VC Funding Worldwide 2005-2015

Web Application Data Visualization Dashboard

Deployed App: https://frozen-dawn-61622.herokuapp.com/

By: Patrick Gendotti, Kanishka Ramanan

Project Motivation - Why?

- What are some major startup hubs around the world - they may be an indicator of tech hubs?
- Is there a large international presence of startup funding?
- Are there factors that contribute to a startup being successful vs. failing that can be determined from this dataset?
- What are the top 10 cities in the world by number of startups? By average startup funding? By total startup funding?
- What percentage of startups are found where in the world - is it evenly distributed, or are there highly concentrated areas with many startups (like Silicon Valley)?

Project Overview & Methodology

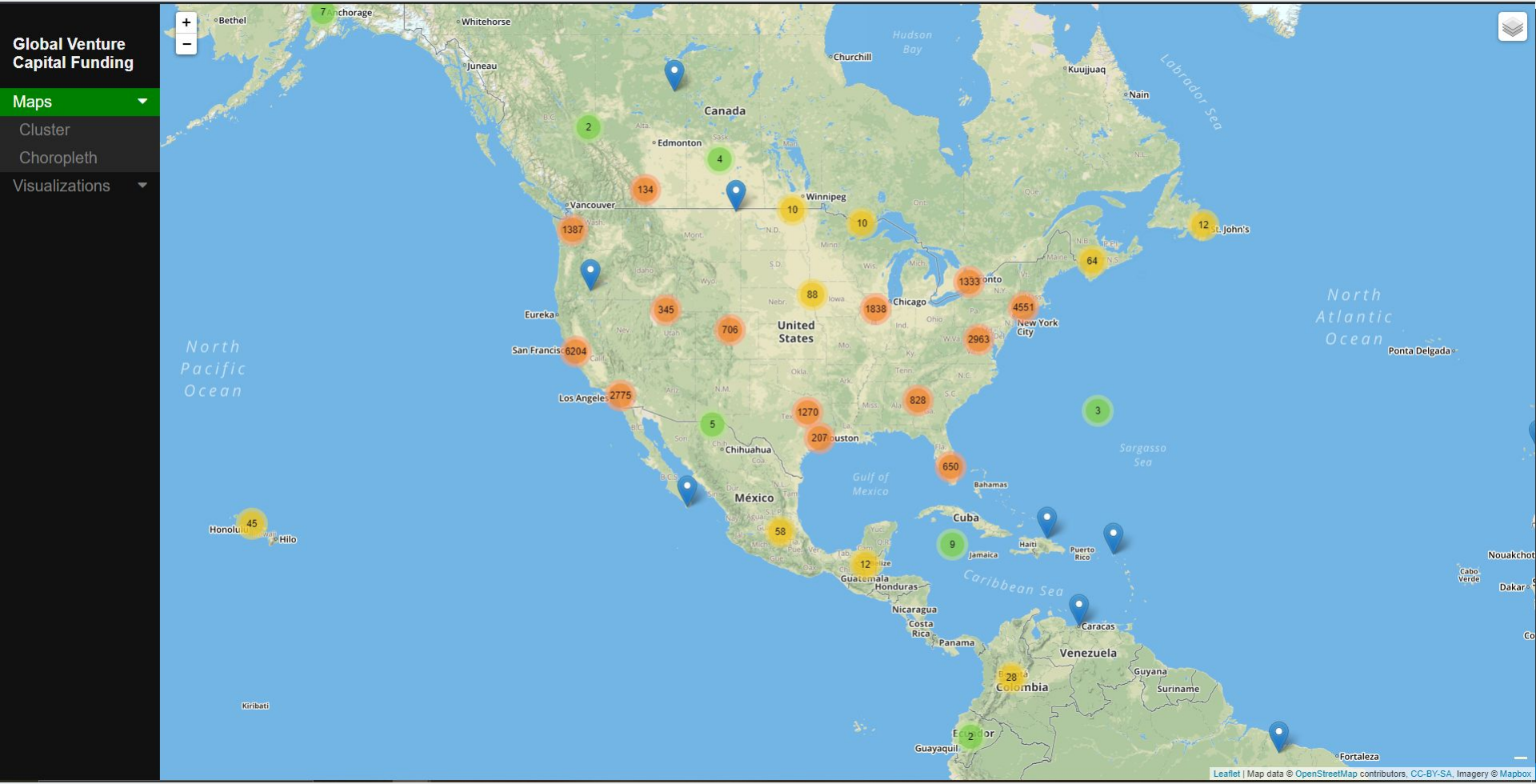
- Data sourced from Kaggle, Crunchbase, and Google Maps Geocoding API
- Data loaded into PostgreSQL using SQLAlchemy and pandas
- Visualizations served as a Flask app
  - Two geographic map visualizations that use Mapbox with the Leaflet framework
  - Four bar chart funding visualizations that uses Chart.js - a JS library from outside this class
- Two RESTful APIs serving the visualizations in this app
  - ./samples.json to provide snapshot of entire, aggregated records of database to geographic visualizations
  - ./api/city/<country_code>/<city-name> to provide single-city data for bar chart funding visualizations
- Project served publicly using Heroku at https://frozen-dawn-61622.herokuapp.com/

