INTERACTIVE VISUALIZATION SYSTEM OF OIL & GAS WELLS SOURCE-ROCK GEOCHEMICAL DATA TO AID HYDROCARBON EXPLORATION

By Skylar Shyu & Pablo Napan

Basic Information

Github Repo: https://github.com/psshyu/dataviscourse-GeochemOilandGas

Team Information:

Name	uNID	emails
Skylar Shyu	u1039726	skylar.shyu@utah.edu
Pablo Napan	u1077731	pablo.napan@utah.edu

Background and Motivation

There has been a recent convergence of conventional and unconventional hydrocarbon resources at the source-rock level and the question of why rich source rocks are located, and where they are in time and space, has become crucial for the global budget of petroleum resources. By having a global/detailed overview of source rocks geochemical data, explorationists can reduce the risk of charge factor and hydrocarbon generation as part of the overall petroleum system analysis.

To aid this research frontier, the E&G institute has been compiling, standardizing and visualizing geospatial & geochemical data, however, the visualization tools that oil companies and research institutes have, are not appropriate for data visualization but mostly for the purpose of geographical data management e.g. ArcGIS, QGIS, WebGIS, etc. which are insufficient for visually interactive data exploration/visualization. More powerful data exploration tools like Spotfire not only have limitations, but also, they lack specialized diagrams/scatterplots rendering for geochemistry data analisis like Van-Krevelen, HI/OI plot, etc.

The main motivation of this project is the need of having an appropriate, powerful and interactive visualization system that allows visual data exploration to aid oil & gas prospection purposes. The geospatial focus of this project will be in the U.S. basins.

Project Objectives

We are trying to accomplish to give our future visualization system users a birds-eye/detailed view of the geochemical system data of source rocks by displaying the data taken from hydrocarbon wells in the U.S.

Also, provide the users with a tool where they can interactively explore the data at different geospatial scales: data point, oil well and basin; and different time spans: from the paleozoic to the Cenozoic. Researchers and operator companies should be able to draw conclusions about: potential conventional/unconventional shale plays in marginal basins, production potential of source rocks, etc. by exploring the data visually and analyzing the specialized geochemical data plots presented.

Data

The data has been being collected and compiled over the last 5 years by the Energy & Geoscience Institute at the University of Utah which data values have been modified for confidentiality purposes.

The basinal data has also been obtained from the EGI project geodatabase, however, it is also possible to download it from this public source:

- https://www.arcgis.com/home/item.html?id=4769216bf0234324881a6764f2979bd5

Data Processing

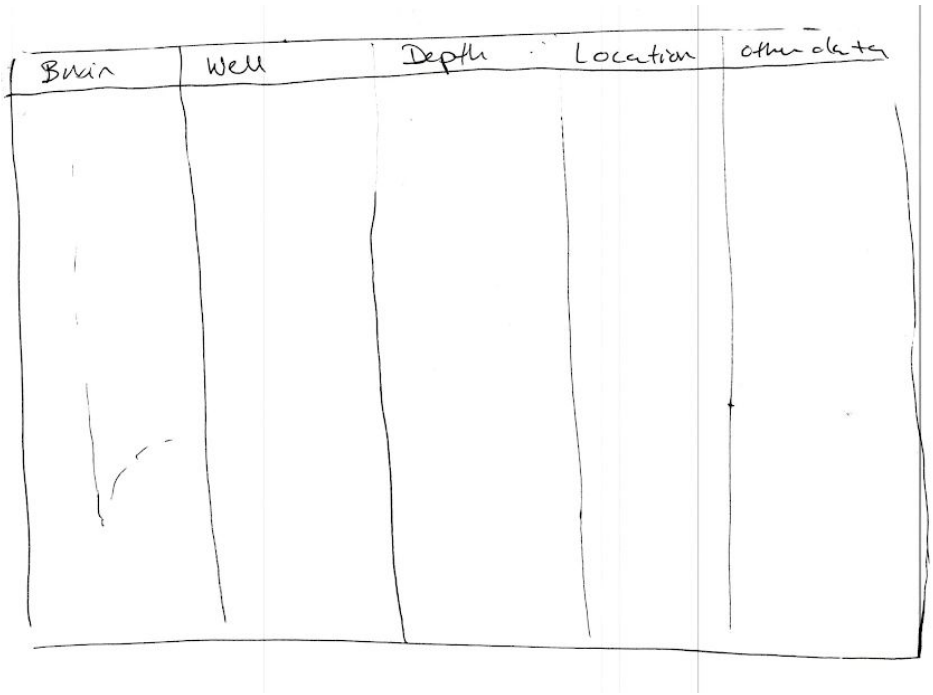
We expect to do a fair amount of data cleanup, as there's some inconsistencies regarding the availability/accuracy of the data for certain basins, groups, and wells and the samples that are extracted from them. This process requires some domain knowledge and discussion to ensure data and results significance.

