

Demand Prediction: EDA

12.08.2017



Nirmal Budhathoki

Garrett Cheung

Jillian Jarrett

Toby Moreno

Orysya Stus

Final Updates

- Locally Based APIs
- Documentation
- Benchmarking API methods
- Data Visualizations

Challenges

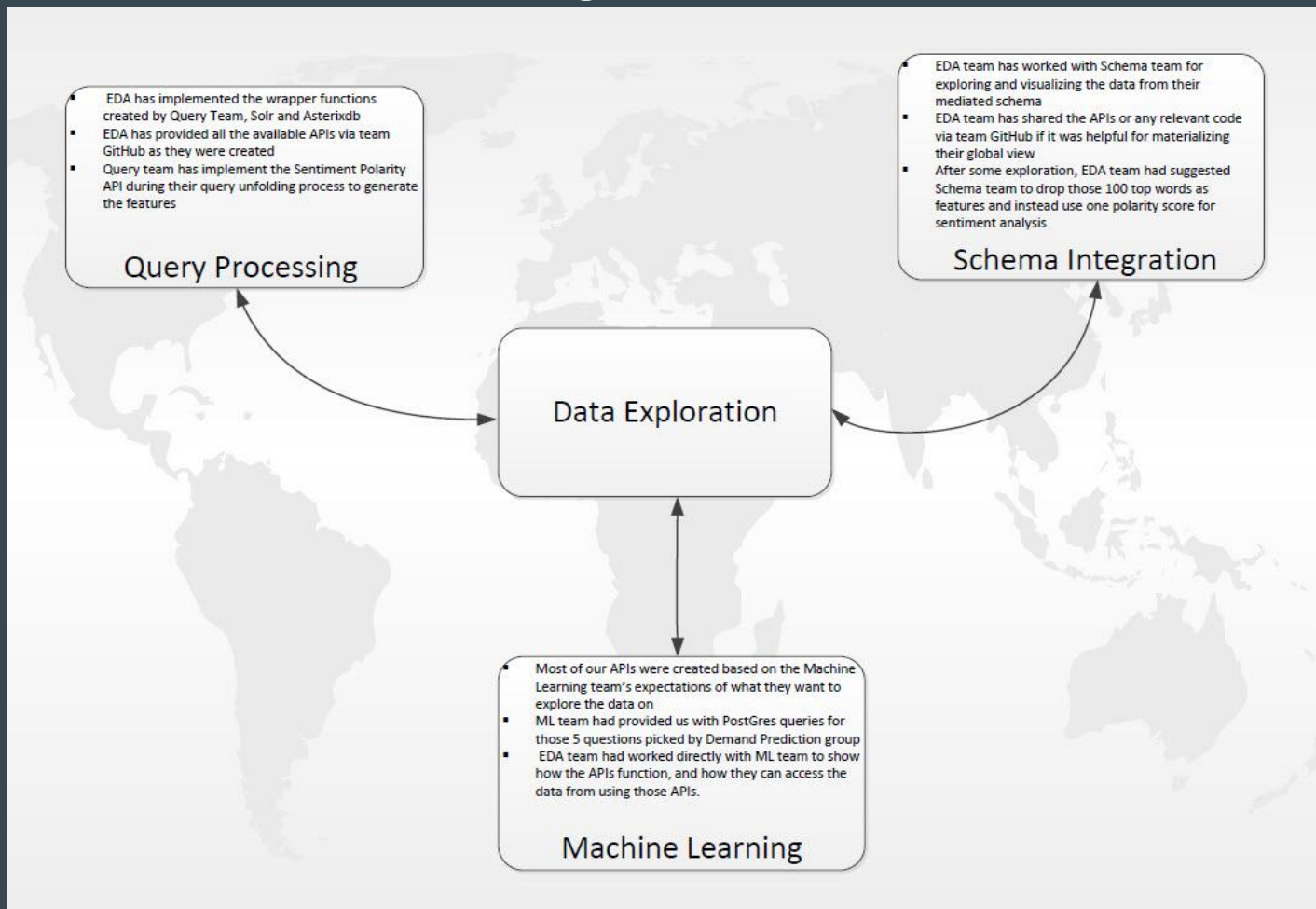
- Data Integration
 - Knowing the data sources, understanding the business objectives and deliverables.
- Data Locales
 - Data migrating from local computers to linode, and to UCSD servers
- Querying Solr
 - Solr wrapper provided by query team
- Working in parallel
 - Providing APIs for other teams as we create; team github; and weekly meetings

