

Auckland Property Scout

How the data is collected, how estimates are built, how the models are tested, and what the numbers can and cannot tell you.

Version 1.1

28 October 2026

Coverage

17 North Shore and Eastern Bays suburbs

Access

Free to read and download · independent research

In brief. This project turns public Auckland listing information into a value range with a stated source, date and error. It is a research tool, not a registered valuation or financial advice.

The Research Lab shows the current numbers. This page explains how they are produced and where the limits are.

1. What this project is and is not

Auckland Property Scout is an independent, non-commercial research project covering 17 suburbs on Auckland's North Shore and Eastern Bays: Browns Bay, Murrays Bay, Mairangi Bay, Rothesay Bay, Campbells Bay, Castor Bay, Milford, Long Bay, Waiake, Torbay, Takapuna, Devonport, Mission Bay, Kohimarama, St Heliers, Glendowie and Ōrākei.

The project shows the evidence behind an estimate. Instead of one price, it gives a range, a source, a date and a record of past errors.

What it is

- A dated snapshot of public listing data and derived signals.
- Ranges and indicators that can be checked and reproduced.
- A record of price, sale-method and description changes, and of listings that leave the market.
- A public release of data, schema and code.

What it is not

- **Not a valuation.** It is not a professional valuation and cannot replace a registered valuer.
- **Not a sale-price forecast.** The main target is a third-party automated valuation, not a future sale price.
- **Not financial, legal or tax advice.** Check the title, get a builder's report, and take legal advice before buying.
- **Not live.** It refreshes daily. Every number has a timestamp, and listing pages can change between captures.

In one sentence: this is a property research tool that shows its range, source and error. It cannot tell you what a property will sell for.

2. Data collected and when

The collector reads public Trade Me search results and listing pages for the 17 suburbs. It stores the facts needed for research, with a timestamp for each capture.

What is captured

GROUP	EXAMPLES	WHAT IT IS USED FOR
Identity and location	Listing ID, address, suburb, latitude/longitude, property type	Remove duplicates, map the listing and compare nearby properties.
Physical facts	Bedrooms, bathrooms, parking, land area, floor area, build decade	Size and type adjustment.
Price and campaign	Asking price, guide/auction/tender display, sale method, deadline, listing date	The advertised signal and how the vendor is asking buyers to respond.
Third-party estimates	Homes.co.nz AVM range, its date, council capital value (CV), land/improvement split	Outside estimates and inputs for the fallback model.
Nearby sales	Sold price, date, distance, property type, size, CV	Comparable-sales evidence.
Text	Cleaned public description wording	Topic and wording analysis.
History	Repeated captures, asking-price states, method changes, description digest	Price and sale-method changes.

Refresh and retention

- The job runs daily at **03:15** local time. If the public site blocks a request, a slower browser fallback tries again.
- Recent snapshots are kept in full. Older ones are reduced to the first and last capture per listing per month, which keeps the database small.
- When a listing disappears from a complete 17-suburb search, it is marked **delisted** with a date. If it returns, it is marked **relisted**. Its descriptions, prices and model outputs stay in the database.
- Raw descriptions are kept privately so the work can be rebuilt. They are **not** redistributed because they are third-party content.

Listing universe and cleaning

Public counts and medians use **standard residential listings that are currently active**. Relocatable homes, tiny homes, rentals and other non-standard entries are excluded by default. A toggle shows them. Leasehold properties are kept but flagged, because a leasehold price is not comparable with a freehold estimate. Tenure comes from a Homes record where available, with listing text as a fallback. It is not a title check.

At the 28 October 2026 snapshot, the database holds just over 1,000 captured listings. About 1,000 are active and standard. The app shows the current counts.

3. Tenure, build year and sale history

Trade Me often leaves out tenure, build year and previous sales. We collect them from Homes.co.nz property records when there is a match.