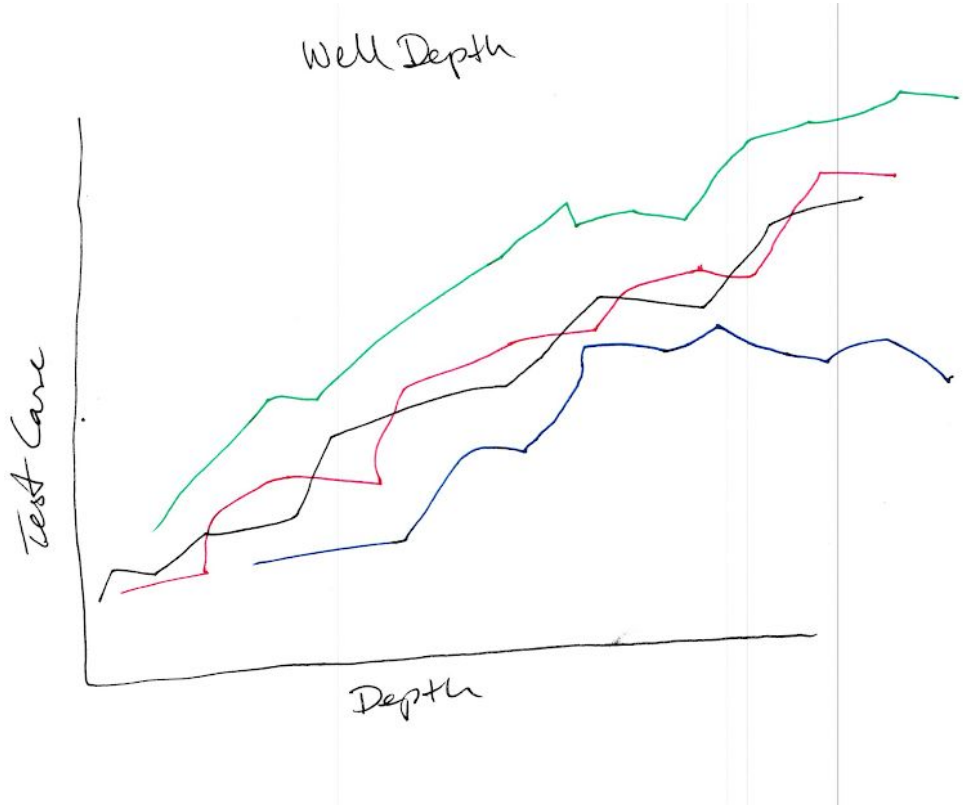
From the tables, we are planning to derive:

