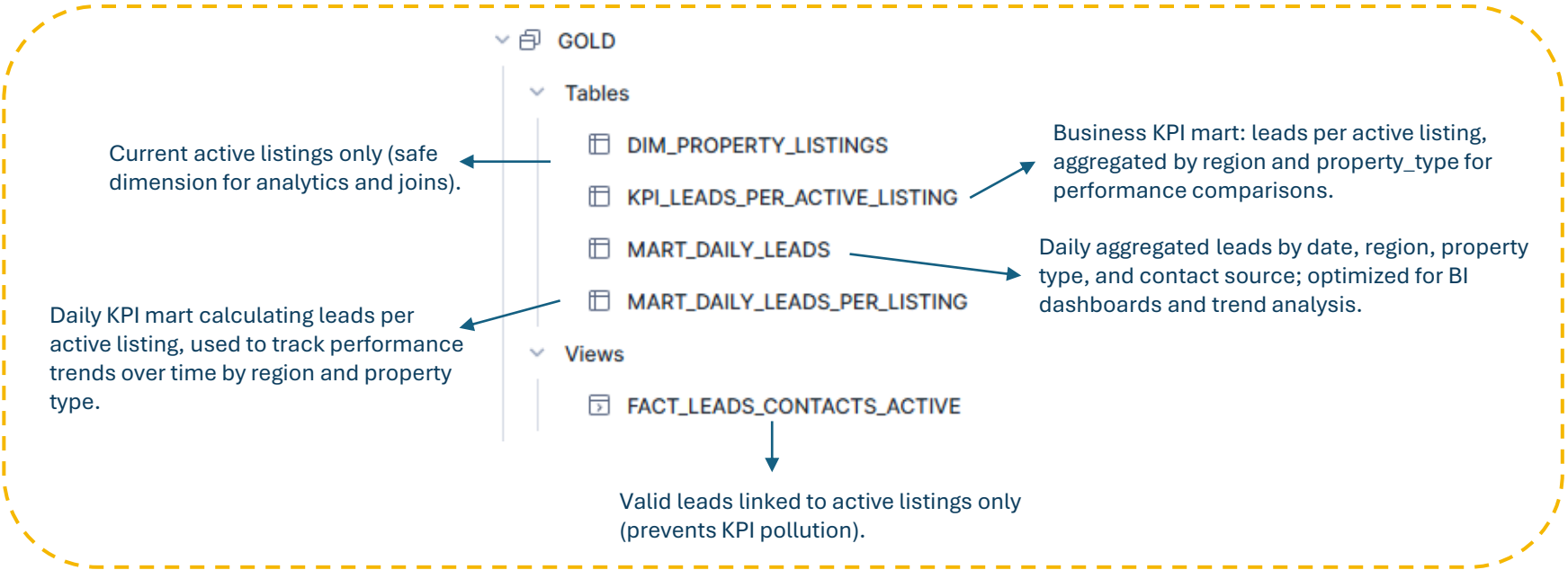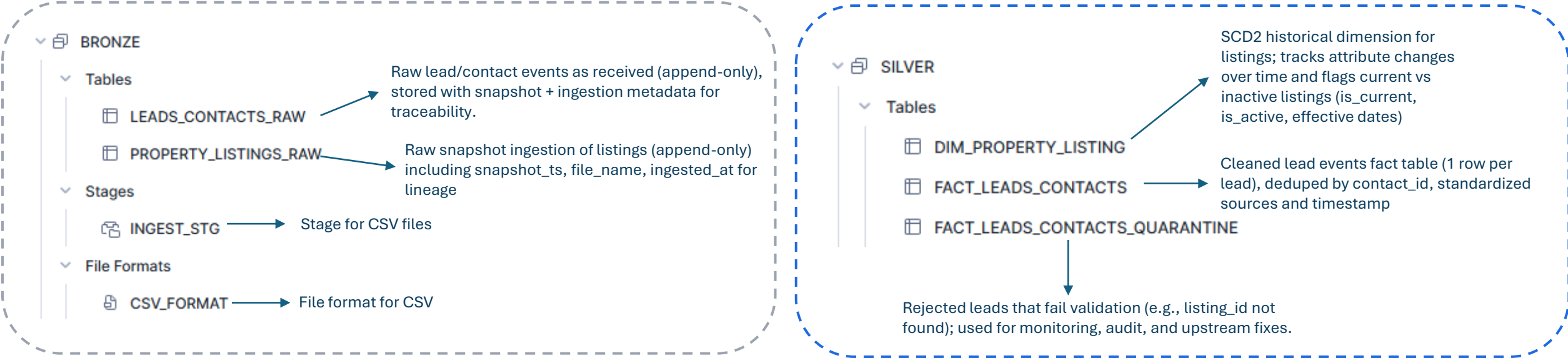# Data Platform Architecture