Tenure

The tenure values are **freehold**, **cross-lease**, **leasehold**, **unit title** and **unknown**. Freehold usually means the owner holds the land and buildings, subject to the title. A cross-lease shares one title between homes, with a lease for each unit. A unit title gives ownership of a unit plus a share of shared property, often through a body corporate. Leasehold means the land is leased and ground rent may apply. Listing text is used only when there is no Homes record.

Unknown is not freehold. A missing tenure stays missing. The Research Lab reports how much of the sample has a source.

Build year

Homes gives a build decade and, for many properties, a built date in its timeline. We store an approximate build year and its source, and show it as c. 1960. It is not an exact construction date. Older homes may need more maintenance or have different renovation needs, but age alone is not used to predict price.

Sale history

The Homes timeline records previous sales: date, price if disclosed, record type and council confirmation. These records are stored in `property_sales` and shown in the listing drawer. A previous sale is history. It is not the current asking price.

PROFILE FIELD	SOURCE	AT THE 28 OCT 2026 SNAPSHOT
Tenure	Homes.co.nz ownership record; listing text fallback; otherwise unknown	69% of standard active listings known
Approximate build year	Homes built event/decade	49% known
Previous sale records	Homes property timeline	1,466 records across 458 listings

We report coverage instead of guessing. The daily refresh adds new matches, and a random audit compares the database with the source.

4. Prices on the map and in listings

Two prices are easy to confuse:

- **Asking price:** what the vendor or agent advertises.
- **Value estimate:** what a model or valuer thinks the property may be worth.

The map and listing cards use the same rule:

DISPLAYED PRICE RULE

effective price = asking price, if it exists
otherwise: estimate midpoint → comparable midpoint → value-guide midpoint

A tilde means there is no asking price, so the number is a value guide: ~\$1.40m. A plain \$1.02m means the listing asks that amount. The map bubble, listing card, price filter and price sort all use this rule.

Example. A property asks \$1,023,000. The map bubble and card both show \$1.02m; amounts are rounded on the map. If the same property went to auction with no asking price, the bubble would show ~\$1.40m, with the tilde marking it as a guide.

Both matter. The asking price shows what the vendor wants. The value guide shows what the evidence supports. The map uses the asking price; the listing detail keeps the estimate range.

5. How a value range is built

Signals are used in a fixed order. Each one is shown with its source and confidence. A later method is used only when the earlier one is missing or weak.

Step 1: Find the advertised signal

We read the asking price, guide ("enquiries over"), auction date or deadline from the listing. This shows how the vendor is asking buyers to respond. It is not a valuation.

Step 2: Use the Homes.co.nz AVM when it exists

Homes.co.nz publishes an automated valuation model (AVM) range for many properties. We use its low, middle and high estimates, plus the update date, as the main range when available. It is an independent estimate, but still a model output, not a sale price.

Step 3: Build a range from nearby sales

When there is no AVM, we use recent nearby sales of a similar property type. Each sale is converted to a ratio of its price to its council CV. We apply those ratios to the subject property's CV. This makes small units and large houses easier to compare.

COMPARABLE-SALES CALCULATION

```
weighti = exp(-distancei / 1.4) × exp(-agei / 11) × type_weighti ×  
same_suburb_factori  
ratioi = sale_pricei / CVi  
range = [Q25(ratio) × CV_subject, Q50(ratio) × CV_subject, Q75(ratio)  
× CV_subject]
```

A sale 0.5 km away and a few months old counts more than a sale 2 km away and two years old. The 1.4 km and 11-month figures are model settings, not sharp cut-offs. A sale of the same property type gets more weight than a very different one, and a same-suburb sale gets a small extra weight. Q25, Q50 and Q75 are weighted quartiles: the low, middle and high points of the evidence.

