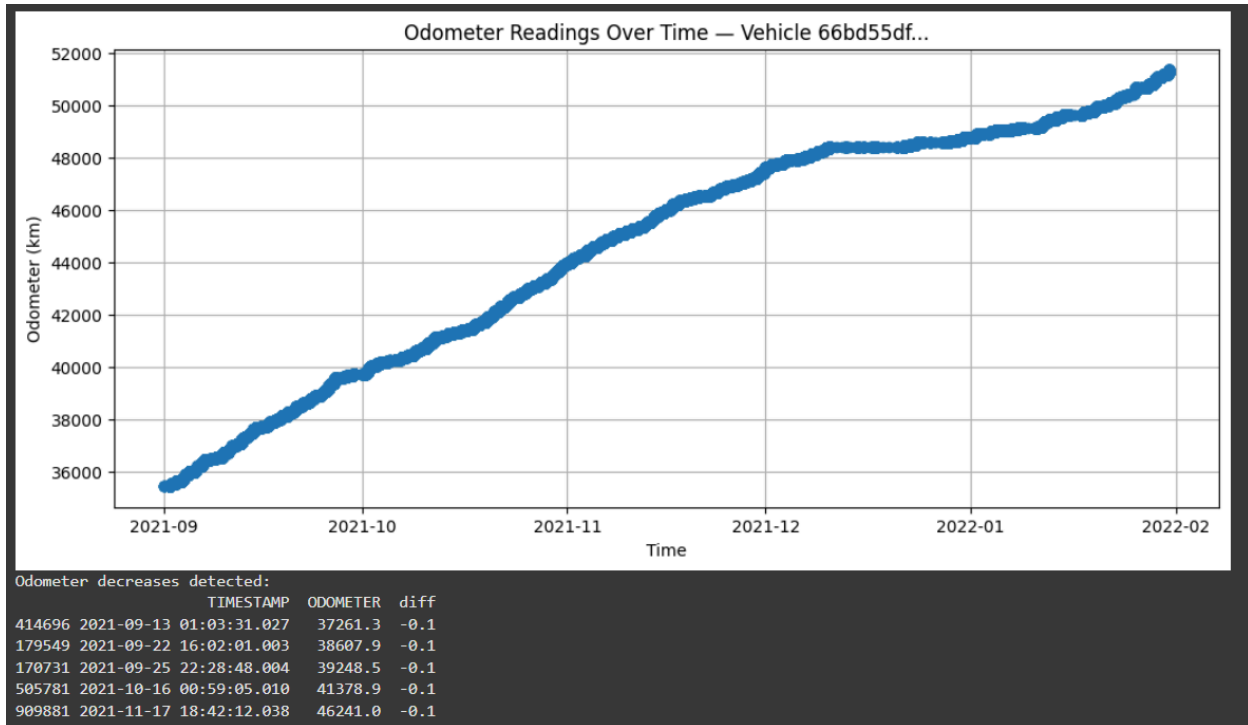


## General Data Quality

- **High missingness in TLM:**
  - ~97% of rows lack ignition status.
  - ~79% lack speed values.
  - ~79% lack EV battery readings.
- **TRG mapping gaps:** ~65.9% of TRG rows couldn't be mapped to a `vehicle_id` via MAP (flagged as `UNKNOWN`).
- **MAP complexity:** Many TRG `PNIDs` map to a single TLM `vehicle_id` through lists of `PNIDs`. Some vehicle IDs had no mapping available.
- **SYN coverage:** Only 11 vehicles included; all events are ignition OFF.

## Range and Plausibility Checks

- **Battery % values:** 0–110%. Slightly above 100% likely due to sensor calibration error or reporting quirk.
- **Speed values:** 0–96 km/h, reasonable for city or short-trip vehicles.
- **Odometer values:** 30,207–66,270 km. Mostly plausible; however:
  - **5 cases of odometer decreases** (drops of 0.1 km), likely resets or rounding issues.



## Duplicate & Noise Issues

- **Duplicates:**
  - ~9,700 duplicates in TLM.
  - ~9,600 duplicates in TRG.
  - Deduplicated during preprocessing.
- **Ignition flickers:** Ignition status sometimes toggled ON ↔ OFF rapidly.
  - Applied a **60-second debounce rule** to avoid double-counting.
  - Noted that if a car actually had electrical flickers, these would be masked — but treated as noise per project brief.

## Ignition Event Findings

- **Total Ignition Events:** 34,911

- TRG: ~30,880
- TLM: ~3,620
- SYN: 411 (all OFF)
- **Balanced ON/OFF overall:** ~17.9k OFF vs ~17k ON.
- **Source nuance:**
  - TRG and TLM provide natural ON/OFF cycles.
  - SYN provides authoritative OFF events only, reinforcing shutdown coverage.
- **Preview misleadingness:** In some slices (e.g., SYN-heavy vehicle\_id), the output appears as “all OFF” even though the overall dataset is balanced.
  - Documented why this occurs to avoid reviewer confusion.

