



# **TOWN PLANNING USING MACHINE LEARNING**

**IBM DATA SCIENCE COURSE ON COURSERA**

Nikhil Verma

# Objective

*Identify areas in Sydney suburbs where new schools are required*

If new schools cannot be built, then these identified areas need better connectivity to existing schools which might be at a distance from the main community centres

## **Motivation for this project:**

Town planning authorities need to better understand the local requirements of their citizens to plan expansion of services, infrastructure layout and utility services. It is a critical exercise to allocate limited resources for maximum benefit.

# Methodology overview

**Step 1:** Get data of Sydney suburbs (name, longitude and latitude)

Source: Wikipedia, GeoPy

**Step 2:** Get data of schools around each of the Sydney suburbs

Source: Foursquare.com

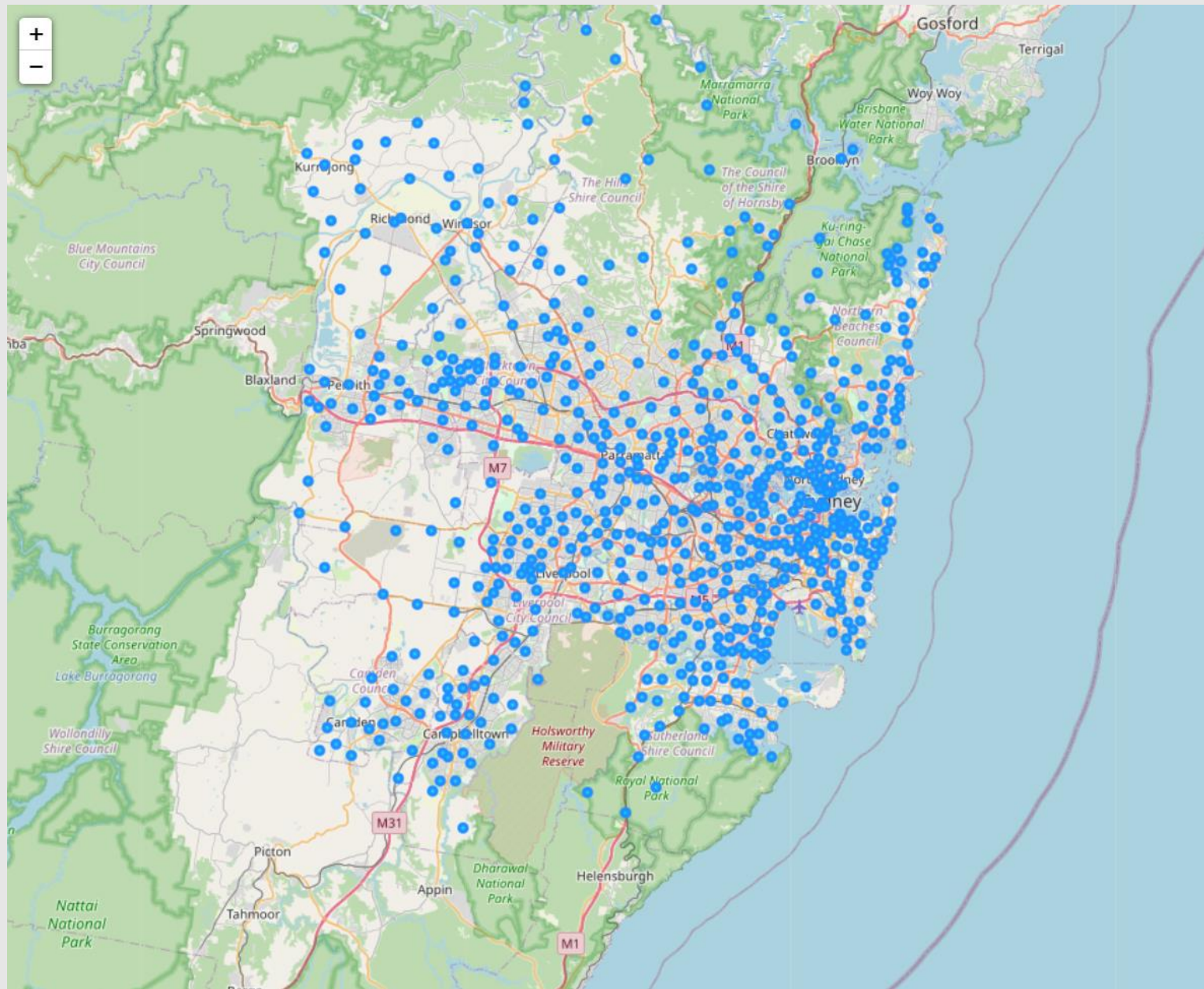
**Step 3:** Assign every school to its nearest suburb using GeoPy.distance

