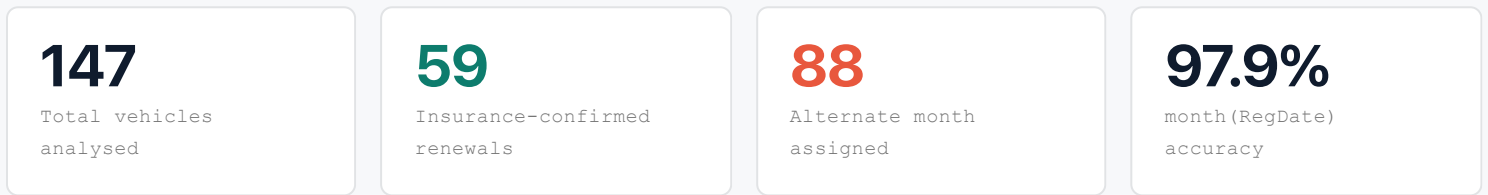


CLIENT	DATE	DATASET	RECORDS
Alivedata (AUST) Pty Ltd — Internal	1 June 2026	Pentana_Matched_20260527.parquet	147 vehicles

CLASSIFICATION
Proprietary · Internal Use Only

EXECUTIVE SUMMARY

Overview



An analysis of 147 Pentana DMS records was conducted to determine the most reliable insurance renewal month for each vehicle. The key findings are:

- The `Matched` flag confirms insurance-verified renewal months** for 59 of 147 records. When `Matched = true`, the `RenewalMonth` field in Pentana is confirmed correct by the insurance company's database.
- `month(RegDate)` is the correct renewal proxy** — it agrees with the insurer-confirmed renewal month in 100% of `Matched=true` records, and in 94% of `Matched=false` records where a clean registration date exists.
- All 88 unmatched records** have been assigned an alternate renewal month (`AltRenewalMonth`) using `month(DeliveryDate)` as the anchor. This is warranted because when delivery and registration fall in different calendar months, the insurance policy is written at the point of vehicle handover (delivery), not at registration.
- A corrected `DeliveryMonth` field** was identified and resolved: 12 records in the source data had `DeliveryMonth` set to `month(RegDate)` rather than `month(DeliveryDate)`. These are corrected in `pentana_matched_clean`.

1. What the Matched Column Means

The `Matched` field was matched against an automotive insurance company's policy database using VIN and/or registration number.

Value	Meaning
<code>true</code> (59 records)	The vehicle's policy was found in the insurance database and the renewal month in Pentana agrees with the insurance record. The <code>RenewalMonth</code> field is confirmed correct .
<code>false</code> (88 records)	Either the vehicle's policy was not found in the insurance database, or the renewal month in Pentana does not agree with the insurance record. The <code>RenewalMonth</code> field is unverified or incorrect .

Important

`Matched = false` does not mean the vehicle is uninsured. It means Pentana's renewal month could not be confirmed against the external database for that record.

2. Core Pattern — What Makes a Record Match

Analysis of the relationship between `DeliveryDate` and `RegDate` for all records reveals a single dominant pattern:

A record is `Matched = true` if and only if delivery and registration occurred in the same calendar month.

The rule holds at 100% precision for all records with a delivery-to-registration lag of 0–31 days:

Lag	Same calendar month?	Matched=true rate
Same day (0 days)	Yes	100% (12/12)
1–7 days	Yes	100% (33/33)
1–7 days	No	0% (0/1)
8–31 days	Yes	100% (4/4)
8–31 days	No	0% (0/10)
> 31 days	Mixed	17% (3/18) — re-registration coincidences only
Reg after delivery	Same month	100% (6/6)
Reg after delivery	Different month	9% (1/11)

The four `Matched = true` records that appear to cross calendar months are all vehicles where the original `DeliveryMonth` field in Pentana was corrupt — it stored `month(RegDate)` rather than `month(DeliveryDate)`. These are corrected in `pentana_matched_clean`.

Interpretation

The insurance database uses delivery month as the policy start anchor. When a vehicle is registered and delivered in the same month, Pentana's registration-based `RenewalMonth` agrees with the insurer's delivery-based anchor — hence `Matched = true`. When they fall in different months, Pentana and the insurer diverge — hence `Matched = false`.

