Week of:	Description of Activities:
8/24	TOTAL 2-3 hours
	Met with Brandon, Viv, Mike. Introductions. Will meet weekly on Tuesdays 7:30am PT
	o Teams meeting link for Tuesdays
	o Reviewed Student Project FY21 Notes (mission areas, project description, etc.)
	<ul> <li>Don't have access to <a href="https://code.usgs.gov/bserna/usgs-dash-project-">https://code.usgs.gov/bserna/usgs-dash-project-</a></li> </ul>
	generator (need login)
8/31	TOTAL 5-6 hours
	Tuesday 1 hour
	<ul> <li>Met with Brandon. Will connect on Slack, install CookieCutter, Goals: open</li> </ul>
	dashboard code template and play around with sample data, share a GitHub
	project with Brandon by the weekend.
	Sunday 1 hour
	<ul> <li>Read about CookieCutter, tried to unzip and open the usgs-dash-project-</li> </ul>
	generator-master.zip.cpgz file without success.
	Set up Slack #usgs_science_data_internship
	Set up GitHub https://github.com/PortfolioSQA/USGS Catalog Dash
	Monday 3-4 hours
	Set-up environment and became more familiar with dash
	<ul> <li>Created a simple dashboard using sample data using my own code. Will</li> </ul>
	request new cookiecutter template file again Tuesday morning.
9/7	TOTAL 7 hours
	Tuesday 1 hour
	Met with Brandon. Created Slack Channel. Resent CookieCutter file and
	opened together. Goals for the week: check out the dash template, look at xml
	from Science Base Catalog (project -> more details) and think about summary
	graphics/tables.
	<ul> <li>Notes: 8 USGS regions, 6-7 mission areas, many science centers! Connect with</li> </ul>
	elastic search index to fill contents (makes aggregation harder), new elastic
	search package works with pandas dataframe.
	Friday/Sunday 2 hours
	Downloaded the cookiecutter template
	<ul> <li>Played with the dashboard app</li> </ul>
	<ul> <li>Looked through Science Data Catalog xml files to determine which data is</li> </ul>
	available
	<ul> <li>Proposed a new idea for dashboard to Brandon- Query/Summarize by map</li> </ul>
	boundaries:
	<ul> <li>User first filters by Mission Area, Science Topic, or none (dropdown</li> </ul>
	menus)
	<ul> <li>User then demarcates geographic boundary requirement on</li> </ul>
	interactive US Map (default is entire US)
	Return record count and all records using geographic location and
	other filters
	<ul> <li>User may filter results further by Science Center, date, or sort by</li> </ul>
	most frequently accessed records
	<ul> <li>Monday 4 hours</li> <li>Read about Dash, Mapbox, &amp; Leafly</li> </ul>
	o Installed Shapely, Created Mapbox account

	<ul> <li>Comments: Polygons are not easy to work with for the user, could choose by</li> </ul>
	selecting geographic boundary by county, state, USGS region, center, or select
	point and radial distance.
	<ul> <li>What to include in sample data: name, keyword, center, lat/long, dates</li> </ul>
9/14	TOTAL 11 hours
	Tuesday 3 hours
	<ul> <li>morning meeting canceled – will meet next Tuesday</li> </ul>
	<ul> <li>Installed faker and created sample dataset to use in dashboard</li> </ul>
	<ul> <li>Spent time understanding the bounding coordinates in the xml files</li> </ul>
	 bounding>
	<westbc>-123.60443114739</westbc>
	<eastbc>-97.764587398423</eastbc>
	<northbc>45.032773794668</northbc>
	<southbc>38.897445871348</southbc>
	Sunday/Monday 8 hours
	Worked on sample data & dashboard
	Questions for Brandon/to figure out:
	How to use a styled map from MapBox
	<ul> <li>Pricing on MapBox – Feasible for USGS?</li> </ul>
	Other options
	Discuss best way to map boundaries of data collection (& data collected for
	entire US)
	<ul> <li>Which summaries are most important for USGS? Why? (record count by</li> </ul>
	mission area/science center/popular/keywords)
	Layer USGS regions on map?
0/21	Use rows from data table without displaying all need columns (join?)
