EARTH SCIENCE ANALYTICS

GEUS Geological Survey of Denmark and Greenland

NTNU

PGS

DNO

Somehow Unsupervised

Dan Austin (Geo Data Scientist)
Steve Purves (R&D Director)
Daniel Coronel (Geophysics Master student)
Mads Lorentzen (Geophysics PhD student)
Kenneth Bredesen (Reservoir Geophysicist)
Laurent Olivier Feuilleaubois (QI Geoscientist)
Claire Sena (Reservoir Geoscientist)

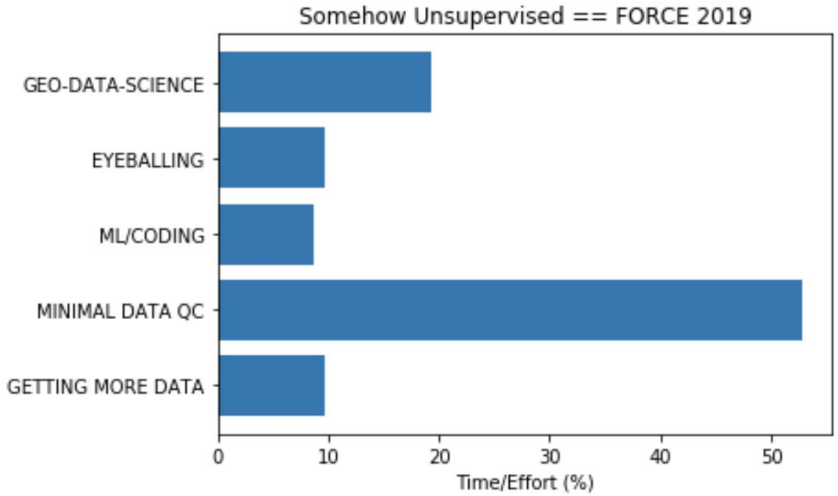
Questions We wanted to Answer

Unsupervised learning

VS

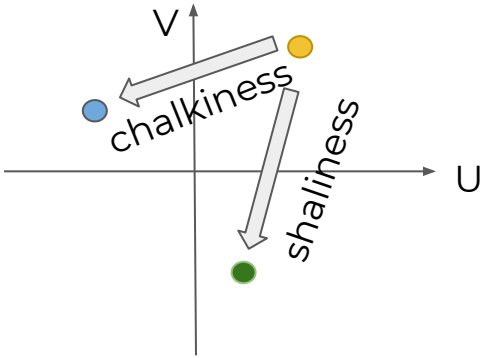
existing interpretation

- dataset > Logs + ~Interpreted Lith Labels
- Can we see depth trends?
- How confident are we in the lithology labels supplied
    - Relative and absolute calibrations
    - Sand reference points
    - Heterolithic classification
- What patterns can the machine see that humans may miss

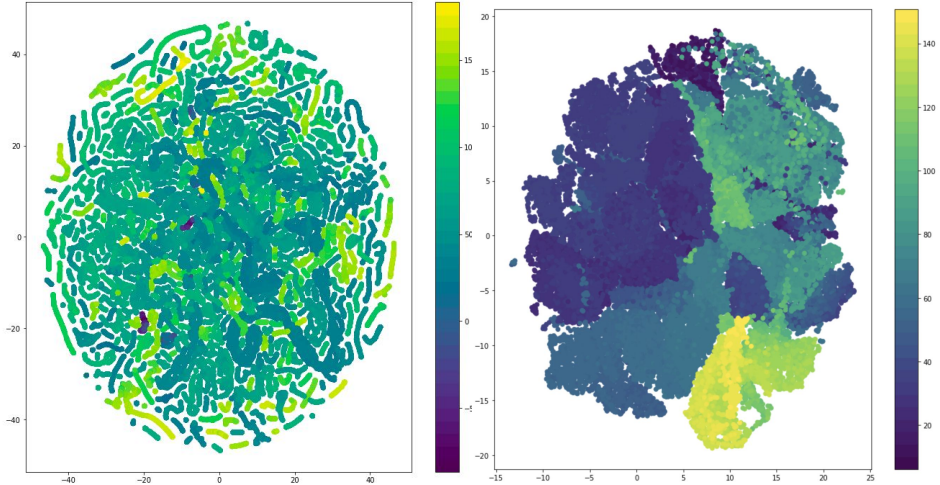


About Manifold Methods

- Dimensionality reduction
- Go from N dimensions (logs) to m (small)



