STACS Data Analyst Case Study

Presenter: Guan

Table of Content

Case Study Description

ESGpedia regularly collects datasets from various data registries and data partners. For this case study, you will be tasked to collect and analyze data from a data registry called NABERS.

NABERS (National Australian Built Environment Rating System) is a sustainable rating for buildings in Australia. A NABERS rating helps building owners to accurately measure and communicate the environmental performance and progress of buildings. It also identifies areas for savings and improvements.

The data is publicly available here: https://www.nabers.gov.au/ratings/find-a-current-rating.

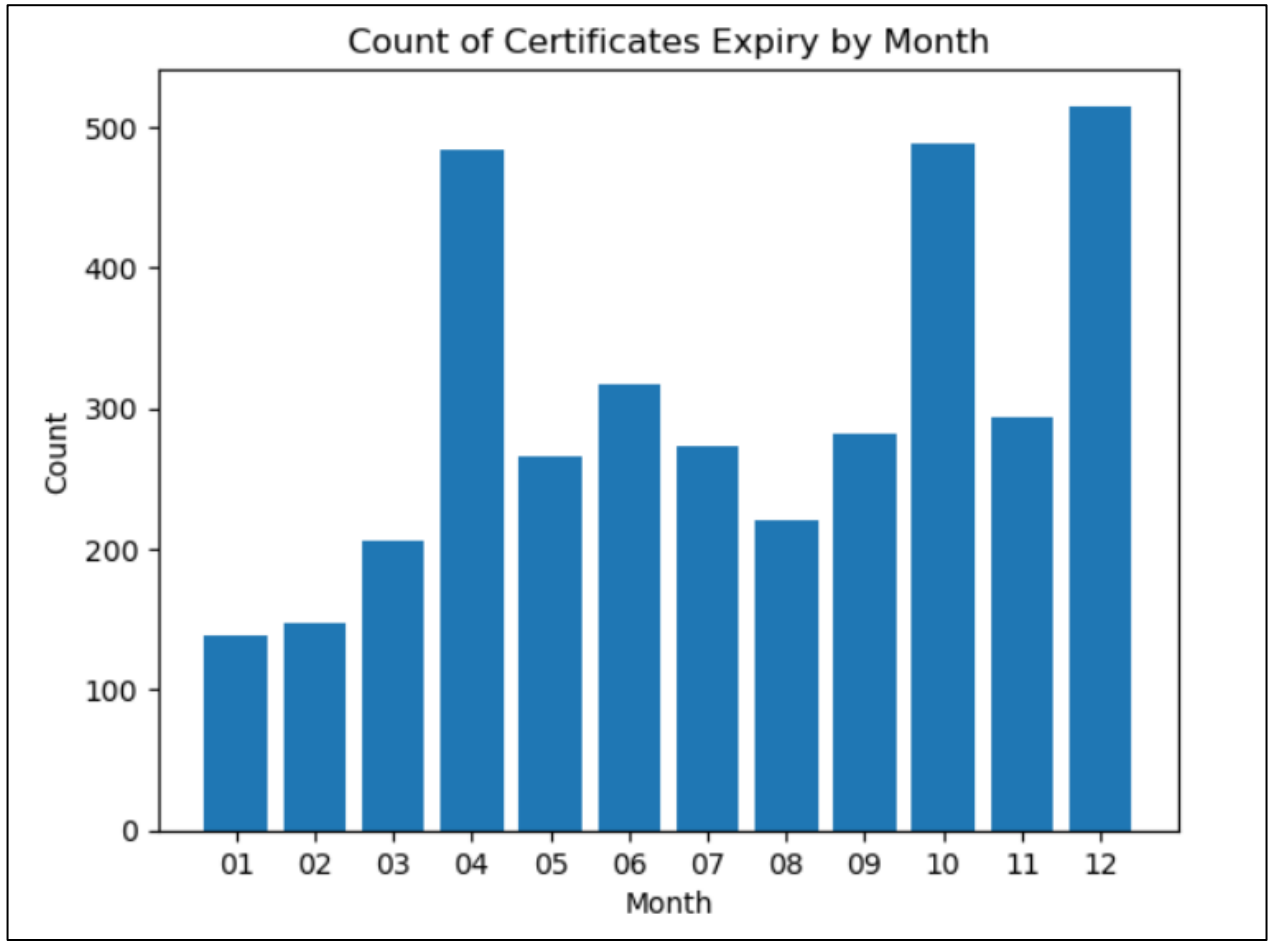
Requirements:

- Explain and show how you would automate the collection of the data from the website.
- Carry out any data cleaning and data processing steps where necessary.
- Present the data in an interactive dashboard to capture any insights you would like to highlight or point out.
- Languages to use: Python/SQL.

Some guiding questions have been provided below to assist you in your analysis. Please note that these questions are not exhaustive and you are highly encouraged to generate additional questions that you believe will be relevant to the task.