Example. Suppose the property has a CV of \$1,200,000. The weighted sales give ratios of 0.82, 0.90 and 1.01 at the 25th, 50th and 75th percentiles. The range is roughly:

EXAMPLE

$$\text{low} = 0.82 \times \$1,200,000 = \$984,000$$

$$\text{mid} = 0.90 \times \$1,200,000 = \$1,080,000$$

$$\text{high} = 1.01 \times \$1,200,000 = \$1,212,000$$

This range is not a statistical confidence interval. It summarises the spread of the comparable sales. With only two or three usable sales, it gets a lower confidence label and should be read more cautiously.

Step 4: Use the calibrated fallback when needed

The fallback is a hedonic model: it estimates a property from its features. We train it on listings that have a Homes AVM. It is used only when there is no AVM, no usable comparable range, and its confidence is acceptable.

HEDONIC MODEL (SIMPLIFIED)

$$\begin{aligned} \log(\text{estimate_mid}) = & \beta_0 + \beta_1 \cdot \log(\text{CV}) + \beta_2 \cdot \text{beds} + \beta_3 \cdot \text{baths} + \beta_4 \cdot \text{parking} \\ & + \beta_5 \cdot \log(\text{land}) + \beta_6 \cdot \log(\text{floor}) + \dots + \text{suburb} + \text{type} + \text{tenure} + \text{error} \end{aligned}$$

The model learns how each feature relates to the AVM after allowing for the others. Logs make percentage differences easier to model and keep predictions positive. It is trained and tested on separate folds, so the error is measured on cases it did not see.

Step 5: Treat the council CV as a cross-check

The council capital value is an official rating value, usually updated every few years. It lags the market. We use it as a feature, as the denominator for comparable ratios, and as a cross-check. It is not a current market estimate.

Which range wins?

1. Homes AVM, when available.
2. Comparable-sales range, when usable sales exist.
3. Calibrated hedonic range, only when confidence is medium/high.
4. Active-suburb guide as a last resort, clearly labelled as the weakest signal.

The listing detail shows the signals used, so you can see how the range was built.

6. Error and validation

A range is only helpful if its error is known. We report three standard measures.

ERROR MEASURES

$\text{MAE} = \text{average}(|\text{prediction} - \text{reference}|)$

$\text{MAPE} = \text{average}(|\text{prediction} - \text{reference}| \div \text{reference}) \times 100\%$

$R^2 = \text{share of variation in the reference explained by the model}$

MAE is the average dollar error. **MAPE** turns that into a percentage, which is easier to compare across cheap and expensive homes. **R²** runs from 0 to 1. A higher R² means the model follows the reference more closely, but says nothing about whether the reference is right.

Example. A model predicts \$1.08m and the reference is \$1.00m. The error is \$80,000, or 8%. At a 7% MAPE, a \$1.0m property is typically about \$70,000 off. That helps research, but is not precise enough to call a valuation.

The 80% interval comes from out-of-fold errors: errors on cases left out during training. On the validation sample, about 80% of actual values fell inside it. It is a check on calibration, not a guarantee for one property.

SIGNAL (28 OCT 2026 SNAPSHOT)	REFERENCE / TEST	TYPICAL ERROR	HOW TO READ IT
Hedonic fallback with CV	5-fold cross-validation against Homes AVM; 497 training cases	MAE \$265k · MAPE 7.0% · R ² 0.973	Most accurate fallback. The target is an AVM, not a sale price.
Hedonic fallback without CV	Same test with CV removed	MAE \$756k · MAPE 24.6% · R ² 0.784	A harder test. Much weaker, with a low-confidence label.
Comparable-sales model	Compared with Homes AVM on the same listings	MAE \$275k · MAPE 7.2% · R ² 0.969	An independent check when an AVM exists. Small samples make it fragile.
Sale-price model	320 recorded sold prices	MAE \$195k · MAPE 13.4% · R ² 0.830	A different target: actual sale prices. It is a cross-check, not the primary range.
Old suburb-guide baseline	Same validation sample	MAE \$1.22m · MAPE 47.5%	Shows the gain over a simple suburb median.

