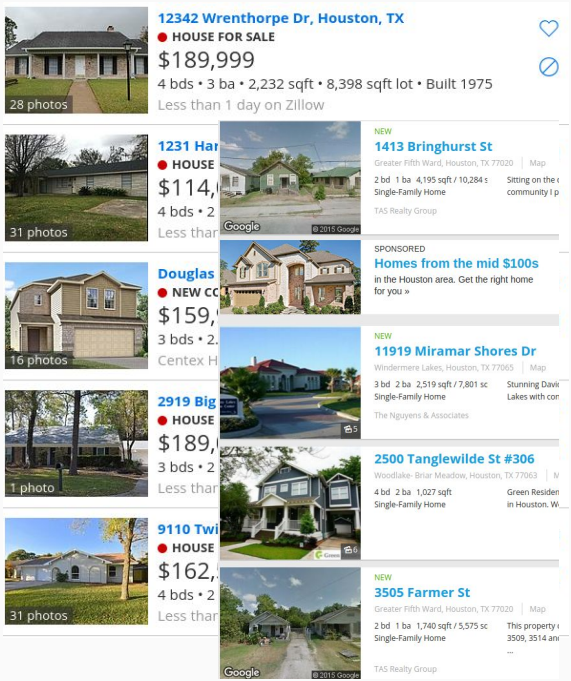
Intelli Invest

Spotting Tomorrow’s Good Investment Today

UCB MIDS W205, Dec 2015 Max & Aruna

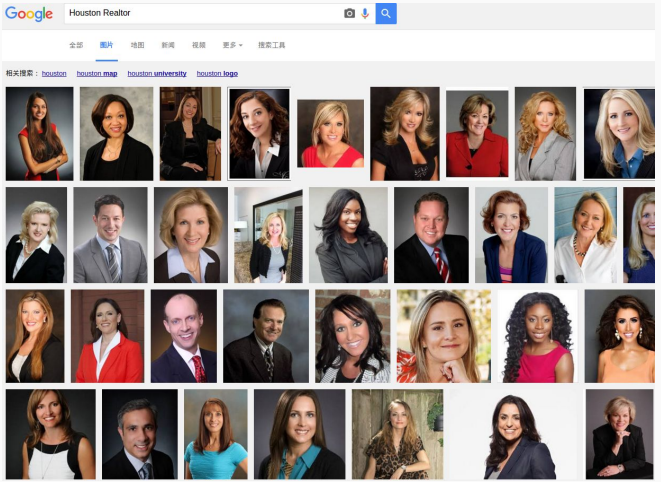
Let’s Start With a Story...



United States™ Census Bureau

America’s Best High Schools

CRIME STOPPERS
CALL 713.222.TIPS



Let’s Start With a Story...

?

What is the Problem?

3000+ Counties

40,000+ Zip Codes

1,000,000+ Listings

0 Time Series Data

0 School Data

0 Crime Data

0 Population / Growth Data

What is the Problem?

Too Big a Decision

Too Scattered the Data

Too Many Choices

Theories of Property Investment

- Focus on “Return on Investment”
- Consistency over time
- Population matters
- A place where people want to live

