

Just Scraping By: Understanding Rents at a Neighbourhood Level Using Web Scraping Technology

May 2024

Boun Inthavong & Klaus Cumani
Social Policy, Research and Analytics
Community and Social Services



Agenda

1. Context and Background
2. Approach and Demo
3. Technical walkthrough
4. Applications and Insights
5. Challenges and next steps
6. Questions

CONTEXT AND BACKGROUND

Lessons from COVID

- **During uncertainty, *timely* data is vital to responding to crisis**
- **Data we rely on during stable times may not be suitable during crisis**

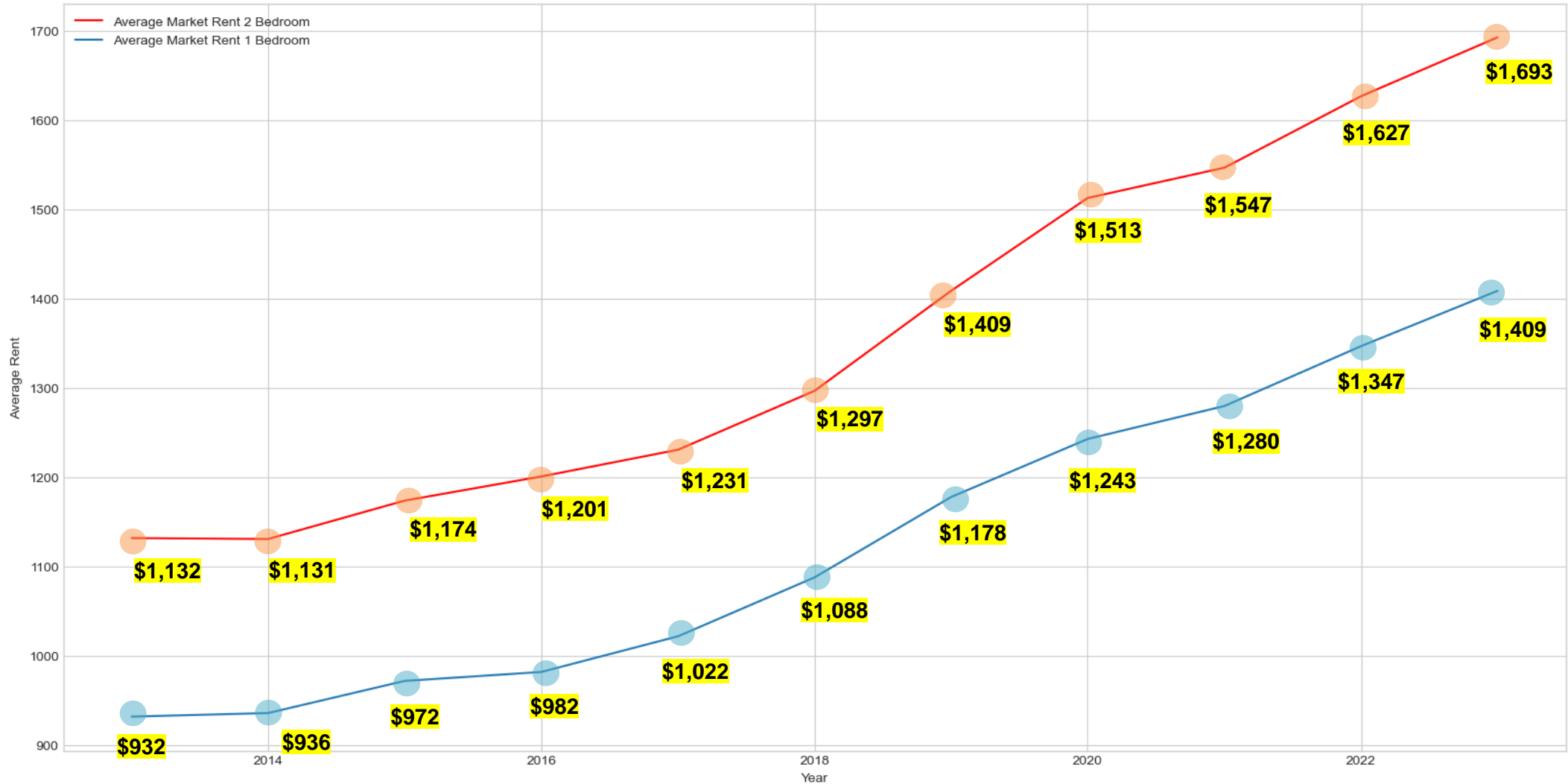
CONTEXT AND BACKGROUND

Affordability Housing Crisis

- **Average Market Rent (AMR) set by CMHC**
- **Vital metric that many benefits program base off**
- **Is it *timely*?**