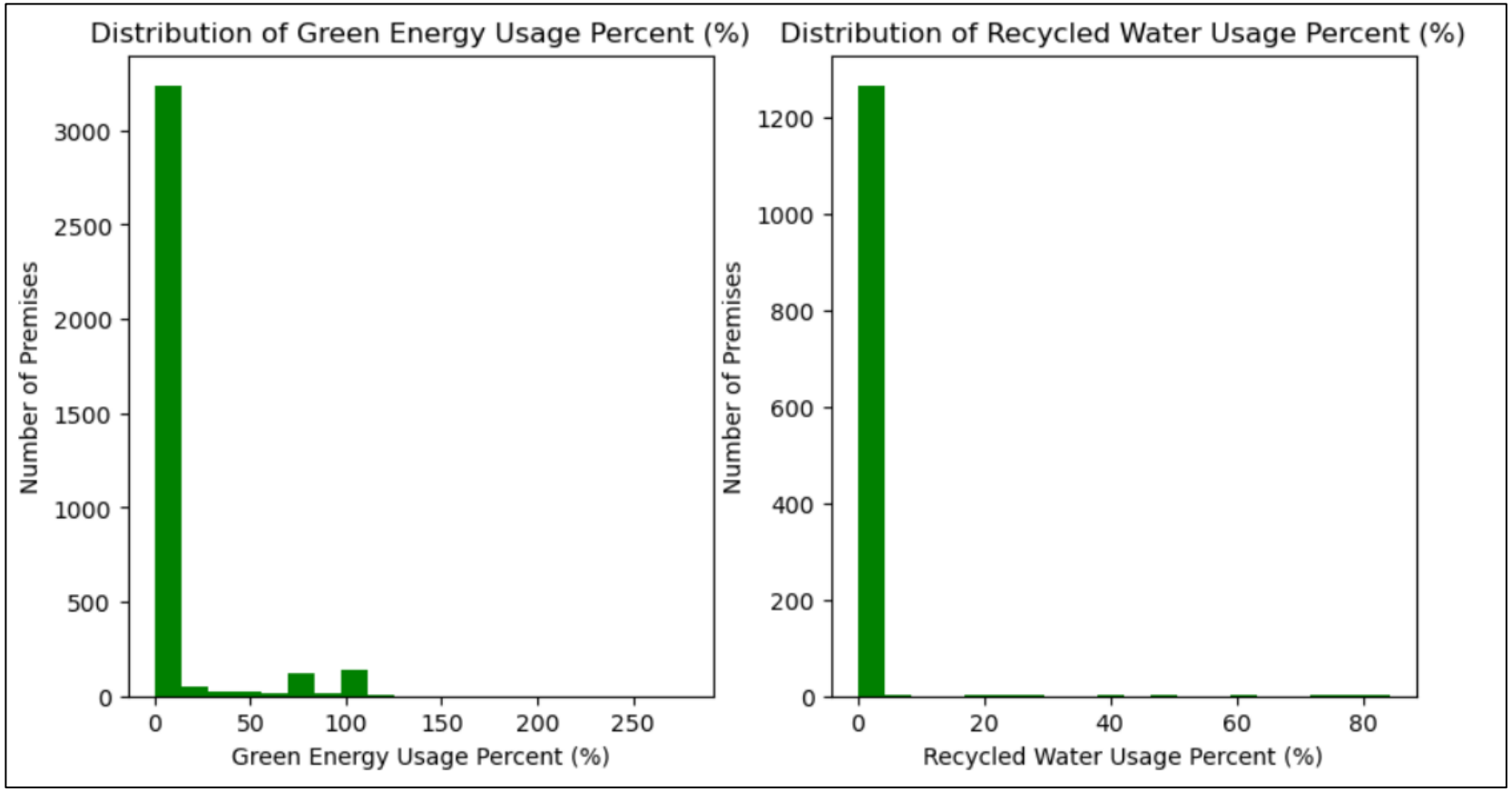
- How often should the data be collected?
- What is the distribution of green buildings in Australia?
- How many certificates are expiring soon?
- What is the average rating value of each building?

How often should the data be collected?



- Based on the bar chart, we would recommend to collect the data monthly, since there are hundreds of certificates expiring in every month (Apr, Oct, and Dec each has around 500 expiry cases).
- The ratings would need to be redone after the certificate expiry.
- Please refer to the Jupyter Notebook to see the bar chart.

What is the distribution of green buildings in Australia?



