Master of Business Analytics

Sobey School of Business
Saint Mary's University

SOUTHWEST PROPERTIES
HALIFAX

DATA ANALYTICS

Property Management & Visualization

Presented By

Pallavi Kumari
Akshita Singhal
Yashwanth Balaji Krishnamurthy
Md Akmam Ul Haque
Nitin Chitharanjan
Iram Shaikh

DISCUSSION POINTS

QUICK OVERVIEW

- **Problem Statement**
- **Data Scraping**
- **Data Cleaning**
- **ML Model - WIP**
- **Dashboards:**
  - **Power BI**
  - **ArcGIS**
  - **UI/UX React**

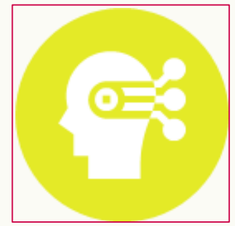
Submission Links

Link to Demo Website

https://661dc52f0677a45e758a5567--zingy-centaur-56e897.netlify.app/

Link to Code Repo

https://github.com/Pallavi7597/SouthwestProject

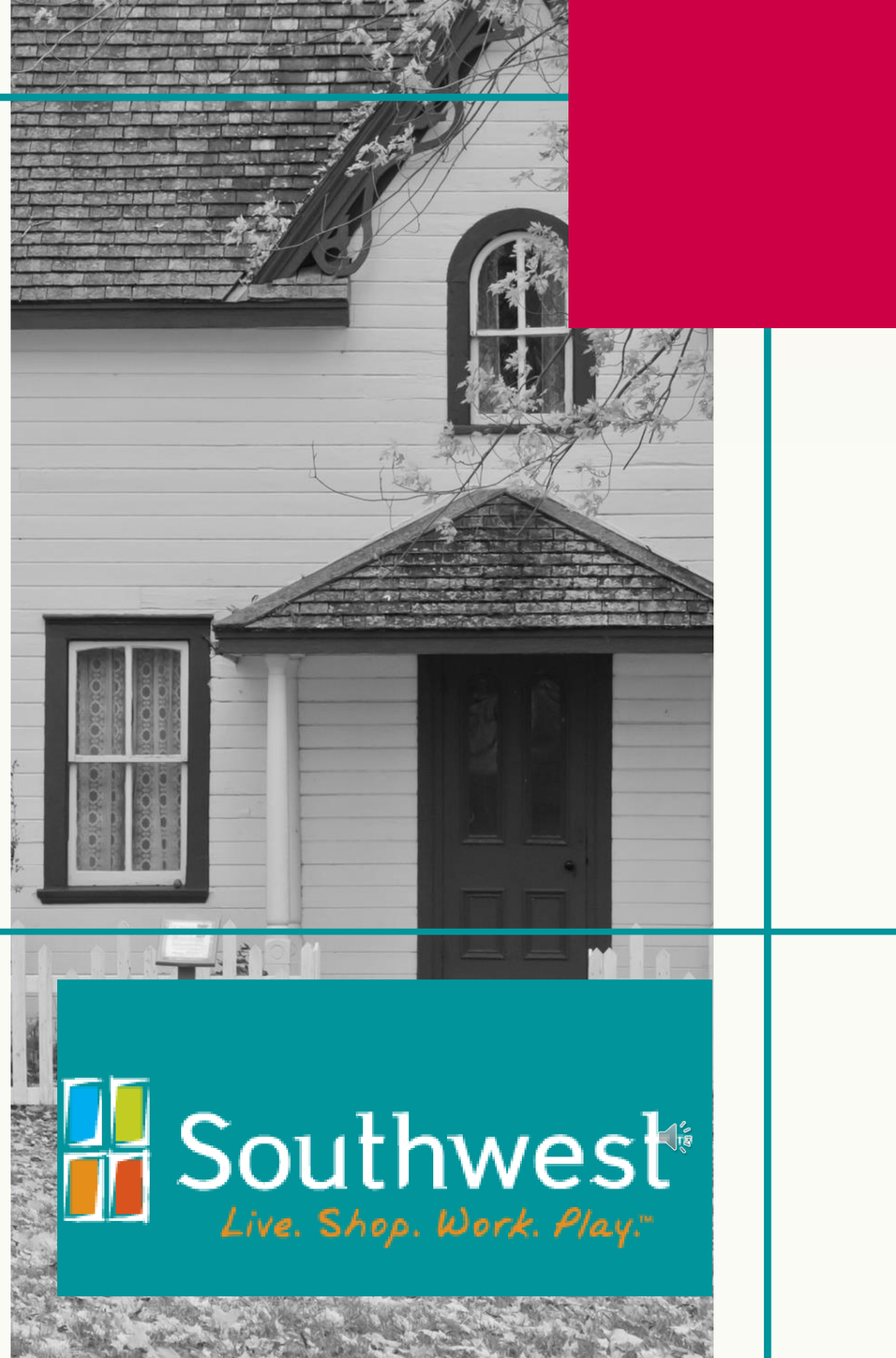


PROBLEM STATEMENT

- ✓ Southwest Properties seeks a centralized platform for accessing and analyzing up-to-date rental market data in Halifax, including rental rates.

- ✓ The platform should scrape web data, organize it by geographic location, and include details such as property management firm, building amenities, square foot, and monthly rents. It should also capture residential parking rates and utilities.