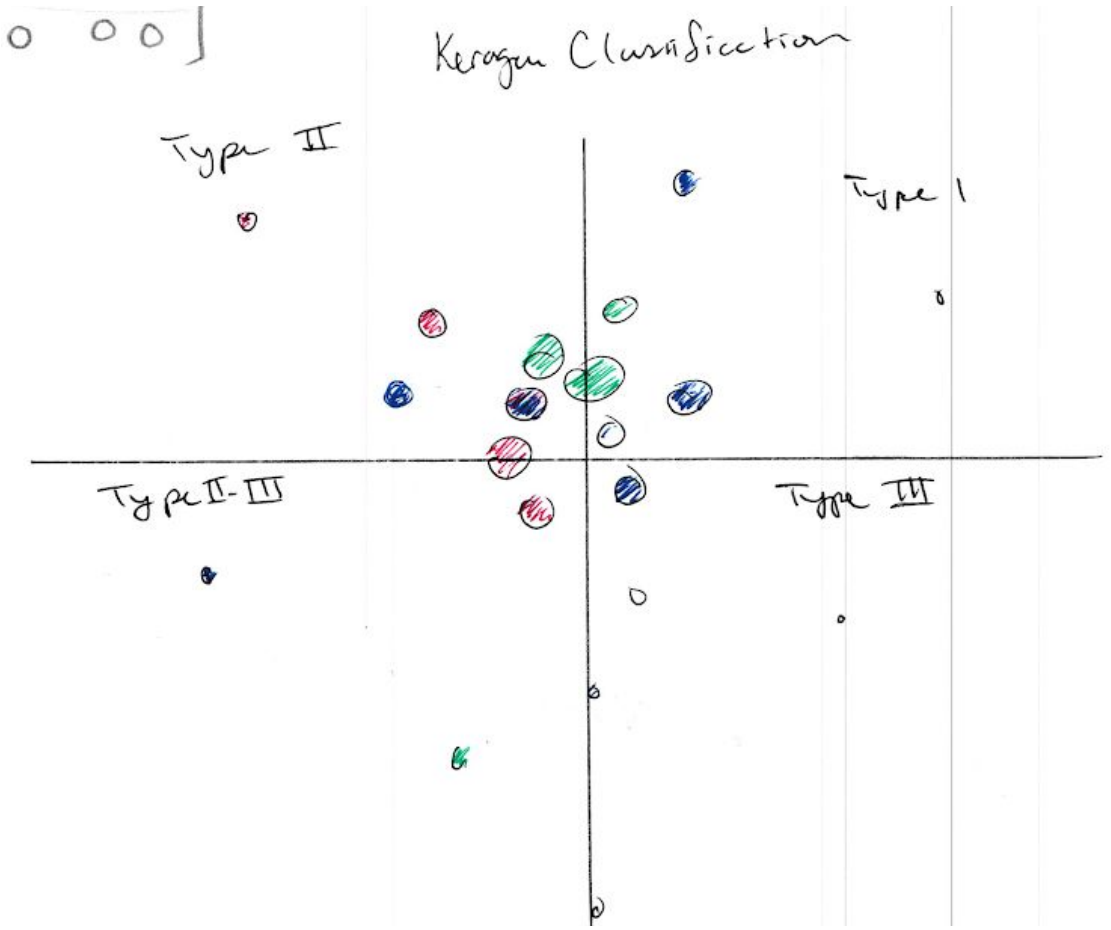
- Average age of the samples.
- Coordinate transformations.
- Production index.
- Averages (TOC, Ro, HI, OI, etc.).