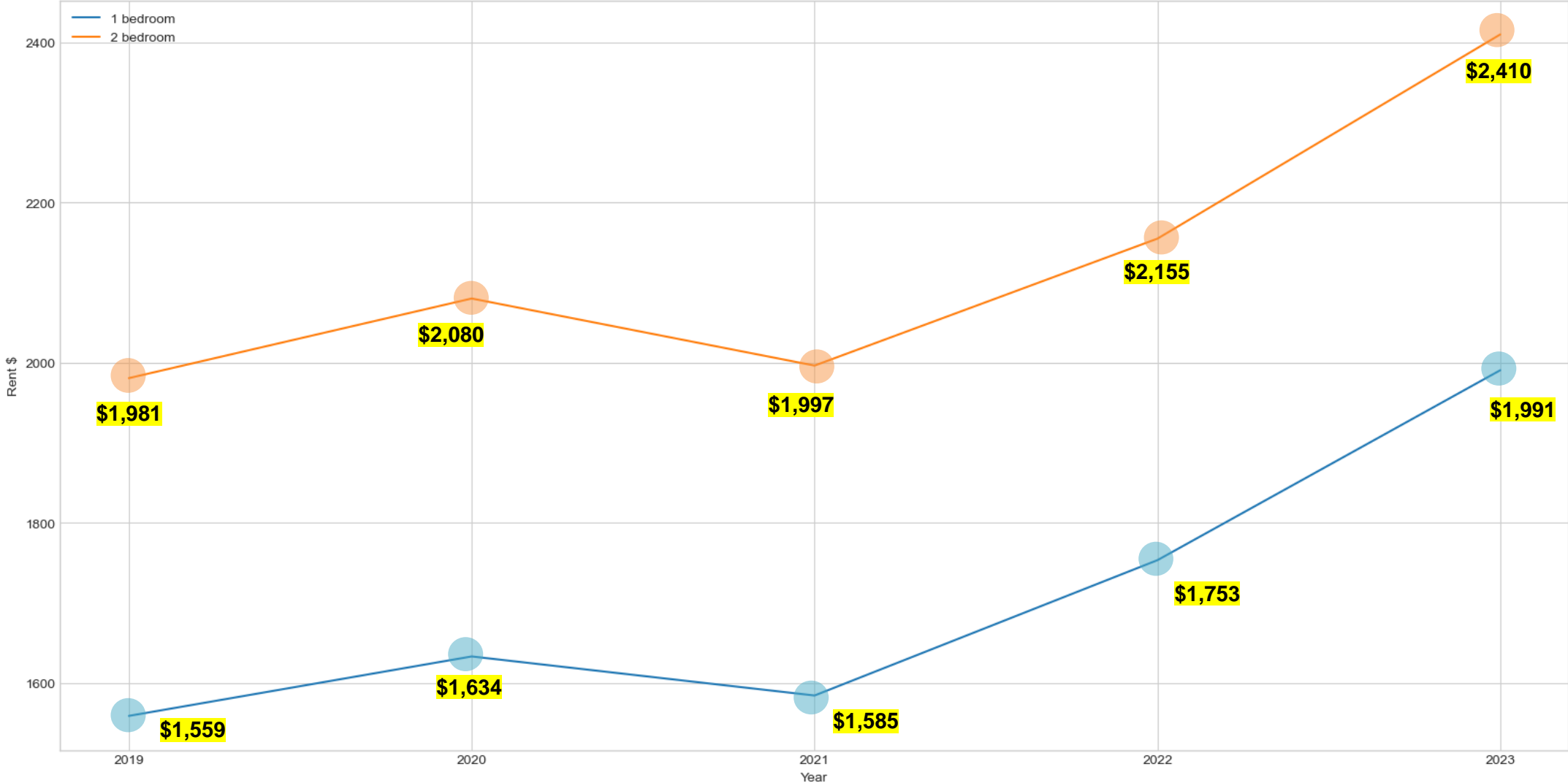
CONTEXT AND BACKGROUND

Apartment Average Market Rent with 1 bedroom and 2 bedrooms from 2013-2023. SOURCE: CMHC Ontario — Rental Market Statistics Summary by Metropolitan Areas, Census Agglomerations and Cities



CONTEXT AND BACKGROUND

Average Monthly Rent per Year in Ottawa | Source: rentals.ca and SPRA calculation



CONTEXT AND BACKGROUND

Late 2022

- **Director of Housing challenged SPRA to find more timely rental data**
- **How can non-admin and non-statistical agency data fit and compliment our data work?**

APPROACH

Can we do it?