- ✓ The dashboard should display current market rentals and planned developments, with predictive analytics to inform rental pricing decisions and maintain competitiveness.



DATA SCRAPING

DATA FOR THE PROJECT WAS PULLED FROM VARIOUS WEBSITES MENTIONED BELOW

Apartments.com

Kijiji.ca

findallrentals.ca

Zillow.com

DATA SCRAPING

The purpose of the scraping script is to scrape data from various websites that list apartment rentals. The code fetches information such as rental prices, addresses, property details, and more from different rental listing platforms.

Libraries Used

Requests: HTTP requests to fetch web pages.

Beautiful Soup: Parsing HTML and XML documents.

Pandas: Data manipulation and analysis

Code Functionality

Step 1 Collector script fetches the URLs of apartment listings from a specific website.

Step 2 Scraping data from each listing URL, extracting details such as title, rent, address, etc.

Step 3 Store scraped data in Pandas dataframe

Step 4 Process dataframe, such as reordering columns according to the specified order

Step 5 Store the dataframe in a file as per the requirement

DATA CLEANING

DATA CLEANING

1. **Merging:** Merged all the scraped data into one single file.
2. **Unique Identifier:** Added a unique identifier for each listing marked by "Sl" column in the dataset.
3. **Out of Scope Check:** Removed all the listings which are outside Halifax Regional Municipality
4. **Duplicity Check:** Removed all Duplicate listings(listings at same address with same rent)
5. **Invalid Rents:** Removed all listings with no rent info
6. **Address Simplification:** Standardized addresses having multiple listings and ranged rent info
7. **Address Geo Location:** Transformed Address line into Street Number, Street Name, Province Name, Country info for Arc GIS mapping
8. **Text to Numeric Conversion:** Converted all text values of Bed Number, Bathroom Number, Rent data into numeric values
9. **Data Pre-processing**: Pre-processed the cleaned data set for ML algorithm application with Python.

MS Excel for Data Cleaning
Python for Data Pre-processing

ML MODEL

Data Loading

Load rental data from Excel.
Check for data quality

Exploratory Analysis

Analyze rental feature distributions. Examine correlations between variables.

Regression Modelling

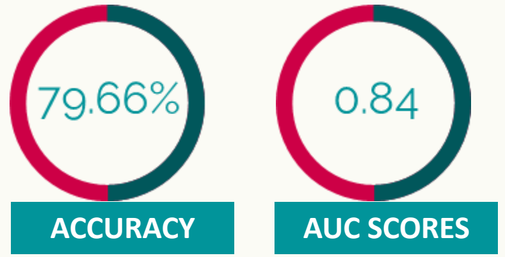
Utilize linear, ridge, and lasso regression. Predict monthly rent prices. Moderate performance achieved.

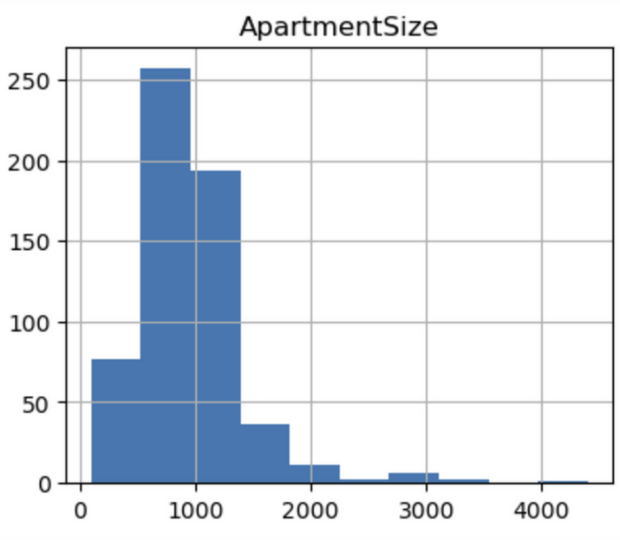
Classification Modelling

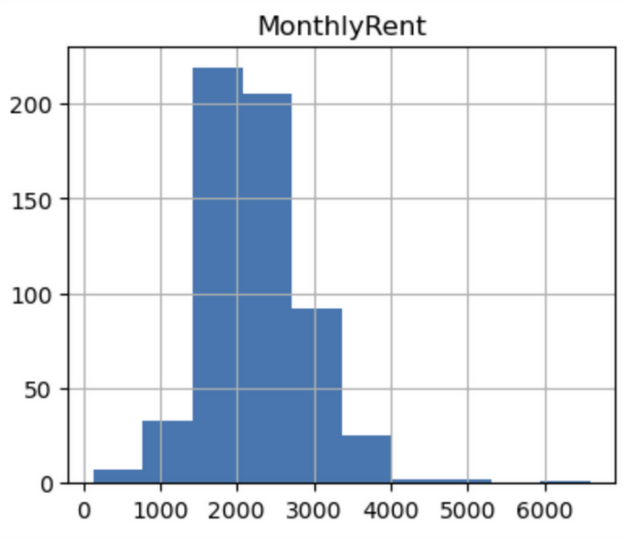
Implement logistic regression, decision tree, random forest, and SVC. Categorize rentals into high/low rent segments. Achieve satisfactory accuracy rates.

