Utilizing Public Data Effectively

Alex Izydorczyk, Founder & CEO, Cybersyn

cybersyn

Agenda

- My background

- An alternative data story about hedge funds

- Alternative data examples

- Why now?

- Assessing public datasets for your enterprise

- Demo

Alex Izydorczyk
CEO & Founder

- **~6.5 years leading Data Science at Coatue**
  - Started data science practice; led team of 40+ people
  - Coatue is a major user of external data; 70bn+ investment manager
  - Private equity data infrastructure investor (including in Snowflake)

- **Started something new last year**

A story about the hedge fund industry

What do hedge funds do?

Discretionary Funds

- (Excel) models to predict business performance

- Extremely qualitative

- Long-term orientation

Quantitative funds

- Models to predict financial stock prices as stochastic time-series

- Extremely quantitative

- Short-term orientation

cybersyn

The rise of alternative data

FT Financial Times

Hedge funds see a gold rush in data mining

Investors seek an advantage from tradable intelligence but legality of alternative data is in question.

FP Financial Post

How CPPIB is tapping 'alternative data' to refine its investment processes

There was a time when the Canada Pension Plan Investment Board didn't even make "alternative" investments. Now, Canada's biggest pension...

Mar 18, 2019

The New Gold Rush? Wall Street Wants your Data

January 17, 2017
AI, Big Data, Fintech, Startups

Twitter | LinkedIn | Facebook

MARKETS | FINANCIAL SERVICES | HEDGE FUNDS

Tiny Satellites: The Latest Innovation Hedge Funds Are Using to Get a Leg Up

The latest technological innovation for data-hungry hedge funds is a fleet of five dozen shoebox-sized satellites

By *Bradley Hope* [Follow]

Aug. 14, 2016 4:37 pm ET

What Is Alternative Data and Why Is It Changing Finance?

Sometimes messy, sometimes over-aggregated, alt data is still taking over.

HOME > FINANCE

Wall Street is finding ways to make money off your every move
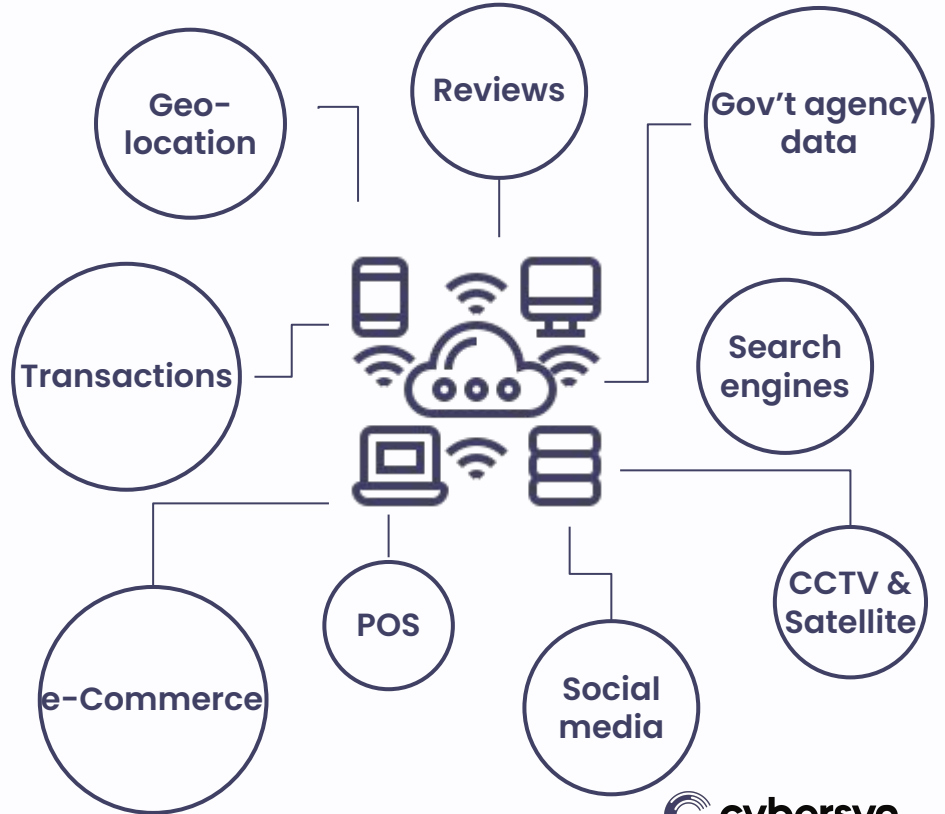
Rachael Levy Nov 21, 2016, 11:25 AM EST

