Weather/Emergency Response Topic

Amy Unruh
Google
@amygdala
@amy on **bigdatahubs** slack

Eddie Fuller
Department of Homeland Security

“How might we improve transportation safety in extreme weather? In cases of crises?“

About me

- **Google Cloud Platform** Developer Relations Engineer (& lapsed academic)
- Focus on ML, Data Science, big data and analytics (and Google OSS)
  - Things like: TensorFlow, Dataflow/Apache Beam, BigQuery, Cloud ML APIs, Kubernetes
- You can find me **@amy** on the bigdatahubs Slack, amyu@google.com

what are some of the things that make this a really interesting problem?

Many relevant data sources…

- weather (both historic and alerts),
- traffic flow, taxi, bus, and bike data, traffic incidents/fatalities, ‘smart city’ data..
- fire, earthquake, stream flow, flood hazards, ..
- social media, ‘unstructured’ data: Twitter, Instagram, …
  - (callout to Cloud ML APIs: Vision, NLP…)
- census and population data, zip codes —> lat lon, …

Crowdsourcing platforms like: https://github.com/ushahidi/platform

Data Characteristics

- Need to integrate multiple heterogeneous information sources
- Data sources updated at differing frequencies— including real-time streams
  - … and reporting may be delayed
- Temporal information
- Geospatial information
- Unstructured data

Analytics considerations

- Lots of historic data —> stats and ML for finding patterns and making predictions
- Incident data will arrive in stages, possibly “out of order”, and might be used to tune earlier predictions
- It may be necessary to incorporate large-scale, high-qps RT data streams
- Actions/mitigations are time-sensitive

Data Sources

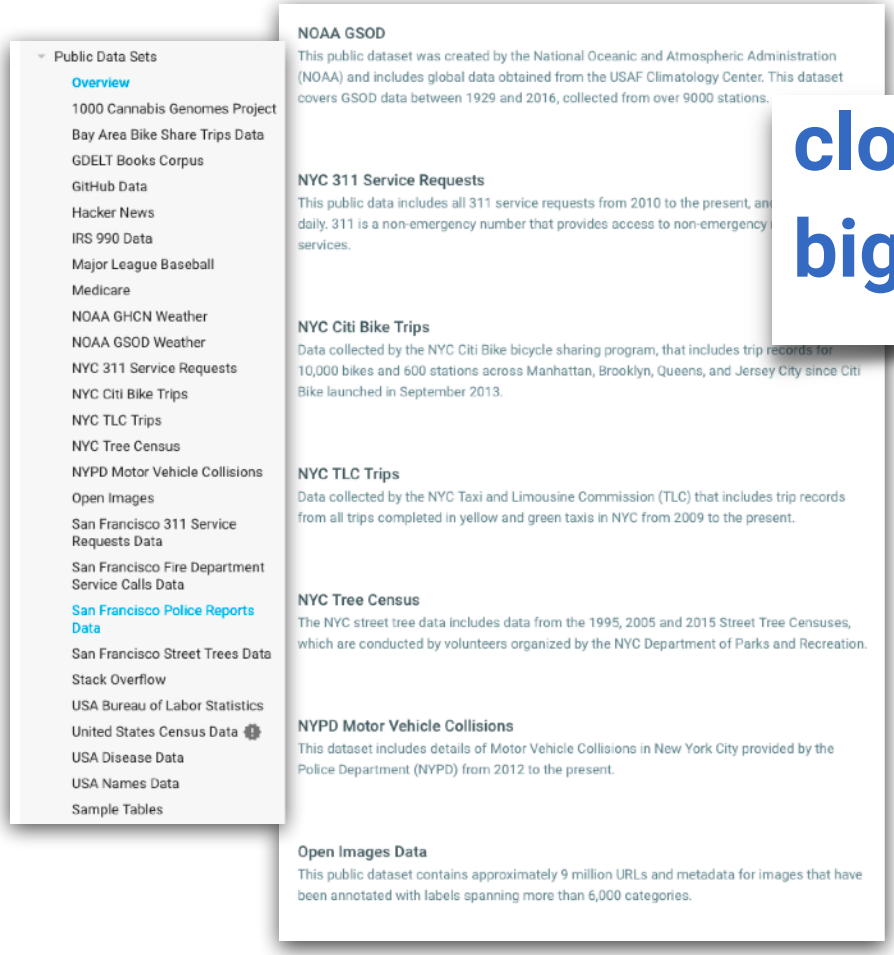
(we'll create a doc for these)

Google BigQuery's public data sets

BigQuery

Managed **big data analytics warehouse**

- **Fast**: terabytes in seconds
- **Simple**: SQL
- **Interoperable**: Java, Python, Tableau, R…
- Instant data sharing
- Free monthly quota - 1 TB