**Step 4:** Cluster the schools in 5 clusters based on distance from nearest suburb

*Please refer to the source code for explanations and annotations:*

Github: [https://github.com/nikroyal/Coursera\\_Capstone/blob/master/Capstone%20Project%20W4%20W5.ipynb](https://github.com/nikroyal/Coursera_Capstone/blob/master/Capstone%20Project%20W4%20W5.ipynb)

IBM Watson Notebook: [https://eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/d18fc033-5709-466b-b4c7-7aeb88250c15/view?access\\_token=d09c1ae065203f24cf9795d1145a8227cc952dea882569e1bdb405c4382c01e8](https://eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/d18fc033-5709-466b-b4c7-7aeb88250c15/view?access_token=d09c1ae065203f24cf9795d1145a8227cc952dea882569e1bdb405c4382c01e8)

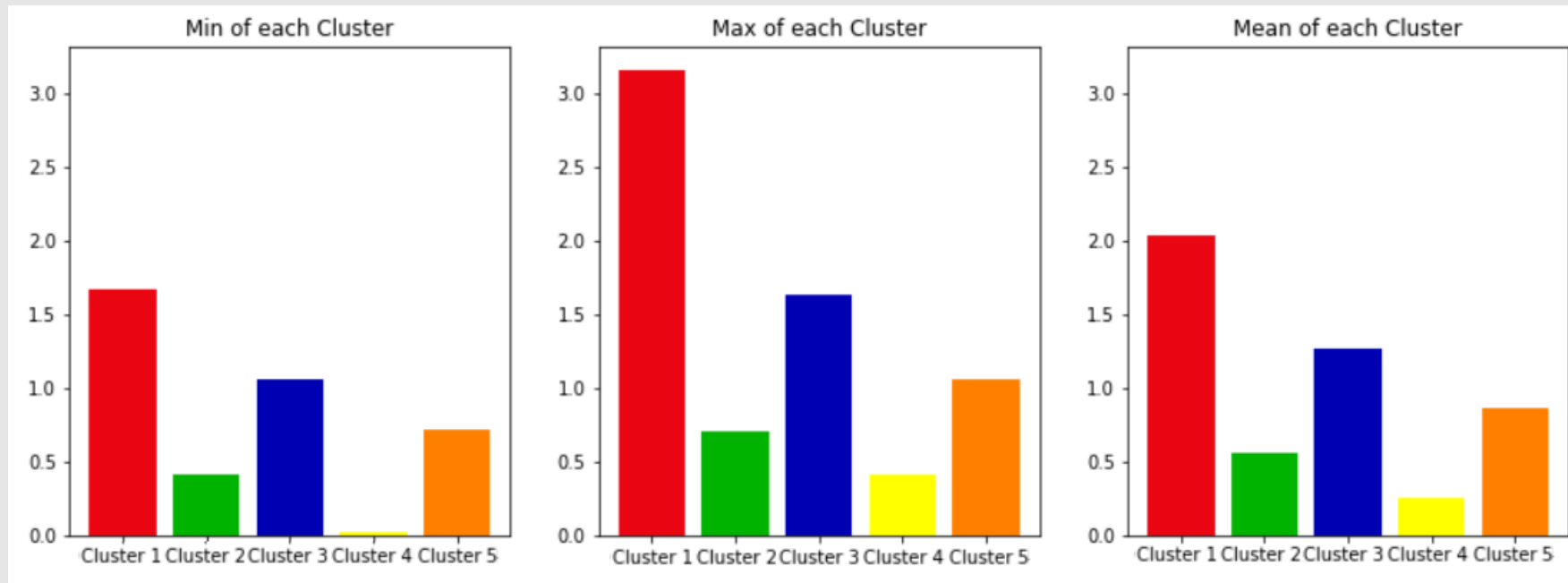


Sydney area with all 692 suburbs marked in blue

Sydney has over 600+ suburbs and over 1400 schools.

Schools are classified as Elementary School, School, High School, Nursery School, Middle School, and Private School.

Distance in Kilometers	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Minimum	1.7	0.4	1.1	0.02	0.7
Maximum	3.2	0.7	1.6	0.4	1.1
Mean	2.0	0.6	1.3	0.3	0.9
Cluster size	55	459	201	376	345
Color on map					