## Visual Insights

- **Bar Chart:** Confirms SYN supplies only OFF events, while TRG and TLM together create balanced ignition patterns.
- **Sample Vehicle Timeline:** Shows alternating ON/OFF cycles over months. Around Jan 2022, red SYN OFF markers appear, supplementing TRG coverage.

Done So Far:

## Data Sanity Pass

- **Timestamps:** Parsed successfully across TLM, TRG, MAP, and SYN.
- **Ranges checked:**
  - Speed: 0–96 km/h (plausible).
  - Battery: 0–110% (slightly over 100% but plausible with sensor calibration).
  - Odometer: ~30,207–66,270 km.

- **Anomalies flagged:**
    - ~65% of TRG rows unmapped to a vehicle\_id (kept with flag as **UNKNOWN**).
    - ~97% of TLM rows had no ignition status.
    - 5 odometer decreases detected (possible resets/rollovers).
  - **Duplicates removed:** ~9,700 from TLM, ~9,600 from TRG.
  - **Decision:** Keep anomalies flagged rather than drop — to preserve maximum event coverage.
- 

## Ignition Event Extraction

Extracted ON/OFF events from three sources:

- **TLM (Telematics):**
  - High-frequency sensor snapshots.
  - Cleaned + debounced to remove flickers.
  - Final count: ~3,620 ignition events (balanced ON/OFF).
- **TRG (Trigger Logs):**
  - Based on **IGN\_CYL** values.
  - Final count: ~30,880 ignition events.
  - ~32% mapped to a known vehicle via MAP.
- **SYN (Synthetic Overrides):**
  - 411 curated **ignitionoff** moments.
  - Highest priority (trusted OFF events).

**Combined IgnitionEvents:**

- **Total:** 34,911
- Schema delivered: `vehicle_id | event | event_ts`
- Saved as CSV & Parquet.

