



ZYTE PAPER

Fueling real estate's big data revolution with web scraping

Four use cases demonstrating the power of web scraped data in transforming the real estate industry.



Introduction

Real estate has changed a lot in the last thirty years. The proliferation of mobile phones and fax machines, trivial to today's realtors, transformed the entire industry, changing the once tepid pace of the industry into a fast-paced sprint.

With advancements in online advertising, document verification and e-signing, these changes only accelerated change within the industry. Now, however, the advent of big data stands to drive a real revolution in the industry, likely to an extent far greater than all changes yet experienced.

By accessing and incorporating web scraped real estate data into their databases and decision making, aof this data for phenomenal effect. Given the culture of M&A within the industry and fierce competition online, accessing this data first may be a matter of survival for smaller firms, and a bulwark against disruption for larger ones.

Real estate is a data-hungry enterprise. Thankfully, web scraping is a powerful solution providing firms of every size a useful tool for monitoring supply/ demand, market trends and online public sentiment. Simply put, this data can only be properly aggregated to any useful extent if generated in quantities too large for manual input, and web scraping such data has thusly become best practice across the world's many public-facing industries.

At Zyte, ten years of experience scraping billions of sites has revealed the nature and true value of certain data types to those who consume them - whether product data, sentiment data, or the likes. However, among the wide scopes and briefs we deal with, real estate data is especially profitable and is uniquely situated to drive transformative change within the industry.

Realtors and their firms exist in a day-to-day race to stay ahead of competition and stay relevant. Most listing services use web scraping to generate leads for their clients, and some have begun operating as brokerages themselves. Arming your team with the most robust data available, and informing important decisions with the rich data available is going to determine who wins - and loses - in this quickly technologizing industry.

In this white paper, we're excited to explore the following use cases for real estate data:



Appraising Property Value



Monitoring Vacancy Rates



Estimating Rental Yields



Understanding Market Direction

Appraising property value

The real estate market moves quickly, and being even half a step ahead of the competition can be incredibly valuable in maximizing value for both buyers and sellers. If firms receive the same commoditised data from a single MLS, brokerages can differentiate their strategy (and bolster their earnings) by incorporating web scraped data into their decision making.

Thanks to the highly structured nature of real estate listing data, scraping it from the web is extremely effective and provides high quality data - at any scale. By examining aggregated, high-dimensional listing data, real estate agents can recommend pricings and offerings with a greater level of confidence and clarity.

Appraising a property's value is a multifaceted process, but one that can be greatly improved by the incorporation of the web's rich data offerings. By using web scraping, the components necessary for determining a property's value - which obviously vary greatly - can be aggregated and integrated directly into a real estate agent's market analysis.

Sellers can use these comparables to justify asking prices, and buyers can use them to confidently position their offers. It is much easier to deal with an unfortunate appraisal when you know the market's comparables in advance.

A little more than location

Buy side and sell side appraisals require a lot of information. Regardless of the property's intended use, having a collation of all the facts - and from all comparables - empowers agents and firms to make competitive, informed decisions.

Data used in the process of building a comparative market analysis is made accessible to an extraordinary degree by web scraping, and examples range from the following to many others:



Time of sale and proximity

In the past, potentially convoluted public records were scoured to pull comparable sales, but with intelligent web scraping, only data related to sales from within a certain time period can be retrieved, saving realtors from potential headaches or simply irrelevant information.

Information on houses outside of a particular neighborhood or area are simply unusable for comps, and by using web scraping to quickly determine location information, a more accurate perception of a property's value can be understood.



