Name

Aarzoo Mali

Project Name

Title Tracker: Real Estate Market + Sentiment Insights

GitHub Repository Link

https://github.com/Aarzoomali/sql_project/tree/main

Job Description

This project is inspired by the **Data Analytics Intern** role at **First American**, a Fortune 100 real estate company. The internship involves analyzing historical data, prioritizing business initiatives with analytics, and building dashboards to guide decisions. The role uses SQL, dashboard tools like Domo (comparable to Tableau/Power BI), and focuses on real estate trends — all directly mirrored in this project.

Why this job?

I'm passionate about data-driven decision-making in real estate and marketing. This role aligns with my goal to apply SQL and analytics skills in an industry setting. The opportunity to build dashboards and gain insight from real data supports my long-term interest in tech-enabled storytelling and business intelligence.

Problem

Business Problem: What real estate markets are heating up, and how does online sentiment correlate with price trends?
This problem is relevant to the job, since interns are expected to analyze market data and build visual tools that inform product strategy.
It's solvable with SQL for querying and for transforming Zillow data and Reddit sentiment scraped into AWS RDS.

Data Sources

API – Zillow Neighborhood Data API

- **Info:** Price trends, inventory levels, median listing prices
- **Source:** https://www.zillow.com/research/data/
- **Method:** Authenticated GET requests using Python `requests`
- **Relevance:** Tracks market pricing and demand trends

Web Scraping – Reddit (r/RealEstate)

- **Info:** User discussions and sentiment around buying/selling homes
- **Source:** https://www.reddit.com/r/RealEstate/
- **Method:** Python `BeautifulSoup` + `requests`
- **Relevance:** Captures public perception and market fear/confidence patterns

---

Solution

Data from Zillow API and Reddit will be extracted using Python and loaded into AWS RDS. Using a layered architecture (raw → staging → warehouse), I'll clean and model the data with QL.
I'll perform two key types of analysis per source:

- **Descriptive:** What cities are seeing the highest price growth?
- **Diagnostic:** Is online sentiment correlated with these price shifts?

Final results will be visualized with Tableau and presented in dashboards for stakeholders — exactly the kind of work expected in the job posting.