- Resulting space unitless but may encode useful trends



TSNE UMAP

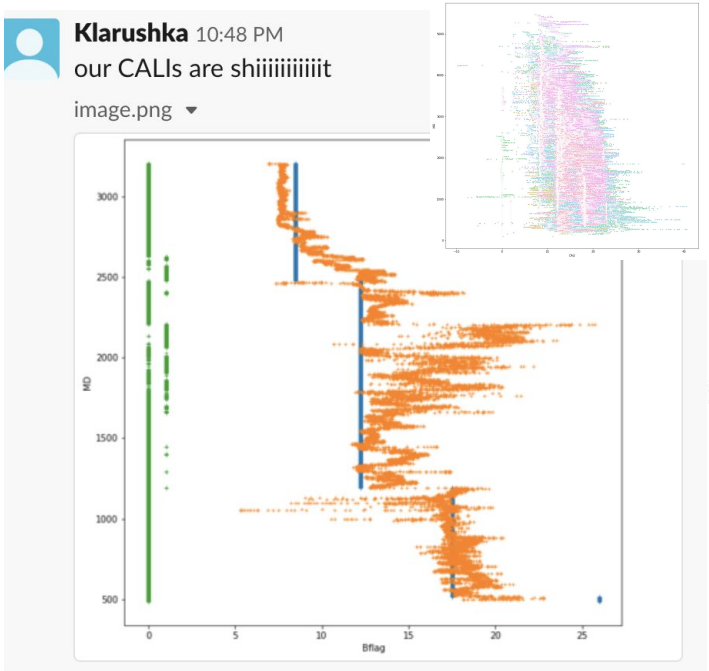
```
In [28]: n_neighbors= 8# default 15
         min_dist=0.99 # defult 0.1
         n_components=2
         metric='minkowski' #'euclidean' #'minkowski'
```