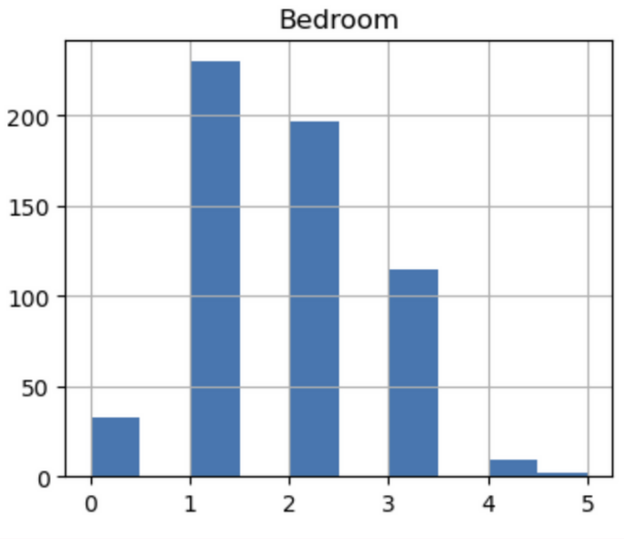
Comprehensive Analysis

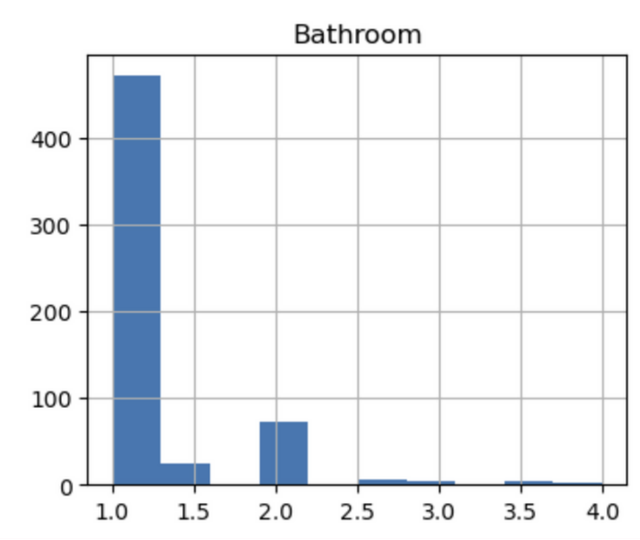
From preprocessing to advanced modeling. Support decision-making in property pricing. Aid in market segmentation insights.

