Authors: Sam Chu, Jimmy Le, Kevin Lee, Natasha Lee, Sean Shelton
Team Name: Bob's Builders
Project Name: Kombucha

Introduction

Problem:
Currently, Procore collects terabytes of user data & analytics. However, none of that important information is visible to the client.

Project Specifics

- Data Visualization Tool: Apache Superset
- Database: PostgreSQL
- Web Hosting Service: Heroku
- Languages: Python, Javascript, SQL
- Work management tools: Trello
- Version Control: Git/Github
- Communication: Slack, Facebook Messenger, GoToMeeting

Team Goals/Objectives

Use Apache Superset to transform user data into charts and graphs
Superset allows us to take in data from our database (PostgreSQL) and create data visualizations. This allows our users to get a better understanding of how their company is making use of Procore's tools, and is central to our targeted consumer.

Populate dashboards with relevant charts and graphs specific to a user role
By giving specific charts to certain users, users will not be as overwhelmed by an over-complicated dashboard. They'll be able to view only the information that they need in a clear, concise manner.

Create a query system to allow users to make their own dashboards
Users have large databases with hundreds of ways they can query their company data, but not all users will be interested in the same analytics. Thus it is important to allow users to customize their queries to make their own dashboards.

Create a way to share dashboards (via URL, CSV, or PDF)
Users should be able to share their usage analytics with others involved with their project. This will allow our users to add charts and visuals to presentations, bring them to meetings, etc. to have better communication overall with their stakeholders.

Host dashboards online using Heroku

By having dashboards online, it allows the clients to easily access the dashboards/information whenever they want to see a dashboard instead of having to ask Procore to take a screenshot or send files to them.

Allow admins to control permission access to specific data
By granting admins the ability to assign granular permissions to specific people or roles it allows the company to control information flow and keep certain confidential information secure.

Connect Procore's data pipeline to a database in order to display info in real time
Having real time info will allow companies to make informed decisions that are based on current, not outdated data.

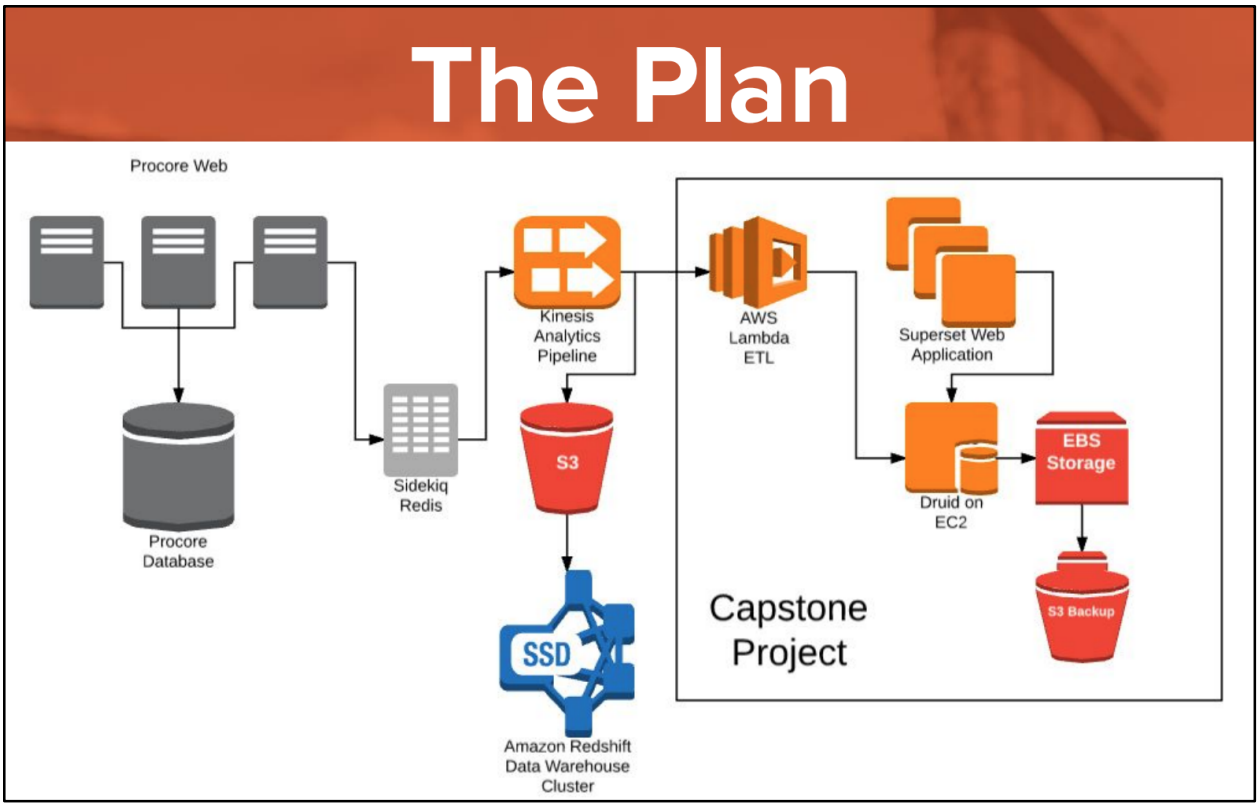
- Background
    - **What problem the project is solving?**
      Procore collects terabytes of data on the usage of their tools. Currently, clients don't have a way to view this data. In other words, they don't know which tools are being utilized in their organization. Thus, we want to innovate the way data is organized and redirected to the customer so that information can be ingested intuitively and quickly.