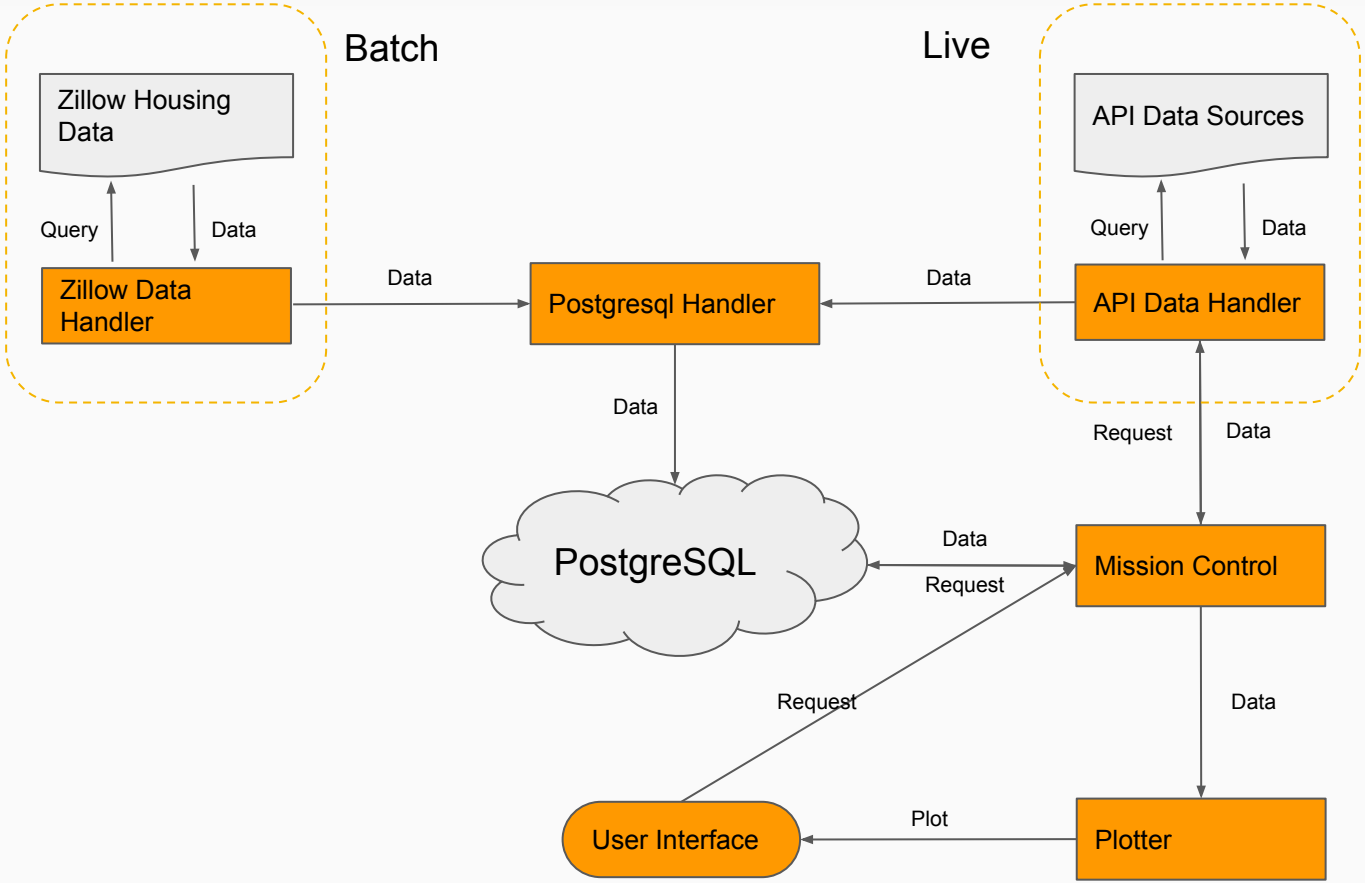
➢ Zillow Price and Rent Data

➢ Historical Price / Rent Data (1996 ~ 2015)

➢ US Census 2010 & Forecast 2013

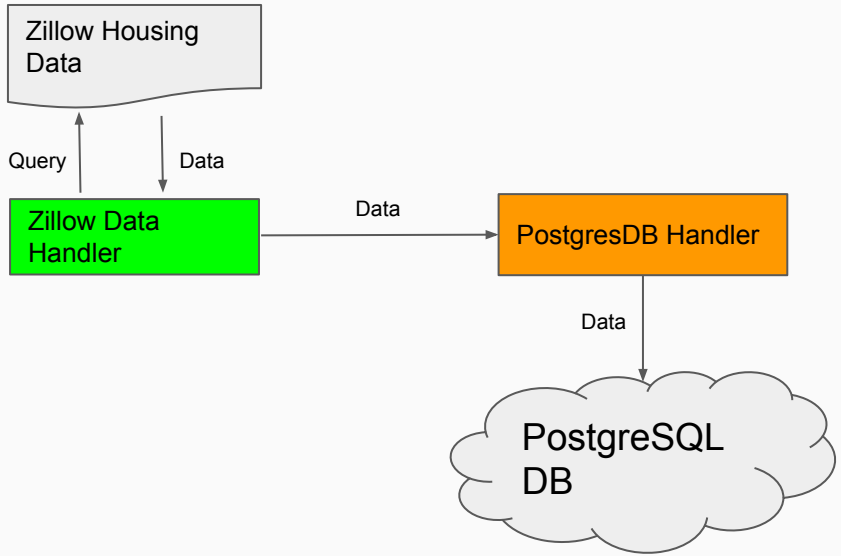
➢ GreatSchools.org, Walkscore etc.