9/21	TOTAL 16 hours
	Tuesday 1 hour
	Met with Brandon
	o To do: Get pricing information for Mapbox, get map working with filtering, fix
	table, complete aggregate graphs, & place code in template
	Next week: Get SDC data to play with mapping, use styled maps
	• Friday 3 hours
	<ul> <li>'Learning App' Got map filtering working with sample data, fixed table,</li> </ul>
	aggregated by visits
	Posted mapbox pricing  Fundamental SDC data
	o Explored SDC data
	Saturday/Sunday 10 hours
	<ul> <li>Graphed SDC data, but need to clean it up to be able to filter by date/keyword</li> </ul>
	etc.
	o ETL SDC data
	<ul> <li>Dropped NA for missing spatial (8 rows)</li> </ul>
	Replaced NA for dates, kept year (see questions)
	Combined keyword columns
	<ul> <li>'Sketch' of where to place map/table/graphs in template</li> </ul>
	Monday 2 hours
	o 'learning_app' committed to Github – has SDC data graphed in map by
	date/science center, ETL python file in folder
	o 'SDC_Map_Dash' template committed to Github - wireframe:
	• Will start working on this version once we decide on layout, filtering
i .	<ul> <li>Filter by date, map, filter by keyword (not sci center)</li> </ul>

	■ In each tab — Datatable + Summary Stats
	<ul> <li>mappable US data (not nationwide)</li> </ul>
	<ul> <li>Nationwide datasets – mapped?</li> </ul>
	<ul><li>world/earth/international datasets – mapped?</li></ul>
	• Questions:
	<ul> <li>If nan data is filtered by date, we lose about 10% (600+ rows). Lose the rows?</li> </ul>
	Save them? Include Map/graph? I replaced nan with 1900-present
	<ul> <li>Best way to determine global, continental, local data (using PlaceKeyword).</li> </ul>
	Agree with three tabs?
	<ul> <li>What to do with keywords?</li> </ul>
	<ul> <li>Could color markers by</li> </ul>
	<ul> <li>Moved away from selecting map area to fill table – table fills map (zooming</li> </ul>
	allows user to determine if data is available in a particular area). Do we want
	map selection ability? Seems messier.
	<ul> <li>Using integer for year – to_datetime doesn't like 1500s</li> </ul>
	<ul> <li>Next steps</li> </ul>
9/28	TOTAL 7-8 hours
	Tuesday 1 hour
	Met with Brandon
	o Notes:
	<ul> <li>Viv may send invitation to bi-weekly meetings Mondays 9am PT –</li> </ul>
	have not received invite
	<ul> <li>Filter by Science Center &amp; USGS Thesaurus Keyword (not by date) -</li> </ul>
	done
	<ul> <li>Figure out a 'pretty' way to select all key words/all science centers –</li> </ul>
	note: dash doesn't seem to have one, we can select all, but shows all
	science centers
	<ul> <li>For now, separate global, US, mappable data by place keyword in</li> </ul>
	three tabs - problematic
	<ul> <li>Create data table and aggregate statistics for each tab</li> </ul>
	<ul> <li>Color markers by science center</li> </ul>
	<ul> <li>Other: Fix map so when table returns nothing, no markers are mapped, when</li> </ul>
	global or US data tab is selected – map shows what? Need to ignore rows in
	data table without lat/lon for the map graph data? Nan Science Center?
	Tuesday 2 hours
	<ul> <li>Updated journal, separated keywords, created unique USGS thesaurus</li> </ul>
	keyword and science center list for dropdowns
	Monday 4-5 hours
	Broke the app trying to add keyword search and spent a couple hours trying to
	fix it. Started from scratch.