- **Skill** ✓
- **Tools** ✓

How would operational staff use the data? ?

How would management consume data? ?

Legal and privacy considerations? ??

APPROACH

Internal Engagements

- **Front line case workers needed a better tool to help find and match housing for their clients**
- **Management wants real time insights on costs by neighbourhood and types**

APPROACH

Legal and Privacy

- **“Grey Zone” but not completely unfamiliar**
- **Concerns of PII and liability**
- **Supportive of our idea and solutions**

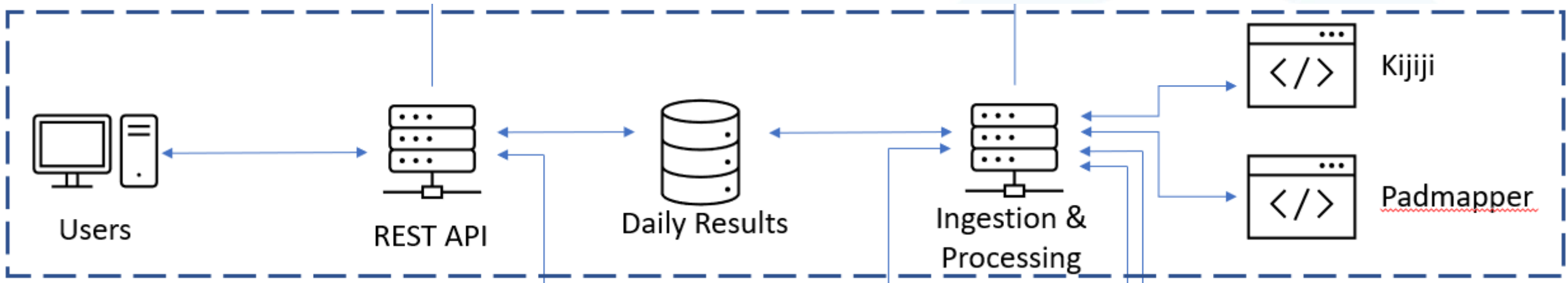
APPROACH

Ottawa Neighbourhood Study (ONS)