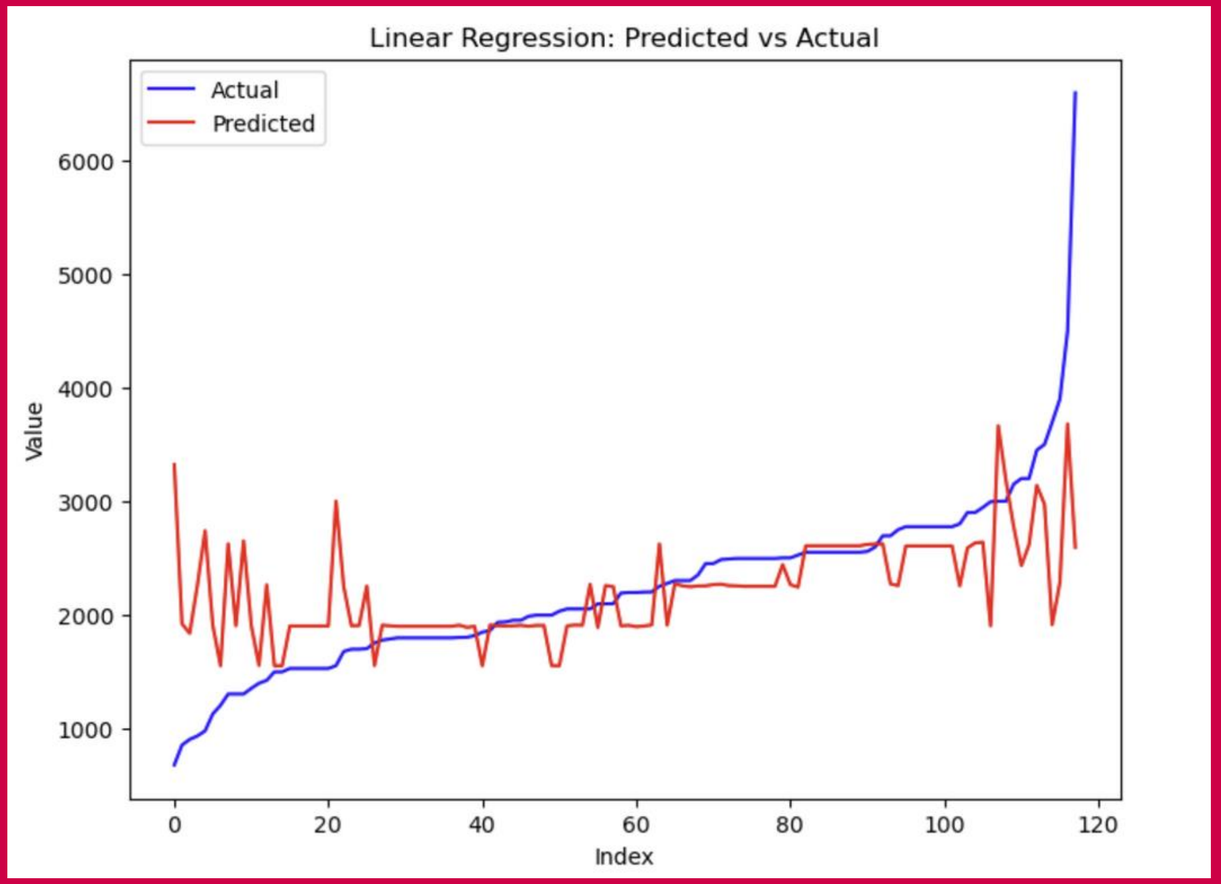


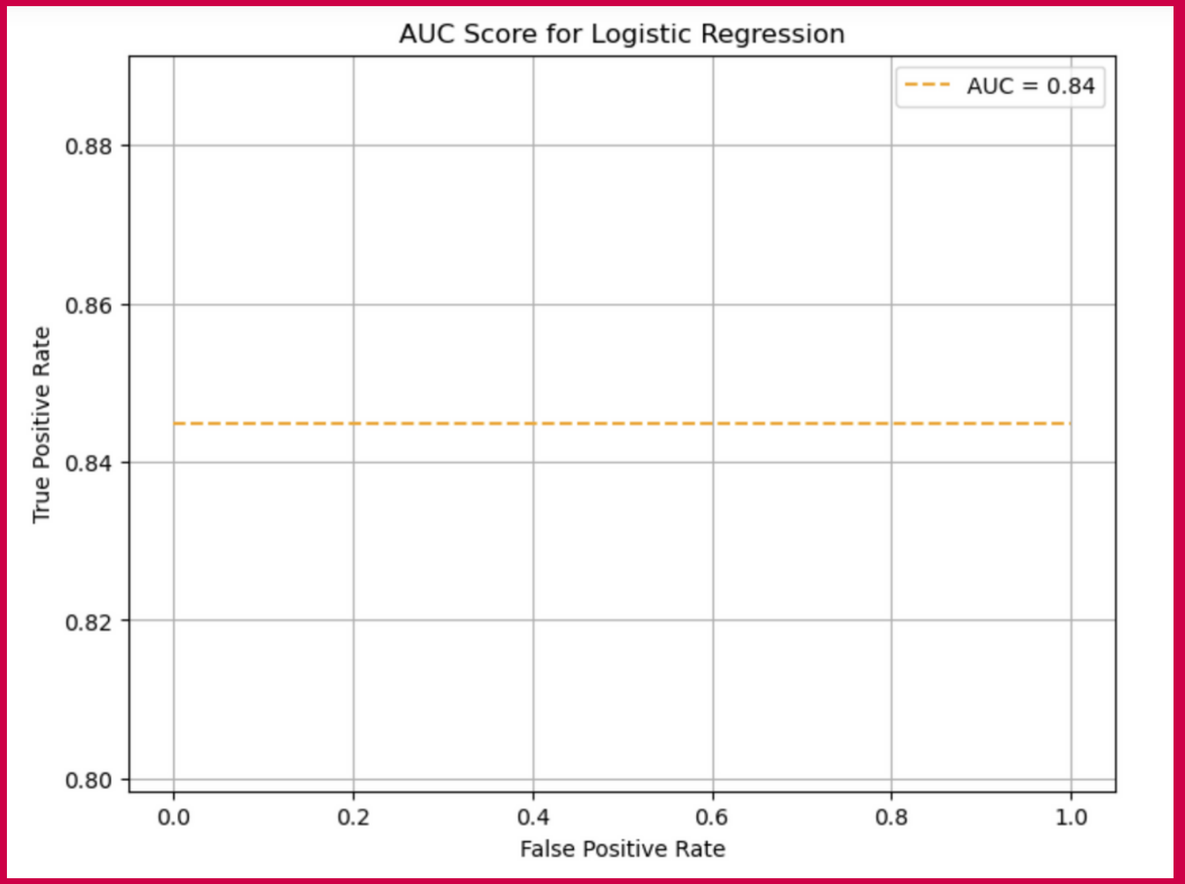

ApartmentSize
0
50
100
150
200
250
0
1000
2000
3000
4000


MonthlyRent
0
50
100
150
200
0
1000
2000
3000
4000
5000
6000

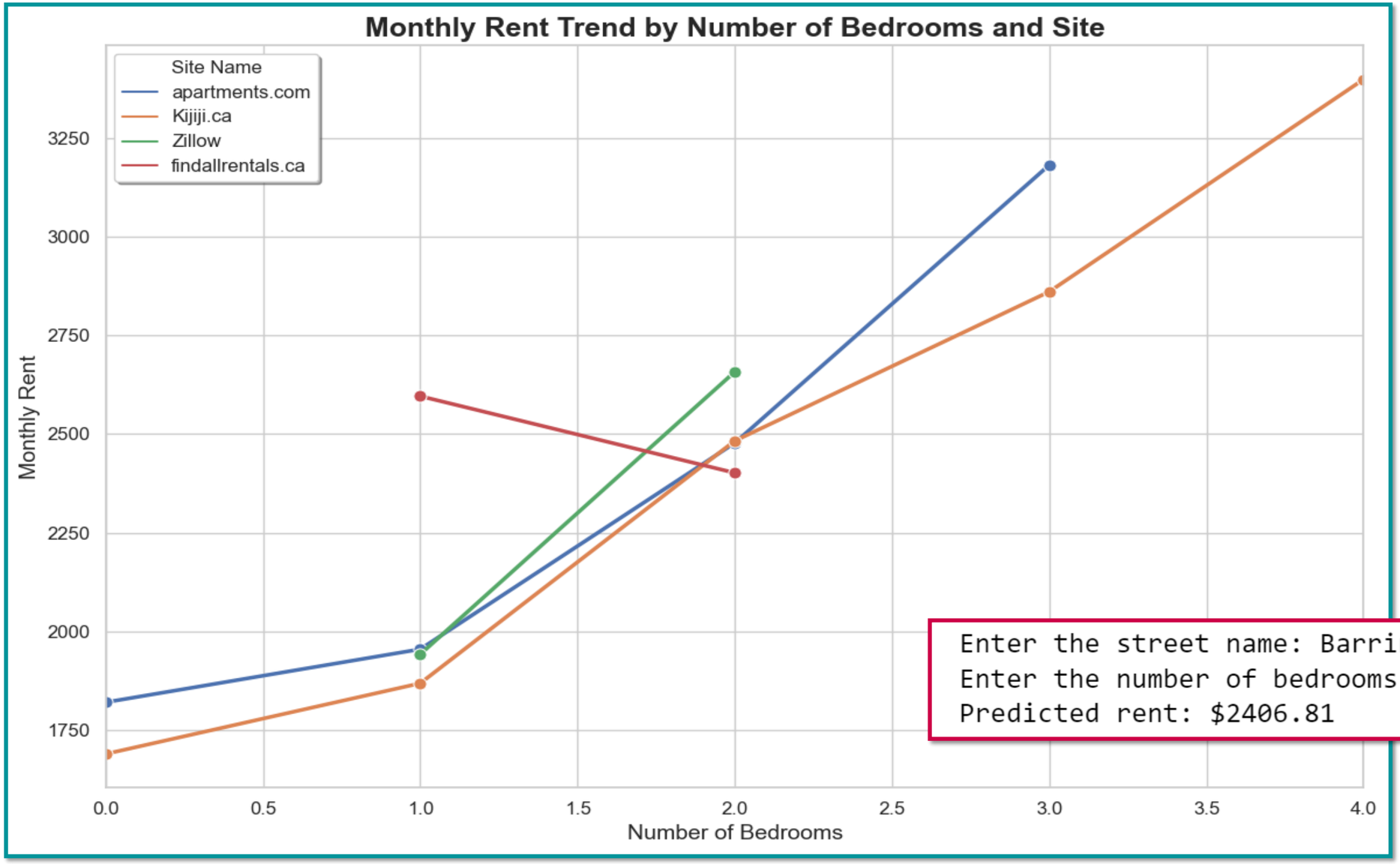

Bedroom
0
50
100
150
200
0
1
2
3
4
5


Bathroom
0
100
200
300
400
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Linear Regression: Predicted vs Actual
Actual
Predicted
Value
Index
1000
2000
3000
4000
5000
6000
0
20
40
60
80
100
120


AUC Score for Logistic Regression
AUC = 0.84
True Positive Rate
False Positive Rate
0.80
0.82
0.84
0.86
0.88
0.0
0.2
0.4
0.6
0.8
1.0

RENT PREDICTION




```
Enter the street name: Barrington Street
Enter the number of bedrooms: 2
Predicted rent: $2406.81
```

