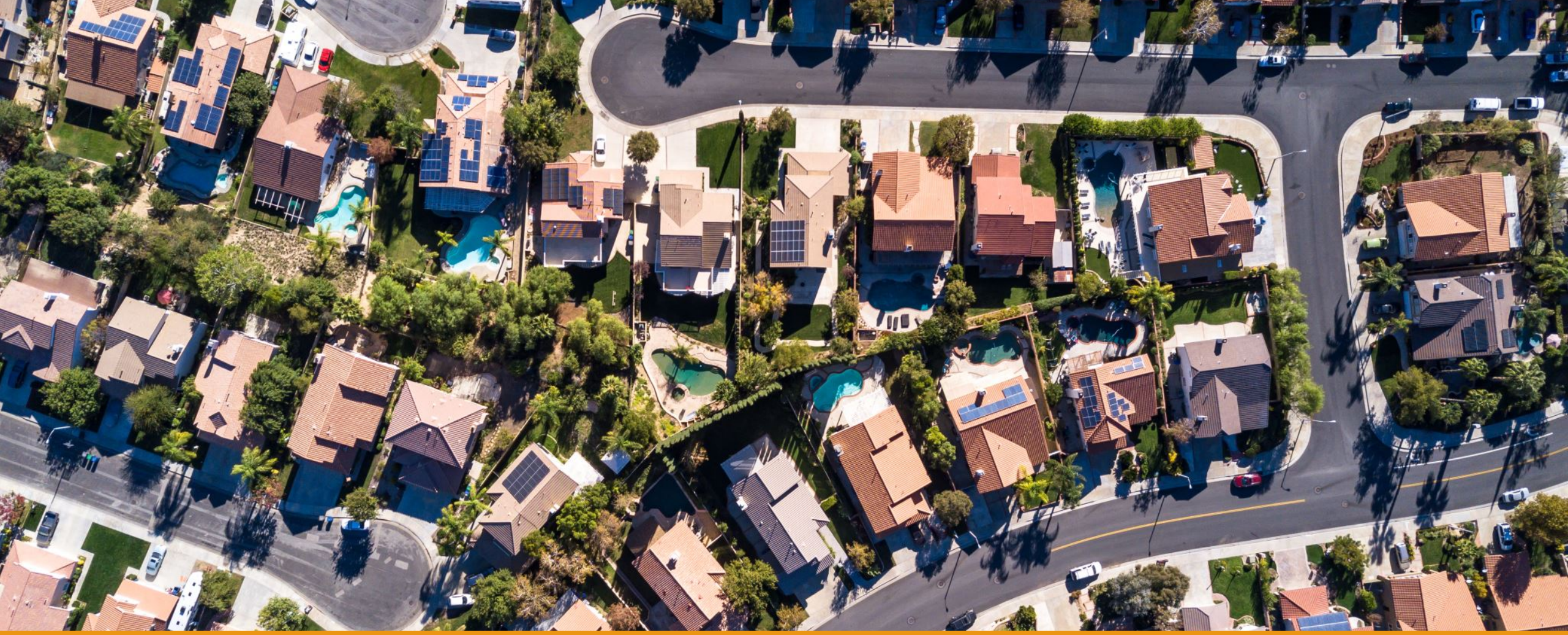
DATA ANALYTICS COMPETITION

INTERNATIONAL JOINT CONFERENCE ON ROUGH SETS 2024 (IJCRS 2024)

https://ijcrs24.cs.smu.ca/

Sponsored by:

- Southwest Properties, Halifax
- Saint Mary's University, Halifax
  - M.Sc. In Computing and Data Analytics
  - Master of Business Analytics



PROBLEM STATEMENT

PROVIDED BY SOUTHWEST PROPERTIES

PROBLEM STATEMENT

- The property management and apartment rental industry in Halifax lacks a comprehensive and centralized platform for accessing and analyzing up-to-date market rental rates (rent and parking).
- There is a notable absence of a system that efficiently scrapes or mines the web for rental market data and organizes it into a meaningful dashboard model for our management and leasing team. A tool that could assist us in making informed decisions on current competitive rental rates and setting future rates and increases.
- Southwest Properties would like to have a tool that captures, sorts by geographic location within the Halifax Regional Municipality and includes the property management firm, apartment building name, building amenities, apartment suite types, monthly rent, square footage and rate per square foot. Additionally, the tool would separately capture any residential monthly parking rates by location, and (if available) whether utilities are additional to the monthly rent.
- The dashboard should include current rental product in the marketplace and any planned or under construction developments within the Halifax Municipal geographic market. Key areas would be downtown / Southend Halifax, Central Halifax, Clayton Park, Rockingham, and the Larry Uteck area.
- Furthermore, this platform should incorporate predictive analytics to determine the factors influencing rental prices and provide a user-friendly dashboard to compare competitor rates for Southwest to remain competitive with current rents and for forecasting future rental rates for new developments.