Architecture



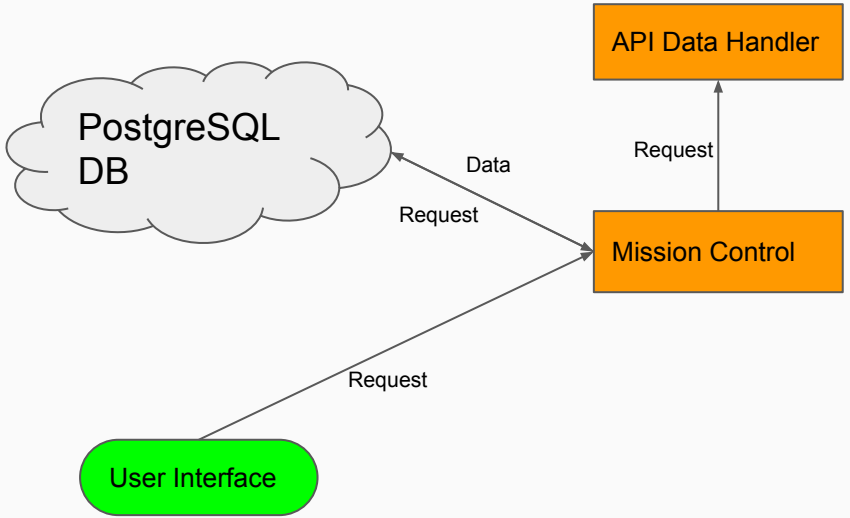
Zillow Data Handling

- Periodic Update (Monthly)
- Original Data in CSV on remote servers
- Size ~ Hundreds of MB
- Preprocessing required



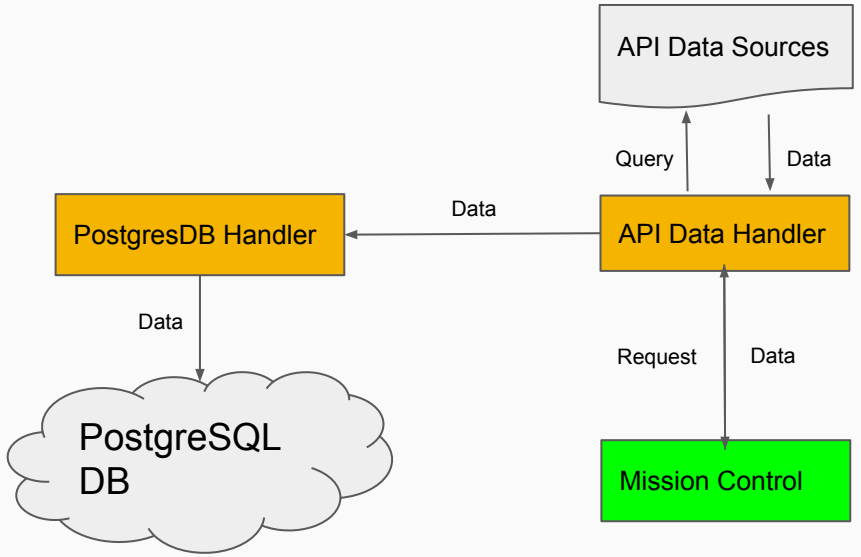
User Request Handling

- Completely reactive
- Always check required data in database first
- Triggers required API only needed