Full-snapshot ingestion • Historical tracking in Silver • Business-ready Gold KPIs

Ingestion CSV Snapshots

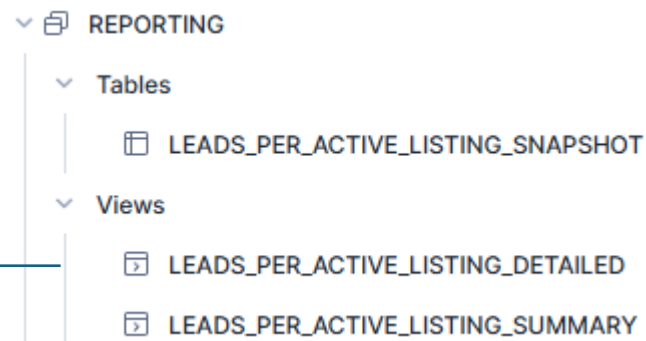- Full snapshot of active listings (can run multiple times/day)
- snapshot_ts derived from filename (YYYY_MM_DD_HHMMSS)
- Loaded via Stage + COPY INTO

Bronze

Raw tables

- Immutable raw tables (no updates, no deletes)
- Snapshot + ingestion metadata for lineage
- Enables deterministic reprocessing

Silver

Curated & historical

- Listings modeled as SCD Type 2 dimension (is_current, is_active, effective from / to)
- Leads modeled as fact events (one row per contact)
- Data quality enforced (invalid leads quarantined)

Gold

Business Ready

- Active only listings and leads
- Clean joins, no invalid references
- Precomputed KPIs / Marts

Reporting

- BI Dashboards (region, property type, channel)
- Ad-hoc SQL by analysts
- Stable inputs

Design Principles

- Full snapshots, not partial updates
- Raw data is never changed
- History is kept, nothing is overwritten
- Only validated data reaches reporting

Data Model & Key Tables - COMPUTATION

BRONZE

- **Tables**
  - LEADS_CONTACTS_RAW → Raw lead/contact events as received (append-only), stored with snapshot + ingestion metadata for traceability.
  - PROPERTY_LISTINGS_RAW → Raw snapshot ingestion of listings (append-only) including snapshot_ts, file_name, ingested_at for lineage
- **Stages**
  - INGEST_STG → Stage for CSV files
- **File Formats**
  - CSV_FORMAT → File format for CSV

SILVER

- **Tables**
  - DIM_PROPERTY_LISTING → SCD2 historical dimension for listings; tracks attribute changes over time and flags current vs inactive listings (is_current, is_active, effective dates)
  - FACT_LEADS_CONTACTS → Cleaned lead events fact table (1 row per lead), deduped by contact_id, standardized sources and timestamp
  - FACT_LEADS_CONTACTS_QUARANTINE → Rejected leads that fail validation (e.g., listing_id not found); used for monitoring, audit, and upstream fixes.

GOLD

- **Tables**
  - DIM_PROPERTY_LISTINGS → Current active listings only (safe dimension for analytics and joins).
  - KPI_LEADS_PER_ACTIVE_LISTING → Business KPI mart: leads per active listing, aggregated by region and property_type for performance comparisons.
  - MART_DAILY_LEADS → Daily aggregated leads by date, region, property type, and contact source; optimized for BI dashboards and trend analysis.
  - MART_DAILY_LEADS_PER_LISTING → Daily KPI mart calculating leads per active listing, used to track performance trends over time by region and property type.
- **Views**
  - FACT_LEADS_CONTACTS_ACTIVE → Valid leads linked to active listings only (prevents KPI pollution).