Two points to keep in mind. The with-CV model is accurate partly because the AVM and CV are related. And the test uses random folds, not a time-based backtest. Daily snapshots will allow a time-based test later. Until then, these are internal validation figures, not an external audit.

7. Sale method and pressure

Sale method follows a six-step ladder:

STAGE	METHOD	INTERPRETATION
0	Auction	Vendor is testing the market and may accept a pre-auction offer.
1	Tender	Buyers submit offers by a deadline.
2	Deadline sale	Similar to tender, with a stated deadline.
3	Price by negotiation	No public number; a range may be discussed.
4	Price guide / enquiries over	A public expectation, often expressed as a floor.
5	Asking price	An explicit advertised number.

The 0–100 **sale-pressure score** is a motivation proxy, not a prediction:

SALE-PRESSURE SCORE

$$\begin{aligned} \text{score} = & 50 \times (\text{stage} / 5) \\ & + 30 \times \min(\text{days_on_market}, 180) / 180 \\ & + \min(20, 8 \times \text{price_drops} + 10 \times \text{method_changes}) \end{aligned}$$

Half the score comes from how explicit the price signal is. Another 30% comes from time on market. Up to 20 points come from price cuts or method changes in the last 180 days. Scores of 40–64 are medium; 65+ is high. The weights are fixed by hand, not estimated from outcomes.

A single-snapshot test found no significant stage effect after controlling for suburb, type, size and other factors. The app therefore calls this a **motivation or urgency proxy**, not a price predictor. With enough price and method changes, we can test each listing over time.

8. Descriptions and text signals

Descriptions are public marketing copy. We test whether the wording carries information about price after controlling for the property and its location. Adjectives are not treated as facts.

Cleaning

We strip HTML, contact details, agency names, viewing details, tender dates and repeated developer copy. Suburb names are removed before topic modelling, so “Milford” or “St Heliers” does not become a topic that just repeats location.

Topics

TF-IDF turns words and two-word phrases into weighted counts. NMF (non-negative matrix factorisation) groups terms that often appear together into 12 topics. Each listing gets a topic share that adds to one. A listing can be partly about family living and partly about renovation.

TOPIC SHARES

W = topic weights from NMF (non-negative)

$$\text{share}_i = W_i / \sum W$$

NMF finds recurring word groups without being given the topics in advance. The labels are editorial summaries. The top terms in the Research Lab are the better guide to each topic.

Lexicon signals

Fixed word lists score urgency, luxury, renovation, views, family/schools and investment language. These scores are descriptive and are not added to the main value range.

What the tests found

- Text features improved a no-CV model in cross-validation. Adding text to the full model with CV made it slightly worse.
- After Benjamini–Hochberg correction, **no individual topic or word list was significant at 5%** for the AVM/CV premium or the gap between asking price and estimate.
- In this snapshot, no individual text effect was statistically significant. Topic shares can help compare copy over time, but are not used for valuation.

What this means. Marketing language is correlated with price in raw comparisons. After controlling for location, size, type and method, it does not reliably explain the premium in this sample.

9. Price changes over time

From 28 October 2026, every listing has a dated asking-price history. We store each **distinct observed price state**, not a row for every day.

What is recorded

- **first_observed:** the first price state seen for that listing.
- **price_change:** a move from one advertised number to another.
- **price_set:** a move from no asking price (auction, tender, POA) to a number.
- **price_removed:** a number is withdrawn for a no-price campaign.
- **price_display_change:** the wording changes without a numeric move, such as "enquiries over" to "asking".
- **description_change:** the copy changes between captures. We store a digest and character counts, not the full private text.

Each row stores the date, price, display text, sale method, previous price, dollar and percentage change, and days since the last observation.

CHANGE CALCULATIONS

```
change = new_price - old_price
```

```
change_percent = change / old_price × 100%
```

```
days_between = date(new observation) - date(previous observation)
```

Example. A St Heliers listing was first seen at \$1,690,000 on 23 October 2026. On 28 October it was \$1,550,000. That is $-\$140,000$, or -8.3% , over 4.34 days. It is what was observed, not evidence that the property will sell at \$1.55m.