KEY CHALLENGES

- Currently Southwest creates google sheets of competitor analysis in the marketplace. Information is captured manually. It is then integrated with an online platform service to generate a visual and database dashboard. This content is not updated frequently and is time consuming to capture content and input.
- We would like to expand this dashboard approach and remove the human element of creating the spread sheet manually.

GOOGLE SHEET
https://docs.google.com/spreadsheets/d/1Z83S7xvq3lbQ-_hevw6DO_VxaNN7s3MMMqSE8qaYWVQ/edit#gid=0

MAP INTEGRATION
https://app.nocodemapapp.com/app/vUzSWRYbJlXGnf2oPHeI

TASKS

1. Scrape the web for information on rental market in Halifax Regional Municipality – current apartment buildings and buildings under development or in approval process
2. Clean the data, organize it using a meaningful database model – location, building type, building amenities, suite types, rental rates per suite type, price per square foot, parking – surface or internal, etc.
3. Provide spatiotemporal data visualization.
4. Identify important variables that determine rental prices – geographic location, amenities etc.
5. Predictive analytics for determining rental prices.

PROJECT GOALS & OUTCOMES

Upon successful completion of this project, Southwest users should have access to a reliable, online dashboard, showing trends & forecasting:

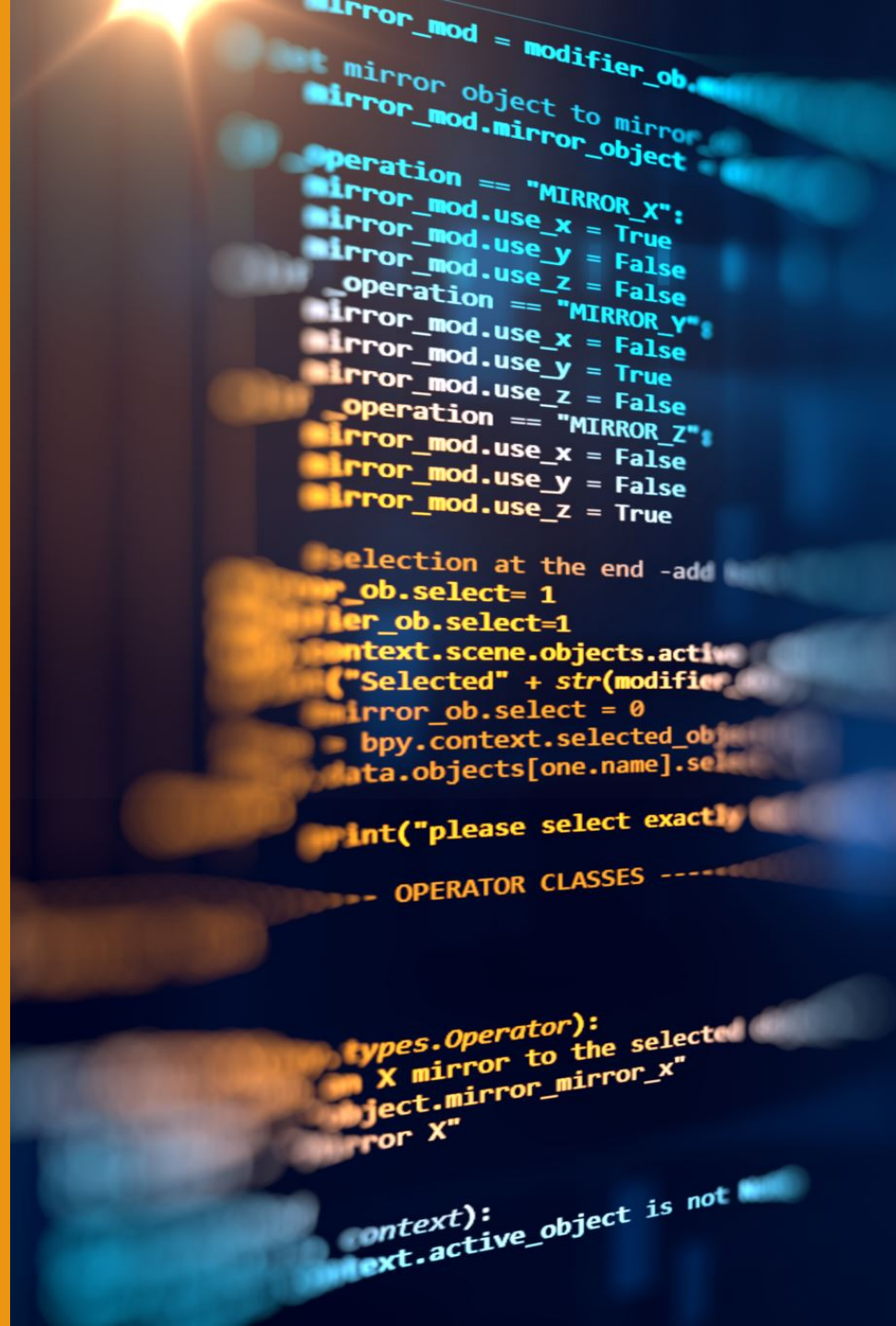
- Pull both parking and rental rates available online from HRM property management and rental companies or other available sources.
- Be an accessible and visual dashboard.
- Allow the most updated data to be exported to MS Excel.
- Increase our productivity by removing human elements and increase our market intelligence by having access to more data, in real-time.
- Potentially increase profitability by assisting in setting rental increases and rates.