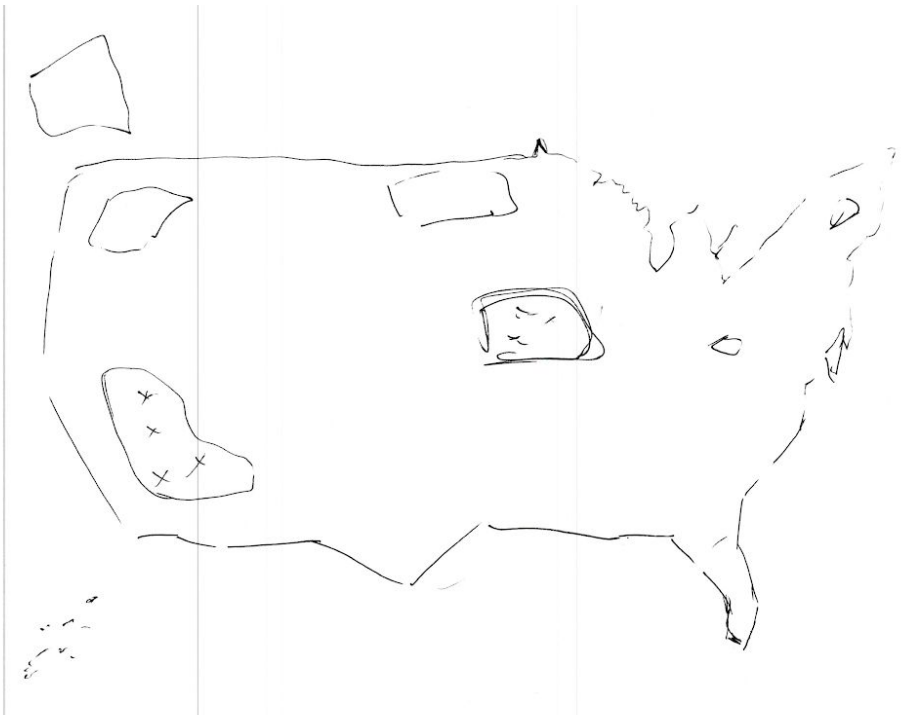
Visualization Designs

Brainstorming visualizations

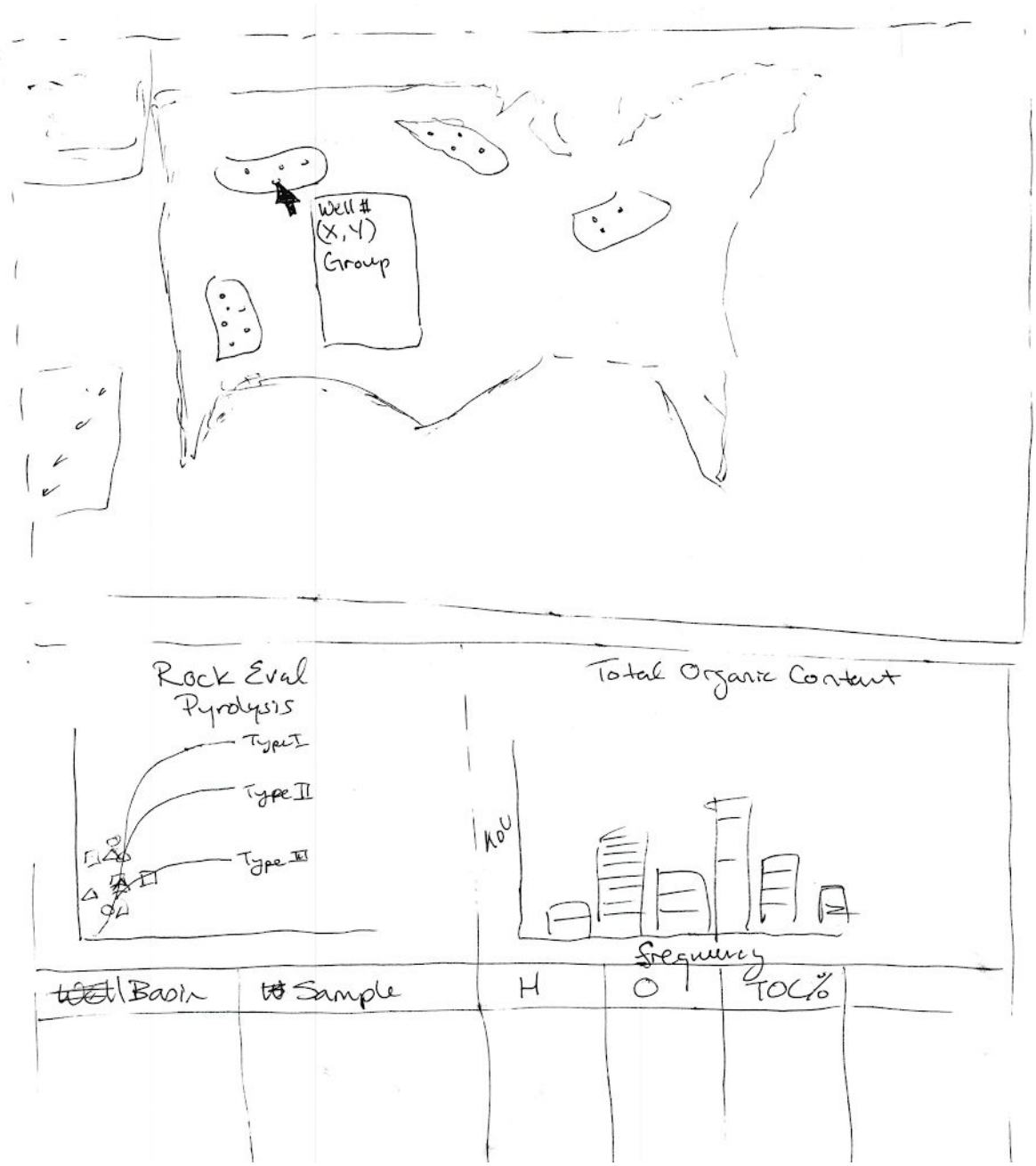





Kerogen Classification
Type II
Type I
Type II-III
Type III

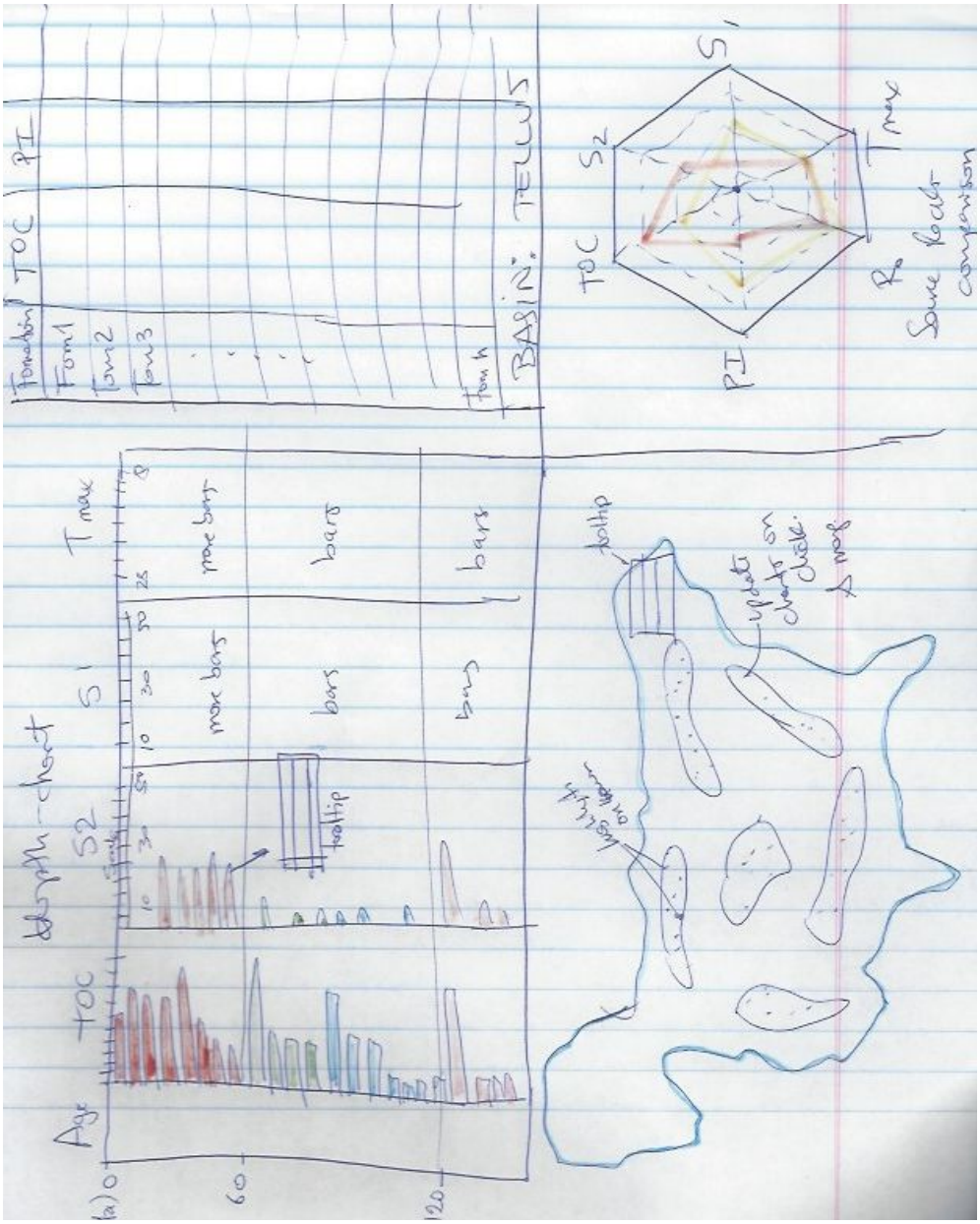