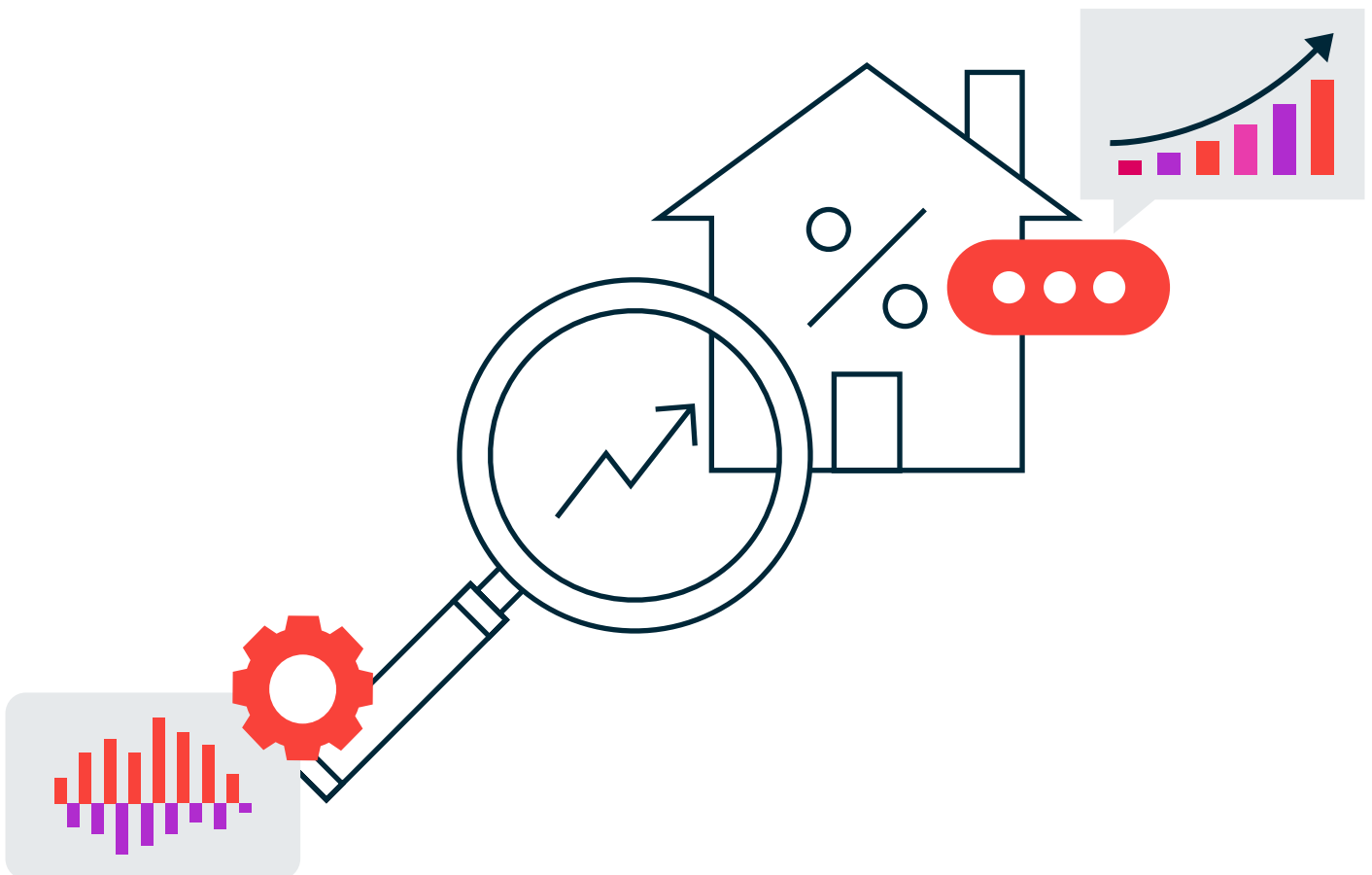
Square footage

Collating hard numbers - like square footage - against other data points with relation to pricing and sales history can provide insights into properties which potentially underwent a remodeling or addition. Square footage may seem a simple item to scrape, but is extremely useful in determining home value when pulling comparables.



Age, construction, and lot size

Accurate market analyses typically exclude properties possessing disparate comparable qualities, and variables like age, build, and lot size can be sifted and intelligently excluded from web scraped data pulls, by incorporating both discrete filters and AI-driven search constraints. This means a McMansion won't depreciate the value of a larger, architecturally superior home a few neighborhoods away.



Monitoring vacancy rates

Every property market in the world is subject to a natural vacancy rate (NVR) representing the tipping point for a market's rental growth. By analyzing market cycles between negative and positive rental growth, brokerages can powerfully position themselves against trends in the market - and there's no better way to fully understand and establish market vacancy rates than by incorporating web scraped data into your analysis.