	<ul> <li>Have not successfully separated international data (see graph) so the tabs don't make sense right now. I could work on separating by lon/lat if we want to</li> </ul>
	do this.
	science center choices inside each tab to be able to change the map/selection  So many Science Centers can use a select all/deselect etc. need to change
	•
10/5	font size. How should we display this?  • TOTAL 10 hours
10/3	Tuesday 1 hour
	Tuesday 1 nour     Dropdown with select all – return all data
	Remove tabs – or use for summaries
	<ul> <li>Zoom into populated marker area</li> </ul>

	Color markers
	Fill null Science Data Center with 'Undetermined'     Counts (Counts)
	Counts/Graph  Kayananda
	○ Keywords
	o <a href="https://data.usgs.gov/modelcatalog/">https://data.usgs.gov/modelcatalog/</a>
	o <a href="https://sciencebase.gov/datarelease/summary/">https://sciencebase.gov/datarelease/summary/</a>
	<ul> <li>Word cloud for the datatable keywords (% top 50 terms)</li> </ul>
	Friday 4 hours
	<ul> <li>Combined all data (instead of mappable, US, global)</li> </ul>
	<ul> <li>Fixed dropdown &amp; added All Science Centers</li> </ul>
	<ul> <li>Added keyword search</li> </ul>
	o Removed tabs
	<ul> <li>Fixed null SC -&gt; Undetermined</li> </ul>
	o Return Count
	Saturday 4 hours
	<ul> <li>Fixed &amp; formatted datatable</li> </ul>
	<ul> <li>WordCloud</li> </ul>
	o Next Steps
	<ul> <li>Zoom to populated markers</li> </ul>
	<ul> <li>Color markers by science center</li> </ul>
	<ul> <li>Make sure key words are searched lower case</li> </ul>
	<ul> <li>Speed up wordcloud</li> </ul>
	<ul><li>Heroku</li></ul>
	Sunday 1 hour
	Researched zoom to markers and marker color
10/12	TOTAL 9 hours
	Tuesday 5 hours
	<ul> <li>1-hour meeting with Brandon. Notes:</li> </ul>
	<ul> <li>10/19 check-in biweekly informal 8-8:30 am PT</li> </ul>
	<ul> <li>10/26 SDM Biweekly 9-10:30 PT</li> </ul>
	<del>■ Text changes (see github)</del>
	- place filters in same search box
	- update datatable headers
	- remove browse these data
	- cache wordcloud for 'all'
	<del>- keywords – lowercase, remove punctuation</del>
	<del>- controlled terms – USGS, etc.</del>
	■ bug – CA/iso, count 2, returns 1??
	<ul> <li>Color by center/zoom into markers</li> </ul>
	- update button to download CSV
	<ul> <li>Place 'downloading state' into dt and we</li> </ul>
	<del>- Heroku</del>
	<ul> <li>4 hours - completing above list</li> </ul>
	Friday 4 hours
	<ul> <li>2 hours – completing above list</li> </ul>
	<ul> <li>2 hours - Connect with git SSH, fix files and attempted to deploy with Heroku</li> </ul>
10/19	TOTAL 2.5 hours
	Monday 1.5 hours
	o Meeting 0.5 hours
	<ul> <li>8am Join Microsoft Teams Meeting</li> </ul>
	<ul> <li>CDI: <a href="https://my.usgs.gov/confluence/display/cdi/Home">https://my.usgs.gov/confluence/display/cdi/Home</a></li> </ul>
1	<ul> <li>Model catalog: https://data.usgs.gov/modelcatalog/</li> </ul>

	Sciencehare data releases:
	Sciencebase data releases:
	https://www.sciencebase.gov/catalog/items?q=&filter=systemType%
	3DData+Release&filter=browseCategory%21%3DData+Release+-
	+In+Progress
	1 hour – tried to get the app to deploy on Heroku  Adams on the 2 hours by the 2 hours by the 4 222
	• Manage.py file? heroku ps:scale web=1???
	Tuesday – 1 hour
10/25	Met with Brandon, deployed on Heroku
10/26	TOTAL 4 hours
	Monday 3 hours
	<ul> <li>Join Microsoft Teams Meeting 9am</li> </ul>
	<ul> <li>Missed the meeting this morning. I'll be there next time.</li> </ul>
	3 hours
	<ul> <li>Fixed loading state for count</li> </ul>
	<ul> <li>If I place a loading state on map, it doesn't map correctly!? Can you help me</li> </ul>
	figure out why that would be?
	<ul> <li>sci_center colors – works with static data!!! Talk about how to fix this, if we</li> </ul>
	want it.
