



Zillow Data Scraping Documentation

Maine Pad Finder Project

COS 457 Database Systems

Jeffrey Fosgate

Yunlong Li

1 Overview

This project is a simple web scraper that collects rental property listings in Maine from Zillow and exports them into a structured CSV file for later analysis and database usage.

The scraper:

- Iterates through Zillow result pages.
- Visits each property page and each unit page for apartment complexes.
- Extracts name, address, rent, bedrooms, bathrooms, apartment number, and square footage.
- Outputs all data to `zillow-properties.csv`.

2 Files

`Fosgate_ZillowScraper.py`

Main scraper script:

- Builds Zillow pagination URLs
- Extracts property links
- Detects single properties or apartment complexes
- Extracts unit-level details
- Saves CSV as output

3 Requirements:

- Python 3
- `pip install requests`
- `Beautiful Soup 4`

4 How To Run

1. Install dependencies:

```
pip install requests beautifulsoup4
```

2. Run the script:

```
python Fosgate_ZillowScraper.py
```

3. The output `zillow-properties.csv` will appear in the same directory.

5 How It Works

5.1 Page Loop

The script loops pages until Zillow shows no more results.

5.2 Property Detection

Each link is checked:

- If no unit table exists → single property.
- If unit cards exist → apartment complex (unit card view).
- If simple table exists → apartment complex (table view).

5.3 Single Properties

Extract:

- Address.
- Rent.
- Beds and baths.
- Square footage.

5.4 Apartment Units

For unit cards and tables:

- Copy base building info.
- Extract unit-specific rent, square footage, bed/bath counts, and apartment number.

5.5 Final Export

Writes all data as rows into `zillow-properties.csv`.

6 Data Format

- **NAME**: Property or unit name.
- **ADDR**: Address stored in a list-like string.
- **RENT_COST**: Monthly rent.
- **BEDROOMS**: Bedroom count.
- **BATHROOMS**: Bathroom count.
- **APT_NUM**: Unit number (0 for single properties).
- **SQFT**: Square footage.