    - **Why the problem is important?**
      Without utilizing this data, clients have no idea which tools their organization or company is using. Knowing what tools are being utilized and what tools aren't is essential in understanding the value of Procore's products.

    - **How the problem is solved today?**
      Currently the problem is solved by organizing this data and presenting it to clients in a way that lets them quickly pinpoint trends and important points so that they can make decisions going forward based on that data. However, this is inefficient because the clients have to consult Procore each time they want to get analytics on their product usage.

- Assumptions
    - The user has an internet connection to view the webpage
    - The user is a customer of Procore and has Procore log-in credentials
    - The user will have a managerial or administrative position in the company
    - Procore will provide the user data in JSON format for us to ingest

System Architecture Overview



Requirements

User Stories / Use Cases

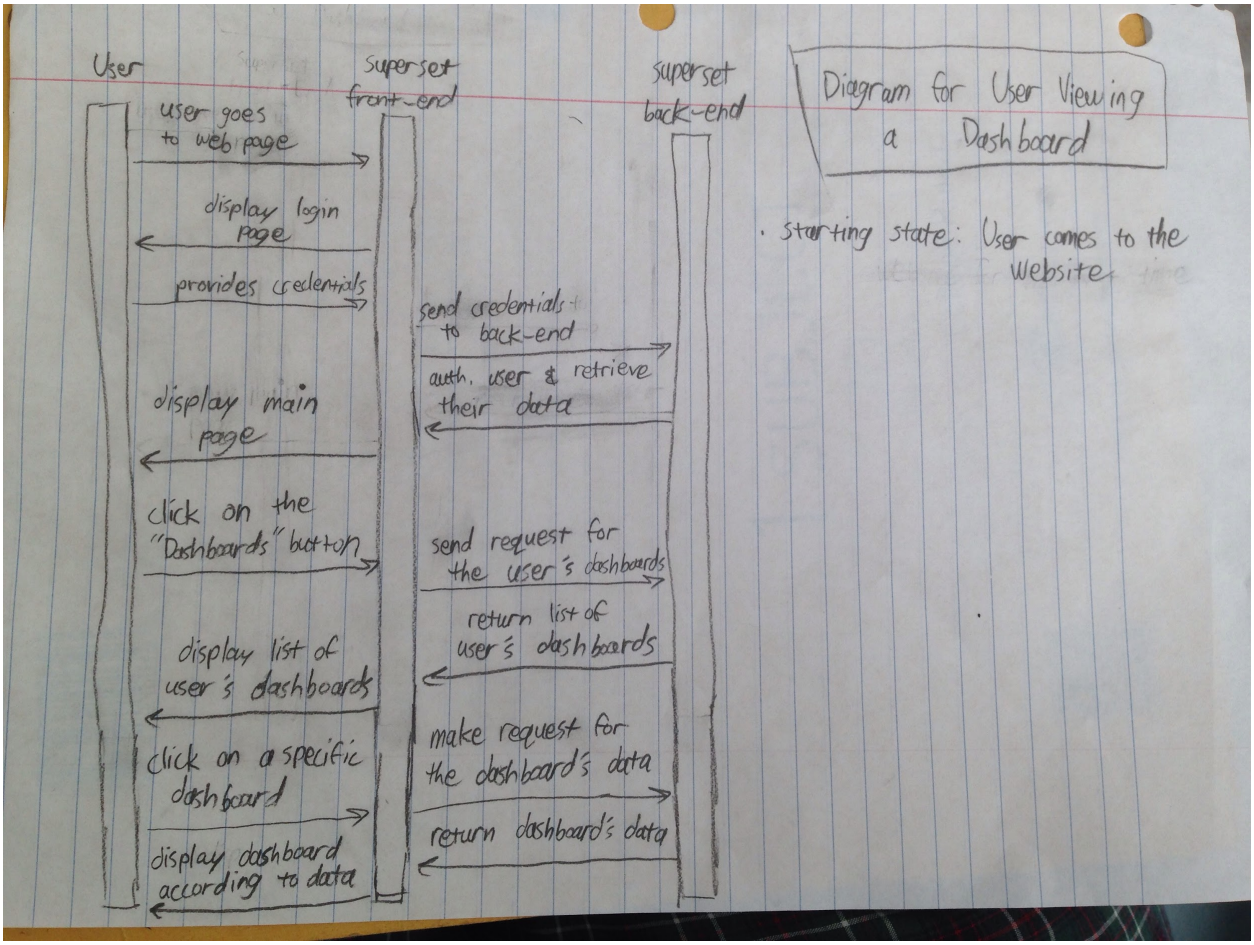
1. As an admin, I want to view my organization's usage of Procore's tool suite to learn how much value I'm getting out of the software.
    - **Acceptance Test:** We can visit the website and see tool usage statistics displayed on the site in a visually intuitive way.
    - https://github.com/procore/superset/tree/json_to_posgresql
2. As an admin, I want to see which tools are NOT being used to investigate / take action to increase usage.
    - **Acceptance Test:** We can visit the website and see a chart of the least used tools under the organization.
3. As an admin, I want easy access to dashboards so that I can view charts relevant to my organization.
    - **Acceptance Test:** Open the webpage and see one single dashboard that contains multiple graphs and charts.
4. As an admin, I want to control the type of data that my employees can see so that no one has access to data that they shouldn't be able to see.
    - **Acceptance Test:** Create two "dummy users" (User A and User B) with differing roles, give access to User A but not User B, then check that User A can access the data and User B cannot.

5. As an admin, I want to create my own graphs and charts so that I can have full control over what graphs are in my dashboards and how they're laid out.
    - **Acceptance Test:** Create a "dummy dashboard" with a set of charts, then check that the dashboard is saved with all customizations.