Single-Page, Full-Stack App

- Single-page web application design gives users a GUI for accessing the data
- Deployment on Heroku enables access anywhere with an internet connection, on all devices - including smartphones and tables
- Simple, intuitive user controls for viewing data visualizations
- Popular, small-footprint JavaScript frameworks used for creating the data visualizations quickly and effectively
- Dynamic links from the Flask app to Heroku-hosted Postgres database enables dynamic updates upon database changes through RESTful APIs

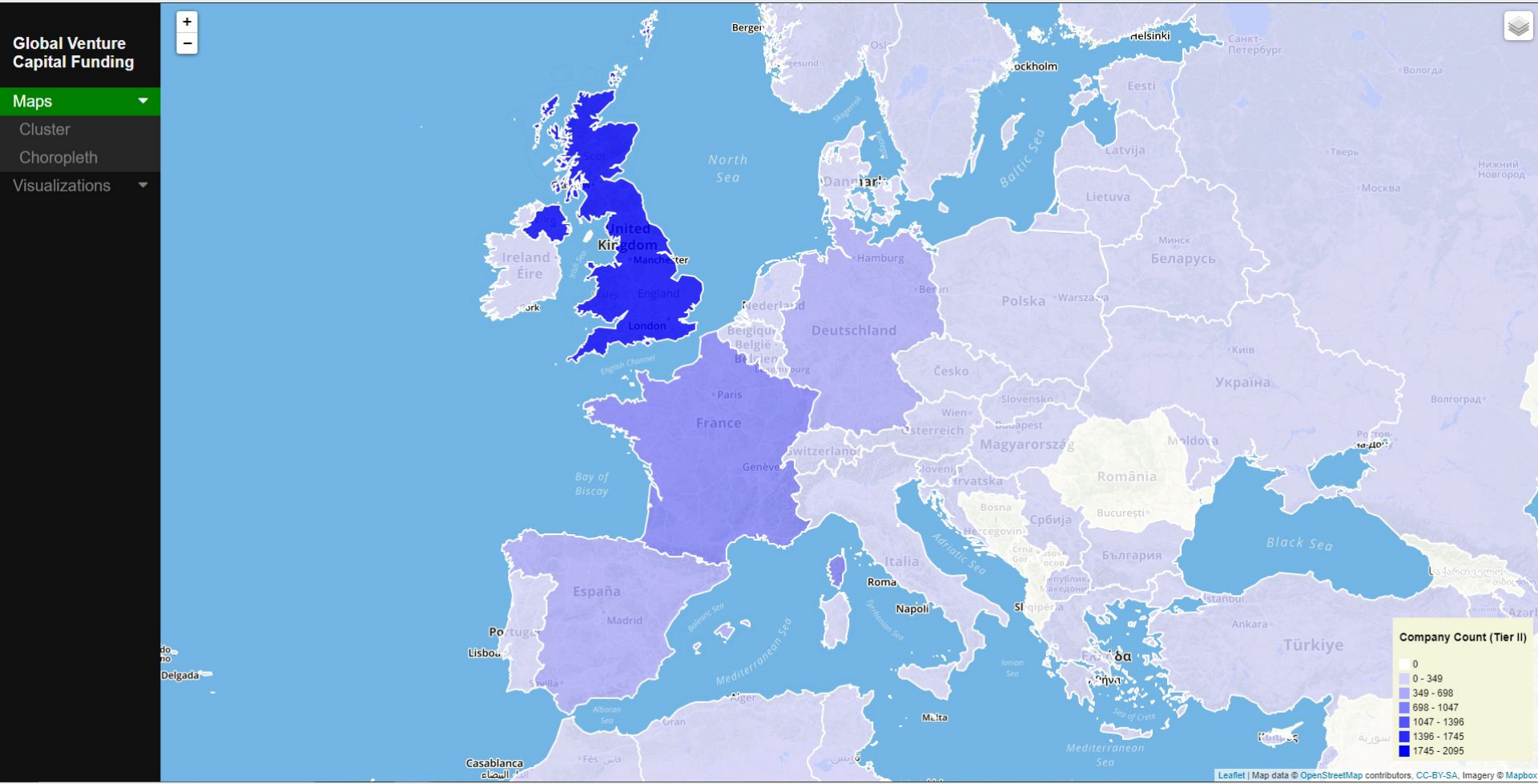
Cluster Map Visualization

- Five different “layers” of markers showing different data
  - Startup count by city
  - Total funding amount by city (all startups within city)
  - Average funding amount by city
  - Percentage of number of startups in one city vs. total amount of startups worldwide in dataset
  - Percentage of total startup funding in one city vs. total startup funding worldwide in dataset
- Different marker styles and legends for each layer to make the map easier to understand
- Cluster map enables high levels of detail when zoomed in, and combined, generalized information when zoomed out
- Can look for overall trends as well as examine some smaller geographical areas for specific information about startups


Global Venture Capital Funding
Maps
Cluster
Choropleth
Visualizations
Canada
United States
México
Cuba
Venezuela
Colombia
Guyana
Suriname
Nicaragua
Honduras
Guatemala
Belize
Costa Rica
Panama
Jamaica
Haiti
Puerto Rico
Bahamas
Gulf of Mexico
Caribbean Sea
Sargasso Sea
North Atlantic Ocean
North Pacific Ocean
Hudson Bay
Labrador Sea
Gulf of Alaska
Bethel
Anchorage
Whitehorse
Juneau
Churchill