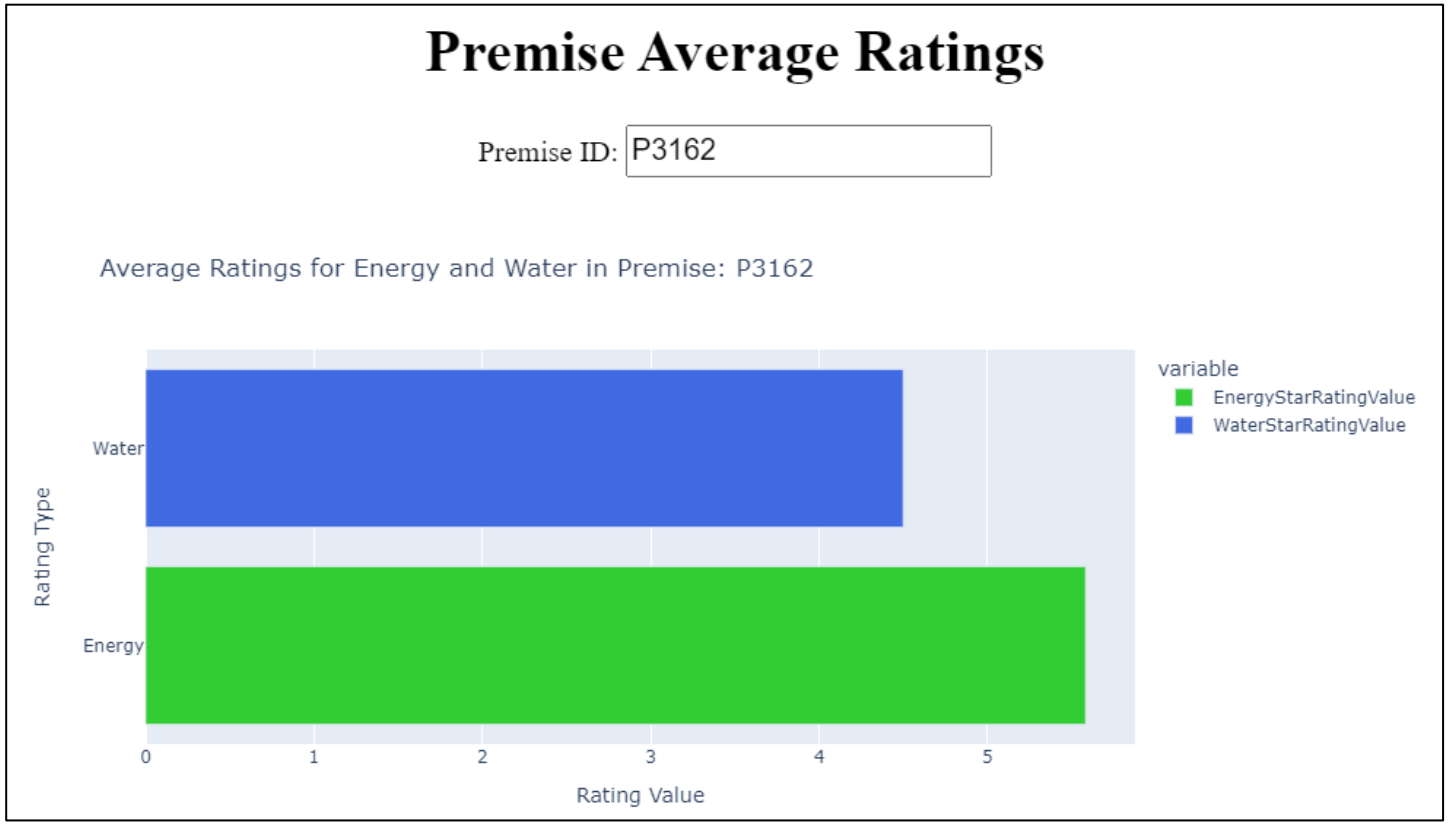
- We assume the green buildings are classified by the percent of green energy usage or the percent of recycled water usage for a premise.
- We assume the premises with greater than 80% of green energy usage can be deemed as green buildings, so 260 premises are green buildings per green energy usage in Australia.
- We assume the premises with greater than 60% of recycled water usage can be deemed as green buildings, so 7 premises are green buildings per recycled water usage in Australia.
- Please refer to the Jupyter Notebook to see the data analysis.

How many certificates are expiring soon?

	index	PremiseID	FullAddress	Latitude	Longitude	RatingType	CertificateValidTo	ExpiryInDays
0	255	P0201	165 Fitzroy St ST KILDA VIC 3182	-37.858574	144.979905	Energy	20/02/2024	5
1	2550	P4497	34 Queen Street MELBOURNE VIC 3000	-37.818016	144.962370	Energy	20/02/2024	5
2	2570	P4572	8 Exhibition Street MELBOURNE VIC 3000	-37.815693	144.972659	Energy	20/02/2024	5
3	2664	P4880	47 Kings Avenue BARTON ACT 2600	-35.304292	149.135326	Energy	20/02/2024	5

- We made a Certificate Expiry Checker in the Jupyter Notebook, it can help users to check the certificates that are expired in an entered date.
- The checker can provide the information on the premises' address, certificate validation, map location, etc.
- You can enter a recent date (e.g. 20/02/2024) to see how many certificates are expiring on that date in the checker.

What is the average rating value of each building?



- We develop an interactive dashboard for users to check the Premise Average Ratings for every premise.
- Users can enter a Premise ID and see the average ratings per energy and water.
- The dashboard is built with Plotly and Dash.

Project - STACS Data Analyst Case Study GitHub URL

- URL:https://github.com/guanp2023/Project_STACS_Data_Analyst_Case_Study/tree/main

The End
Thanks