6. As an admin, I want to be able to share my dashboards so others can see the graphs I've created.
    - **Acceptance Test:** Create a "dummy dashboard," use our Share feature, then check whether or not the same exact dashboard is shown to the person that the dashboard was shared to. (by checking properties, visuals, etc.)
7. As an admin, I want to be able to clone dashboards and charts in order to make slightly different iterations easily.
    - **Acceptance Test:** Create dummy dashboards and charts, use our "clone" function, then check whether or not the dashboards and charts created by the clone function are equivalent to the original. (e.g. check properties, check visuals, etc.)
8. As an admin, I want to be able to change the time period of my charts so that I can view recent as well as long term trends.
    - **Acceptance Test:** Go to a dashboard and choose a chart. Then click on the edit button for the chart and change the time period I want to show. Go back to the dashboard to see if the chart was updated to the new time period I chose.
9. As an admin, I can login with my Procore credentials and have my data pulled from my account.
    - **Acceptance Test:** Create a mock User with a fake username/credentials. Check that they are able to login successfully into the Procore ecosystem.
10. As an admin, I want to access to navigate to the webpage and still feel like I'm within the Procore ecosystem.
    - **Acceptance Test:** Design a webpage that follows Procore's official design guidelines.
    - https://github.com/procore/superset/pull/2
11. As a data analyst, I can use the data visualizations in order to easily create better data-driven decisions.
    - **Acceptance Test:** Make comparisons between the charts generated by our software and charts generated by other software in order to see which one has better clarity
12. As a Procore developer, I want to easily add or remove specific data that gets put into Postgresql for clients to see so that I don't need to dive into code in order to control data flow.
    - **Acceptance Test:** Have a text file that lists out the metrics to import into PostgreSQL. The code reads from this file to decide what gets imported.
13. As a user, I want to be able to access Procore's data online without needing to download any extra software or download any data locally so that access is easy and convenient.
    - **Acceptance Test:** We can access Superset online and see all the data on the Heroku instance.