Kuujjuaq
Nain
Edmonton
Vancouver
Winnipeg
Chicago
New York City
Eureka
San Francisco
Los Angeles
Chihuahua
Houston
Honolulu
Hilo
Kiribati
Guayaquil
Caracas
Fortaleza
Ponta Delgada
Cabo Verde
Dakar
Nouakchott
St. John's
2
4
134
1387
10
10
12
64
1333
88
345
706
1838
4551
2963
6204
2775
1270
207
828
650
5
3
58
12
9
45
28
2
Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

Country Choropleth Visualization

- Three different “layers” of information available on this map
  - Number of startups per country (US dominant, so needed additional layer to show second-tier countries for startups)
  - Number of startups per country excluding the US
  - Average funding for startups in each country
- Enables viewing relative count of startups by country/continent
- Clear US dominance in startup count, which when multiplied by avg. startup value, results in highest influx of VC funding (according to Crunchbase sources)
- Promotes understanding of different average funding levels for startups by country, broken into three categories - high funding, some funding, and no funding


Global Venture Capital Funding
Maps
Cluster
Choropleth
Visualizations
Company Count (Tier II)
0
0 - 349
349 - 698
698 - 1047
1047 - 1396
1396 - 1745
1745 - 2095
Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

VC Funding Bar Charts

- Has a hard-coded array of the Top Cities; then fetches the data through our API
- Due to America's dominance in the Dataset, the charts are split into Top 5 American Cities and Top 5 Cities Excluding America
- Visually Distinguishes Funding by Equity and Debt
- Total Funding
    - Breaks up Fundings by Type and Round
- Average Funding Per Company
    - Splits funding from Equity and Debt

