeed, Facebook)
 - Satellite data (eg. lights, parking lots, tanker and crop heights)
 - Sensor data (smart phones, IOT devices)
 - Crowd-sourced data (web, mobile phones)
 - Online data (eg. Billion Prices Project, PriceStats)
- The 5 Vs
 - Volume (inc. high-frequency) → all
 - Velocity → online, sensor, search, satellite
 - Variety → online, sensor, search, crowdsourced
 - Veracity → ?
 - Value → ?



Comments

- High frequency \neq real-time (velocity).
 - We need both, but real-time is much harder to get right.
- Post-Covid:
 - quick progress in volume, velocity, and variety
 - more uncertain about veracity and value
- Veracity \rightarrow can we trust the data or the source?
 - Measurement biases and noise
 - Data provider s methods, incentives, longevity

\rightarrow Important role for public institutions, including national statistical agencies and central banks
- Value \rightarrow what is it useful for?
 - complement or replace traditional data
 - ad-hoc analysis or permanent monitoring of the economy