14. As an admin, I want to be able to update my view of all the graphs at once so that I don't need to update each graph individually if I want to make a common change to all of them.
    - **Acceptance Test:** We have a component on a dashboard that is capable of updating all the graphs at once.
15. As a user, I want to be able to view drop-down menu options in a visually appealing, intuitive way so that I can quickly make options
    - **Acceptance Test:** Make a graph (without SQL Lab) and open up the drop-down menus in the sidebar. See that the options are written in plain English rather than something like "TOT_SUM_AVG_ACTION_NAME"
16. As a user, I want to see some documentation that describes each available metric so that I can understand the data that is available to me.
    - **Acceptance Test:** Have access to documentation somewhere within the Superset webpage to read and learn about what type of metrics and data points Procore provides within Superset.
17. As a user, I want to see a set of default charts when I open Superset for the first time so that I get an idea of what kind of charts can be made and how they're done.
    - **Acceptance Test:** Log into Superset and see a set of default charts that have already been created.
18. As a user, I want to be able to customize the location of my charts easily by dragging them to the location I want.
    - **Acceptance Test**: Go to a dashboard and be able to click and drag a chart to a different location.
19. As a Procore developer, I want to stream data directly into PostgreSQL using a data streaming tool such as Kafka or AWS Kinesis/Lambda so that I can see live updates to my data.
    - **Acceptance Test**: See automatic, live updates to Procore data within the database through a Kinesis/Lambda or Kafka data stream.
20. As a user, I want to search the dashboards by name to I can quickly find the dashboard I want.
    - **Acceptance Test**: Go the the dashboard list page and type in a dashboard name in a search bar and have the table filter the dashboards.