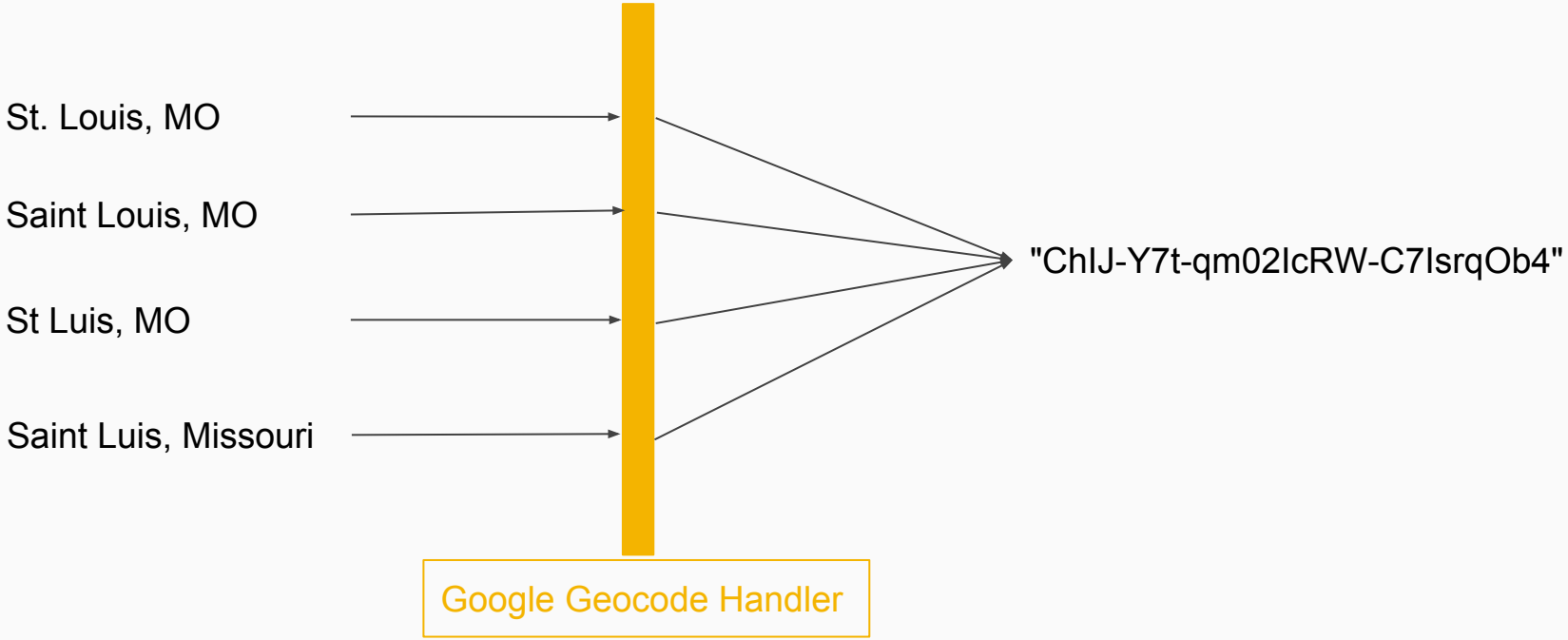


API Data Handling

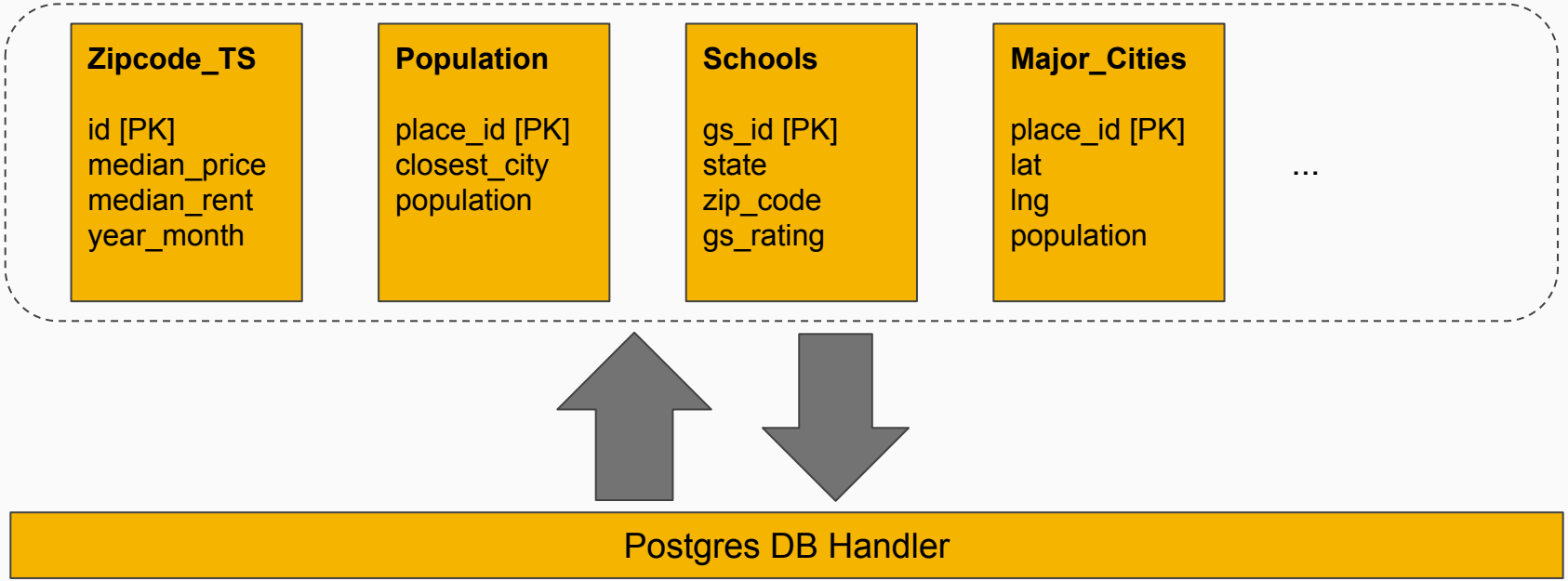
- On demand query
  - Always checks internal DB first
  - Always updates DB with API data
- Common API Data Handler layer for parsing
- Small in size, but big in variety