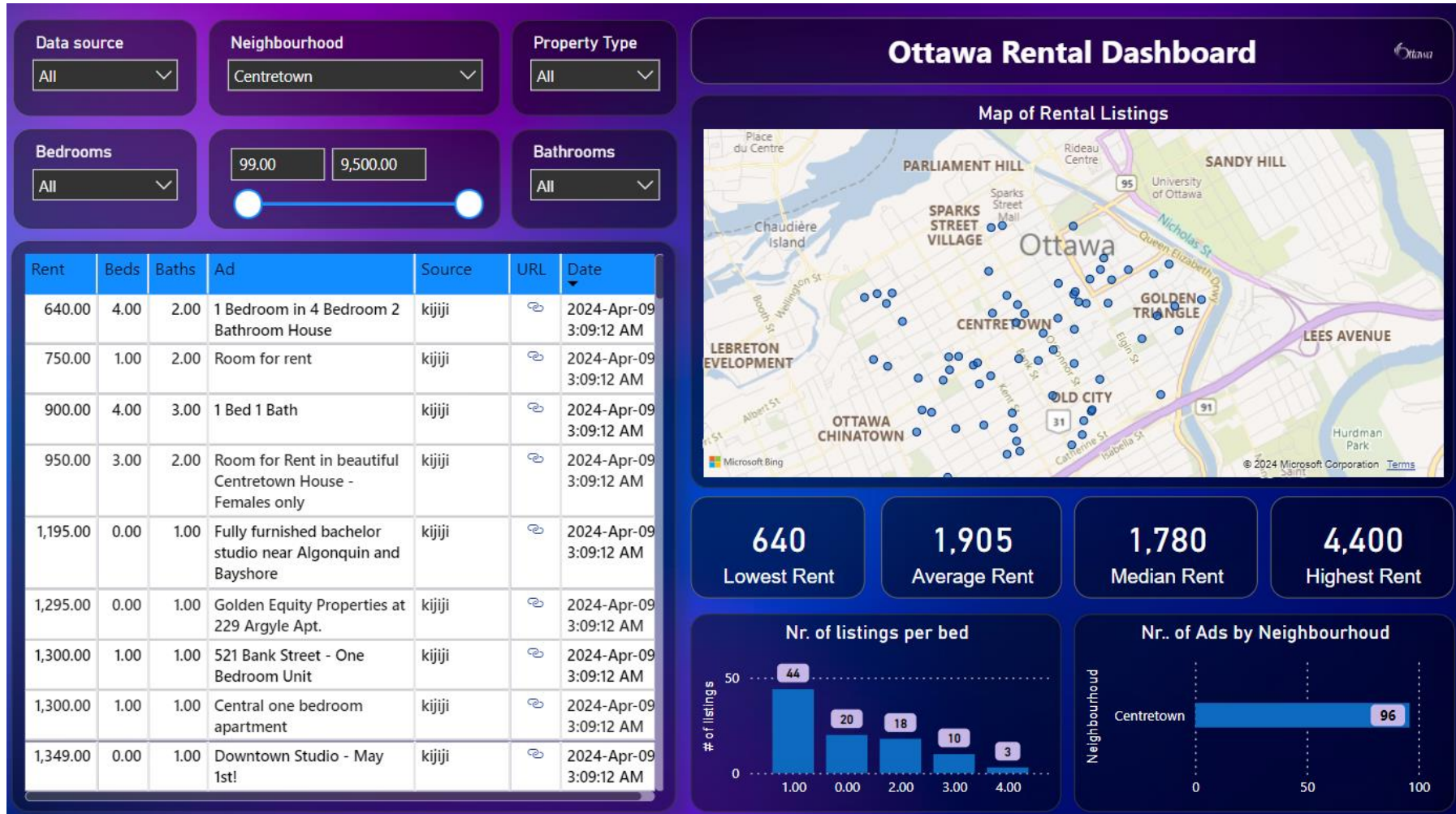
- **Not for profit research group out of University of Ottawa**
- **Neighbourhood boundaries and socio-economic profiles**
- **Interested in private rental market**

APPROACH

Current Architecture



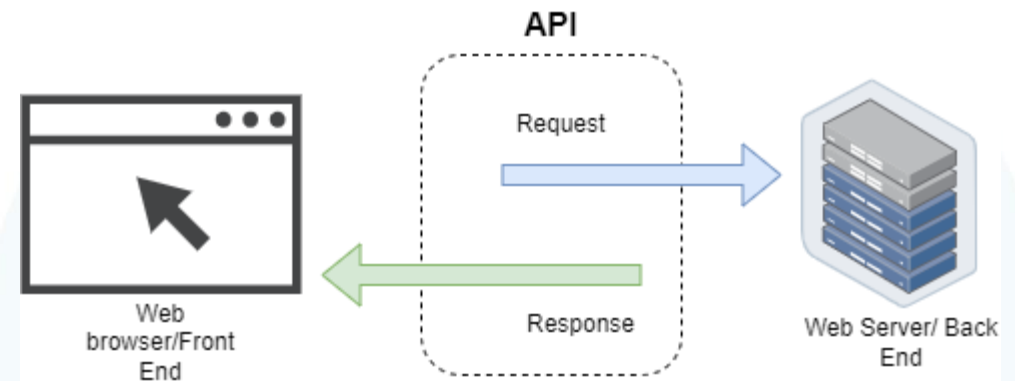
DASHBOARD – DEMO



TECHNICAL WALKTHROUGH

Modern websites are sophisticated software/applications that run in a web browser

