MICHIGAN STATE UNIVERSITY

Project Plan

Technology Peripheral Inventory Predictor

The Capstone Experience

Team Humana

Gabe Apaza

Siru Chen

Linda Duong

Katie Sydlik-Badgerow

Brendan Vande Kieft

Department of Computer Science and Engineering
Michigan State University

Spring 2019

From Students…
…to Professionals

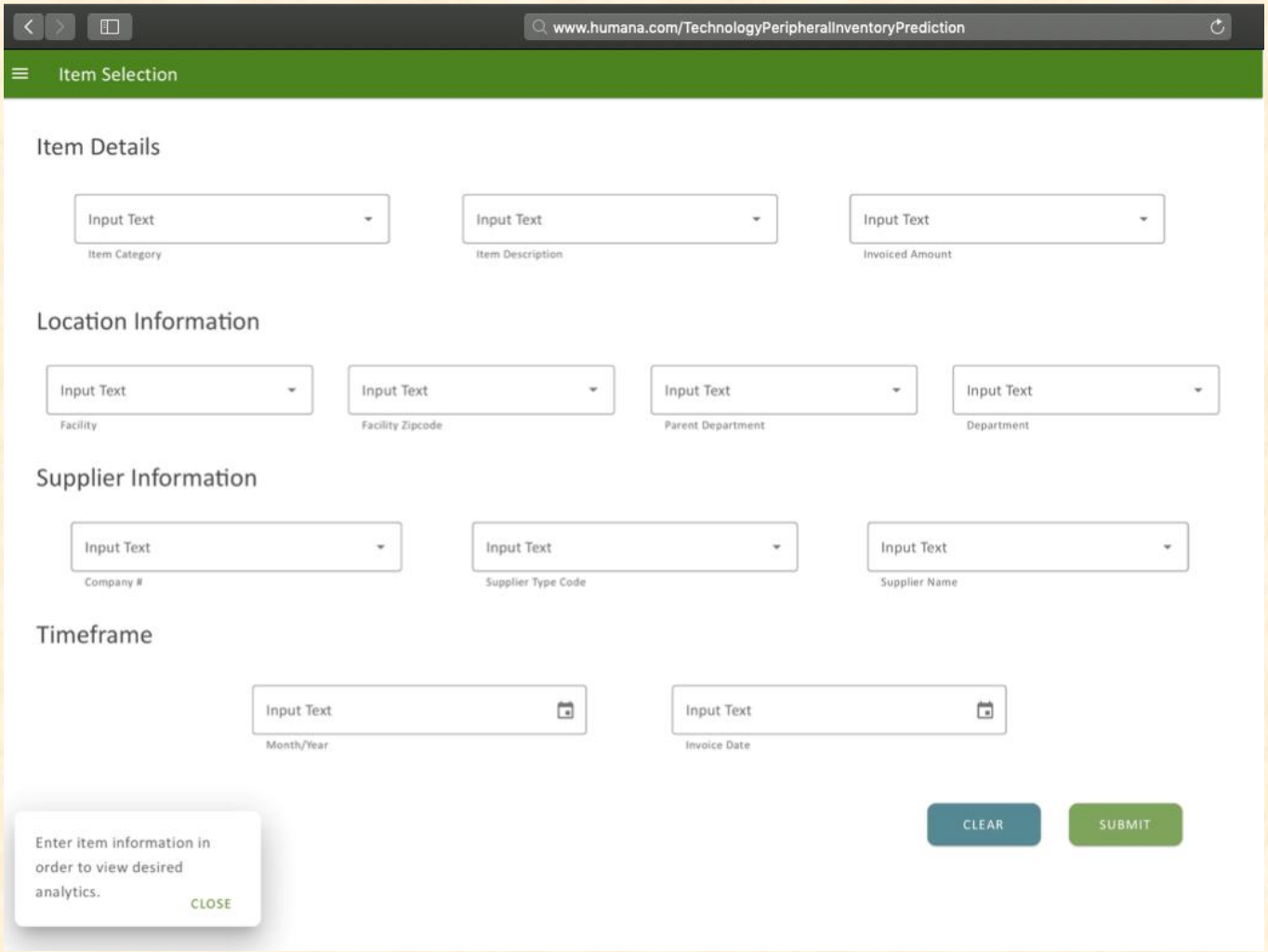
Functional Specifications

- Peripheral vending machines
- How to efficiently stock them?
- Machine learning to optimize purchases
- Metrics to inform purchasers of purchase history
- Cost analysis to see expected purchase cost

