

DATA2001

Data Science, Big Data
and Data Variety

Week 8
Geo-Spatial Data



THE UNIVERSITY OF
SYDNEY

Outline

- Introducing and Ingesting Spatial Data – 45 mins
- Querying Spatial Data – 45 mins
- Visualising Spatial Data and APIs – 15 mins

Dataset



Beyond simpler data types we've encountered so far exists the world of **geospatial** data – allowing us to represent the points, lines, polygons and more that comprises the world in which we live.

Today we'll use a shapefile containing international **country** boundaries, and a spreadsheet detailing the coordinates of capital **cities**, before experimenting with other sources of spatial data.



Spatial Data

- Stored in various forms (**shapefiles**, geoJSON, etc)
- Requires a few upgrades to our typical packages
 - **GeoPandas** for interpreting dataframes with spatial information
 - **PostGIS** as a spatial database extension to PostgreSQL
- Download the files from Canvas to get started
 - **Week 8 Tutorial.ipynb**: contains the main notebook
 - **Capitals.csv**: a spreadsheet containing capital cities of the world
 - **world.zip**: contains the shapefiles for country boundaries
- Ensure to install **PostGIS** on your local database
 - In the “Application Stack Builder” app that was installed with your pgAdmin, install “Spatial Extensions → PostGIS”
 - To enable PostGIS, then proceed to run the below query in pgAdmin:
CREATE EXTENSION PostGIS

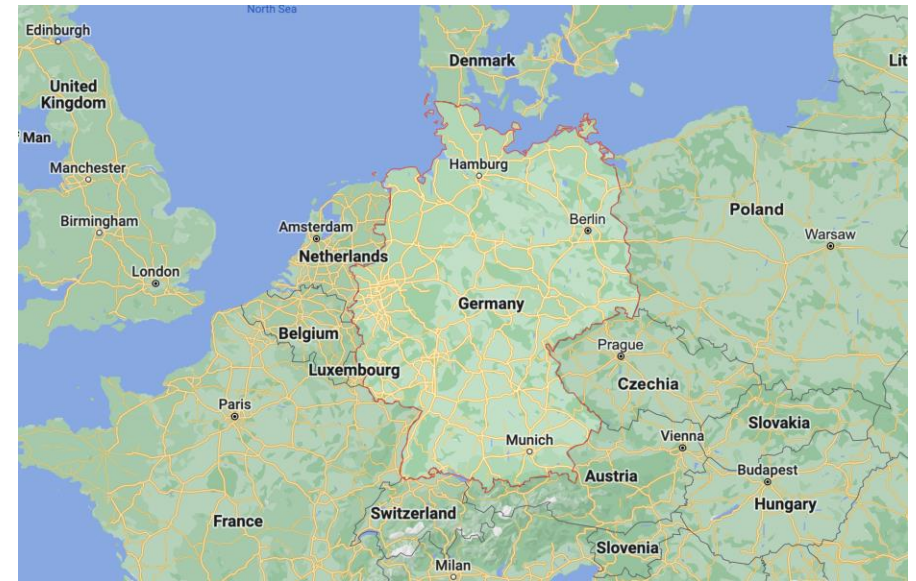


Spatial Query Tasks

Head to the notebook for this week – the below maps may be useful reference for Section 3.



Data quality consideration for part b) of Section 3.2.



Germany's neighbouring countries for part a) of Section 3.3.

School of Computer Science

For comparison with the Web API results in Section 4.2...a pretty good match!

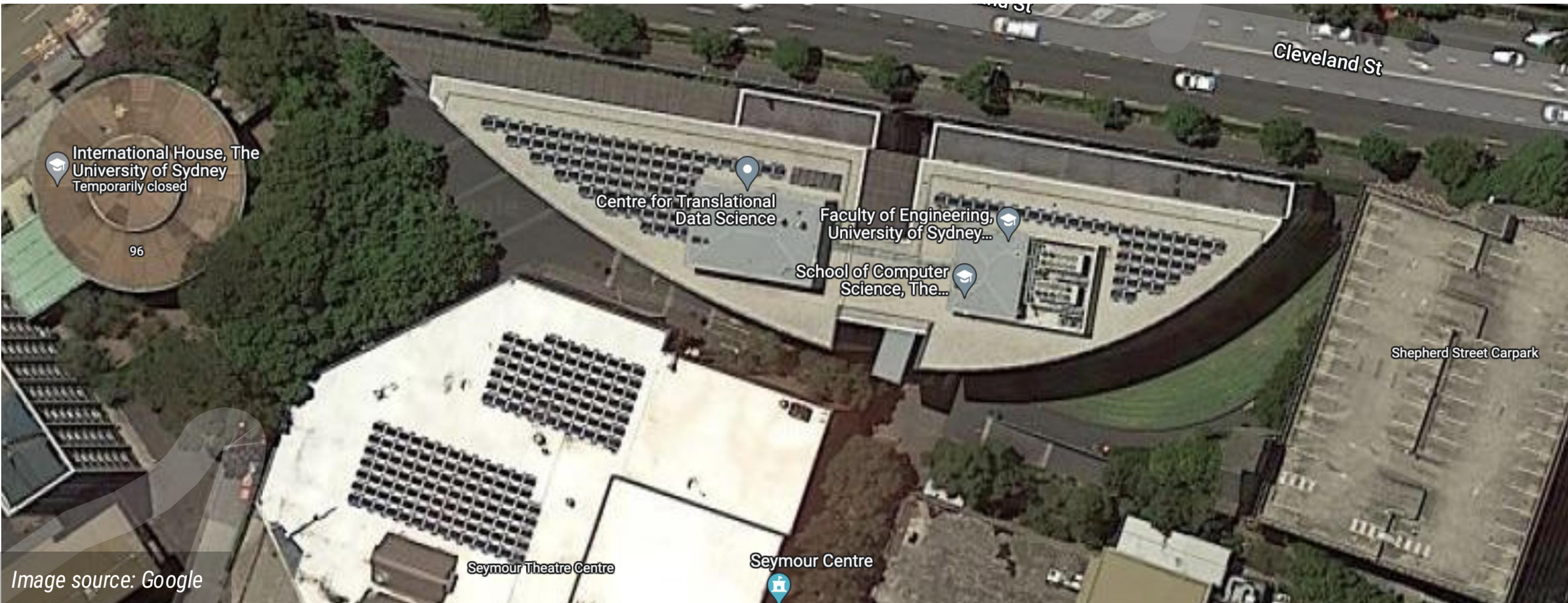


Image source: Google

Group Assignment

Remember to register your group on Canvas if you haven't already! Specifications are available there.

- Groups must be formed with **3 people** from your same enrolled tutorial (exceptions made where necessary)
- **Register** your group on Canvas using the "Group Assignment" tab in the "People" section
 - if you're unsure which class number you're in, you can confirm using the "[Class Allocations](#)" Canvas page
 - feel free to select any group number within your lab
- We'll soon create an **Ed megathread** for any questions that require further clarification



Enjoy the mid-semester break!