The per-listing summary adds the number of changes, days to the first change, average gap between changes, cumulative change, largest drop and last direction. These measures can later show whether reductions cluster at a certain listing age, whether a description edit comes before a price cut, and how quickly vendors respond to a slow campaign.

Early data. The current snapshot has only a few days of daily price-change history.

The counts and medians are real, but too small to estimate how fast Auckland vendors move in general. The picture will improve as daily data builds up.

10. Sales and model feedback

A prediction can be tested when the property sells. The feedback layer compares the disclosed sale price with the estimate that existed **before the sale**. It records the error and recalibrates the model over time.

Matching rule

A Homes.co.nz sale is matched to a listing only when the sale date is on or after the listing start (the advertised date, or first observation). A sale before the listing is history, not feedback. The prior estimate is the last Homes AVM point before the sale, or a pre-sale snapshot if that is missing. Today's refreshed value is not used as a forecast.

OUTCOME CALCULATION

```
sale_error = actual_sale_price - prior_estimate
absolute_percent_error = |sale_error| ÷ prior_estimate × 100%
inside_interval = prior_low ≤ actual_sale_price ≤ prior_high
```

Calibration

For each group (all listings, suburb, property type and tenure), the layer calculates the median ratio of actual sale price to the prior estimate:

CALIBRATION RATIO

```
ratio = actual_sale_price / prior_estimate
calibrated_mid = base_value_mid × median(ratio)
```

A ratio above 1 means prior estimates were, on average, too low; below 1 means too high. It is applied as a **separate sale-feedback signal**, not a silent replacement for the Homes AVM or comparable range. At least ten pre-sale outcomes are needed before a factor is applied. Until then the status is “warming up”.

Machine-learning feedback

At 30 outcomes, a ridge regression is fitted on actual sale prices. Ridge regression is a linear model that keeps its weights from becoming too large. Inputs include the prior value, asking price, CV, size, bedrooms, build year, days to sale, tenure, type, suburb and sale method. The target is $\log(\text{sale price})$, which models percentage errors. Predictions are stored as

sale_feedback_ml and tested on outcomes not used to fit the model. They are not the primary range.

RIDGE FEEDBACK MODEL

$$\log(\text{sale_price}) = \beta_0 + \beta^T x + \varepsilon$$
$$\text{predicted_sale} = \exp(\beta_0 + \beta^T x)$$

Current snapshot. Four same-property sales are available. Their median absolute percentage error is about **5.5%**, and 50% fell inside the prior 80% range. Four cases are not enough to recalibrate every property, so the layer is still warming up.

ADDRESS	SALE DATE	ACTUAL	PRIOR ESTIMATE	ABSOLUTE ERROR
101 Glamorgan Drive	25 Mar 2026	\$1.50m	\$1.44m	4.5%
1/36 Thatcher Street	30 Mar 2026	\$1.312m	\$1.205m	8.9%
40B Waiake Street	26 Aug 2026	\$1.73m	\$1.851m	6.5%
2A, 2B, 2E Rangitoto Terrace	4 Sep 2026	\$1.40m	\$1.419m	1.4%

What we do not do.

One good sale does not prove the model, and one bad sale does not disprove it. The layer waits for a minimum sample, reports coverage and error, and keeps each outcome checkable. That avoids fitting the model to a handful of sales.

11. Reproducibility and downloads

The work is set up so someone else can check it, not just trust the interface.

- The SQLite schema is public, including price history, events, predictions and text tables.
- The release has the database, schema, selected CSV files, analysis code, manifest and SHA-256 checksums.
- A verification command checks the manifest, hashes, SQLite integrity and row counts.
- Raw descriptions and photos are left out of the release. Derived features, metrics and code are included, and the scraper can collect the source data.