Sheet 1 (Not using):



We are not going with this design since it's fairly static. The table format creates a lot of redundant information, as one basin may have many wells that share similar features. While the tooltip is great for elaborating on a well's information, we are concerned that the dots may be too dense for users to hover over accurately. Furthermore, there isn't a lot of space here for us to elaborate on data specific to each basin.

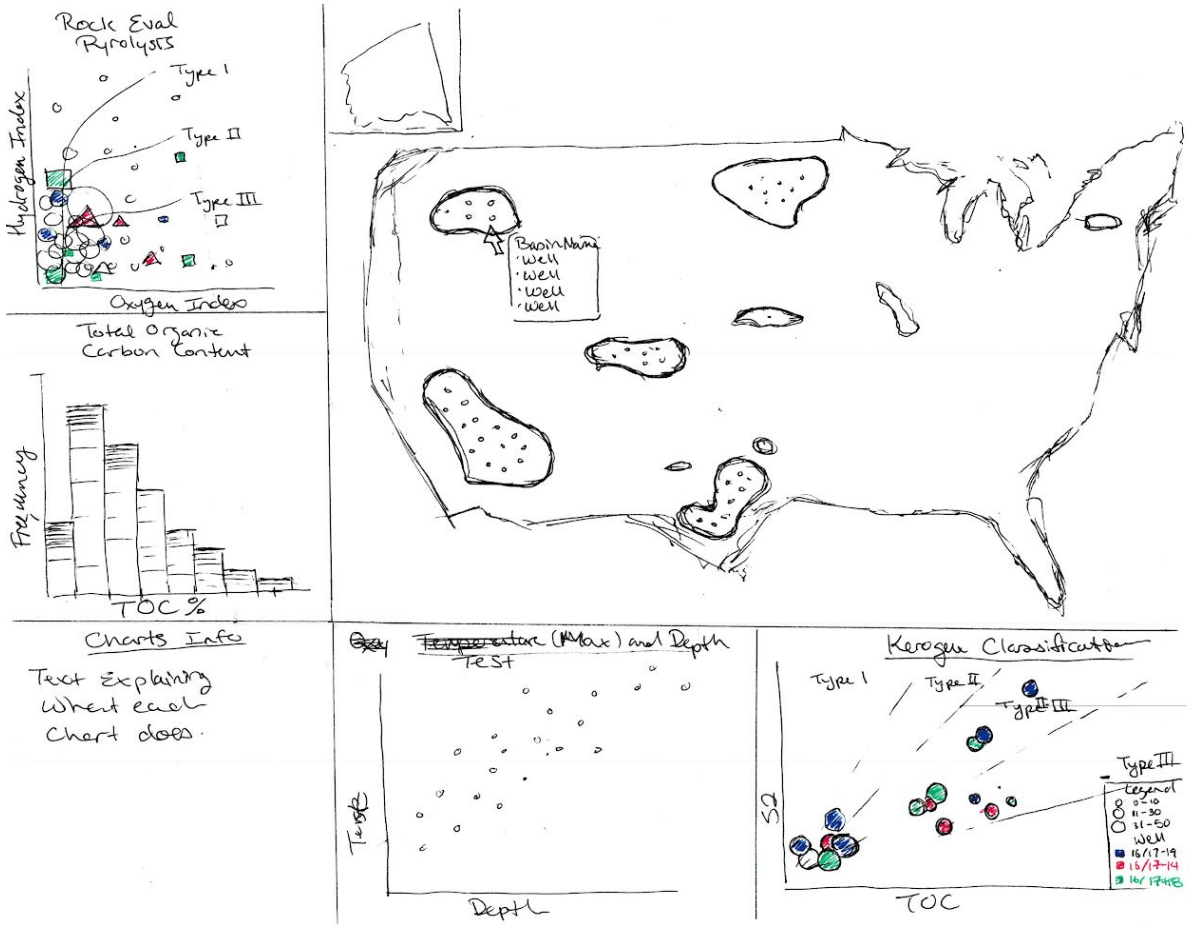
Sheet 2 (Not using):