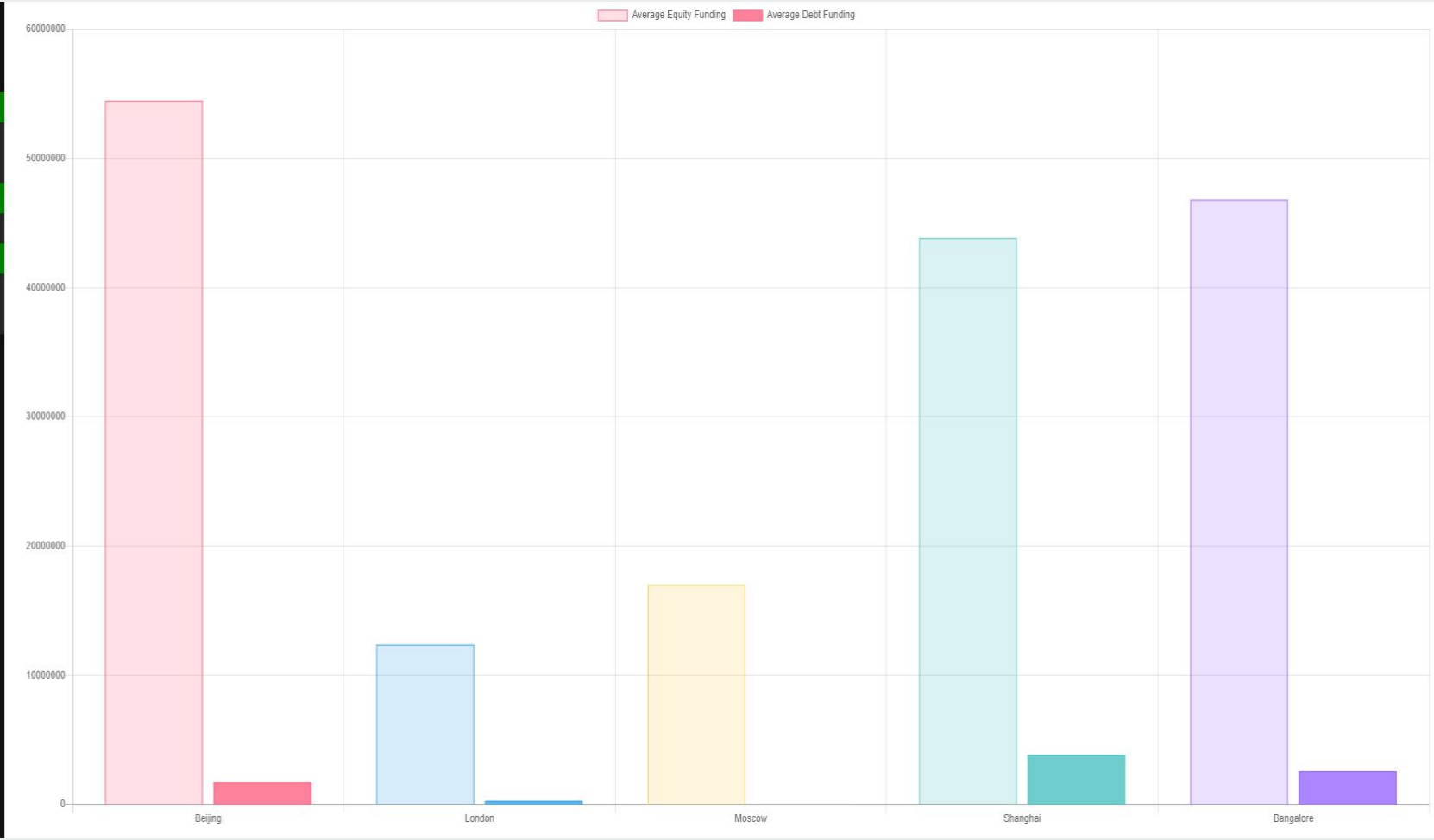
Used Charts.js to create

Global Venture Capital Funding

- Maps
  - Cluster
  - Choropleth
- Visualizations
  - Total Funding
    - Top 5- US
    - Top 5- Not US
  - Avg Funding

Global Venture Capital Funding

- Maps
  - Cluster
  - Choropleth
- Visualizations
- Total Funding
- Avg Funding
  - Top 5- US
  - Top 5- Not US



Caveats about Dataset

- Data has some issues with accuracy
  - Had an entry for the city “Albert, Germany” but the company website showed a company in Great Prairie, Alberta, Canada
- Data is older, and may be out of date with current trends
  - Only startups that received funding between 2005 and 2015 are shown - a good next step in this project would show similar visualizations with date filters
- Data may be incomplete
  - The funding was in USD, and most of the startups were headquartered in the US, Great Britain, and Canada
  - There may be foreign startups or funding sources that are not reported or known to Crunchbase, especially in countries with web censorship and vague accounting rules like China and Russia

Questions?

Libraries Used

- Data Cleaning, Scraping, & SQL Loading (Python)
  - pandas
  - requests
  - json
  - pycountry
  - JQuery
  - time
  - gmaps
  - os
  - SQLAlchemy
- Dashboard App - Backend (Python)
  - Postgres
  - SQLAlchemy
  - Flask
  - os
- Dashboard App - Frontend (JS)
  - Chroma.js
  - Choropleth.js
  - ISO 3166-1 alpha-2 country code to country name conversion
  - Leaflet
  - Leaflet.markercluster
  - os
  - Chart.js