“If you are working on something for more than 5 hours, you are doing it wrong” -Prof. Gupta

TEAMWORK



Collaborating With Other Teams



API Documentation

- Used Sphinx to generate documentation for our API functions
- Allows for easy navigation
- Easily rewritable if docstrings or functions change
- Available on our github or through our REST API local server:

<http://localhost:5000/static/html/index.html>

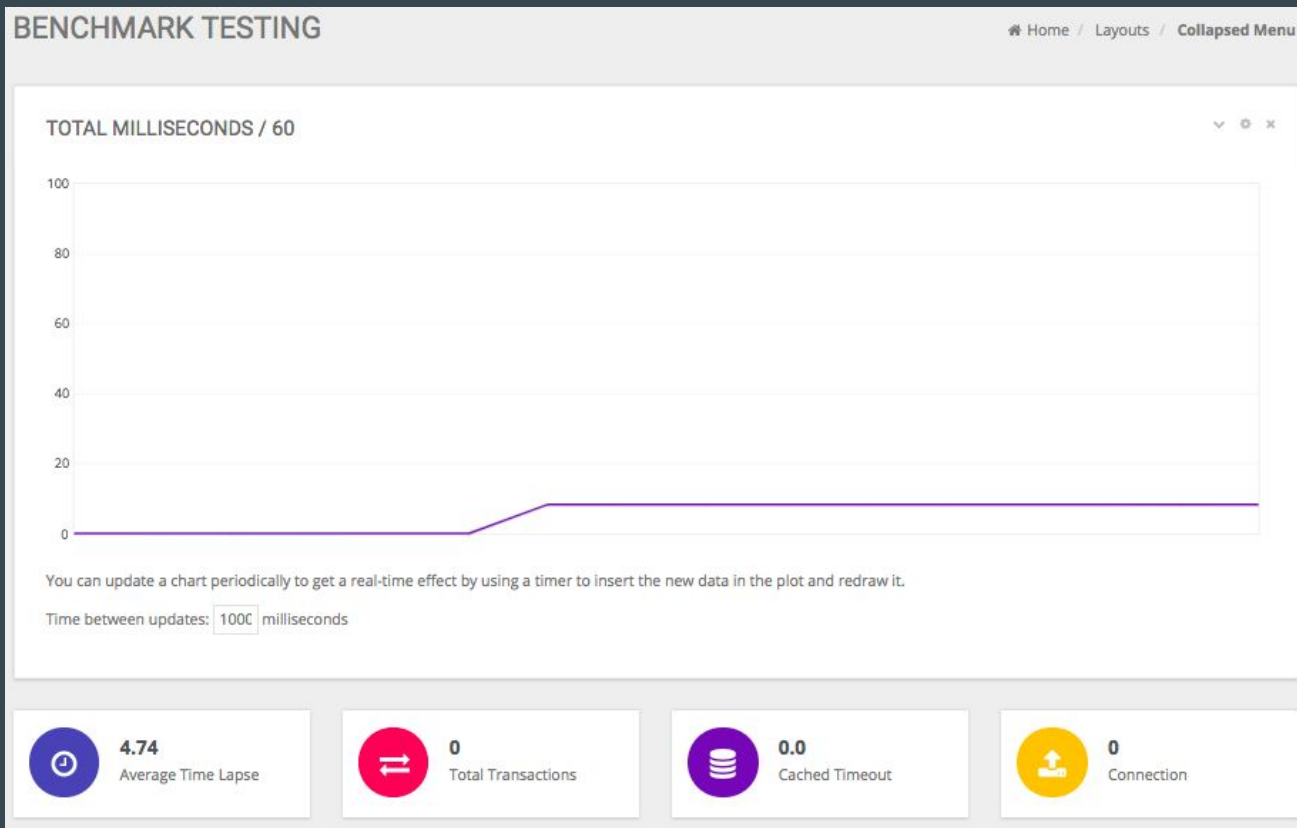
API DEMO

<http://localhost:5000/static/html/index.html> PYDOCS

<http://localhost:8888> API Sample Code

Benchmark and Optimization

http://localhost:5000/static/level1/benchmark_client.html



REST API Visualization

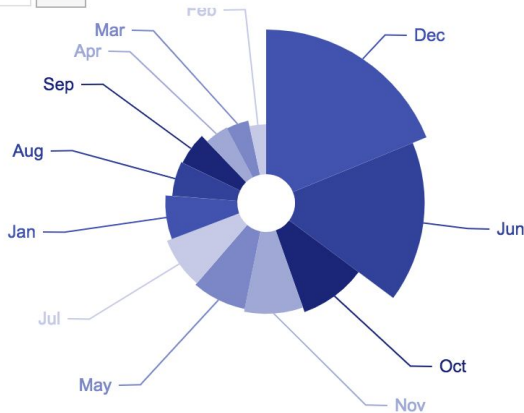
<http://localhost:5000/static/level1/SalesByCategory.html>

<http://localhost:5000/static/level1/SentimentPolarity.html>

SALES ANALYSIS

MONTHLY SALES BY CATEGORY

Programming 12 Go



SOLR REVIEW TEXT ANALYSIS

SENTIMENT POLARITY SCORE

Education 3 Go

NodeID	Sentiment Polarity	Score(0 to 100)	Strength
10633	0.2568417127	62.84	<div><div></div></div>
10680	0.2572916667	62.86	<div><div></div></div>
10687	0.2173214286	60.87	<div><div></div></div>