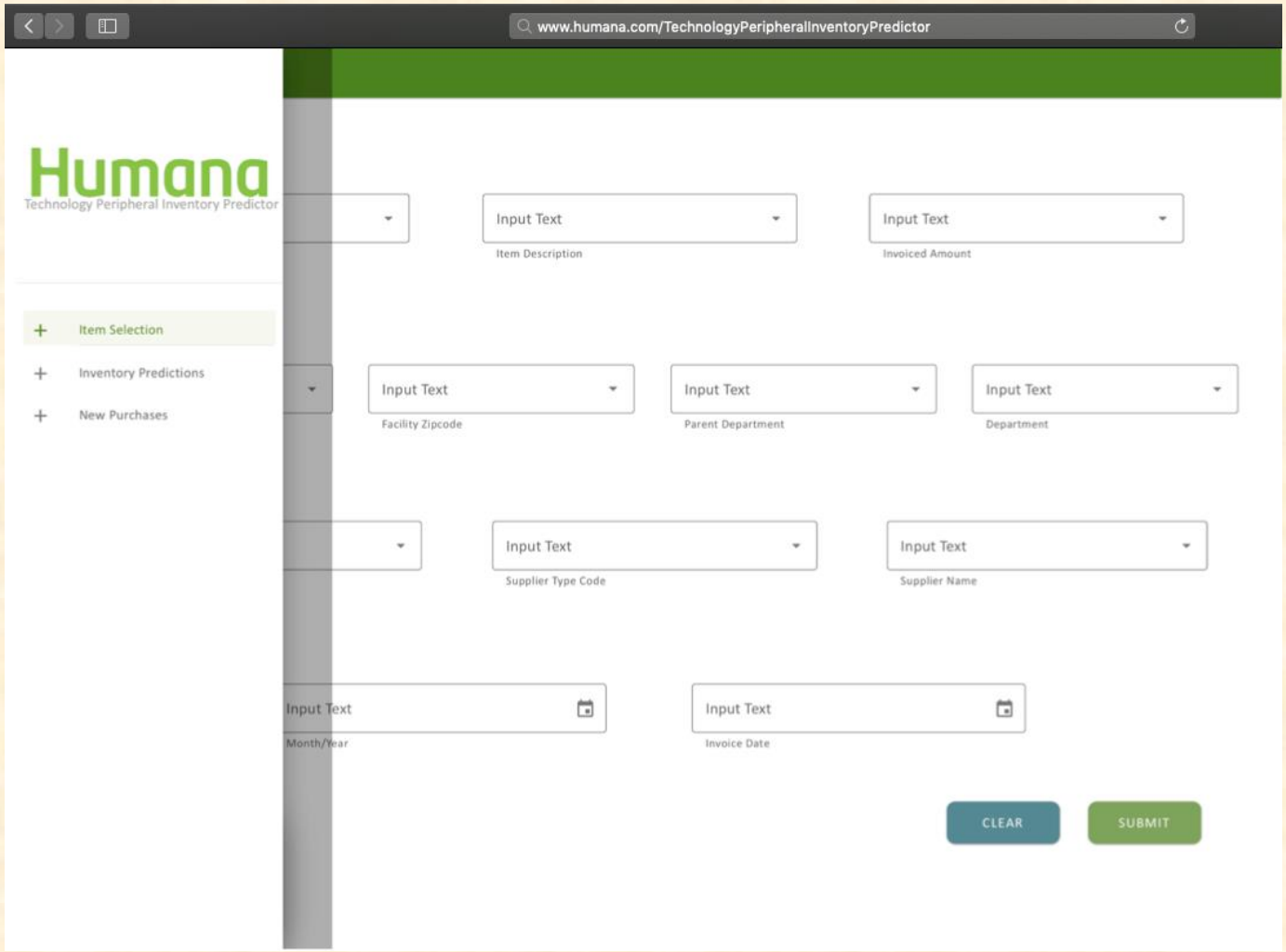
Design Specifications

- Web app split into frontend/backend
- Item selection interface to filter peripherals to only what the user wants to see
- Analysis/Prediction interface for seeing what the algorithm has predicted for future purchases
- New Purchase interface to incorporate new purchases into the system