3. Renewal Month Proxy Accuracy

Using the 95 records with a clean registration date (year ≥ 2015, `RegDate` not null):

Method	Correct	Base	Accuracy
<code>month(RegDate)</code>	93	95	97.9%
<code>DeliveryMonth</code> field (corrected)	55	95	57.9%
<code>month(DeliveryDate)</code>	55	95	57.9%

`month(RegDate)` is by far the strongest proxy — it fails on only 2 records (`YGL68Z` and `K63DD`), both of which have an unexplained +2 month offset in the Pentana system (possibly a configurable grace period or DMS anomaly — requires confirmation with Pentana).

`month(DeliveryDate)` scores 57.9% because it fails on all 40 records where delivery and registration fall in different months. However, for those same cross-month records, `month(DeliveryDate)` is what the insurer uses — making it the correct alternate for unmatched records.

4. Rules for Assigning the Best-Estimate Renewal Month

All logic is implemented in the `pentana_renewal_corrected` table with three new columns:

Column	Type	Description
<code>AltRenewalMonth</code>	Integer 1–12	Best-estimate renewal month for all 147 records
<code>AltRenewalConfidence</code>	String	Confidence tier (see rules below)
<code>AltMonthChanged</code>	Boolean	Whether the alternate differs from Pentana's original

— Rule hierarchy — first matching rule wins

Rule 1 — Confirmed

confirmed 59 records

Condition `Matched = true`

Action Use `RenewalMonth` as-is. `AltRenewalMonth = RenewalMonth`

The insurance company's database has verified this renewal month. No override is needed.

Examples

Rego	DeliveryDate	RegDate	RenewalMonth	Action
<code>2ES6IU</code>	2026-04-14	2026-04-13	April	Keep — confirmed
<code>FMX64</code>	2022-06-16	2022-06-15	June	Keep — confirmed
<code>AIJ092</code>	2016-06-27	2023-06-27	June	Keep — confirmed (re-registration, same month)

Rule 2 — EOM Cross-Month

High confidence

12 records

Condition `Matched = false`, clean `RegDate` (year ≥ 2015), lag between -7 and +28 days, and `month(RegDate) ≠ month(DeliveryDate)`

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_eom`

End-of-month registrations — vehicle processed in Pentana on the last days of one month, physically handed over to the customer in the following month. Insurance is written at handover. Pentana's `RenewalMonth` follows registration; the insurer anchors to delivery. **Using Pentana's month would outreach one month too early for all 12 vehicles.**

Examples

Rego	DeliveryDate	RegDate	Lag	Pentana month	Alt month
2ES5HF	2026-04-01	2026-03-31	1 day	March	April
2ES5GU	2026-04-09	2026-03-31	9 days	March	April
2ES5HL	2026-04-14	2026-03-31	14 days	March	April
2ES5IS	2026-04-18	2026-03-31	18 days	March	April
8AA1FV	2026-04-02	2026-03-11	22 days	March	April
8AA1GA	2026-04-09	2026-03-18	22 days	March	April
ANX235	2022-07-08	2022-06-26	12 days	June	July
FBG82B	2023-09-25	2023-08-31	25 days	August	September
2BK5KB	2024-11-28	2024-10-31	28 days	October	November
2ET5HZ	2026-05-04	2026-04-07	27 days	April	May
2ES3AS	2026-03-31	2026-04-01	-1 day	April	March (reverse: delivered last day of March)

Rule 3 — Future or Pending RegDate

High confidence

9 records

Condition `Matched = false`, clean `RegDate` (year ≥ 2015), lag < -7 days (registration significantly after delivery)

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_future_regdate`

Vehicle delivered before registration was finalised in the DMS — plates not yet issued, registration transfer delay, or retrospective DMS update. The delivery date is the only confirmed event; insurance is arranged at handover.

Examples

Rego	DeliveryDate	RegDate	Lag	Pentana month	Alt month
YGL68Z	2026-03-17	2026-05-08	-52 days	July (anomalous)	March
UNREG	2026-04-17	2026-06-30	-74 days	June	April
AKD295	2026-03-05	2026-09-16	-195 days	September	March
AOP451	2017-11-15	2025-05-10	-2,733 days	May	November

Rule 4 — No Registration Date

Medium confidence

48 records

Condition `Matched = false`, `RegDate IS NULL`

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_delivery_only`

No registration date is present. The delivery date is the sole available anchor. This is the largest unmatched group (48 records = 33% of the dataset). **Recommended next step:** backfill `RegDate` for these records via REGO/NEVDIS API using VIN.