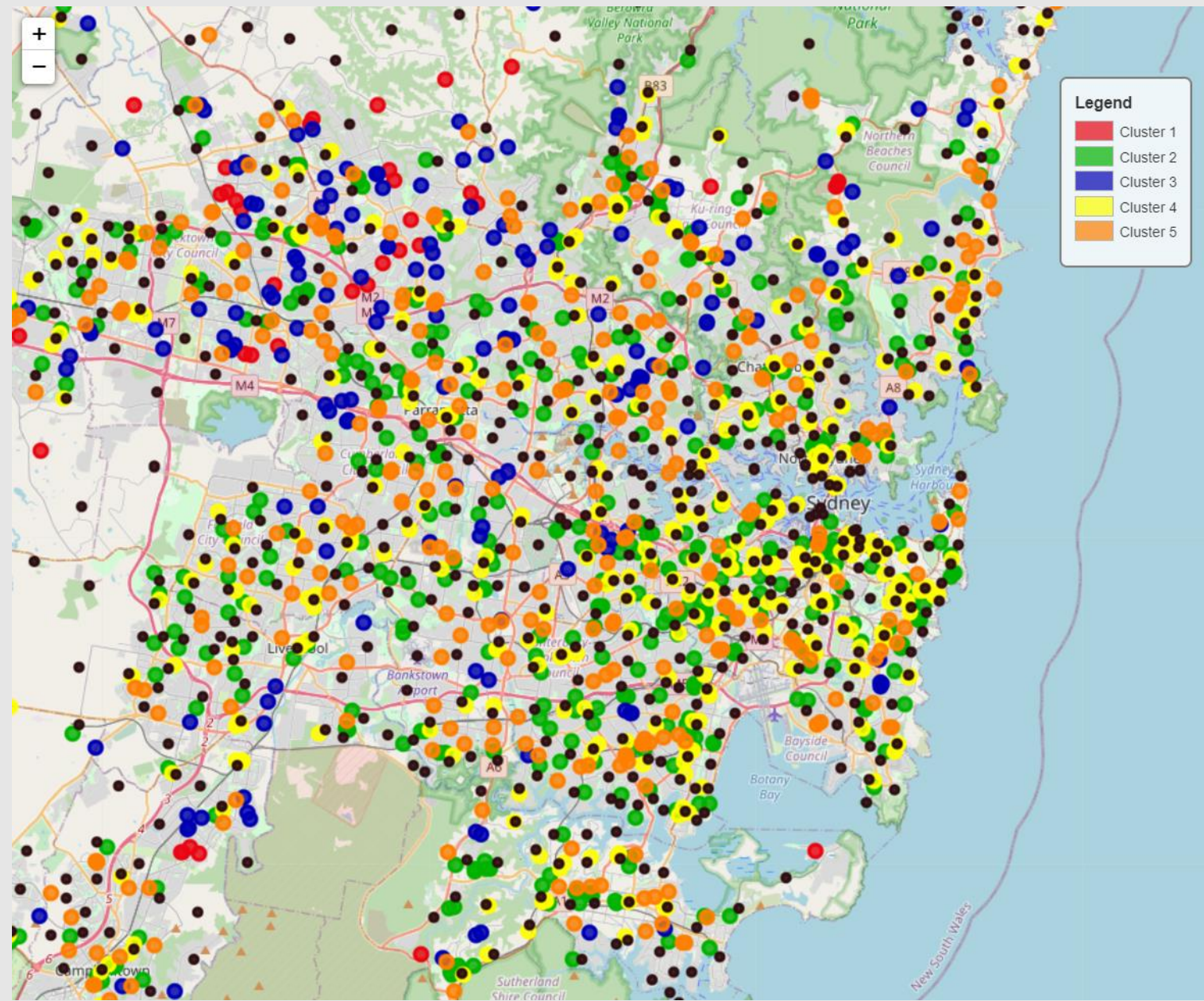


Clusters to focus are:

**Cluster 1:** Represents schools with large distance between them and suburb-center.

**Cluster 3:** These are away from CBD and have some scope and must be explored further to build schools that is easily accessible to the community





Densely populated areas have more yellow and green clusters.

Sparsely and growing areas have more blue and red clusters.

**Cluster 1:** Represents schools with large distance between them and suburb-center. These are typically suburbs around the periphery of Sydney. These are often sparsely but fast growing suburbs and have few schools which are spread away from the suburb-center.

**Cluster 3:** These are clusters that still have large distances from the nearest suburb-center. In such suburbs Town Planning authorities should look to build more schools and/or plan better connectivity/road projects to make existing schools more accessible.

# Thank you

- The Github code is available at: [https://github.com/nikroyal/Coursera\\_Capstone/blob/master/Capstone%20Project%20W4%20W5.ipynb](https://github.com/nikroyal/Coursera_Capstone/blob/master/Capstone%20Project%20W4%20W5.ipynb)
- In case you face issues, the IBM Watson Notebook is available at: [https://eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/d18fc033-5709-466b-b4c7-7aeb88250c15/view?access\\_token=d09c1ae065203f24cf9795d1145a8227cc952dea882569e1bdb405c4382c01e8](https://eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/d18fc033-5709-466b-b4c7-7aeb88250c15/view?access_token=d09c1ae065203f24cf9795d1145a8227cc952dea882569e1bdb405c4382c01e8)