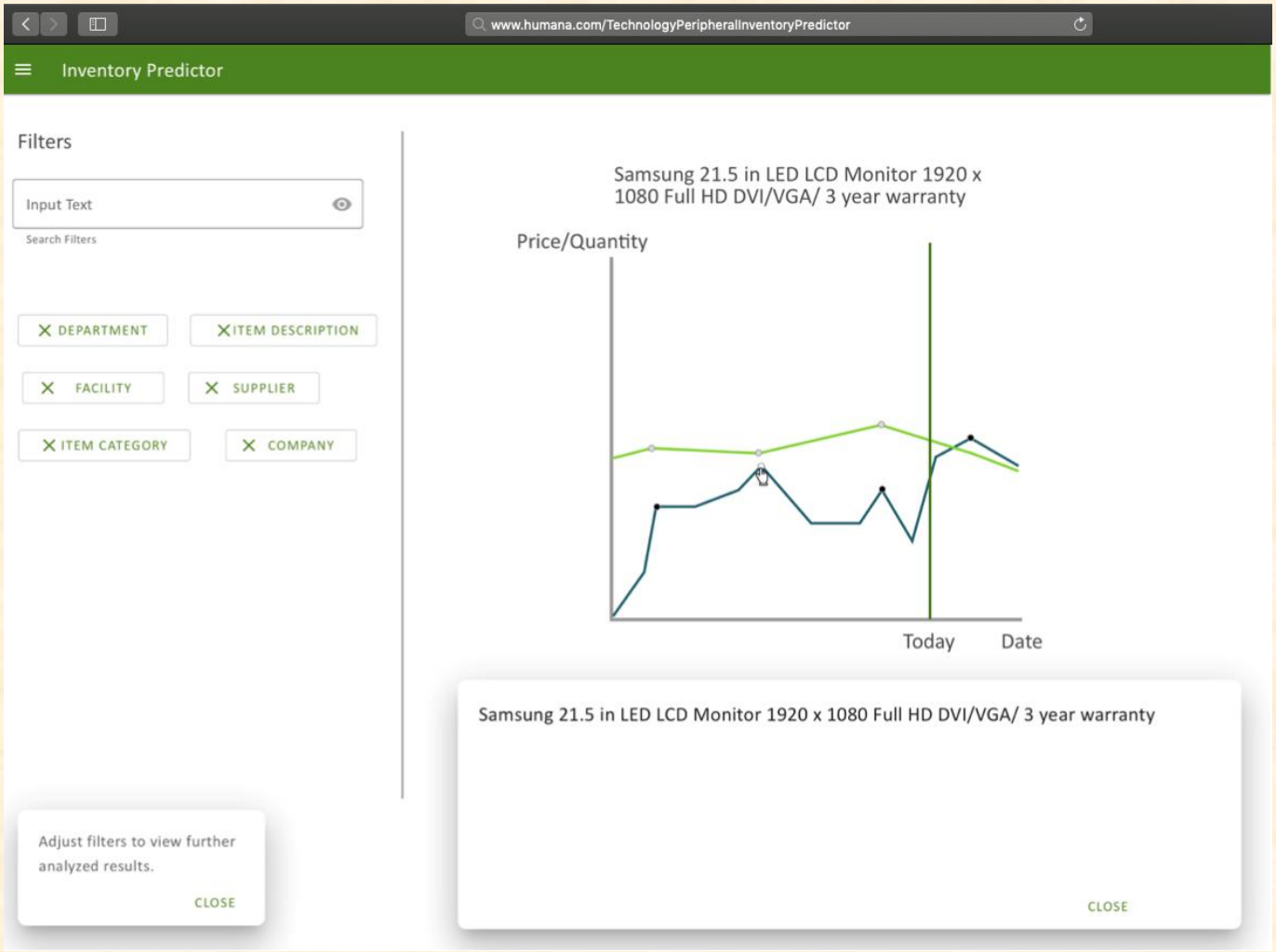
Screen Mockup: Item Selection



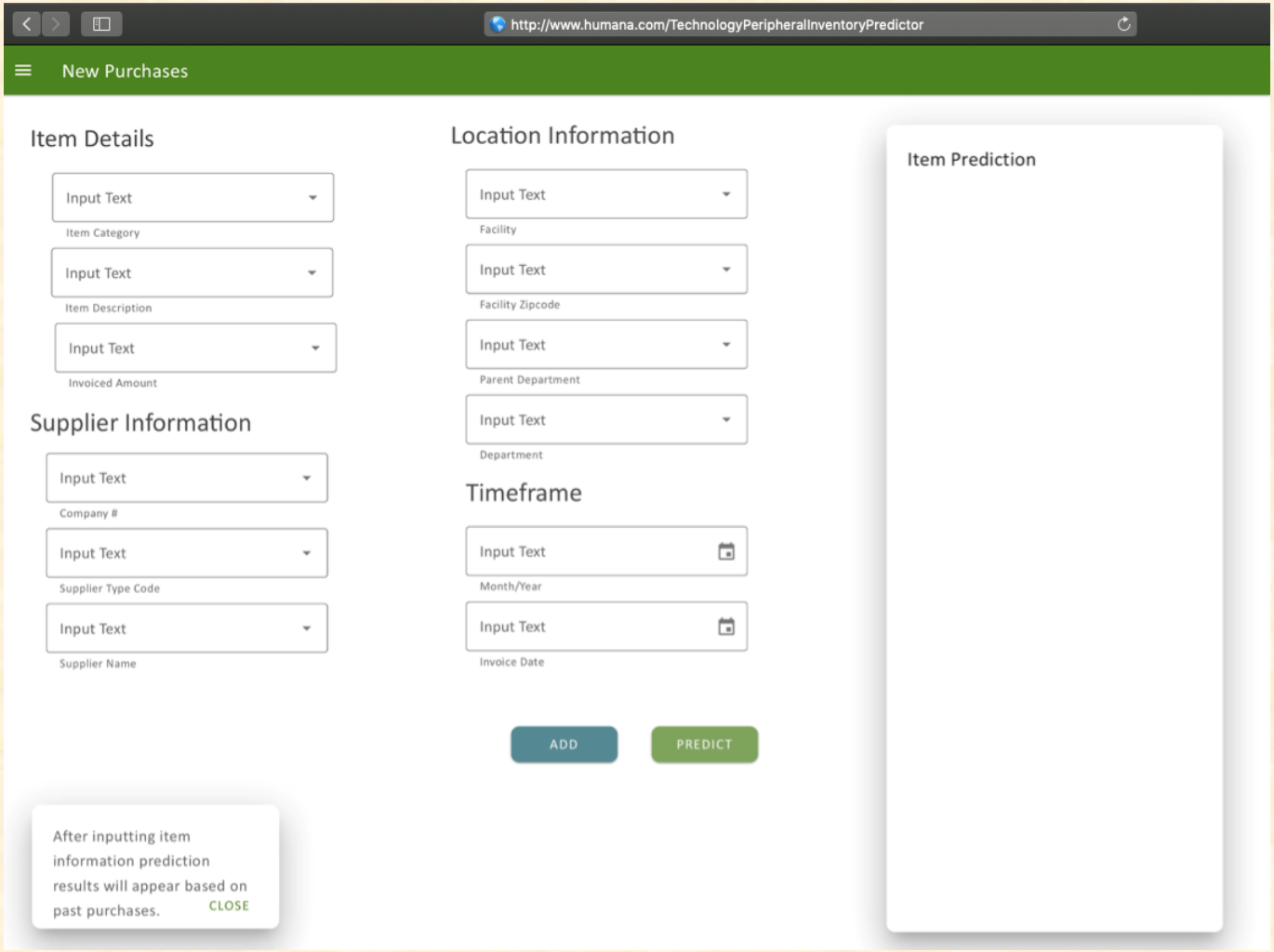
Screen Mockup: Sidebar Navigation



Screen Mockup: Predictor



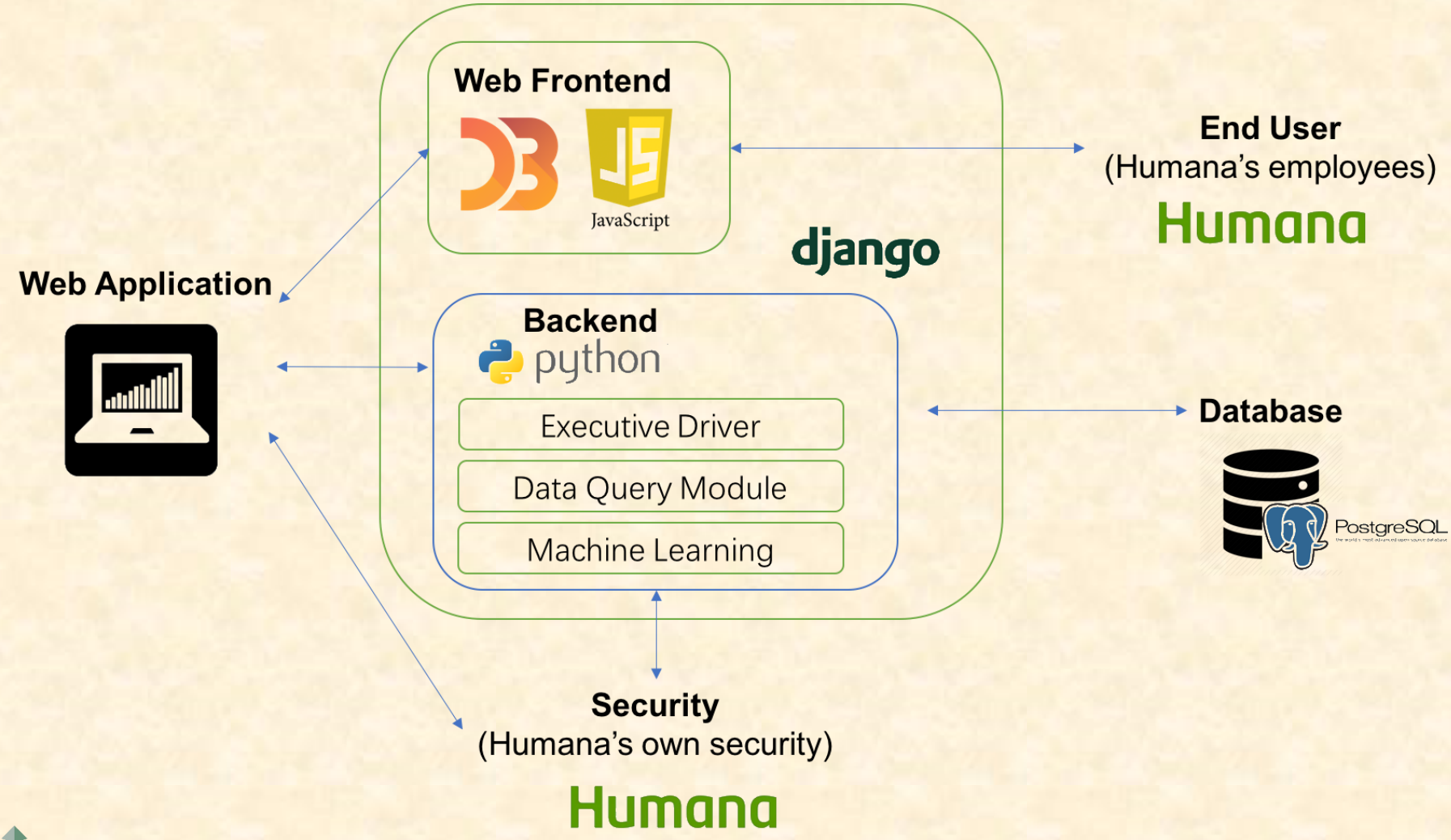
Screen Mockup: New Purchase



Technical Specifications

- Frontend and backend hosted on Django
- Backend connects to PostgreSQL database to access purchase history
- For machine learning, scikit-learn is used in the backend to create predictions
- The requested data is processed in the backend, then the response is visualized with D3.js

System Architecture



System Components

- Hardware Platforms
  - iMacs in lab
- Software Platforms / Technologies
  - Django for hosting web app
  - PostgreSQL for database
  - Scikit-learn for machine learning
  - D3.js for data visualization
  - Material.io for UI

Risks

- Prediction Format
  - Past purchase history cannot predict future demand; it only identifies past trends and extrapolates.
  - Discuss with client; present multiple ways of viewing the data and select whichever one is preferred.
- New Purchases
  - The format in which future Humana employees would like to upload bulk purchase history is unknown.
  - Create a flexible importer for bulk data, and discuss with client to find suitable data format
- Online Machine Learning
  - System must incorporate new data into predictions; our team has no experience with this
  - Discuss the technique with others who are knowledgeable; do research into which frameworks are best suited for the task
- Data Quality
  - The given data may have some duplicate entries and some peripherals are misspelled. This can significantly affect the performance of the algorithm.
  - Explore ways in which the data may be cleaned; isolate the low-quality parts of the data and see if they negatively affect the machine learning algorithm performance.

Questions?