DASHBOARDS

POWERBI DASHBOARD

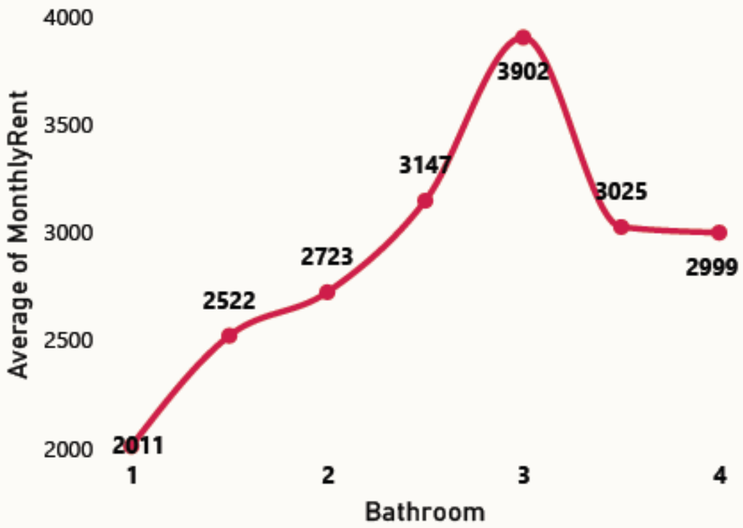
DASHBOARD

Halifax Rental Database

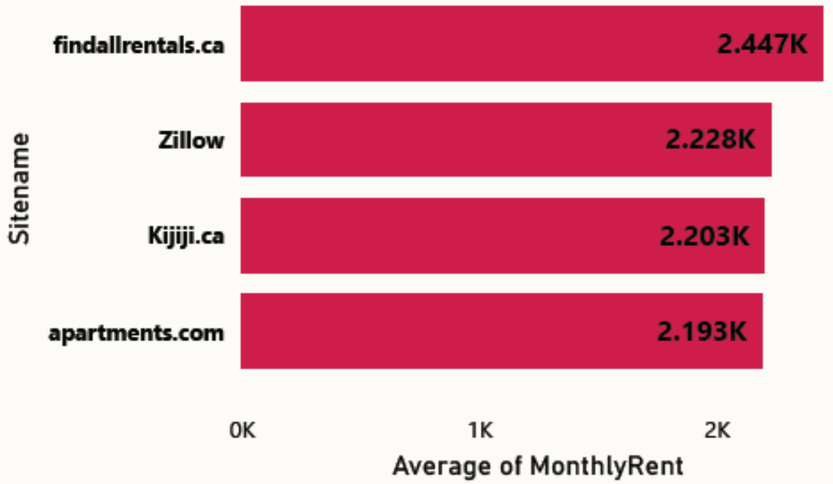
404

Number of Listings

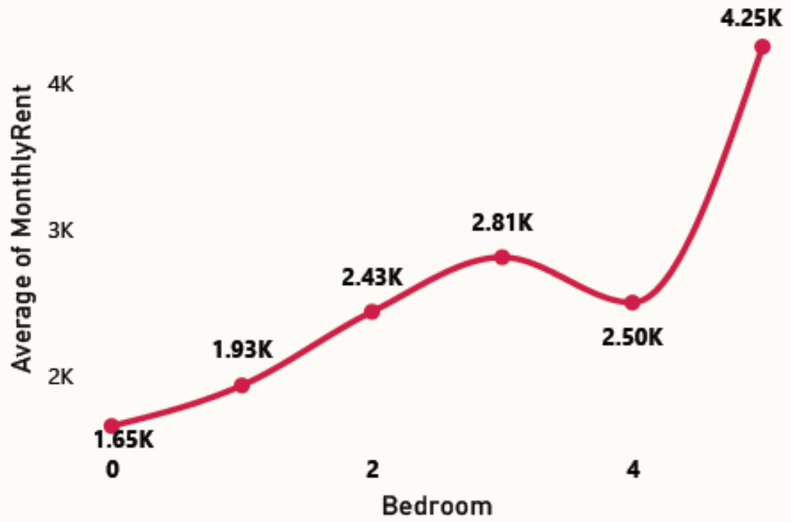
Average of MonthlyRent by Bathroom



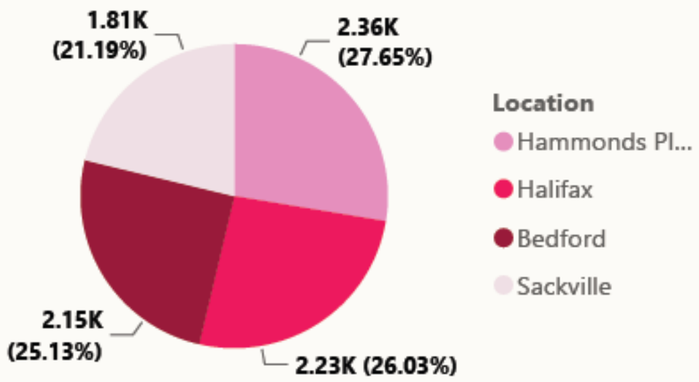