System Models

- Contexts
- Sequences



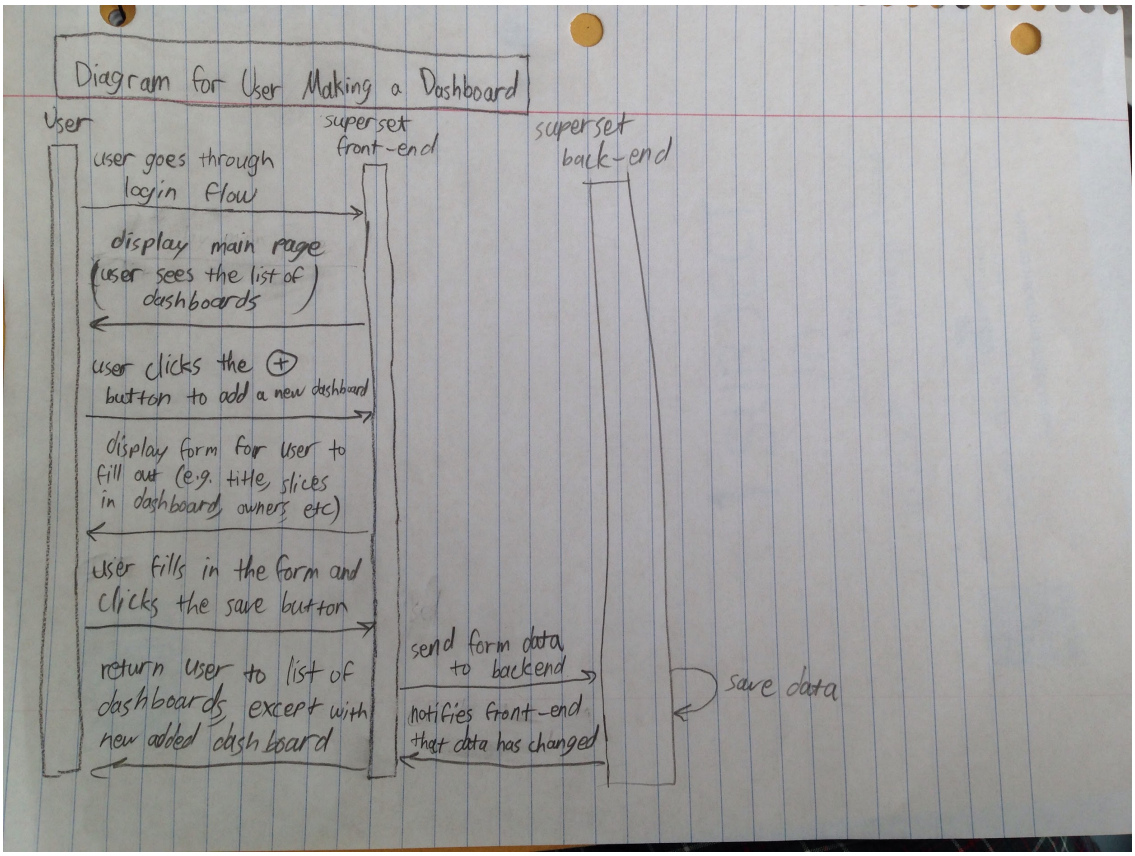
