Cloudant crime visualization demo

Contact info:

Developer: Raj Singh

email: rrsingh@us.ibm.com twitter: @rajrsingh

web: http://ibm.biz/rajrsingh tel: 617-642-92372

Marketing: Doug Flora

email: dsflora@us.ibm.com tel: 901-289-7954

Data engineering

Developer Advocates from IBM Cloud Data Services are building a system to harvest crime data from cities across the country to create the largest up-to-date, freely available information base for analyzing public safety patterns available today. Among the cities available today are Boston, San Francisco, Chicago and Las Vegas, along with all of Great Britain. Thanks to municipal governments that have the foresight to join the open data movement early, businesses can leverage these data to make smarter decisions.

We harvest crime data daily into a single Cloudant database on the IBM Cloud where the crime events are standardized across jurisdictions into 3 major types: street crime, violent, and domestic. This allows customers to create apps and analytics that work everywhere without rewriting a single line of code.

Here are some details about how the data ingest works.

A NodeJS application in Bluemix runs every morning at 5am Eastern and “harvests” data for the prior day into Cloudant. It not only grabs the new data, but uses a lookup table to tag each crime with an aggregate crime type: **CDSSTREET** (street crime); **CDSDV** (domestic violence); and **CDSNV** (non-violent). These are NOT mutually exclusive categories by the way. Since each city has its own categorization scheme for crime data, having these 3 consistent categories across all crimes allows for multi-region analysis and thematic mapping.

The NodeJS application is here:

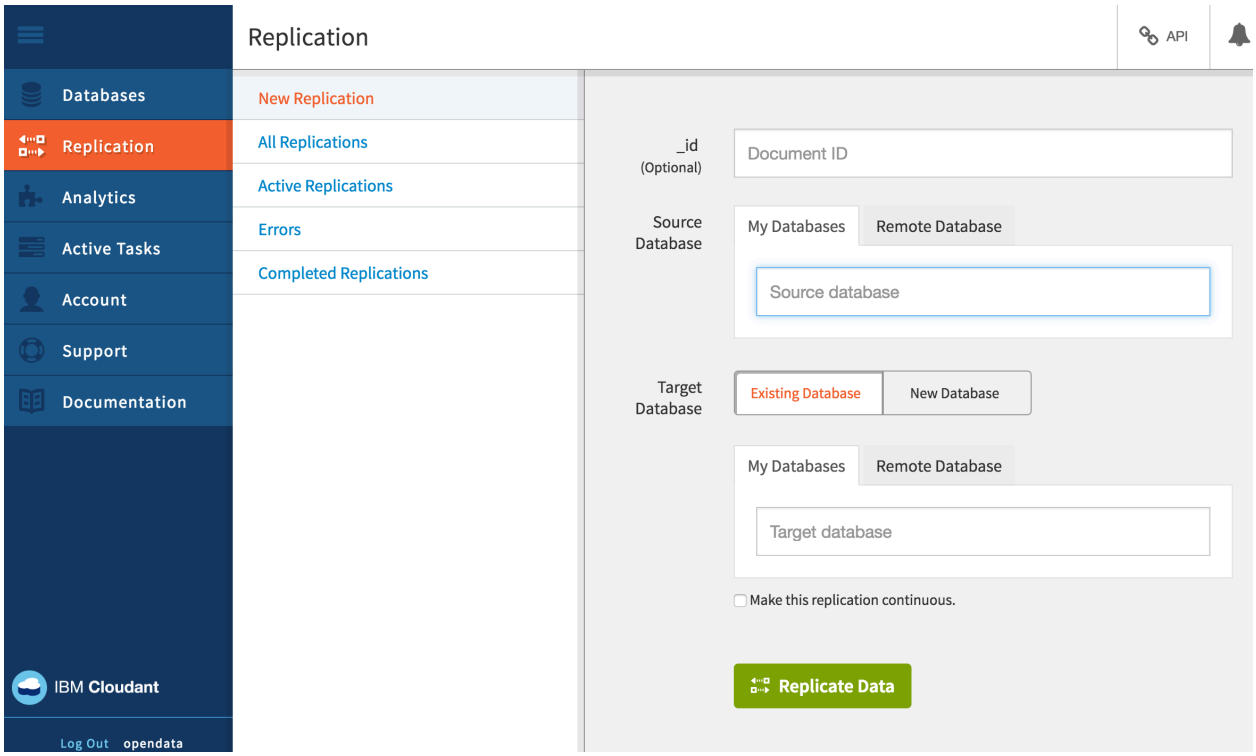
- https://github.com/ibm-cds-labs/crimeharvest

And the lookup tables, which are different for each city, are here:

- https://github.com/ibm-cds-labs/open-data/tree/master/crime

Replicating the database

Let's replicate the Las Vegas data and look at the raw GeoJSON, a standardized flavor of JSON that facilitates mapping and spatial analysis.