Examples

Rego	DeliveryDate	RegDate	Alt month
<code>AHR045</code>	2016-04-23	—	April
<code>AIQ618</code>	2016-05-20	—	May
<code>AIU381</code>	2016-07-11	—	July

Rule 5 — Long Lag (Used Vehicle / Re-registration)

Medium confidence

12 records

Condition `Matched = false`, clean `RegDate` (year ≥ 2015), lag > 90 days

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_long_lag`

A large gap between registration and delivery typically indicates a used vehicle registered by a prior owner. Pentana's `RenewalMonth` reflects the original owner's registration date. The current owner's insurance policy is written at the time of their purchase (delivery date). The delivery month is the correct anchor for this customer's renewal cycle.

Examples

Rego	DeliveryDate	RegDate	Lag	Pentana	Alt	Note
<code>AHT535</code>	2016-08-08	2016-03-31	130 days	March	August	Delayed delivery of new stock
<code>CSB697</code>	2024-04-30	2023-12-04	148 days	December	April	Floor stock, ~5 month lag
<code>3G3UB</code>	2026-03-31	2025-10-14	168 days	October	March	Used vehicle purchase
<code>AJI152</code>	2022-11-01	2016-12-28	3,250 days	December	November	Vehicle re-sold ~9 years later
<code>AHV933</code>	2026-12-01	2016-08-16	3,410 days	August	December	Vehicle re-sold ~10 years later

Rule 6 — Lag 29–90 Days

Medium confidence

3 records

Condition `Matched = false`, clean `RegDate` (year ≥ 2015), lag 29–90 days

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_lag_29_90d`

Registration more than a month before delivery but within 90 days. May be interstate transfers, dealer demo vehicles, or delayed stock intake. Insurance is written at delivery.

Examples

Rego	DeliveryDate	RegDate	Lag	Pentana	Alt
<code>AJI164</code>	2017-02-24	2017-01-03	52 days	January	February
<code>ARH784</code>	2018-02-24	2017-12-07	79 days	December	February
<code>3E6ZP</code>	2025-05-28	2025-04-07	51 days	April	May

Rule 7 — Bad Registration Date (1973 Sentinel)

Low confidence

4 records

Condition `Matched = false`, `year(RegDate) < 2015`

Action `AltRenewalMonth = month(DeliveryDate)`

Confidence `derived_bad_regdate`

A `RegDate` year of 1973 is a Pentana DMS placeholder for an unknown registration date. The value is not real and must be discarded. All four records have `RenewalMonth = 7` (July) — derived directly from `month(1973-07-xx)` — which is incorrect.

Rego	DeliveryDate	RegDate	Pentana RenewalMonth	Alt month
<code>2ER6YZ</code>	2026-04-01	1973-07-17	July (placeholder)	April
<code>ATS369</code>	2023-04-17	1973-07-09	July (placeholder)	April
<code>FPO71J</code>	2026-04-14	1973-07-17	July (placeholder)	April
<code>AIN557</code>	2025-01-18	1973-07-08	July (placeholder)	January

5. Revised Renewal Month Distribution

Using `AltRenewalMonth` across all 147 records:

Month	Total	Confirmed	Derived	Notes
January	4	1	3	
February	6	1	5	2 × lag-29-90d, 2 × delivery only, 1 × long lag
March	10	3	7	Includes 1 EOM, 3 future-RegDate, 2 long lag
April	71	36	35	Dominated by Apr 2026 new-stock batch (dealer intake wave)
May	7	3	4	1 × EOM, 2 × delivery only, 1 × lag-29-90d
June	7	5	2	
July Δ	4	0	4	All derived — 2 anomalies, 1 EOM, 1 delivery only. Use with caution.
August	11	1	10	Driven by used-vehicle deliveries and no-RegDate records
September	6	2	4	
October Δ	5	0	5	Zero confirmed — all derived
November	10	4	6	
December	6	3	3	
Total	147	59	88	

April dominance — sample bias

This Pentana extract was pulled in May 2026 and captures a wave of April 2026 new-stock deliveries (72 records with a 2026 delivery date). A full extract across all months would normalise this distribution.

July caution — zero confirmed records

No July records are insurance-confirmed. The 4 derived-July records include 2 with the known +2-month Pentana anomaly (`YGL68Z` , `K63DD`) and 1 with a no-RegDate. Do not use July as a campaign renewal month from this dataset without further validation.