#### Validation:

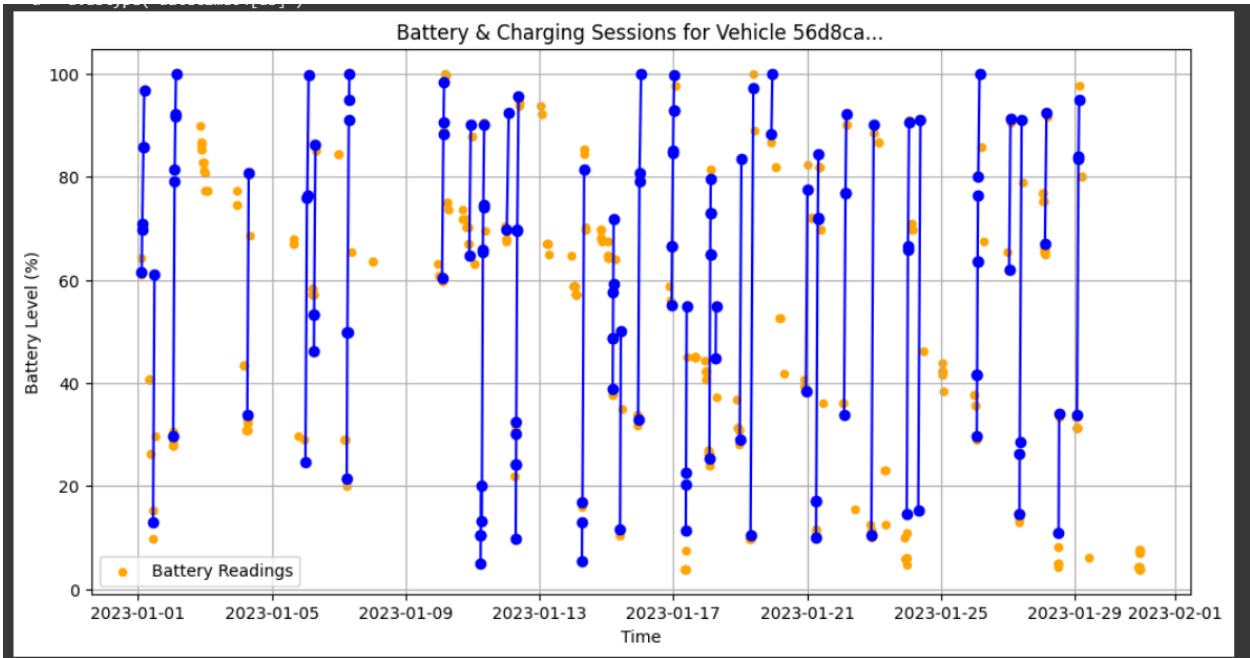
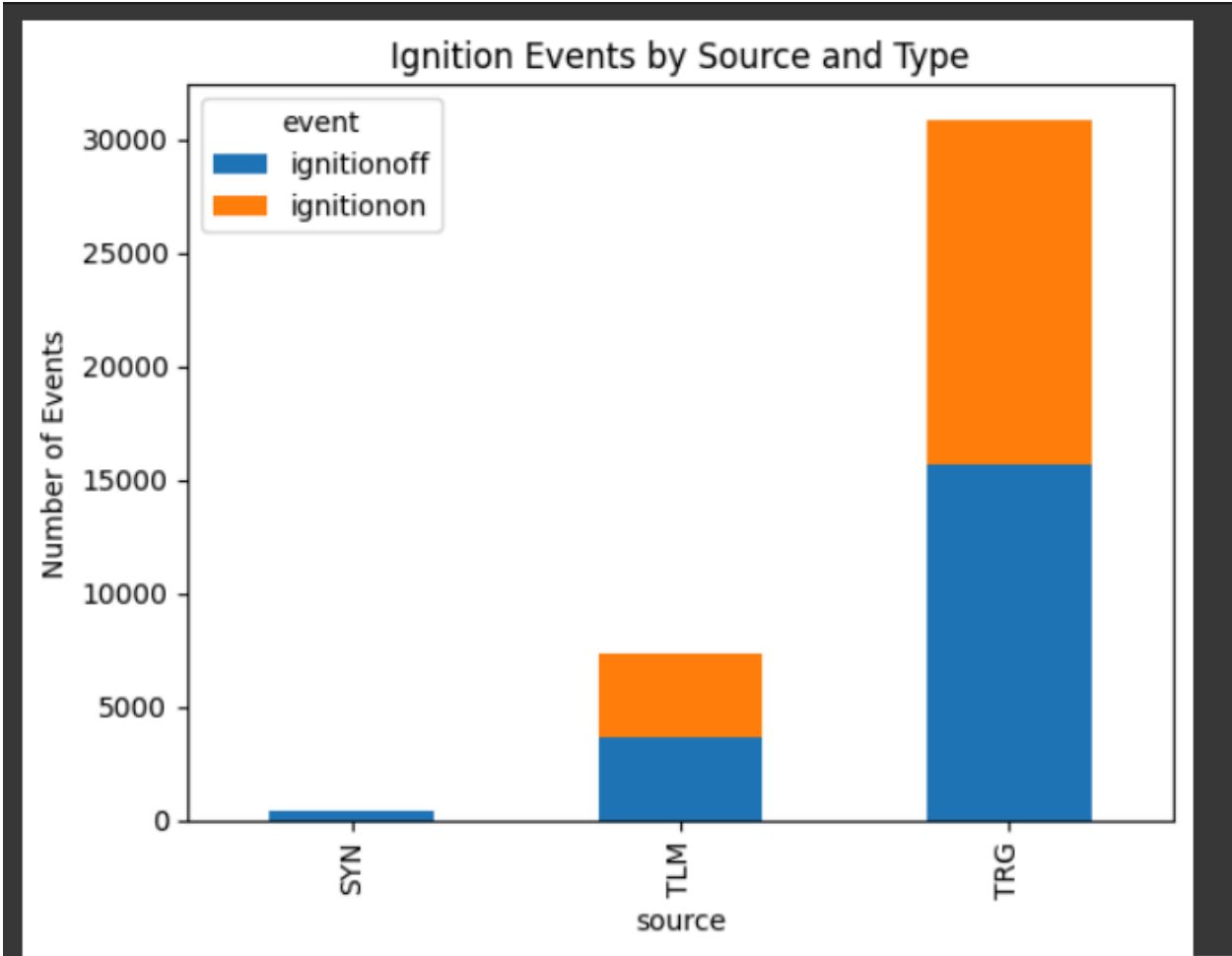
- Distribution balanced overall: ~17.9k OFF, ~17k ON.
- Visual timeline confirms alternating ON/OFF cycles, with SYN OFFs reinforcing critical shutdowns.

---

## ✓ Visualizations

- **Stacked bar chart:** Events by source and type, confirming SYN contributes only OFFs while TRG/TLM supply balanced ON/OFF cycles.
- **Sample vehicle timeline:** Alternating ON/OFF pattern over time; SYN OFFs appear later, ensuring shut-down coverage.

source	ignitionoff	ignitionon	Total
SYN	411	0	411
TLM	3693	3696	7389
TRG	15692	15188	30880
Total	19796	18884	38680



source

### Ignition Timeline for Vehicle 66bd55df...