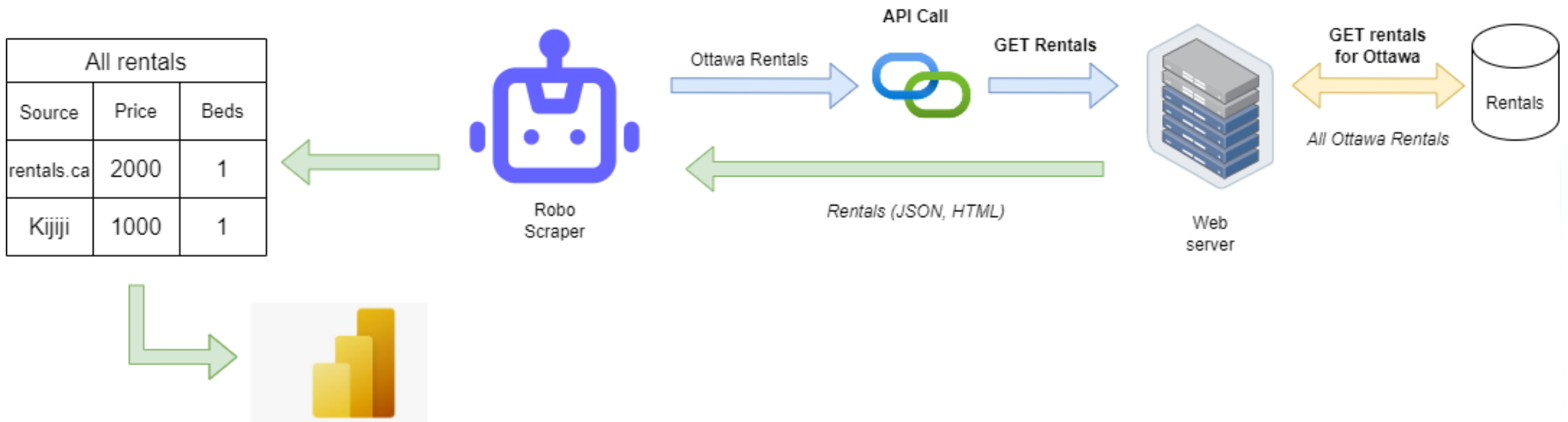
- **Front-end**
 - What we see and what we feel
 - Dynamic, static, or both
- **Back-end**
 - processing of requests
 - retrieval of information and data
- **Typically, the two components interact via an API which manages requests and responses**



TECHNICAL WALKTHROUGH

What if we wrote a program that imitates a web browser?

- Instead of rendering the data for a fancy website, we parse the data (response)
- The response data can be JSON or HTML
- Program then manipulates the response and throws everything into a table
- Table is consumed by Power BI



TECHNICAL WALKTHROUGH

Key tools and libraries

- **R or python (we prefer python)**
- **Power BI for ingestion, data transformation and viz**
- **R: Httr for requests, jsonlite for json parsing, tidyverse for all data manipulations**
- **Python: Requests, cloudscraper, json, pandas**