	Tuesday 1 hour
	<ul> <li>Meet with Brandon and Lisa Zolly</li> </ul>
	o Data needed:
	■ Science center
	■ Latest harvest
	■ Status
	■ Doi citation info
	■ DOI
	o Filter by:
	Science Center
	■ Dates (Beg, End, Updated, Latest Harvest)
	Status – Active/Inactive
11/2	TOTAL 11 hours
11/2	Monday 1.5 hours
	o 8am Join Microsoft Teams Meeting
	Tuesday Meeting (0.5 hours)
	Looked at data (1 hour)
	·
	• To Do:
	Filter by Science Center (datasource)
	<ul> <li>Filter by Status - Active/Inactive (where is this?)</li> </ul>
	<ul> <li>Filter by Dates (avail: first_harvest_date, last_harvest_date,</li> </ul>
	last_mdate_check_date, mdate) which?
	o Filter: doi: None Do we use this?
	<ul> <li>Return Table: SciCenter, # citations, doi, status, date?</li> </ul>
	Return Count: dataset count with filters
	<ul> <li>Return # of citations in the last month (don't have by date)</li> </ul>
	<ul> <li>Return: Pie Chart of Active/Inactive (if not filtered by this)</li> </ul>
	<ul> <li>Flag certain datasets? inactive with citation?</li> </ul>
	o Other Questions:
	How do I know if data is active/inactive? Is this if the data has been
	harvested? What would the row look like if it weren't?
	<ul> <li>For the ORCID availability for retrospective DOI assignment - Do I just</li> </ul>
	check if doi is None?

	Lisa talked about rate of growth of the collection over time (monthly)
	growth? Which dates?) Bar chart
	Do we want to show the related primary publications?
	• Include non-primary pubs for citation count?
	Layout similar to dash template? Use tabs?
	• Friday 3.5 hours
	o etl
	<ul> <li>New file from Cookiecutter template using sample data</li> </ul>
	<ul> <li>Filter by Science Center (datasource)</li> </ul>
	o Filter by Status - Active/Inactive
	<ul> <li>Filter by Dates (beg, end, last updated (mdate), last harvest)</li> </ul>
	<ul> <li>Return Table: SciCenter, # citations, doi, status</li> </ul>
	Saturday/Sunday 6 hours
	o Filter by Dates
	<ul> <li>Return Count: dataset count with filters</li> </ul>
	<ul> <li>Return: Pie Chart of Active/Inactive (if not filtered by this)</li> </ul>
	Questions/Issues:
	<ul> <li>When I align the CSV download link to the right, it breaks! Why?</li> </ul>
	<ul> <li>Using mdate instead of last_harvest_date doesn't work Worked for a while</li> </ul>
	trying to figure out problem with date type, couldn't find it
	<ul> <li>Fix tab1 so it gives the datatable for selection (doesn't return to all). Why do</li> </ul>
	the other tabs retain filter info, but not datatable?
11/9	TOTAL 8 hours
	Monday 2.5 hours
	o 9am Join Microsoft Teams Meeting 9am
	<ul> <li>Fix count and commit to new repository on GitHub:</li> </ul>
	https://github.com/PortfolioSQA/SDC Manager Dashboard
	Tuesday Meeting 1 hour
	<ul> <li>TODO: Fix table (so it doesn't change to default when tabs are changed), Fix</li> </ul>
	<del>mdate</del>
	<ul> <li>Talked about NLP project - creating a custom NER to identify models from</li> </ul>
	epubs warehouse?? for the Model catalog. If time permits start researching
	this.
	<ul> <li>Also talked about volunteering for usability testing for Sophie</li> </ul>
	Saturday 3 hours
	o Fixed mdate
	<ul> <li>Worked on datable persistence for a couple hours I can't figure out what to</li> </ul>
	do to keep the table from reverting to default in tab.
	○ Changed 0/1 => inactive/active
	Sunday 1.5 hours
	Research on how to create custom NER model
	<ul> <li>Tried for another 1-2 hours to debug the table persistence. Used dcc.store,</li> </ul>
	persistence = True, and filtering in another callback. Didn't get it to work 🗵
11/16	TOTAL 4.5 hours
	Monday 0.5 hours
	Monday 8am Join Microsoft Teams Meeting
	Tuesday – meeting cancelled due to power outage
	Wednesday 3 hours
	Deploy to Heroku
	Tab update works in Heroku. Weird!