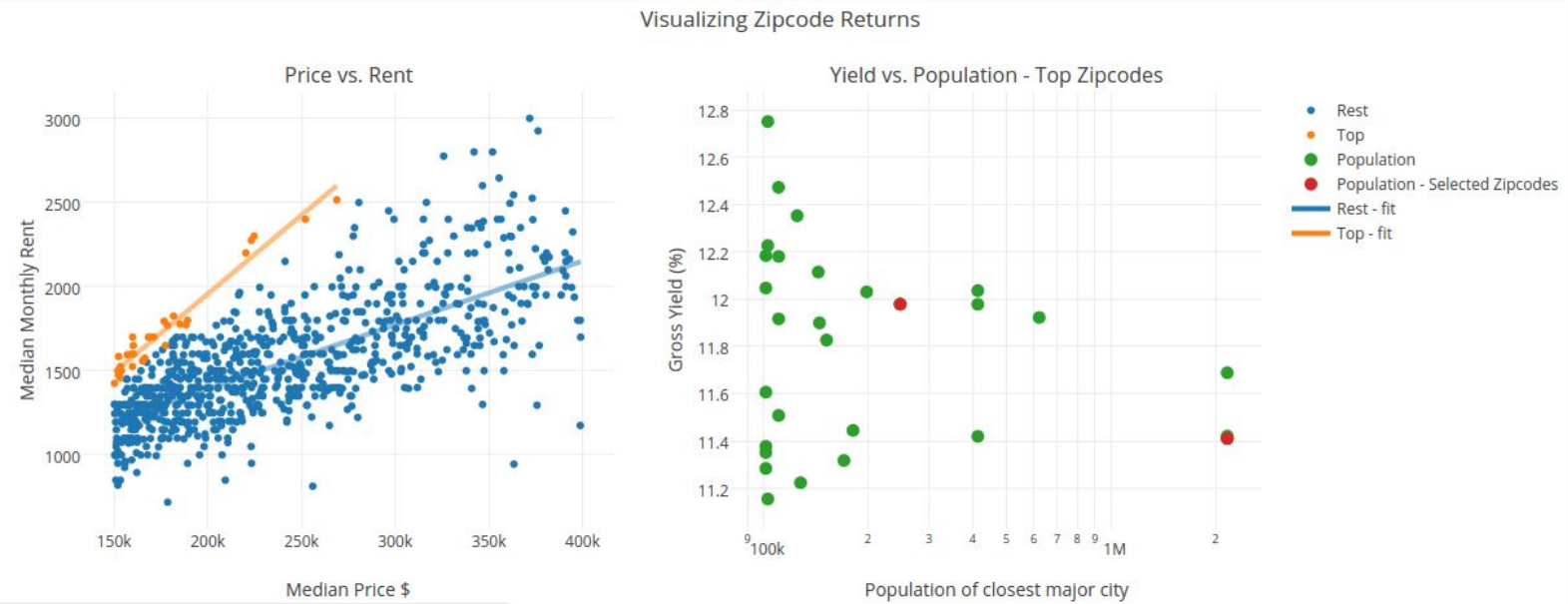
Overcoming Challenges



Database and Tables

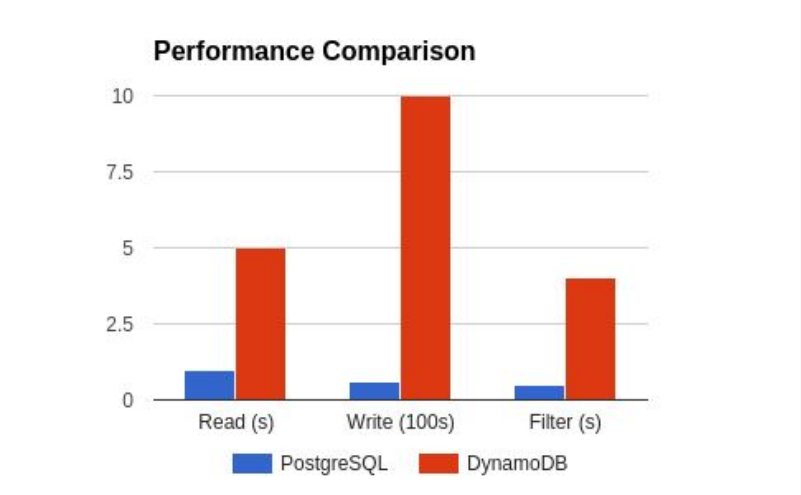


Shall We See Some Results?



Practical Considerations

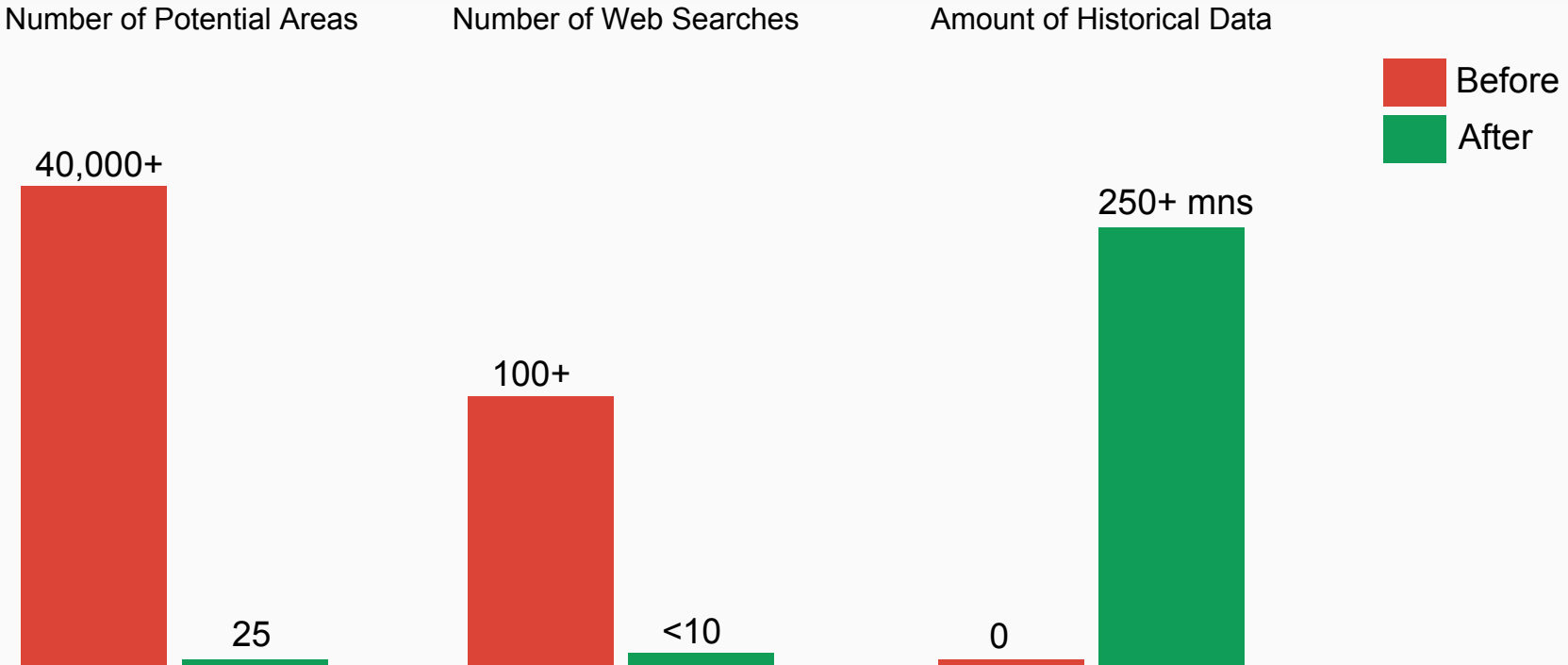
- Postgres vs. DynamoDb
- Mission Control / Plotter



Areas to Improve

Front End Portfolio Analysis	**API Handler** More Data Sources
ElastiCache Large scale concurrency	**DynamoDB** Less reliance on Schema

Practical Impact



Questions?