Judging process and tips

Submissions

Submissions will be in the form of links to

- Video presentation
- PDF of the presentation
- Link to the demo website
- Link to the code repository



Timelines

February 7th: Preliminary problem statement as described in the attached document will appear on the conference website.

February 14th: Further explanation including the existing manual process practiced by Southwest in form of videos and presentations will be provided. Additional details will be forthcoming as needed.

March 15th: Submission site is launched.

April 15th: Last day for first round submissions.

May 3rd: Five shortlisted entries will be announced. They will also receive feedback for their submissions.

May 10th: The shortlisted teams will submit an updated version of their entries. They may also submit a paper that can possibly be considered for publication in the conference proceedings and subsequent journal publication.

May 17-20: The shortlisted entries will present their work during a session in the conference (in-person or remote)

Possible Solutions

- **Solutions can be divided into three categories:**
  1. **Web Scrapping:-** Scrape the web for information on rental market in Halifax Regional Municipality – current apartment buildings and buildings under development or in approval process.
  2. **Visualization:-** A dashboard with current rental product in the marketplace sorted by geographic location and any planned or under construction developments within the Halifax Municipal geographic market sorted by required fields.
  3. **Predictive Analytics:-** Determine the factors influencing rental prices and provide a user-friendly dashboard to compare competitor rates for Southwest to remain competitive with current rents and for forecasting future rental rates for new developments.
  4. **Other:-** Any additional aspects that may be relevant to a real estate company are also welcome.
- An entry may choose to tackle one, two. three or more aspects described above.
- The judging will be a weighted sum of scores from various aspects described above.

Level of Difficulty/Importance

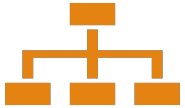
We will assign a level of importance/ difficulty to each aspect of the solution.

For example,

We cannot develop a product/service without data. Hence, data scraping will probably be assigned the highest level of difficulty.

Predictive analytics may require researching of existing methods acquiring sufficient data, training and testing validity of the proposed model. It may be second most difficult aspect.

Number of entries that tackle a given aspect of the problem will also affect the degree of difficulty. More solutions, lower the degree of difficulty.

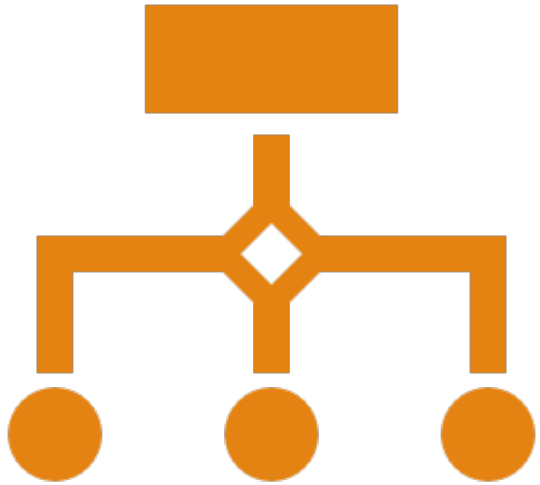


The following is the initially proposed level of difficulty:

Web Scrapping	Visualization	Predictive Analytics
100%	60%	80%

Examples of Level of Difficulty Calculations

- If a team scores 55 in web scrapping, the final score will be 1.0*55=55
- If a team scores 90 in visualization, the final score will be 0.6*90=54
- If a team scores 70 in predictive analytics, the final score will be 0.8*70=56
- If a team tackles two aspects: Web Scraping (scores 40%) and Visualization (scores 60%), final score will be 1.0*40+0.6*60=76.



Possible metrics in determining scores

Web Scrapping	Visualization	Predictive Analytics

• Number of Fields
  • Acquiring all or some of the Fields currently shown in the spreadsheet
  • Acquiring additional fields
• Level of Automation
  • Fully automated with an ability for experts to modify the database
  • Partially automated by providing mechanism for experts to easily identify and add the information to the database
• Create alerts when a new development appears in Halifax
	• UI/UX Design of the dashboard
	  • Can you do better than the existing dashboard
	• Level of Automation in generating the dashboard
		• Background Research on Machine Learning in Predicting Rental Prices
		• Extraction of relevant variables.
		• Modelling
		• Model validation