	Fixed datatable sort
	<ul> <li>Need to fix table so we can sort by status and right align download link</li> </ul>
	1

	Sunday 1 hour
	Created random dates in Excel for sample data
11/23	Thanksgiving Week – WEEK OFF
11/30	Final Exam – WEEK OFF
12/7	TOTAL 12.5 hours
	1.5 hours Monday <u>Join Microsoft Teams Meeting 9am</u>
	1 hour Tuesday meeting
	<ul> <li>Methods: CNN, SVM, etc</li> </ul>
	Tasks: Image classification, image segmentation, etc
	Applications (use cases): precipitation-induced landslide warning, tracking
	rainfall thresholds,
	3 hours Tuesday
	<ul> <li>Create sample data (dates for beg, end, update, harvest), change code to</li> </ul>
	incorporate new 'dates'
	3 hours Wednesday
	<ul> <li>Debug code changes for new sample dates, deploy on Heroku, update journal, e-mail Sophie with updates</li> </ul>
	2 hours Thursday
	<ul> <li>Create notes document for model scraping</li> </ul>
	Review text
	2 hours Sunday
	<ul> <li>extract abstract, identify method, task and application</li> </ul>
	<ul> <li>Create txt files with abstracts</li> </ul>
12/14	TOTAL 10 hours
	• 0.5 hours
	<ul> <li>Monday 8am <u>Join Microsoft Teams Meeting</u></li> </ul>
	• 1 hour
	<ul> <li>2pm - Meet with Lisa, Brandon, and Sophie</li> </ul>
	TO DO LIST:
	Link – learn how metrics are calculated
	<ul> <li>Table – font choice?</li> </ul>
	Org of interest -> Choose your science center/program
	• Select dataset status
	<ul> <li>Explanation of active/inactive</li> <li>Put date selection in same box</li> </ul>
	<ul> <li>Justify radio buttons</li> </ul>
	Colon after headers
	• Remove beg/end dates
	• Last harvested, last updated -> start date?
	<ul> <li>Flexibility in search by date? Calendar year?</li> </ul>
	<ul> <li>Dataset Count - &gt; Active, Inactive, Total, pie chart</li> </ul>
	<ul> <li>Make stacked bar graph for active inactive by date</li> </ul>
	→ Remove pie chart tab
	→ doi – links?
	→ Alternate IDs for datasets?
	→ Delete widgets in graph area
	⊖ Format date label
	Download -> more usable file name     To the control of the c
	Help & documentation (Tool tips)  The part of it's a set was alian and the part of th
	© Errors: If it's not working what happens?
	• 8.5 hours

	Complete TODO list, redesign site, insert links for DOI, alternate identifiers,
12/21	stacked bar chart for counts by date, delete widgets in graphs  WEEK OFF - HOLIDAY
12/21	TOTAL 7-8 hours
12/20	30 min -Monday 8am <u>Join Microsoft Teams Meeting</u>
	https://data.usgs.gov/datacatalog/
	1 hour – Tuesday Meeting with Brandon
	<ul> <li>Questions: title, markdown – open link in new tab/window for DOI, filter by</li> </ul>
	active/inactive?, explanation for active inactive, Errors
	→ TODO for dashboard: table -> dbc, tab (dash -> SDC Dashboard), Sentence -
	bold or large numbers,
	• https://dash-bootstrap-
	components.opensource.faculty.ai/docs/components/table/ - dbc table
	<ul> <li>https://data.usgs.gov/datacatalog/api/docs/v1, swagger API – look at</li> </ul>
	documentation to get data except citation (leave as null for now)
	<ul> <li>TODO for NER: exploratory functional prototype: look at</li> </ul>
	https://paperswithcode.com/ , format text in jsonl, get prodigy (costs \$390),
	{"text": "This is a text"}
	{"text": "This is another text"}
	https://data.usgs.gov/modelcatalog/search Training Data
	<ol> <li>https://pubs.usgs.gov/tm/14/a2/tm14a2.pdf</li> <li>https://pubs.usgs.gov/of/2008/1159/downloads/pdf/OF08-1159.pdf</li> </ol>