Average of MonthlyRent by Sitename



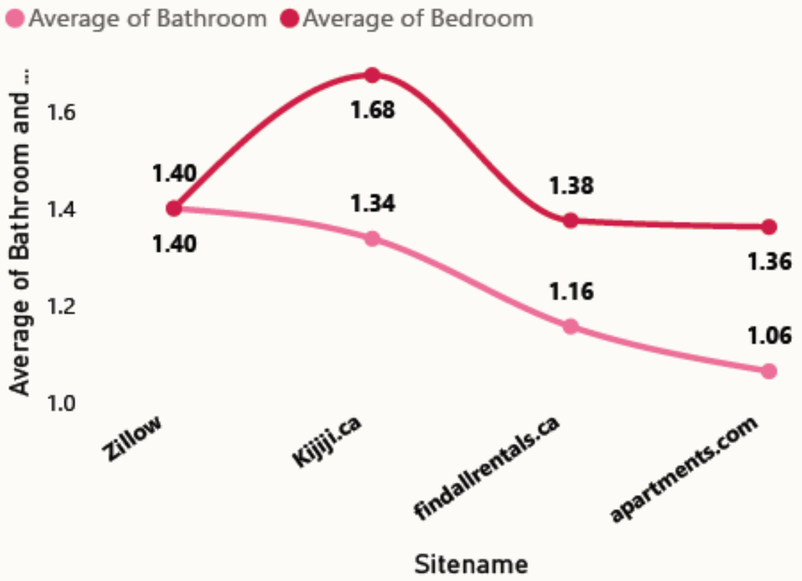
Average of MonthlyRent by Bedroom



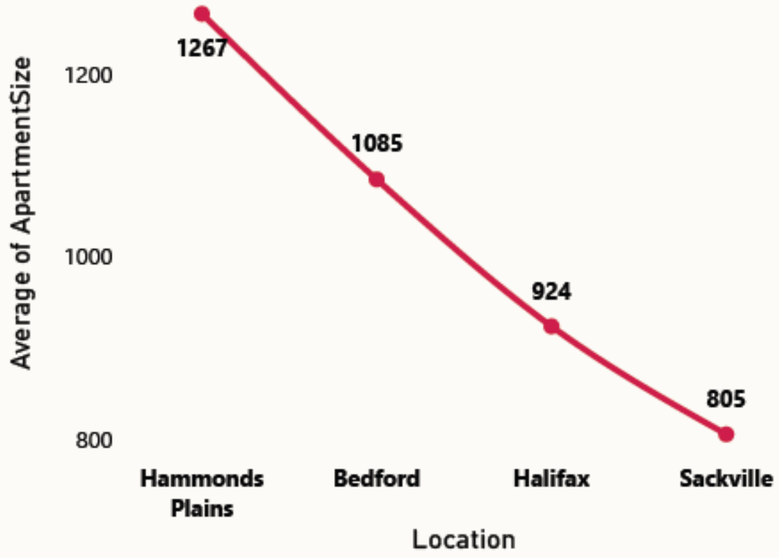
Average of MonthlyRent by Location



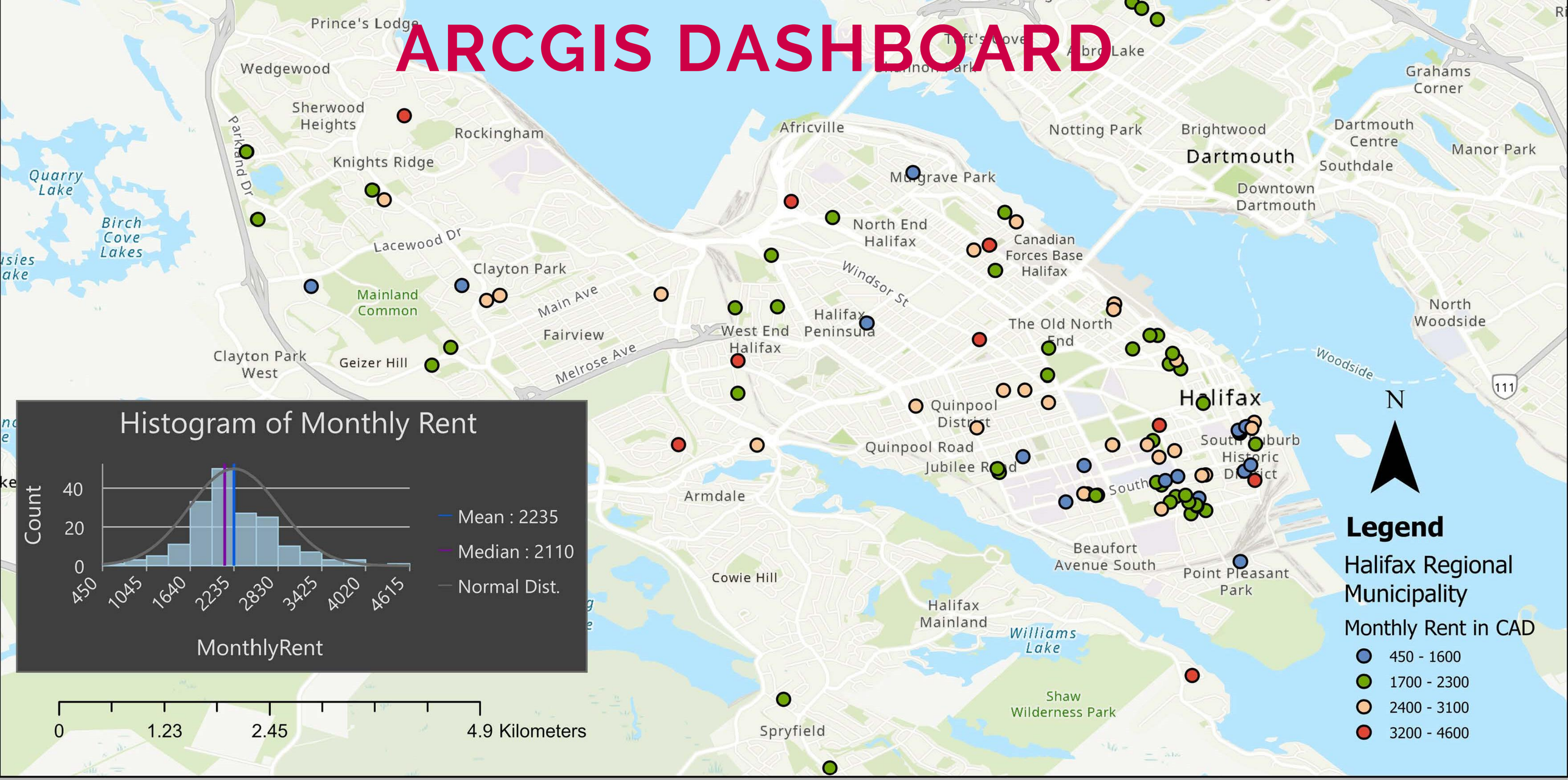
Average of Bathroom and Average of Bedroom by Sitename



