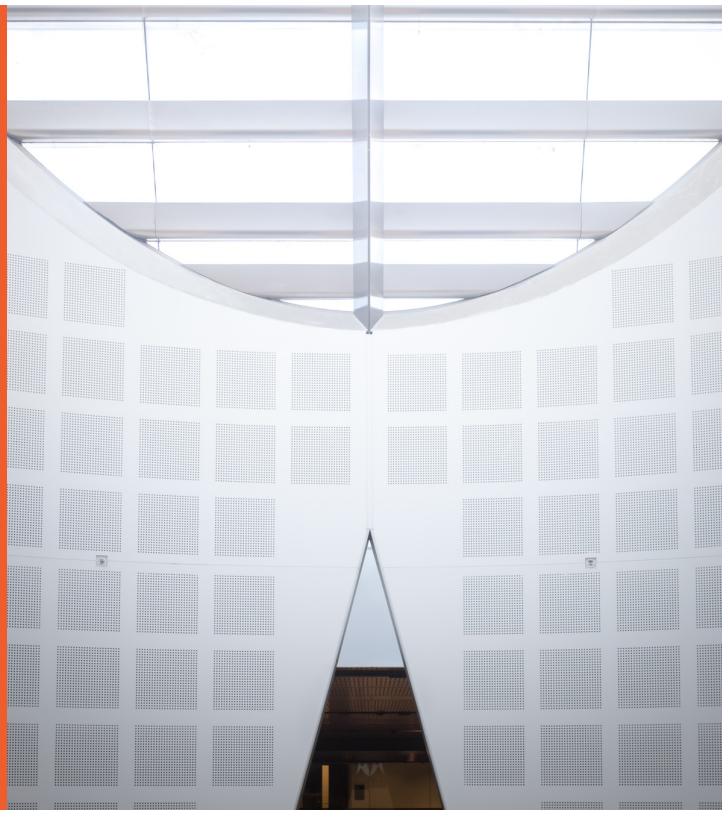
DATA2001 – Data Science, Big Data, and Data Diversity
Week 8: Administrativa

Presented by
A/Prof Uwe Roehm
School of Computer Science

THE UNIVERSITY OF SYDNEY

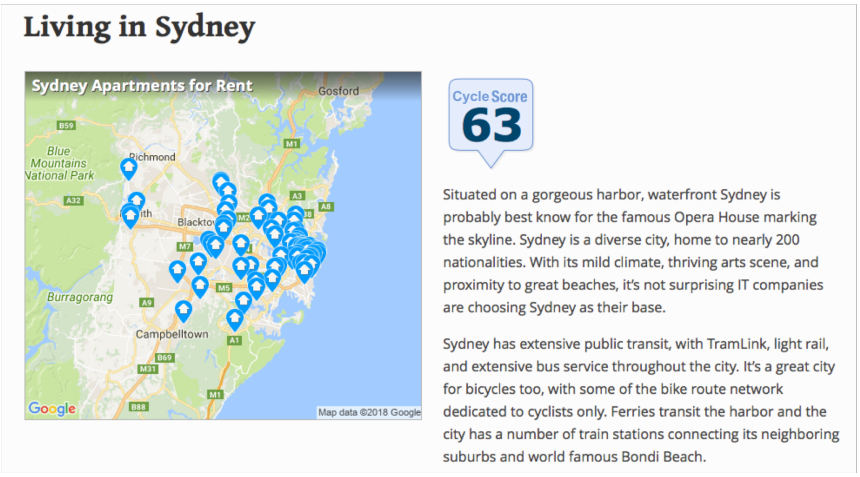


Practical Assignment

THE UNIVERSITY OF SYDNEY

Practical Assignment: Cyclability Analysis

- Assignment to be published end of week (Canvas: Modules –> Assignment)
  - **Worth 20% pf the final grade in DATA2001**
  - **Due in tutorial of Week 12**



[https://www.walkscore.com/AU-NSW/Sydney]

Practical Assignment: Cyclability Analysis

- Goal: Practical experience with data variety, data analysis, and presentation
  - Technologies as covered in this course: Python, Jupyter notebooks, and SQL
- Three tasks:
  - Data import and integration
    - We provided census data and some bike parking data
    - Needs to be combined, eg. via spatial join (to be covered in Week 9)
    - Feel free to extend with own datasets
  - Cyclability Analysis
    - Computation of cyclability score; example formula given
    - When adding other datasets, feel free to adjust formula
    - Correlation of your score with some ABS statistics
  - Documentation and (brief) Report
  - Some additional tasks/options for teams in advanced streams

Provided Datasets (to be published on Canvas)

- ABS Data
  - Census data on *neighourhoods* (SA2-level areas) in Greater Sydney such as population, land_area, number of dwellings
  - Business statistics per SA2-area
  - Income and rent statistics to check for correlation with
- Bike-Sharing "Pods"
  - One example of transport data: names and locations of dedicated bike parking locations ('pods').
- Note that SA2-level data from the ABS does not always match suburbs, and that the bike-pods have a GPS location, but not the neighbourhoods
  - cf. tutorial after break on how to retrieve boundary data for neighbourhoods too
- Adding more datasets from your side is explicitly encouraged.
  - Try different types and forms, not just CSV…

Assignment Rules

- Groupwork
  - teams of 2 (unless odd-size class or other good reasons)
  - All team members should be in the same tutorial
- Deliverables see handout, page 3
- Due on Friday of Week 12
  - Submission page and marking rubric to be published in Canvas
  - Late submissions: -20% of achieved mark per day late
- Demo in Week 12
  - There will be a short demo during the tutorials to tutors
  - Individual grades can be scaled based on participation in project or demo