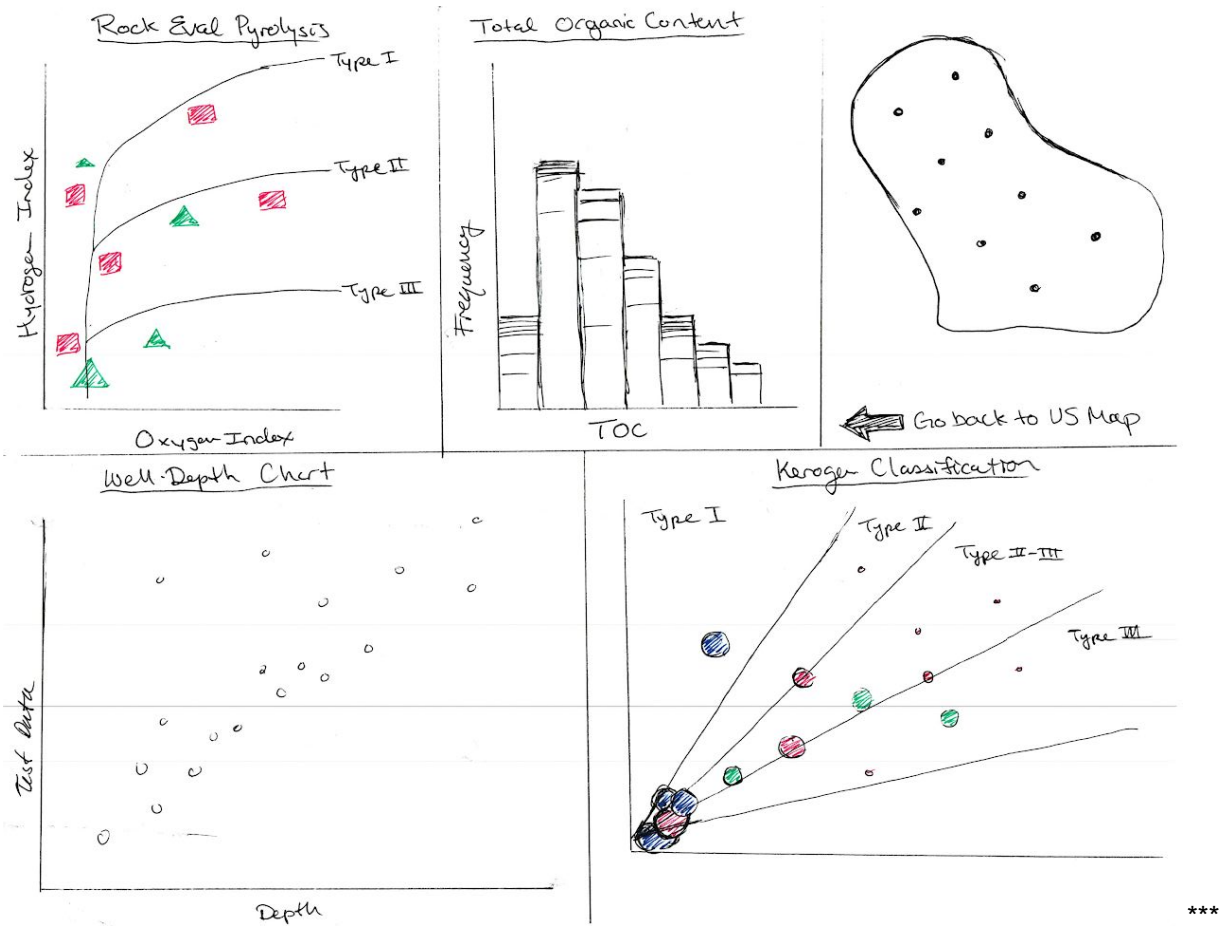
We are not going with most of this design either, because there are wells with so many data points that the bars would look too dense. Also, the hexagonal comparative chart, although it would have been nice to have, it is somewhat complicated to generate as it would need different scales/angles. And the summary table space, as Skylar suggested, could be use that for another visualization, so we have 4 final visualizations. Also, the map may be too small and may show very dense data points.