11764643011	0.13125	56.56	<div><div></div></div>
3048961	0.3302639417	66.51	<div><div></div></div>
3182	0.3184983921	65.92	<div><div></div></div>
3192	0.1802777778	59.01	<div><div></div></div>
355563011	0.2583333333	62.92	<div><div></div></div>
69829	-0.0006781046	49.97	<div><div></div></div>

Tableau Data Visualizations

NodeId-related data views

Tables utilized from mediated schema:

reviews_agg_yrmn

sales_agg_yrmm

https://public.tableau.com/profile/orysya.stus#!/vizhome/EDA_DemandPrediction/Demand

EDA | DEMAND ANALYSIS

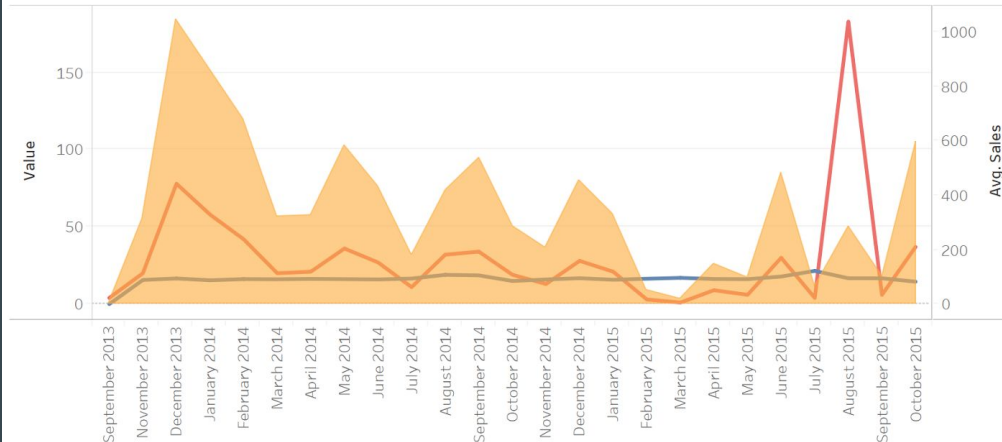
Instructions: Sort Node
Ids based on metric of
interest. Choose Node
Id to view sale-related
metric trends and
associated book
ratings.

Nodeid	Average Sales	Average Volume	Average Sales per Volume
226502	\$134.30	29.78	\$8.61
9432891011	\$2,290.35	29.74	\$67.98
713173011	\$352.53	29.72	\$14.58
16926946011	\$0.00	29.67	\$0.00
3351944011	\$4.95	29.63	\$4.17
16252731	\$27.80	29.00	\$4.63
8951155011	\$379.63	28.32	\$16.04

The selected NodeId: **713173011**

Sales Metrics

■ Avg. Avgprice
 ■ Avg. Sales
 ■ Avg. Vol



1
Total Book
Reviews



5.0
Average
Rating



THANK YOU