REBUILD LOCALLY

```
python3 -m db.build  
python3 -m analysis.research_digest  
python3 -m release.build --out release/out --keep 12
```

Download the data: all data (ZIP) · SQLite database · listing facts CSV · price-change CSV · previous-sales CSV · checksums.

The ZIP bundle is the easiest single download. The checksum file records file hashes and build details. Derived data is CC BY 4.0; code is MIT. Source listing content remains subject to Trade Me's terms and is not redistributed.

12. Limitations

- **The main target is an AVM, not a sale price.** A model can follow Homes.co.nz closely and still miss the eventual sale price.
- **The sample is selective.** Homes AVM matches ordinary houses more often than new builds, units and unusual titles. The fallback is trained on that subset and applied elsewhere. Confidence labels reflect this.
- **The validation is not a time-based backtest.** Random folds test how well the model generalises within this sample. A true time test needs a longer daily panel.
- **Comparable ranges can be thin.** Nearby sales are a small public sample, not a complete sales register. Two or three sales can make a range look more precise than it is.
- **CV changes slowly.** Council values are updated every few years and are not a current market price.
- **Tenure is not a title check.** The flag comes from a Homes record or listing text. Check the title and any ground rent directly.
- **Tenure and build-year coverage are incomplete.** About 69% of standard active listings have a source-backed tenure record and about 49% have a build year in this snapshot. Unknown stays unknown; it is not assumed to be freehold.
- **Some sale prices are withheld.** A transfer can be recorded without a price. We keep the record with an empty price rather than invent one.
- **Price history is young.** The first reliable changes were captured in late October 2026. Early medians will move as the panel grows.

- **Sale feedback starts from a small sample.** Four disclosed same-property outcomes are currently available; the calibration factor and ridge model wait for ten and thirty pre-sale outcomes respectively.
- **Text effects are not significant here.** Topics and word scores are exploratory, not proven price drivers.
- **Association is not causation.** This project reports observed links and behaviour. It does not claim that a feature causes a price, or that a price will move in a set direction.

When a number is uncertain, the interface should show a range, a confidence label, a source date or a caveat. A missing caveat is a defect, not a stronger claim.

13. Glossary

Asking price

The number a vendor or agent currently advertises. It is a negotiating position, not a sale price.

AVM

Automated valuation model. A statistical estimate produced by a data provider, in this project primarily Homes.co.nz.

CV

Council capital value. An official rating value, updated every few years. Used as a cross-check and to compare sales.

Comparable sale

A nearby sold property used as evidence. It is weighted by distance, recency, type and whether it is in the same suburb.

Hedonic model

A model that estimates a value from property characteristics such as size, bedrooms, location and CV.

Out-of-fold error

The error on cases kept out of model training.

Calibration

Adjusting estimates using observed outcomes, such as actual sale prices.

Ridge regression

A linear model that keeps its weights from becoming too large.

MAE

Mean absolute error: the average dollar gap between predictions and reference values.

MAPE

Mean absolute percentage error: the average gap as a percentage of the reference value.

R²

The share of variation in the reference that a model explains. It is not a measure of whether the reference is correct.

80% interval

A range intended to contain about 80% of validation outcomes. Its coverage is checked empirically.

NMF

Non-negative matrix factorisation. A method for finding recurring mixtures of words in a document collection.

BH-FDR

Benjamini–Hochberg false discovery rate. A correction for testing many hypotheses at once. It makes a chance result less likely to look significant.

DOM

Days on market: how long a listing has been publicly advertised.

Price state

A distinct observed asking price, display text and sale method. Repeated daily captures of the same state do not add rows.

Delisted

A listing absent from a complete 17-suburb search. It is retained in the database with a date rather than deleted.

This document is explanatory, not a valuation, not financial or legal advice, and not a substitute for professional due diligence. Current metrics and the live interface: [Research Lab](#). Download: [PDF](#).