cybersyn

Traditional Data
Company Published Sources

- Trading
- Financial Reports
- Press releases & media
- Interviews & Surveys

Alternative Data
Non-traditional Sources

- Geo-location
- Reviews
- Gov't agency data
- Transactions
- Search engines
- e-Commerce
- POS
- Social media
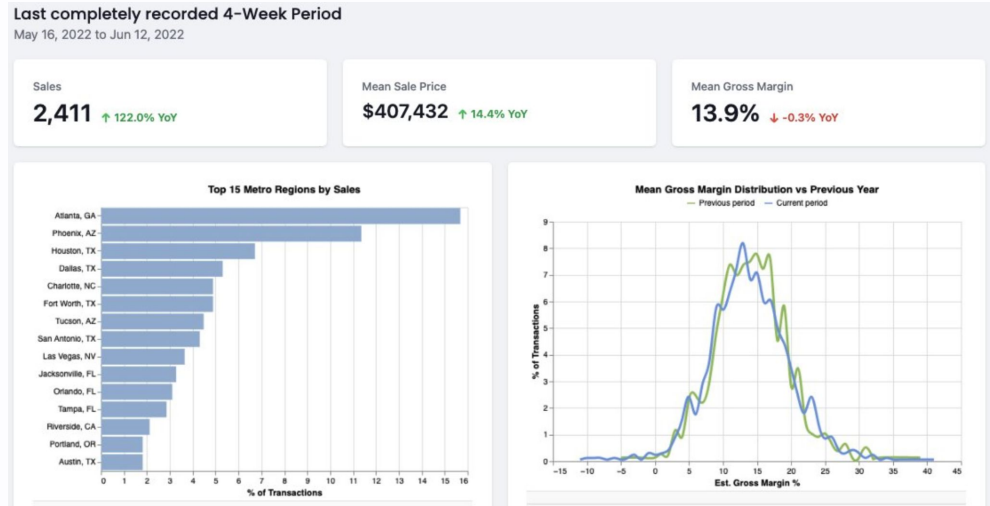- CCTV & Satellite