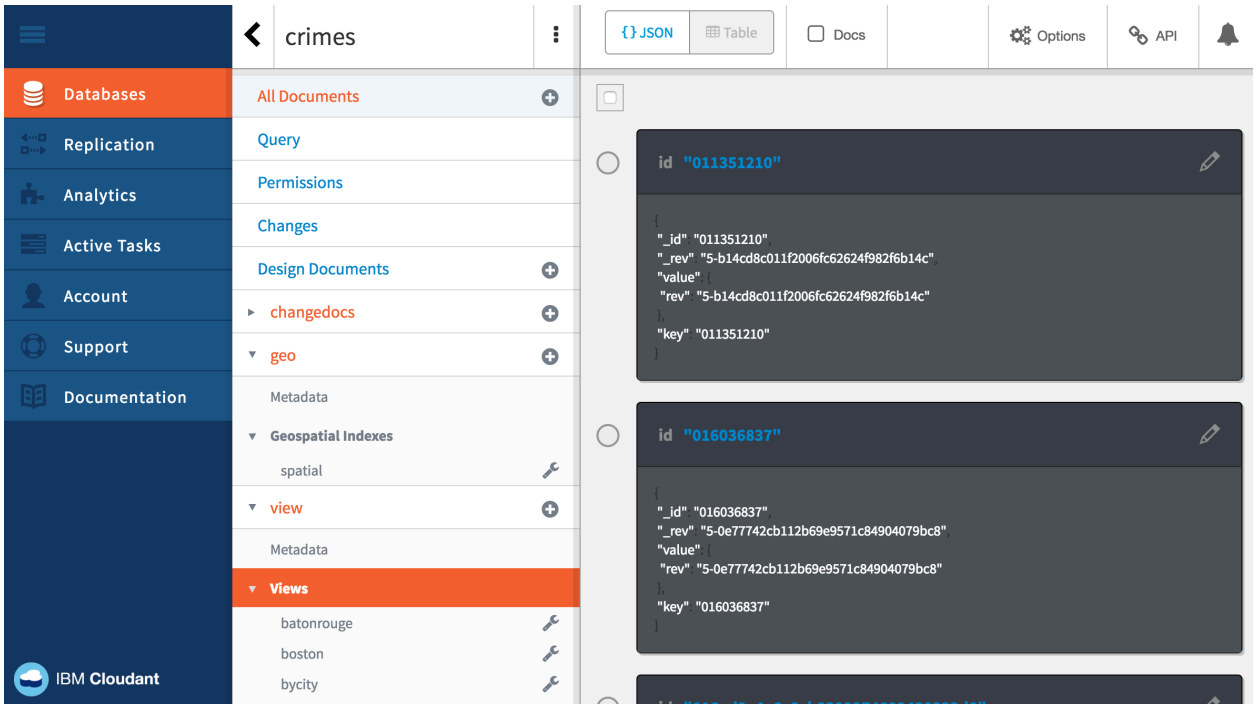


Instructions:

1. In the Cloudant administrative dashboard, click on Replication -> New Replication
2. Under **Source Database** click on **Remote Database** and enter this: https://opendata.cloudant.com/crimes
3. Under **Target Database** click on **New Database** and type crimes
4. Click **Replicate Data**

Database administration exercises

Cloudant's administration tools let you preview your GeoJSON data on a map, supporting pan, zoom and clickable features.



When you replicated the database, you also replicated the indexes. Click on the **Databases** tab and select your **crime** database. Expand the indexes by clicking on the little disclosure triangles next to **geo**, **Geospatial Indexes**, **view** and **Views**. If you click on any of the **Views** indexes you can see JSON documents for just that city. Double-click on the documents on the right side of the window (or click the pencil icon) to see the full JSON document.