Sheet 3 (Using):

Screen 1: US Map



Screen 2: Charts showing basin information (after clicking on a basin in Screen 1)



We're opting to go with Sheet 3 where we can manipulate the sizes of charts with zoom-in shifting translations. Since the data for each chart is relatively dense and a tiny map isn't much use to anyone, we wanted our design to allow the audience to focus on one or the other.

Screen 1: US Map

To do so, our design will initially present a topological map of the US with basin/well overlays as the focal point, taking up about 75% of the screen. This will allow users to select the basin they wish to explore further; upon doing so, the map is no longer as vital since the focus is on the selected basin. Thus, we will zoom in on the selected basin while minimizing the size of the map chart as a whole, then move it to a corner, taking up about ⅛ of the screen. Other charts can be small as they're meant to represent the total data set.

Screen 2: Charts showing basin information

Whenever a basin gets clicked, 4 charts will show up:

- **Van-Krevelen diagram**: HI vs OI scatterplot (TOC encoded in plotpoint size).
- **TOC-chart**: stacked bar chart showing the contribution of each formation to the overall TOC.
- **S2 vs TOC:** scatter plot: showing the kerogen classification.
- **Well-depth chart**: Important for visualizing the how the geochemical varies in the Z dimension.

The channels in each plot (bars, circles) will highlight whenever a well (inside a basin) gets hovered over. Tooltips will be used when necessary, especially when explaining what certain abbreviations stand for, the significance of selected metrics, etc.

Must-Have Features

- An interactive U.S. map that allows zooming in of basins and hover-over of individual oil wells.
- Van-Krevelen diagram: a scatterplot to display Rock Eval Pyrolysis (hydrogen index & oxygen index)
- A stacked barchart: Total Organic Carbon (TOC) chart describing the frequency of the amount of total organic carbon.
- Well-depth charts for each individual well; these would including information such TOC, Tmax, vitrinite reflectance, S2, hydrogen index, etc.

Optional Features

- Provided that we have enough time, we will also allow implement a depth-chart that shows the geochemical data scaled in the Z (depth) axis and/or a slider that allows users to analyze the temporal evolution of the most potential source rocks in the U.S.

Project Schedule

Week	Tasks/Goals
Sat. Oct. 27th - Fri. Nov 2nd	Pablo ● Data gathering, wrangling, integrity assurance, refining, outlier removal and standardizing. ● Initialize Google Doc for Process Book Skylar ● Spin up a basic boilerplate project with empty svgs (representing charts) that can downsize and move to correct places upon clicking spots on a map svg
Sat. Nov. 3rd - Fri. Nov. 9th	(Project Milestone is due on the 9th) Pablo ● Download basins shapefiles, filter and convert them to geojson or topojson. ● Map implementation Skylar ● Implement the topological map ○ Basin overlays as clickable regions ○ Ensure that the coordinates of wells can be returned and are accurate when clicked
Sat. Nov. 10th - Fri. Nov 16th	Pablo ● Basin rendering and first scatter plots implementation (V-K and depth-Chart) Skylar ● Start implementing graphs (TOC-Chart & S2 vs. TOC)
Sat. Nov. 17th - Fri. Nov. 23rd	Pablo ● Heavy work on implementation of any remaining chart and refine details. ● If there's time, slider implementation. Skylar ● Finish implementing graphs
Sat. Nov. 24th - Fri. Nov. 30th	Skylar & Pablo ● Fix any last minute bugs