Data Model & Key Tables - CONSUMPTION

REPORTING

Tables

⊞ LEADS_PER_ACTIVE_LISTING_SNAPSHOT →

Point-in-time snapshot of the "Leads per Active Listing" KPI, stored as a table for stable reporting and historical comparisons.
Why it exists
•Ensures consistent numbers for reports
•Protects dashboards from upstream changes
•Useful for exports and month-end reporting
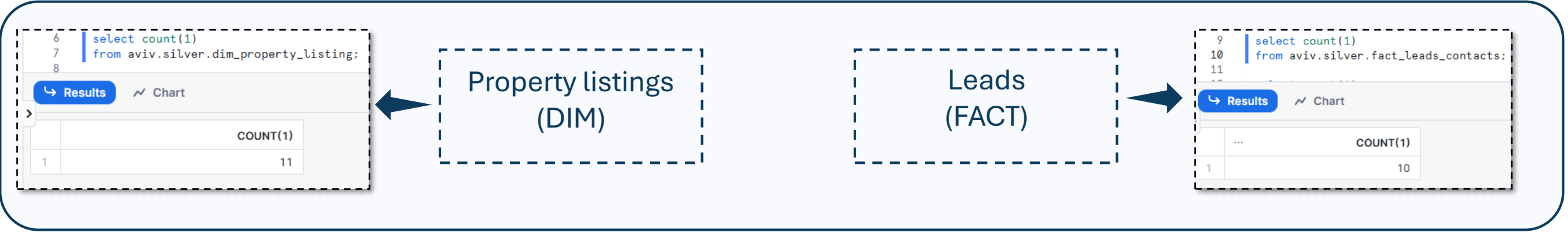
aviv.gold.kpi_leads_per_active_listing

Views

← ⊡ LEADS_PER_ACTIVE_LISTING_DETAILED

⊡ LEADS_PER_ACTIVE_LISTING_SUMMARY

Detailed view of the KPI with additional breakdowns (e.g. date, source), used for deeper analysis and drill-downs.
Typical use
•Analysts exploring trends
•Root-cause analysis

aviv.gold.mart_daily_leads_per_listing

Aggregated view of leads per active listing by region and property type, used for high-level dashboards.
Typical use
Executive or product-level reporting
Compare regions and property types

aviv.gold.kpi_leads_per_active_listing

Data Model & Key Tables – SCD Type 2

```
 6    select count(1)
 7    from aviv.silver.dim_property_listing;
 8
```
Results ∼ Chart

	COUNT(1)
1	11

Property listings (DIM)

Leads (FACT)

```
 9    select count(1)
10    from aviv.silver.fact_leads_contacts;
11
```
Results ∼ Chart

...	COUNT(1)
1	10

SCD2 Type – data updated

```
16    select listing_sk, listing_id, price, source_updated_at, effective_from, effective_to , is_current, is_active
17    from aviv.silver.dim_property_listing
18    where listing_id = 'L1004';
19
```
Results ∼ Chart

	LISTING_SK	LISTING_ID	PRICE	SOURCE_UPDATED_AT	EFFECTIVE_FROM	EFFECTIVE_TO	IS_CURRENT	IS_ACTIVE
1	6	L1004	410000	2025-01-01 10:20:00.000	2025-01-01 11:30:25.000	2025-01-02 11:30:25.000	FALSE	TRUE
2	101	L1004	400000	2025-01-02 11:20:00.000	2025-01-02 11:30:25.000	null	TRUE	TRUE

SCD2 Type – data removed from bronze

```
16    select listing_sk, listing_id, price, source_updated_at, effective_from, effective_to , is_current, is_active
17    from aviv.silver.dim_property_listing
18    where listing_id = 'L1009';
19
```
Results ∼ Chart

	LISTING_SK	LISTING_ID	PRICE	SOURCE_UPDATED_AT	EFFECTIVE_FROM	EFFECTIVE_TO	IS_CURRENT	IS_ACTIVE
1	3	L1009	240000	2025-01-01 13:00:00.000	2025-01-01 11:30:25.000	2025-12-12 05:09:03.086	FALSE	FALSE

Data Quality

```
21    select CONTACT_ID, LISTING_ID, INGESTED_AT
22    from aviv.silver.fact_leads_contacts_quarantine;
```
Results ∼ Chart

	CONTACT_ID	...	LISTING_ID	INGESTED_AT
1	C2010		L1009	2025-12-12 05:52:55.752
2	C2011		L9999	2025-12-12 05:52:55.752