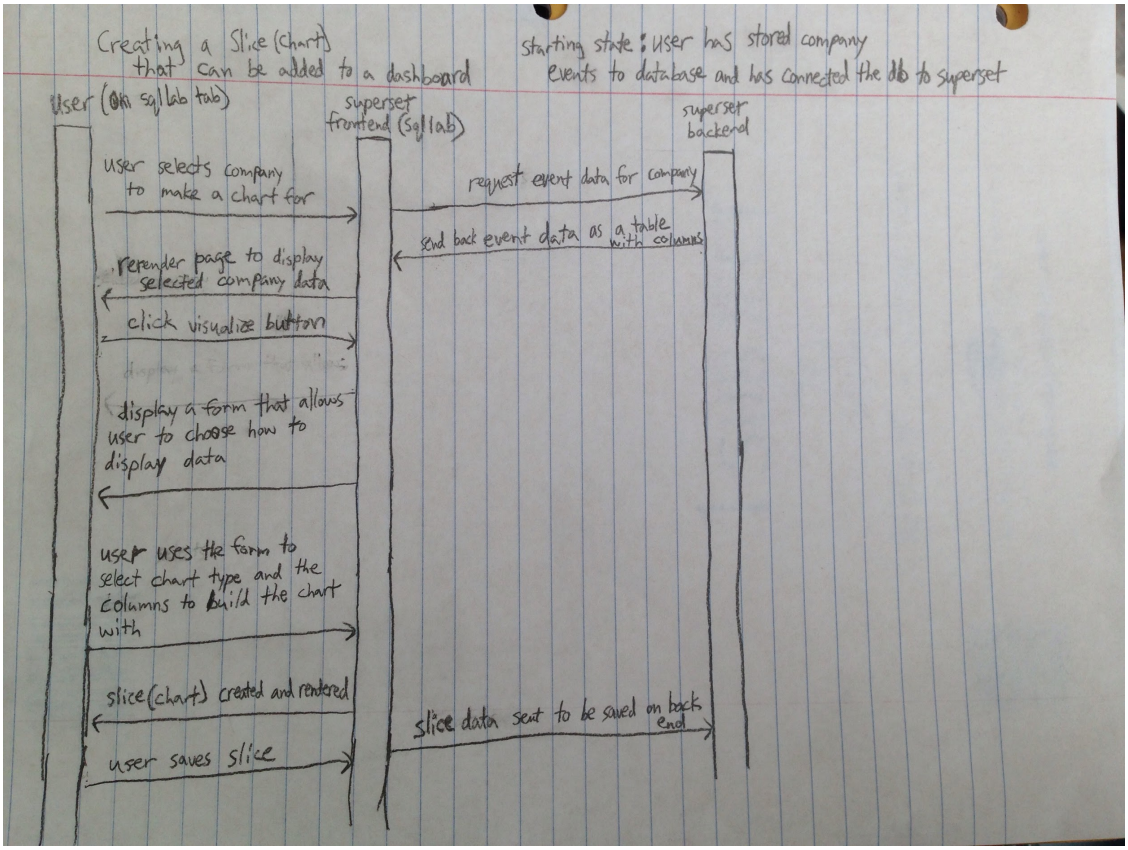
