

## **Real Estate Data Engineering + AI + Geospatial + Blockchain + Power BI**

### **PHASE 1 — DATA COLLECTION & CLEANING**

- Clean raw dataset, parse price, area, location.
- Remove duplicates, missing values.

Deliverables: Cleaned CSV, cleaning notebook.

### **PHASE 2 — SNOWFLAKE DATA PIPELINE**

- Database: RESIDENTIAL\_ANALYSIS\_DB
- Schemas: RAW → STAGING → CURATED → GOLD
- Load data via Stage + COPY INTO.

Deliverables: All Snowflake tables + SQL scripts.

### **PHASE 3 — METADATA LOGGING**

- DATA\_LOAD\_LOG tracks row counts, timestamps, load status.

Deliverables: Metadata table maintained automatically.

### **PHASE 4 — MACHINE LEARNING**

- Feature engineering + model training (RandomForest).
- Export model.pkl + predicted\_prices.csv.

Deliverables: Best model + predictions merged later.

### **PHASE 5 — GEOSPATIAL**

- geopy-based lat/lon extraction.
- Heatmap + price map saved as HTML.

Deliverables: Geo-enriched dataset + visual maps.

### **PHASE 6 — BLOCKCHAIN SMART CONTRACTS**

- NFT contract + Marketplace contract + Audit contract.

Deliverables: 3 Solidity contracts compiled.

### **PHASE 7 — TESTNET DEPLOYMENT**

- Deploy on Sepolia via MetaMask + Remix.

Deliverables: Contract addresses + ABI files.

### **PHASE 8 — FULL INTEGRATION**

- FINAL\_MASTER = Gold + AI + Geo combined.

Deliverables: Final integrated Snowflake table (14.5K rows).

### **PHASE 9 — WEB3 FRONTEND (DAPP)**

- HTML + CSS + JS frontend.

Deliverables: index.html + style.css + ABI integration.

### **PHASE 10 — POWER BI DASHBOARD**

- 5-page Snowflake-connected dashboard.

Deliverables: PBIX + PDF.

### **PHASE 11 — GITHUB RELEASE**

- Final folder structure + README.md

Deliverables: Complete GitHub repository.

**Final Output:**

Enterprise-grade, end-to-end Real Estate Intelligence + Blockchain Marketplace System.