	2. https://pubs.usgs.gov/of/2006/1139/downloads/pdi/OF00-1139.pdf
	4. https://pubs.usgs.gov/of/2007/1088/pdf/of07-1088_508.pdf
	5. https://www.mdpi.com/2073-4441/8/1/17
	6. https://pubs.usgs.gov/wri/1990/4130/report.pdf
	7. https://pubs.usgs.gov/tm/12b1/
	8. https://pubs.usgs.gov/tm/2006/tm6b3/
	9. https://www.mdpi.com/1999-4893/1/2/52
	10. https://data.usgs.gov/modelcatalog/data/5eb4485782ce25b5135abf28
	11. https://data.usgs.gov/modelcatalog/data/5eb4485082ce25b5135abee3
	12. https://data.usgs.gov/modelcatalog/data/5f6240eb82ce38aaa2361498
	13. https://data.usgs.gov/modelcatalog/data/5eb4485e82ce25b5135abf70
	14. <a href="https://data.usgs.gov/modelcatalog/data/5eb4485f82ce25b5135abf86">https://data.usgs.gov/modelcatalog/data/5eb4485f82ce25b5135abf86</a>
	15. <a href="https://data.usgs.gov/modelcatalog/data/5eb4485f82ce25b5135abf7c">https://data.usgs.gov/modelcatalog/data/5eb4485f82ce25b5135abf7c</a>
	16. <a href="https://data.usgs.gov/modelcatalog/data/5eb4485582ce25b5135abf16">https://data.usgs.gov/modelcatalog/data/5eb4485582ce25b5135abf16</a>
	17. <a href="https://data.usgs.gov/modelcatalog/data/5f036b3f82ce0afb2446e04a">https://data.usgs.gov/modelcatalog/data/5f036b3f82ce0afb2446e04a</a>
	18. <a href="https://data.usgs.gov/modelcatalog/data/5eb4486182ce25b5135abfb0">https://data.usgs.gov/modelcatalog/data/5eb4486182ce25b5135abfb0</a>
	19. <a href="https://data.usgs.gov/modelcatalog/data/5eb4485282ce25b5135abef4">https://data.usgs.gov/modelcatalog/data/5eb4485282ce25b5135abef4</a>
	20. https://data.usgs.gov/modelcatalog/data/5eb4486182ce25b5135abfaa
	3 hours – Wednesday
	<ul> <li>Update journal, look at prodigy (\$390 license), get 10 more text files, create</li> </ul>
	jsonl file (see above links)
	3 hoursThursday
	<ul> <li>Dashboard updates, dbc tables are SUPER SLOW - decided to format dash data table (I have another file with the dbc table I can show if you'd like)</li> </ul>
1/4	TOTAL 9 hours
	• 2 hours - Monday:
	Doctor's Appt- Couldn't join meeting Join Microsoft Teams Meeting 9am
	<ul> <li>2 hours Install prodigy, en_core_web_sm, etc. Label text, git commit, etc.</li> </ul>
	• 1 hour - Tuesday
1	·

Meet with Brandon 7:30 - will meet with Mike & Viv soon, Ruby Gem linguist shows which languages in code, TODO: 10-20 unlabeled texts for evaluation data (good and bad examples), run in prodigy (see slack), try again using shorter NEs, use en\_core\_web\_lg, future: establish list of methods, concept of datasets on prodigy?, deploy dbc table on Heroku and see if its faster 3 hours – Thursday o 1+ hours Meet with Sophie to talk about Usability Testing 2 hours Evaluation Data (in excel & jsonl) 3 hours - Friday ○ Train/evaluate ner – 0% accuracy, relabeled training set 0% accuracy, read more documentation, tried with models only, frustrating - next combined to 40 for training set and re-labeled. Didn't help https://doi.org/10.1126/science.aat4723 https://pubs.usgs.gov/of/2001/ofr-01-0002/ https://data.usgs.gov/modelcatalog/data/5ff62dc1d34ea5387df035fa https://data.usgs.gov/modelcatalog/data/5eb4485982ce25b5135abf3c https://doi.org/10.1111/gwat.12397 https://data.usgs.gov/modelcatalog/data/5eb4485482ce25b5135abf0e 6. 