Average of ApartmentSize by Location



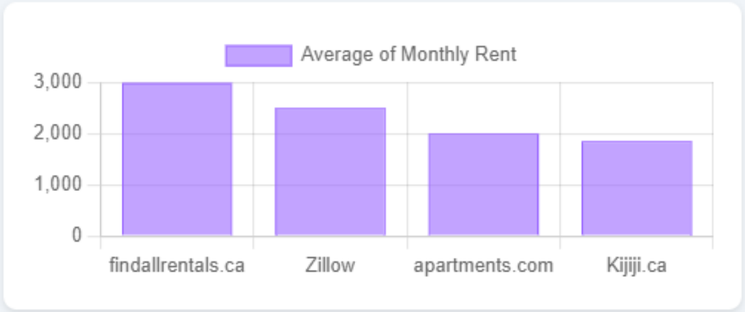
ARCGIS DASHBOARD

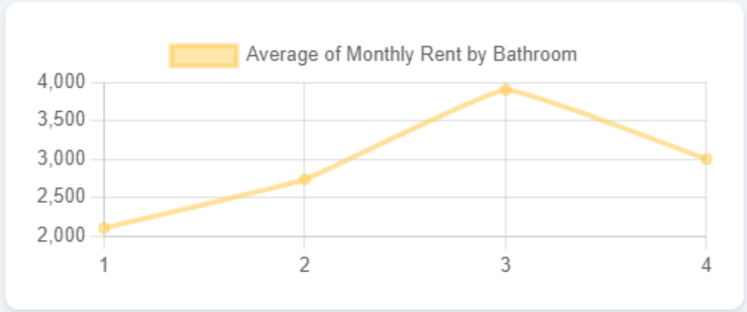


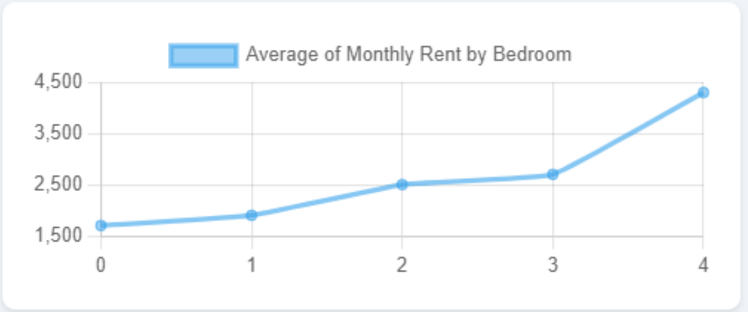
Rent in Halifax

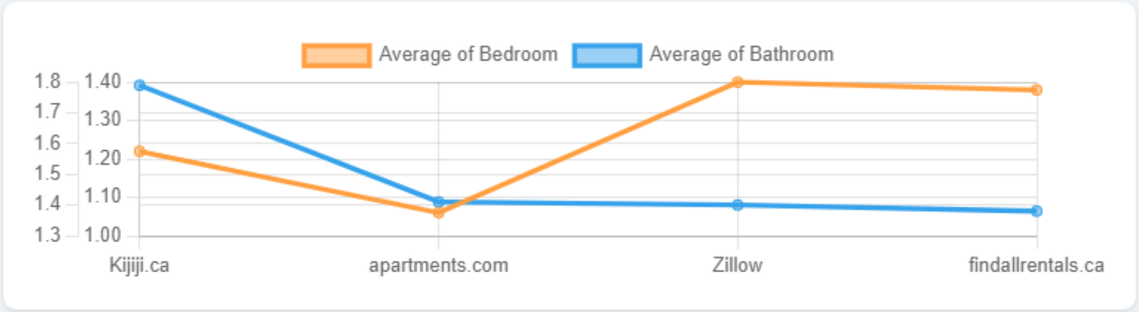
UI/UX REACT

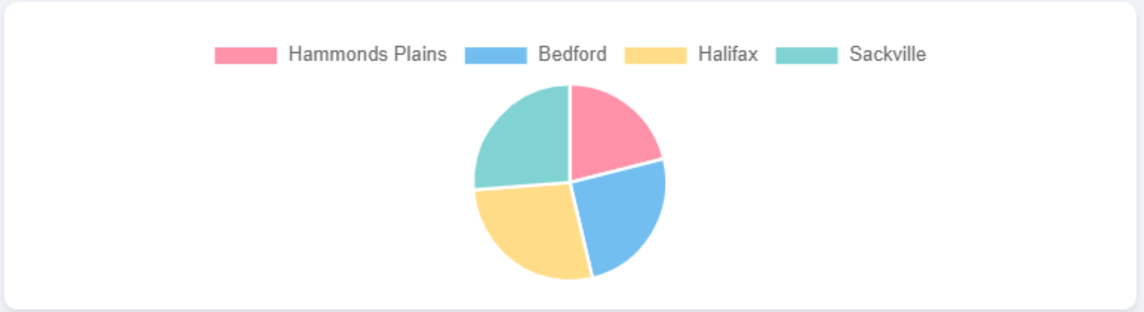
DASHBOARD











THANK YOU!

TEAM MEMBERS

Pallavi Kumari
Akshita Singhal
Yashwanth Balaji Krishnamurthy
Md Akmam Ul Haque
Nitin Chitharanjan
Iram Shaikh