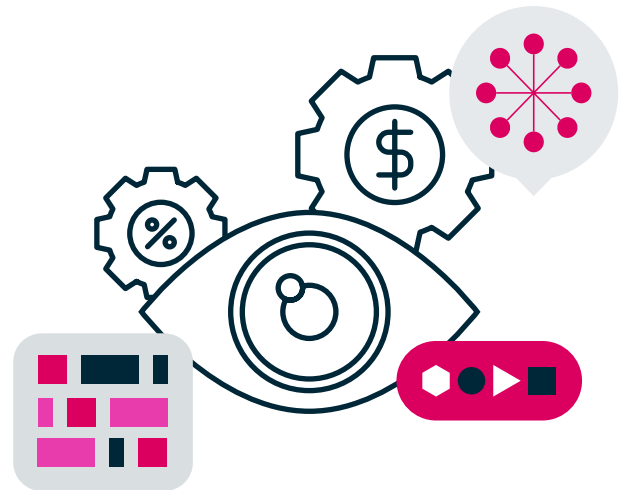
Unfortunately, some econometrics-minded agents will use a static NVR across multiple sectors within real estate - actual NVRs are quite different, varying over historical periods and across commercial, residential, and special use properties. While savvy economists might frequently avail of all the tech big data has made available to them to determine employment trends and market equilibrium, most agents simply don't have the time, and large firms of the recent past didn't have the resources or expertise needed to parse this complex data.

Thankfully, web scraping provides the ultimate solution for scraping and creating high-dimensional data sets, purpose-built for agents across any sector of the industry. Collating this data with other types of important data used in real estate, buyers, sellers, and property managers unlock huge potential in forecasting opportunities and providing their clients with the highest possible return on investment.

Profit at any point of the cycle

One wouldn't walk right into the glass of a revolving door - and the same principle applies to the real estate market's varying and all-important vacancy cycles.

For commercial real estate agents and firms, incorporating this data into your processes enables third parties or in-house intelligence teams to determine the best time to buy and sell this highly sensitive type of property.



For buyers, this data can allude to vacancy opportunities empowering them to negotiate from the strongest position possible, whereas property managers can buttress their expectations and weather potential storms without liquidating assets. This goes for commercial and residential properties alike, and thanks to the extremely customizable nature of web scraping operations, geographic constraints can be set so that only data related to a particular sector or property type, region or cost are represented in the data delivered.

Opportunity knocks

Landlords who time their rentals at the top of the market are sure to profit, and as increasingly standard upwards-only clauses become increasingly prevalent in today's data-driven real estate market, understanding these trends and placing them against an accurately determined NVR is a huge opportunity for growth and profit. The following web scraped data points provide real insight into the market and should be used with other tools for determining action:



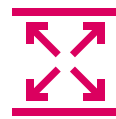
Lease length

Scraping recent records to calculate lease lengths can provide insights into NVR. Shorter leases often indicate high turnover, a signal for higher NVR and frictional unemployment.



Rate of new property construction

A higher number of property completions signals speculative development, and a higher NVR.



Average lot size

A small average lot size may signal speculative development, thus a higher NVR.



Estimating rental yields

It's simple: property owners seek properties that generate an annual income comprising a high percentage of funds invested.

While desired rental yields are in the high single digits, market fluctuations related to any odd number of variables can cause extreme yield volatility in particular markets. By web scraping historical yields confined to comparable properties in the vicinity of a target property, variables like neighborhood, property type, square footage and other metrics can provide real insight for investors and prevent blind decision making.

Furthermore, while gross yield is often a point of frequent discussion, net yield may be an even more crucial metric that is enabled by web scraping.

Vacancy rates, cost of insurance, management fees, to name a few, are all contributors to costs that affect net rental yield.

Web scraping allows the compilation of complex, high-dimensional data sets from numerous real estate websites, so that oft ignored but important criteria can be included in analytics-driven investing.

Property is among the most suitable and stable markets for investment, but until very recently investors have struggled to make informed decisions, with the information necessary to make smart choices irretrievable and historically unavailable. Now, web scraping tools have overcome these trials of the past, and investors are discovering unique, insightful approaches to optimizing their property portfolio.



Making informed investments

Here are a few ways in which web scraping empowers investors to estimate a property's potential and time their investment at the top of the market:



Size of property

With empirical data extracted from rental websites, investors can gain information about which size houses they should buy to expect a given rental yield. If investor wants X % of yield in a certain district, then they can use web scraping to determine what property makeup and sq. footage will provide the highest payout.



Private party leases

Web scraping is not just content specific, but audience specific. Many commercial data sets may give a rough idea of market metrics, but only fine tuned scraping of numerable listings, agreements, etc can give insight into the private rental market that comprises a significant amount of rentals. With this missing data, we are missing a significant and targeted group of investors and renters.



Correlation of rising interest rates and rental yields

Rising interest rates that discourage some homebuyers can drive rental yield up as rental demand grows. A savvy investor will scrape data of historical interest rates and rental yields in an area and invest accordingly to target a high net yield.



Historical trend

Rental yields vary year over year, but scraping historical rental details can help tell a story about a property or neighborhood. Knowing how a property's rental value tracked with the trajectory of its property value provides immense insight into the % yield one can expect.