APPLICATIONS AND INSIGHTS

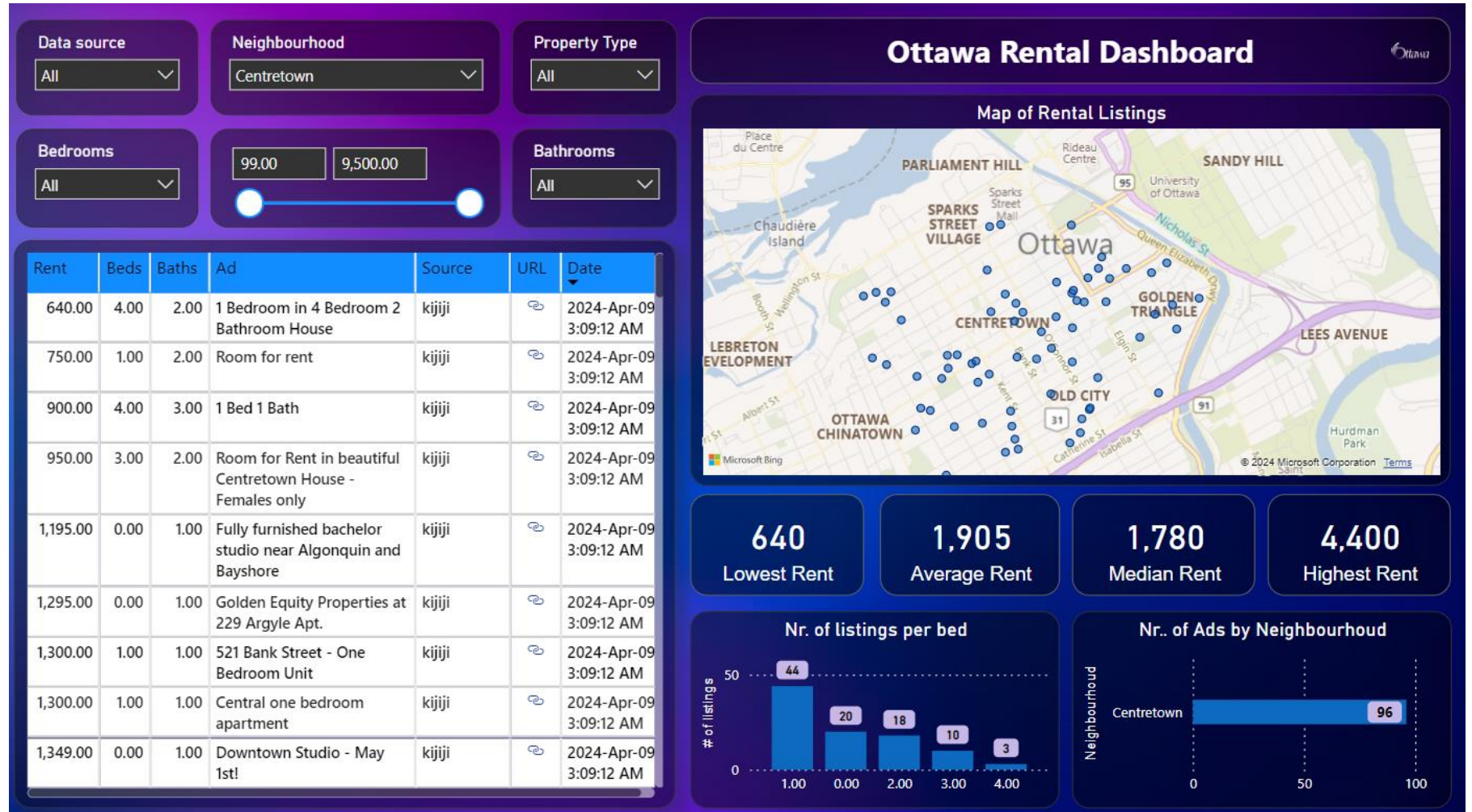
EARLY APRIL 2024 DATA

All Apartments and Basements only		
	Neighbourhood	Metric
Most expensive 1 bed	Chatelaine Village (Orléans)	\$5199
Most expensive 2 bed	Byward Market	\$8500
Most expensive 3 bed	Westboro/Stonebridge-Halfmoon Bay	\$6500
Most affordable 1 bed/studio	Centretown	\$1300
Most affordable 2 bed	Fallingbrook	\$1700
Most affordable 3 bed	Greenboro East	\$1650
Most listings	Centretown	96
Fewest listings	7 neighbourhoods	1
Total number of listings	N/A	1844
Number of 1 bed listing	N/A	511
Number of 2 bed listing	N/A	594
Number of 3+ bed listing	N/A	739
Average rent	N/A	\$2282
Median rent	N/A	\$2195

	2022 AMR	2024 Rental Data
1 bed	\$1347	\$1784
2 bed	\$1635	\$2278

APPLICATIONS AND INSIGHTS

- Internal and external use
- Well over 10 agencies and community partners
- Over 100 users



APPLICATIONS AND INSIGHTS

Rapid Rehousing Pilot

- **Urgent need to housed our Physical Distancing shelter users**
- **What were the barriers to their attaining housing?**
- **Used web scrape dashboard to assess availability of units in 1 bed AMR**
- **Pilot: limited City funds to top-up COHB to help remove cost barrier via EHAB (Enhanced housing allowance benefit)**

APPLICATIONS AND INSIGHTS

EHAB

- **Approx. \$1800 would increase availability of units**
- **January 2024: 74 listings \$0-1349 vs. 1353 listings from \$1350 - \$1800**
- **161 users housed using EHAB**
- **30 participating landlords with 14 new landlords**

CHALLENGES AND NEXT STEPS

- **DATA QUALITY AND RELIABILITY (DUPLICATES, GAMEIFICATION, ROOMS)**
- **POINT IN A TIME—NO ARCHIVING NO HISTORICAL DATA YET**
- **DATA DEFINITIONS (AVG? BEDROOMS?)**
- **HARDENING SECURITY**

CHALLENGES AND NEXT STEPS

- **EXISTING REPO:**
<https://github.com/Ottawa-Neighbourhood-Study/aptafford>
- **Migrating repo to python → improve code to handle hardened security, make it easier to deploy across province**
- **End of 2024: Public repo for other municipalities an use**

THANK YOU!

QUESTIONS?