

GovHack2015 New Zealand

Team: WWSNNZ

Joshua Scarsbrook



What is the Wettest Street Name in New Zealand?

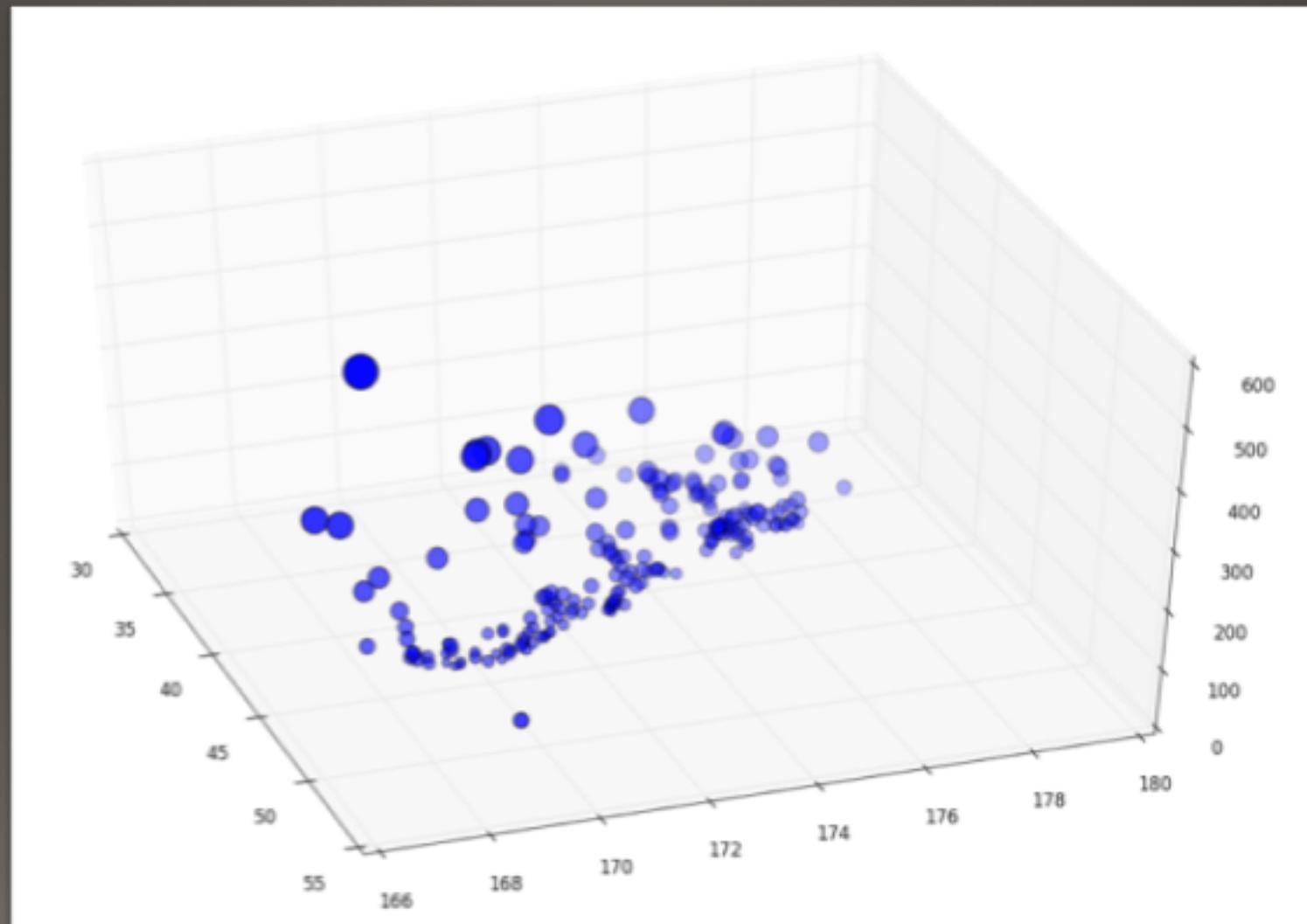
Rainfall Information

Rainfall information is sourced from The National Climate Database (<http://cliflo.niwa.co.nz/>).

Data used is limited to a set of weather stations that have a lot of data and are current active.

The data queried is average rainfall for each month over the past 5 years (2010-2015).

The graph was plotting with the latitude and longitude in X and Y and the mean rainfall over the entire data period as the Z and point magnitude on a scatter graph with matplotlib (<http://matplotlib.org/>) and mplot3d.



Street Address Data

The Street Address database comes from the LINZ Electoral Street Address layer (<https://data.linz.govt.nz/layer/779-nz-street-address-electoral/data/>).

Because of the size of the dataset it was imported into a MySQL database to make querying easier.

Latitude	Longitude	Address
-44.399	168.047	104 Lake Street
-44.4014	168.049	64 Lake Street
-44.4013	168.049	66 Lake Street
-44.3993	168.048	94 Lake Street
-44.3992	168.048	96 Lake Street
-44.7597	168.139	805 Lower Hollyford Road
-44.7412	168.314	314 Routeburn Road
-44.7639	168.342	404 Glenorchy-Routeburn Road
-44.7724	168.342	399 Glenorchy-Routeburn Road
-44.7796	168.349	325 Glenorchy-Routeburn Road
-44.8275	168.353	668 Kinloch Road
-44.8424	168.349	846 Kinloch Road
-44.8424	168.349	848 Kinloch Road
-44.8425	168.349	850 Kinloch Road

Problems encountered initially

- Without interpolating data there's a massive gap in resolution between the street address data and weather station rainfall data.
 - `select count(*) from weatherStation where YEAR(end_date) = 2015 and percent_complete > 90;`
 - The above SQL query was used to filter weather stations to ones with a useful amount of data and returns 219 stations.
 - `select count(*) from streetAddress;`
 - The above query returns the total number of street address's in the country and returns 1,865,589 results.
- This results in a predictable result for the question. What's the wettest weather station in the country and what street address is closest to that station?

The solution.

- Climate data can be interpolated if you want higher resolution but you end up with a bad result unless you implement a climate model.
- NIWA already does this with the VCSN (Virtual Climate Station Network). Public access to that network was suspended on 13-Oct-2014 so I can't use that information for GovHack2015.
- A naive approach to the same problem is just to weight all data points based on distance to stations to provide a better map.

The Solution (Cont.)

- This solution fixes the problem but remains a computationally expensive solution. There are 1.8 Million addresses to calculate and tally the results of, although my code makes use of cached function results even the mean rainfall information for streets takes at least 1 hour to run, changing it to the interpolated approach would mean caching would just eat memory without providing any benefit.

Results

- A process and code for acquiring data from the national climate database and inserting it into MySQL.
- A library for resolving address's to a latitude and longitude based on government data and vice versa.
- A database schema for connecting the data and reading connected information with MySQL.
- Oh and the wettest street name is Sinbad Drive, the weather station that contributes that result is the one in Milford Sound Airport. Because google maps doesn't have a street name for that address the Latitude and Longitude is (-44.6737, 167.922)