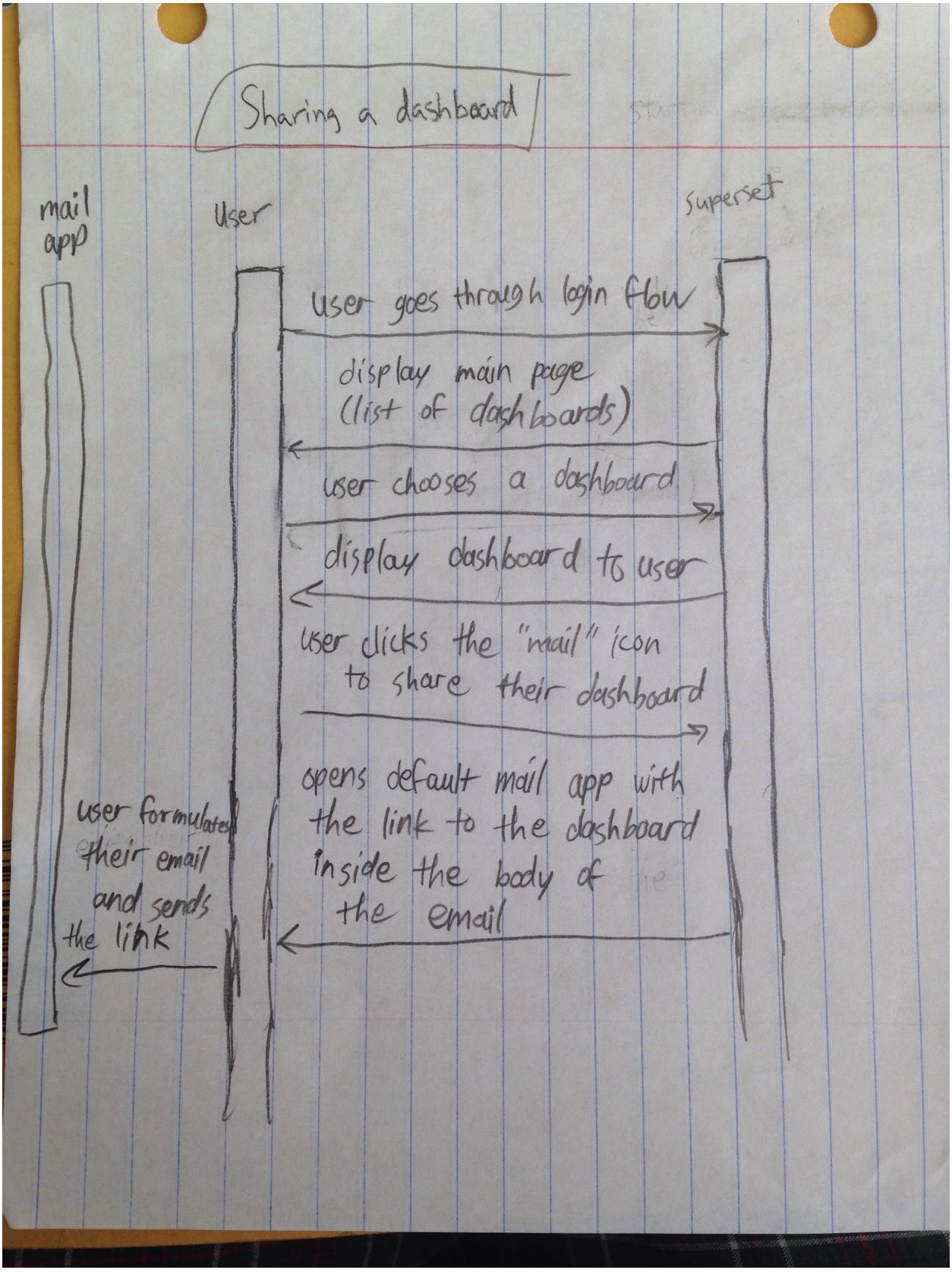
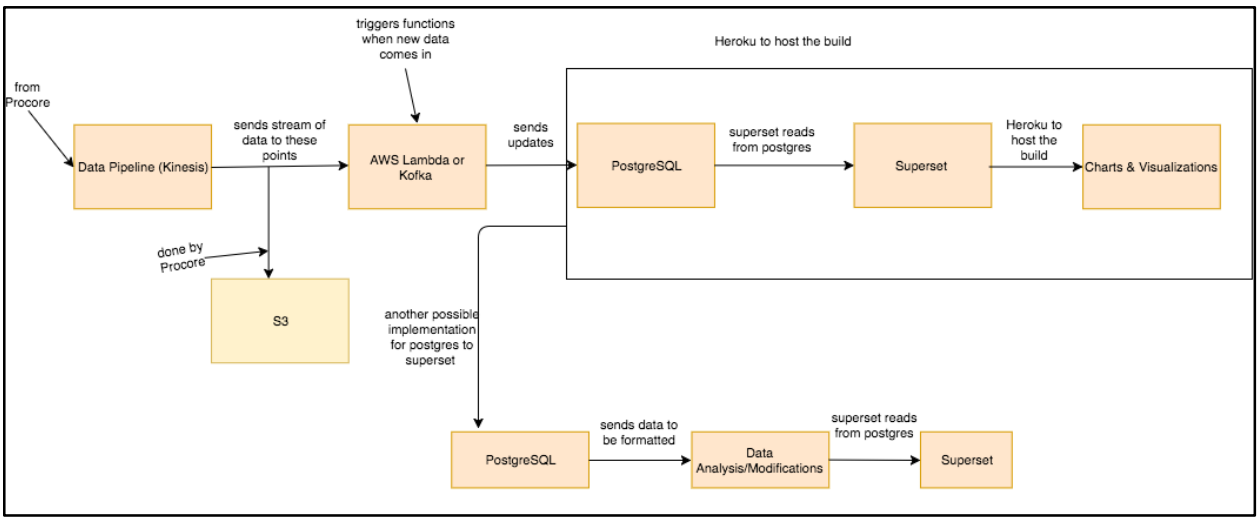