7. https://data.usgs.gov/modelcatalog/data/5eb4485682ce25b5135abf1c https://data.usgs.gov/modelcatalog/data/5eb4485382ce25b5135abefc https://doi.org/10.3133/wri874163 10. https://doi.org/10.1002/2017JC013204 11. https://doi.org/10.1029/2011JB008968 12. https://doi.org/10.3133/wri974022 https://pubs.usgs.gov/tm/tm4f2/ 13. https://doi.org/10.1016/j.ocemod.2010.07.010 15. https://doi.org/10.3133/ofr20151009 16. https://data.usgs.gov/modelcatalog/data/5ef3952782ced62aaae3ef55 17. https://pubs.usgs.gov/tm/tm6a37/ 18. <a href="https://doi.org/10.1002/joc.3625">https://doi.org/10.1002/joc.3625</a> 19. <a href="https://doi.org/10.3133/tm14A1">https://doi.org/10.3133/tm14A1</a> https://doi.org/10.3133/tm6A43 1/11 **TOTAL 9.5 hours** 2 hours - Monday: 30 min. Monday 8am Join Microsoft Teams Meeting 1.5 hours – tried again with prodigy, read a bit on topic analysis, dbc table. No pagination, can't sort, slower 2 hours - Tuesday Meet with Brandon 7:30 o TODO: Tutorials, analyze n-grams, constrain to better examples o Peer review for a "resource review" that Sophie preparing for the CDI Usability Collaboration Area 2 hours - Wednesday Prodigy tutorials https://prodi.gy/docs/named-entity-recognition (Food Ingredient entities) 1.5 hour - Thursday Continued Prodigy tutorial 1 hour - Friday Continued Prodigy tutorial 1 hour – Saturday N-gram analysis of text https://github.com/PortfolioSQA/USGS Catalog Dash/blob/master/ngram tex t analysis.ipynb 1/18 **TOTAL 10 hours** 

	Martin Luther King Jr Holiday Monday
	0.5 hour - Tuesday
	<ul> <li>Meet with Brandon 7:30 – 8:00</li> </ul>
	<ul> <li>TODO: methods NER: try water balance, transport model, etc. separate and</li> </ul>
	then try DE, Lin reg, random forests, etc. (may have to use regex for that), Graphical interface - Keep track of articles that may not be models, New Github directory with files and commands
	2.5 hours – Friday
	<ul> <li>New model for methods (1) Used all methods</li> </ul>
	4 hours - Saturday
	<ul> <li>New model for methods (2) Used geologic models (not stats and ML methods)</li> </ul>
	3 hours – Sunday
	o Model 3 + notes
1/25	TOTAL HOURS
	1.5 hours Monday
	<ul> <li>30 min Monday 8am Join Microsoft Teams Meeting</li> </ul>
1	
	Notes, github commit, ml model
2/1	
2/8	Notes, github commit, ml model
	<ul> <li>Notes, github commit, ml model</li> <li>Monday <u>Join Microsoft Teams Meeting 9am</u></li> </ul>
2/8	<ul> <li>Notes, github commit, ml model</li> <li>Monday Join Microsoft Teams Meeting 9am</li> <li>Monday 8am Join Microsoft Teams Meeting</li> </ul>
2/8 2/15 2/22 3/1	O Notes, github commit, ml model  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am
2/8 2/15 2/22 3/1 3/8	O Notes, github commit, ml model  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15	O Notes, github commit, ml model  Monday Join Microsoft Teams Meeting 9am  Monday Sam Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am
2/8 2/15 2/22 3/1 3/8 3/15 3/22	O Notes, github commit, ml model  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15 3/22 3/29	Monday Join Microsoft Teams Meeting 9am  Monday Bam Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15 3/22	Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday 9am Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15 3/22 3/29 4/5 4/12	Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15 3/22 3/29 4/5	Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting  Monday Join Microsoft Teams Meeting  Monday 8am Join Microsoft Teams Meeting
2/8 2/15 2/22 3/1 3/8 3/15 3/22 3/29 4/5 4/12	Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting 9am  Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday 8am Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting Monday Join Microsoft Teams Meeting