If you click on the **spatial** index, you will see a map. It will initially be centered in the “middle” of the world, latitude=0/longitude=0. You need to manually zoom out and scroll over to North America, then zoom into Las Vegas (this is a product limitation and there's currently no way around it). Once zoomed into Las Vegas, you can begin drawing queries with the 4 options on the right side - box, hex, circle, polygon.

Using the demo app

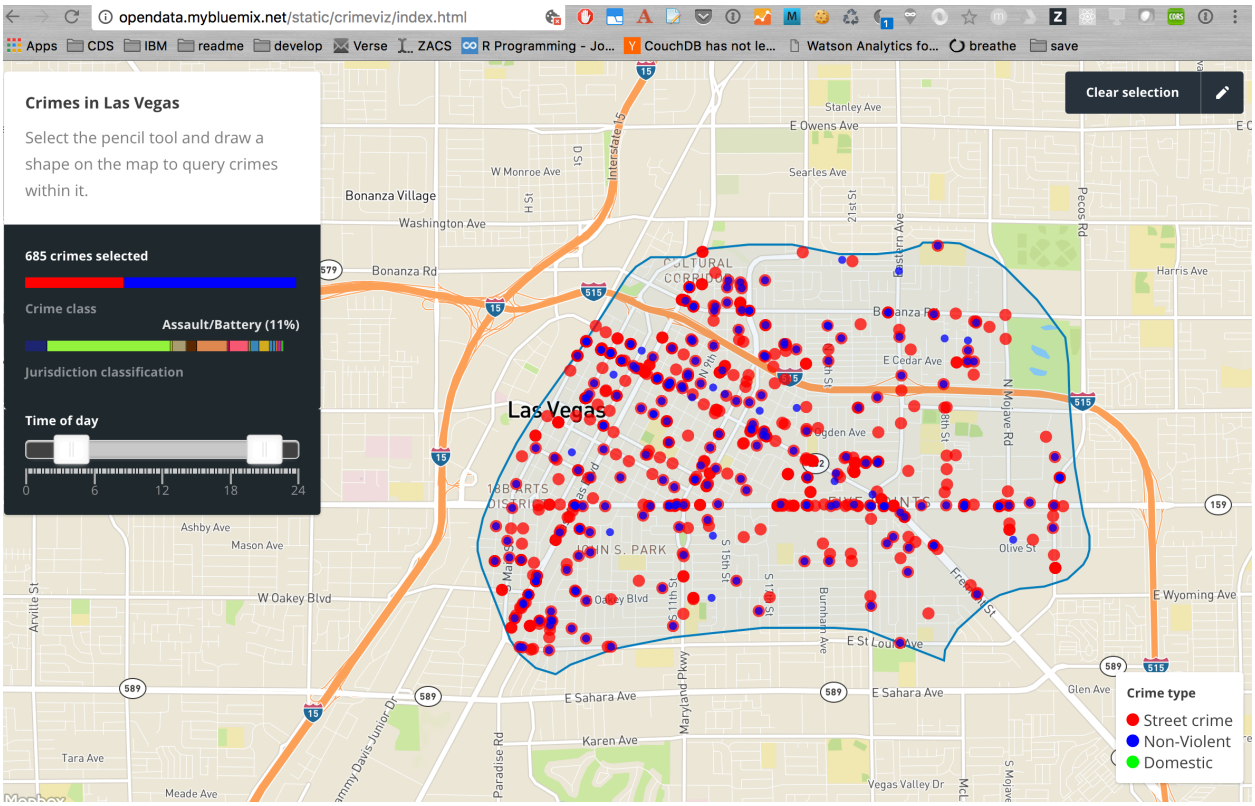
The purpose of the demo app is to highlight spatial and temporal analytics. The URL is:

- http://ibm.biz/crimeviz

These demo instructions are online here:

- http://opendata.mybluemix.net/static/crimeviz/crimevizdemo.pdf

You can select crimes on the map using the pencil icon in the top right, and look at the distribution of crime types, and use the slider to focus in on different times of day.



The legend on the left shows:

1. The total number of crimes in the outlined area
2. **Crime class:** % of crimes falling into the 3 aggregate categories (same as legend in bottom-right)
3. **Jurisdictional classification:** % of crimes by the city's categorization scheme
4. **Time of day:** slider to adjust start time and end time, which restricts the display of crimes on the map and updates the stacked bar charts above.