Feature specific

Net rental yield is also affected by maintenance and repair expenses so it helps investors to see how certain features like AC/Garages/etc factor into rental yield. Informed investors can then use the scraped data to compare the incremental benefit with the incremental cost.

Understaining market direction

The U.S real estate market at large generates key indices that excite investors or evoke fear, but the variability market by market is one of the characteristics of the industry which makes web scraping so valuable.

Web scraping has historically been used to narrowly investigate specific questions.

However, as the real estate market evolves with traditional real estate companies and online vendors and MLSs discovering how they might work together, web scraping might prove a vital tool for a company of any size to find an edge. That edge lies in being able to determine how a market is going to perform in the future.

Understanding market direction is a prerequisite for being a consultative investor or realtor, and is critical for informing decisions empirically with an eye towards what's coming next.

Though many will claim intuition about a market heating up or cooling off, aggregating hard, empirical data is critical for both accuracy and foresight.

Measure what matters



Aggregated property data

As is the case for any company, insights lie partially in the sheer quantity of data that can be compiled and analyzed around certain designations.

Pulling large quantities of current and historical data on properties in a market would be time consuming for an individual, but with web scraping can provide rich insight into the story a market has told over time.



Property investment

The markets in which investor activity is climbing point toward a market consensus that the area is going to appreciate in value.

Web scraping specific parameters like the net growth in investor activity, ratio of investors, ROI, etc., can offer quantitative metrics by which to compare various “hot” markets.



Tracking how long properties are on the market

Many in the industry become concerned when houses start taking over 5-6 months to sell on average. A “buyer’s market” may have some benefits, but in terms of market direction it points toward a cooling market.

As we have seen in recent recessions, there is also a breaking point for real estate bubbles that are the product of overextended lenders. Thus, by web scraping to estimate average time on the market and variance from list price means it’s possible to see correlations between these metrics across a large data set.

Healthy economic growth and a favorable market direction might be evident when properties are closing efficiently close to listing price. A market performing according to plan is more readily quantifiable with web scraped data comparisons.



Monitoring changes in competitor pricing

Real estate is a competitive industry, with companies of all sizes fighting to stay relevant and gain market share. In the race to stay on top of market direction, one’s biggest mistake would be shortsightedness.

Utilizing tools like web scraping to challenge one’s own intuition and analyze how competitor’s are responding to the market is key. Knowing and understanding your competitors’ strategies is half the battle, especially in this industry.

Web scraping listing websites by tags like agent names, company names, etc, datasets that show average listing price, the # of times the price has dropped, time on the market and the agent/company responsible can be produced. These data sets enable competitor tracking and empower firms to deconstruct competitor strategy.



Conclusion

We hope some of these exciting use cases for real estate data have inspired you to implement such data into your own workflows and to innovate in your industry.

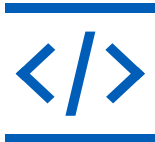
Investors, brokers and property managers alike can avail of the broad insights brought by this data to realize their fullest potential and endow their operations with the most effective, world-class solutions currently available.

As the internet continues to grow, the amount of data it generates grows with it, opening new opportunities for all types of organizations to improve their processes and make more informed decisions.

Therefore, we firmly believe that now is the best time to act, and that by incorporating such data streams into your organizational processes you can ensure your organization is disruption-proofed and fully prepared for the world of tomorrow.

Zyte is proud to offer the most robust web scraping solutions on the market, so if these use cases inspired you or you're curious about how web scraped news and article data can benefit your operation, request a free consultation with our solution architecture team today.





At Zyte we turn websites into data with industry leading technology and services.

Our solutions include:

- **Data Extraction Service**

Let our web scraping experts build and manage the bespoke data extraction solution for your business needs.

- **Smart Proxy Manager (formerly Crawlera)**

Forget about proxy lists. We manage hundreds of thousands of proxies, so you don't have to.

- **Automatic Extraction powered by AI**

Instantly access accurate web data through our user-friendly interface or various Extraction APIs and save time getting the data you need.

- **Data extraction platform**

Access developer tools, data extraction APIs and documentation, built and maintained by our world-leading team of over 100 extraction experts.



It's yours. The web data you need.

Access clean, valuable data with web scraping services that drive your business forward.

[Talk to us](#)

www.zyte.com

Copyright 2021 © Zyte

Cuil Greine House, Ballincollig Commercial
Park Link Road, Ballincollig / Co. Cork, Ireland