cybersyn

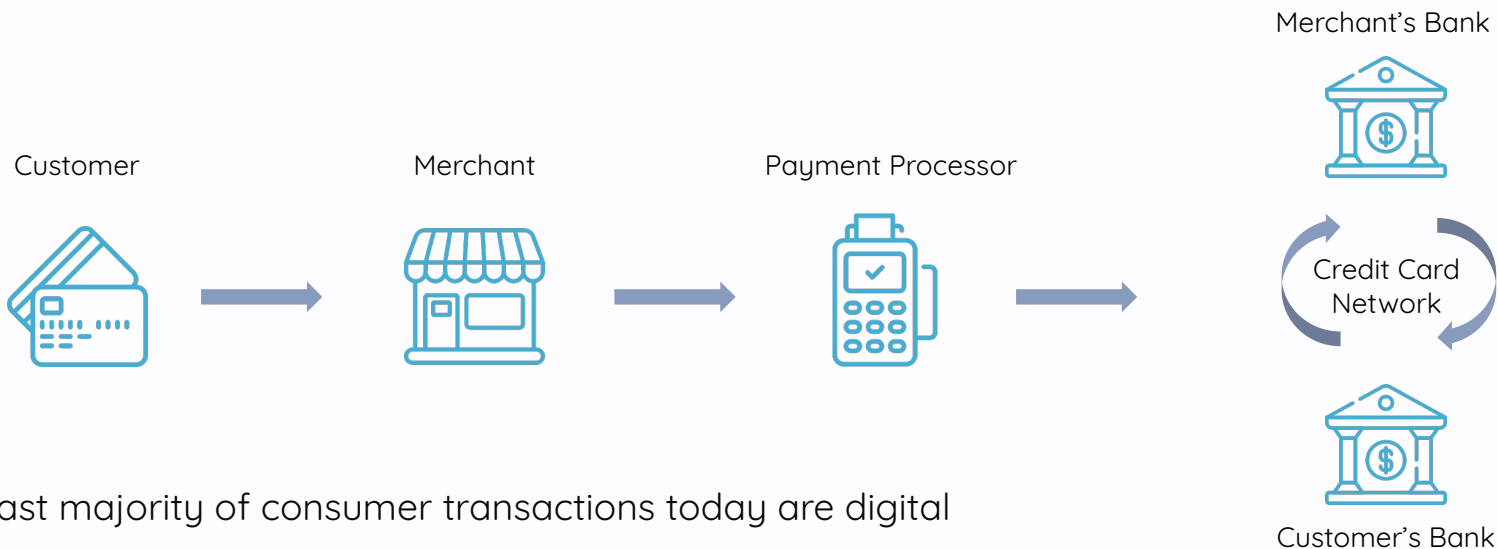# "Hidden" public datasets

Datasets from the government often have incredible detail yet <u>ostensibly public</u> and hard to find, access, or parse

Did you know every home sale in the US is *public information* – including the price and mortgage issuer?

cybersyn

Real-time consumer data

Customer   Merchant   Payment Processor

Merchant's Bank

Credit Card Network

Customer's Bank

- Vast majority of consumer transactions today are digital

- Multiple vendors involved so multiple sources of data that are shaped very differently

- **BUT** need to connect product-level

cybersyn

Why external data now?

1. **Digitization:** Economy has digitized, every transaction is recorded.

2. **Modern Data Stack:** Organizations are centralizing data assets into cloud warehouses and lakes, making access possible.

3. **Privacy Changes:** IDFA changes, GDPR, etc. have made first party data collection harder. Compliant external data solves this.

cybersyn

Hedge funds are advanced but not entirely unique from other industries

- You can build a moat with <u>proprietary data</u>

- But unless you are Google, you have only <1% of the data about your customers

- **<u>If you combine internal data with external, you have something unique!</u>**

cybersyn

I'm so convinced I quit my job!

- **Founded '22,** Cybersyn was built to provide you with market intelligence data that you can act on quickly.

- Our Mission: Our mission is to **make the world's economic data transparent** to governments, businesses, and entrepreneurs and enable a new generation of decision makers.

─── OUR INVESTORS ───

snowflake®

SEQUOIA

COATUE

cybersyn

Cybersyn's Strategy

Get Data and Start Analysis in Minutes
One centralized location for data. Combine disparate data sources.
Always up to date. No engineering required.

Public Domain Data

Proprietary Data

Unstructured Data

Snowflake

Streamlit

via
snowflake | MARKETPLACE

Assessing public datasets for your enterprise

Getting Started

Detailed docs with straightforward data dictionaries and SQL examples

Cybersyn public datasets are free, documented, and joinable

Engineering Effort

Data updated as fast as it is released - your team should focus on data analysis not building pipelines

Cybersyn delivers analytics-ready data to your Snowflake instance

Pricing and Compliance

Transparent pricing via enterprise-wide licenses, detailed restrictions on data usage

Cybersyn provides simple terms and pricing for enterprises

cybersyn

Free public datasets

- Global economic data

- US-specific inflation, sales inventory

- SEC company filings

- Unemployment rates, openings, and new hiring in the US

- Real estate prices at the state, city, MSA, and county level

- US addresses

- GitHub archive

- US patent grants

- US government contracts

cybersyn

Demo

Cybersyn Government Essentials Streamlit App

We are hiring!

- Analytics Engineer

- Analytics Engineer Lead

- Data Science Lead

- Product Manager - Data Visualization

Learn more and apply here

cybersyn

Questions?

Check out our products: **Snowflake Marketplace**
Follow me: **Twitter**, **Blog**
Sign up for our bi-weekly newsletter: **Subscribe**
Follow Cybersyn: **LinkedIn**, **Twitter**