6. Data Quality Issues Requiring Remediation

Issue	Count	Recommended action
No <code>RegDate</code>	48	Backfill via REGO/NEVDIS API using VIN
<code>RegDate</code> year = 1973 (sentinel)	4	Investigate with Pentana DMS team; obtain real RegDate
<code>DeliveryMonth</code> field \neq <code>month(DeliveryDate)</code> (corrected)	12	Corrected in <code>pentana_matched_clean</code> . Raise with Pentana: field should always reflect physical delivery month
<code>RenewalMonth</code> anomaly — +2 month offset	2	<code>YGL68Z</code> , <code>K63DD</code> . Confirm with Pentana whether a configurable offset applies to certain vehicle types
Future <code>RegDate</code> (> DeliveryDate by > 7 days)	7	Confirm with Pentana: pending registrations, re-registrations, or data-entry errors?
Rego = NULL or UNREG	4	Cannot be matched by rego. Use VIN-only matching
VINs not matched in auto_insights DB	49	Load 2025–27 model-year VINs into auto_insights

7. Rules Summary — Decision Logic

Complete rule set in priority order. First matching condition wins.

```
-- AltRenewalMonth assignment logic (priority order)

CASE

-- Rule 1: Insurance-confirmed
WHEN Matched = true
  → AltRenewalMonth = RenewalMonth          -- confirmed

-- Rule 2: EOM cross-month batch
WHEN Matched = false
  AND RegDate IS NOT NULL AND year(RegDate) >= 2015
  AND lag BETWEEN -7 AND 28 days
  AND month(RegDate) != month(DeliveryDate)
  → AltRenewalMonth = month(DeliveryDate)   -- derived_eom

-- Rule 3: Future / pending RegDate
WHEN Matched = false
  AND RegDate IS NOT NULL AND year(RegDate) >= 2015
  AND lag < -7 days
  → AltRenewalMonth = month(DeliveryDate)   -- derived_future_regdate

-- Rule 4: Bad RegDate (sentinel)
WHEN Matched = false
  AND year(RegDate) < 2015
  → AltRenewalMonth = month(DeliveryDate)   -- derived_bad_regdate

-- Rule 5: No RegDate
WHEN Matched = false
  AND RegDate IS NULL
  → AltRenewalMonth = month(DeliveryDate)   -- derived_delivery_only

-- Rule 6: Lag 29-90 days
WHEN Matched = false
  AND RegDate IS NOT NULL AND year(RegDate) >= 2015
  AND lag BETWEEN 29 AND 90 days
  → AltRenewalMonth = month(DeliveryDate)   -- derived_lag_29_90d

-- Rule 7: Long lag - used vehicle / re-registration
WHEN Matched = false
  AND RegDate IS NOT NULL AND year(RegDate) >= 2015
  AND lag > 90 days
  → AltRenewalMonth = month(DeliveryDate)   -- derived_long_lag

END

-- Universal fallback for all unmatched cases:
-- AltRenewalMonth = month(DeliveryDate)
-- Insurance policies are written at the point of vehicle handover.
```

8. Recommendations

1. Use `pentana_renewal_corrected.AltRenewalMonth` for all campaign scheduling. Filter by `AltRenewalConfidence` to manage risk:
 - `confirmed` + `derived_eom` + `derived_future_regdate` → include with full confidence
 - `derived_delivery_only` + `derived_long_lag` + `derived_lag_29_90d` → include with a review flag
 - `derived_bad_regdate` → exclude until `RegDate` is resolved
2. Do not use the `Matched` field as a campaign audience filter. Filtering to `Matched = true` only removes 88 records with derivable renewal months, reducing campaign reach by 60% unnecessarily.
3. Do not use Pentana's `DeliveryMonth` field directly. 12 records have it set to `month(RegDate)` rather than `month(DeliveryDate)`. Always derive delivery month as `month(DeliveryDate)`.
4. Backfill `RegDate` for the 48 no-`RegDate` records via REGO/NEVDIS API. These are the only records where the alternate is a best-guess with no cross-check available.
5. Confirm `RegDate` semantics with Pentana: Does `RegDate` always reflect the current owner's registration, or can it carry a prior owner's date? Clarifying this affects how long-lag used-vehicle records are treated.
6. Investigate the +2 month offset on `YGL68Z` and `K63DD`. If this is a Pentana system configuration, it may affect a larger population in a full extract.
7. Request a full Pentana extract. At 147 records, April is over-represented due to a single new-stock wave. A full historical extract is needed for reliable seasonality modelling.