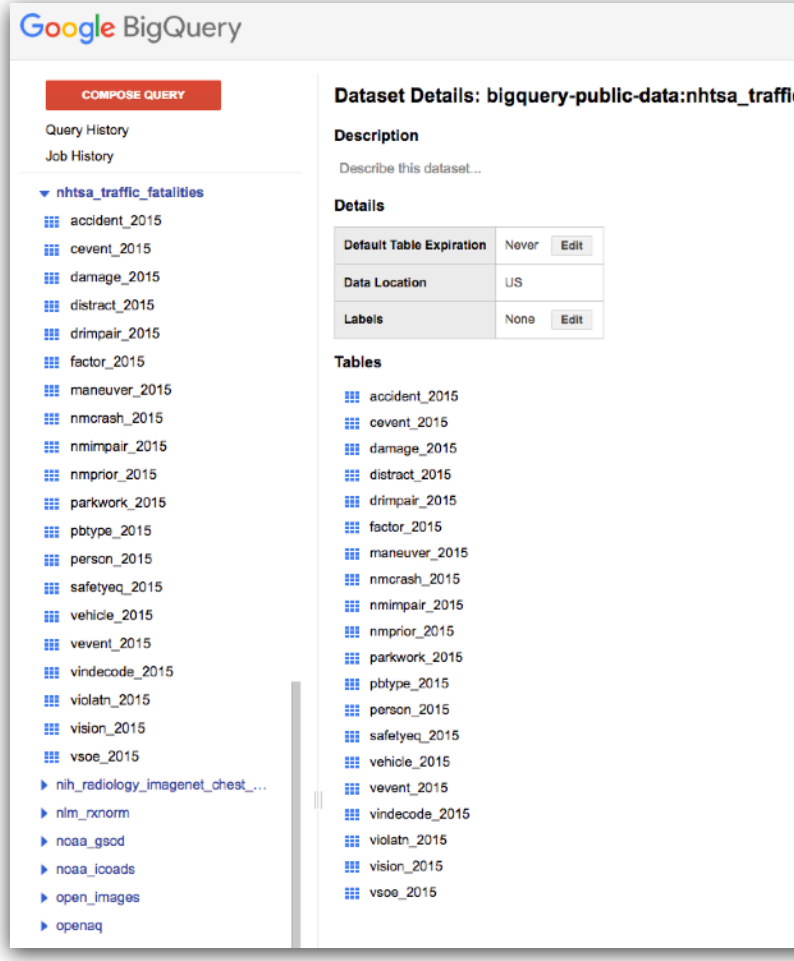
cloud.google.com/bigquery/public-data/

Just added as a BigQuery public dataset!: **NHTSA Traffic Fatality Data**

https://cloud.google.com/bigquery/public-data/nhtsa

This public dataset was created by the Unites States Department of Transportation's National Highway Traffic Safety Administration (NHTSA) and includes 20 tables that describe numerous aspects of traffic accidents that resulted in fatalities.

Aspects of traffic accidents include: the types of cars and roads, the maneuvers that preceded the accident, and the involvement of pedestrians and cyclists.



BigQuery NHTSA Traffic Fatality Data: example query

```
#standardSQL
SELECT
  state_name,
  COUNT(consecutive_number) AS accidents,
  SUM(number_of_fatalities) AS fatalities,
  SUM(number_of_fatalities)/COUNT(consecutive_number) AS fatalities_per_accident
FROM
  `bigquery-public-data.nhtsa_traffic_fatalities.accident_2015`
GROUP BY
  state_name
ORDER BY
  fatalities_per_accident DESC;
```

state_name	accidents	fatalities	fatalities_per_accident
North Dakota	111	131	1.18018018
South Dakota	115	133	1.156521739
Vermont	50	57	1.14
Iowa	282	320	1.134751773
Nebraska	218	246	1.128440367
Texas	3124	3516	1.125480154
Arkansas	472	531	1.125
Wyoming	129	145	1.124031008
Mississippi	604	677	1.120860927
New Mexico	269	298	1.107806691

And more…

UtahDOT:
udottraffic.utah.gov
udot.numetric.com
udot.bt-systems.com

Weather:
NOAA Storm Events Database– https://www.ncdc.noaa.gov/swdi/#Intro
Census American Community Survey – https://www.census.gov/programs-surveys/acs/
Census Business Patterns – https://www.census.gov/programs-surveys/cbp.html
FEMA Flood Map Data - https://msc.fema.gov/portal
National Flood Data - https://www.data.gov/floods/
National Flood Hazard Data Map Layer -
https://catalog.data.gov/dataset/national-flood-hazard-layer-nfhl

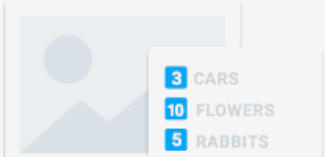
what’s next?

What's next?

- Shared Topic doc to be created
  - Let's collect links to data sources
  - (Let's find some more good candidates for BigQuery public data sets)
- Regular video meetings
- Slack channel
- (talk to me if you want to use Google Cloud Platform, or are interested to learn more about its Data Science and ML products)

thank you

bridge.jpg



Insight From Your Images

Easily **detect broad sets of objects** in your images, from flowers, animals, or transportation to thousands of other object categories commonly found within images. **Vision API improves over time** as new concepts are introduced and accuracy is improved.

Try the API

Google, headquartered in Mountain View, unveiled the new Android phone at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.

ANALYZE

See supported languages

Insights from your customers

Extract actionable insights on product reception or user experience from customer conversations in email, chat or social media by using entity detection and sentiment analysis.

Multimedia, Multi-lingual Support