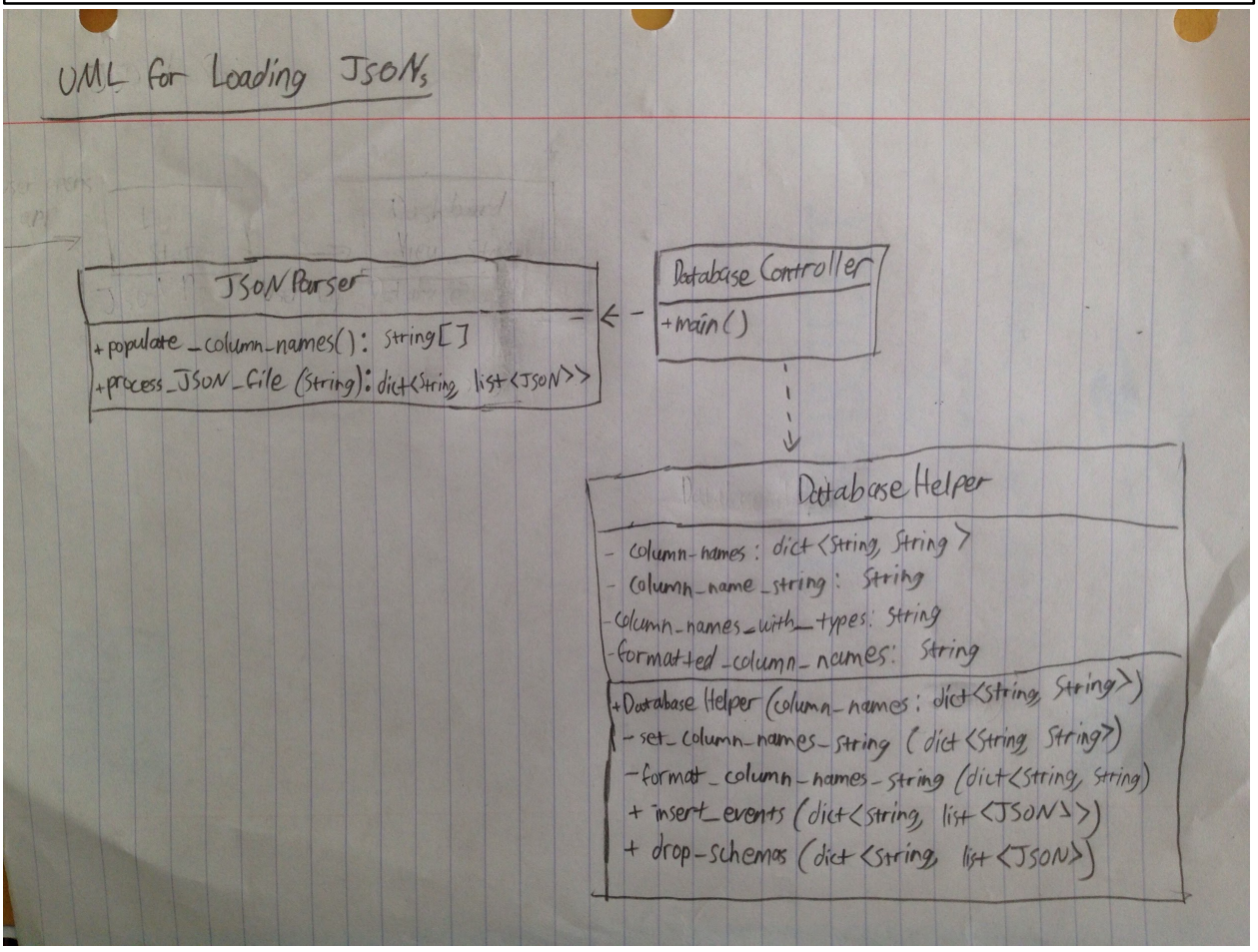
Diagram for User Making a Dashboard
User
superset front-end
superset back-end
user goes through login flow
display main page (user sees the list of dashboards)
user clicks the + button to add a new dashboard
display form for user to fill out (e.g. title, slices in dashboard, owners, etc)
user fills in the form and clicks the save button
send form data to backend
save data
return user to list of dashboards, except with new added dashboard
notifies front-end that data has changed


Creating a Slice (chart) that can be added to a dashboard
starting state: user has stored company events to database and has connected the db to superset
User (on sqllab tab)
superset frontend (sqllab)
superset backend
user selects company to make a chart for
request event data for company
send back event data as a table with columns
rerender page to display selected company data
click visualize button
display a form that allows user to choose how to display data
user uses the form to select chart type and the columns to build the chart with
slice (chart) created and rendered
slice data sent to be saved on back end
user saves slice


Sharing a dashboard
mail app
User
Superset
User goes through login flow
display main page (list of dashboards)
user chooses a dashboard
display dashboard to user
user clicks the "mail" icon to share their dashboard
opens default mail app with the link to the dashboard inside the body of the email
user formulates their email and sends the link

- Behaviorial / UML





- State

Appendices

- Technologies
  - PostgreSQL
  - Superset (Private)

- Superset (Public)
- AWS: Kinesis
- AWS: Lambda
- AWS: S3