- Some big runs (100+ wells)
- Some small runs (Troll)
- Some parameter tuning

Data Selection and QC

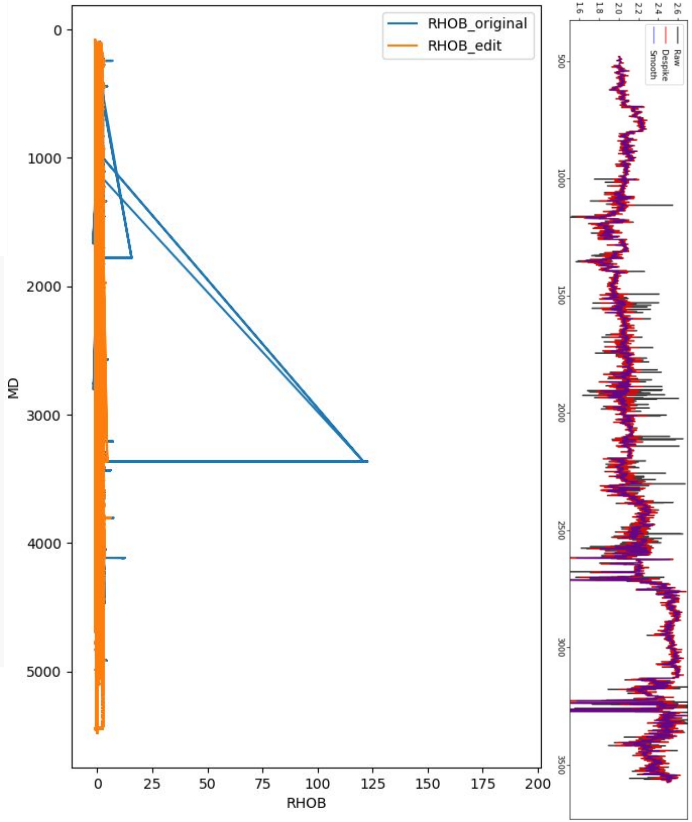
Dataframe - Data Presence

Flag - Caliper QC

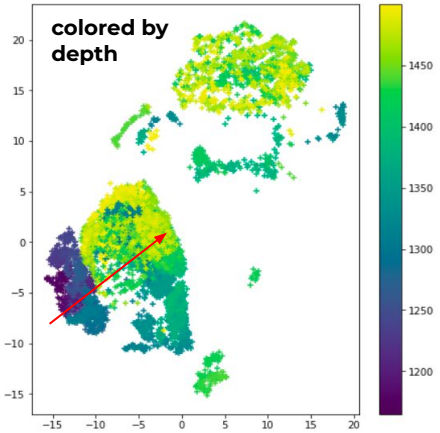


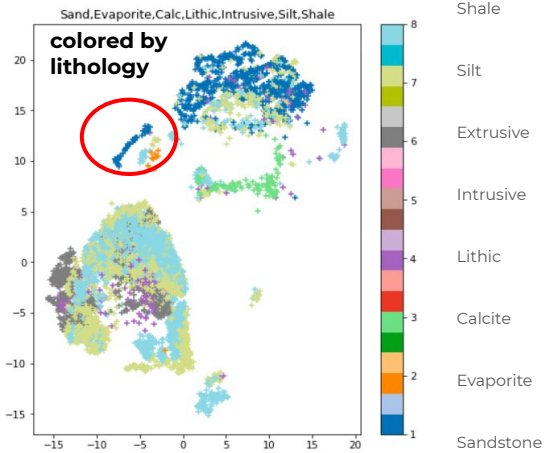
Bad hole flag
Borehole size
Caliper

Conditioning - Log De-Spike

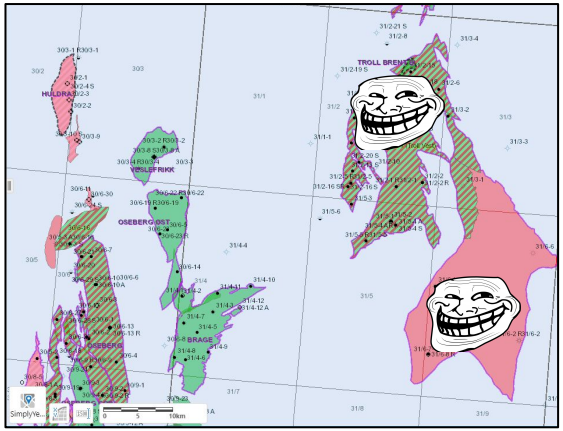


Lets Focus on Trolls....

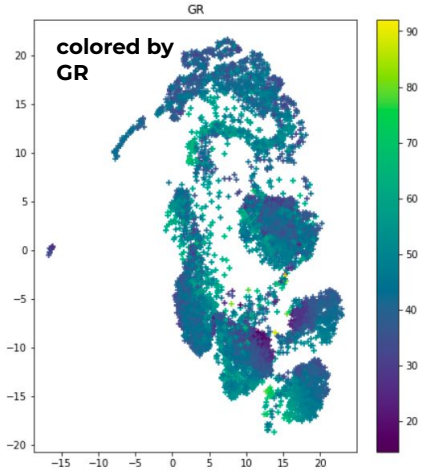


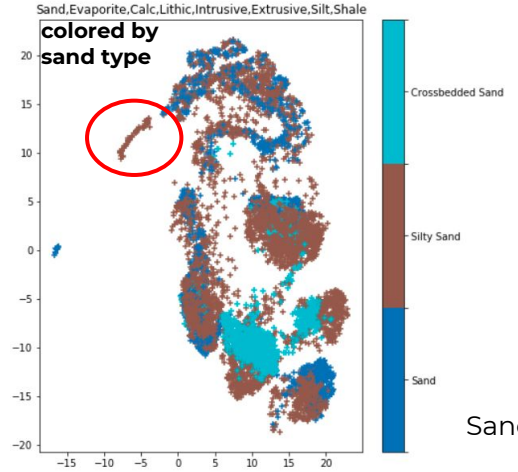


1000-1500 m

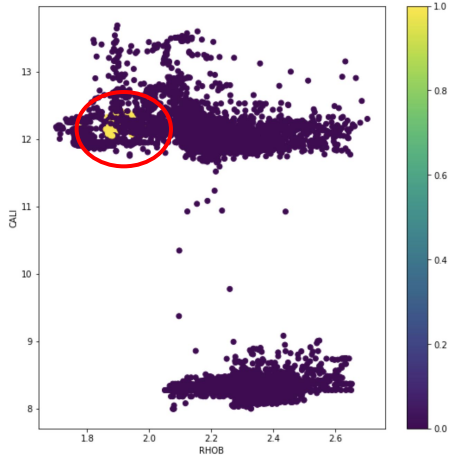


Density vs caliper colored by outliers



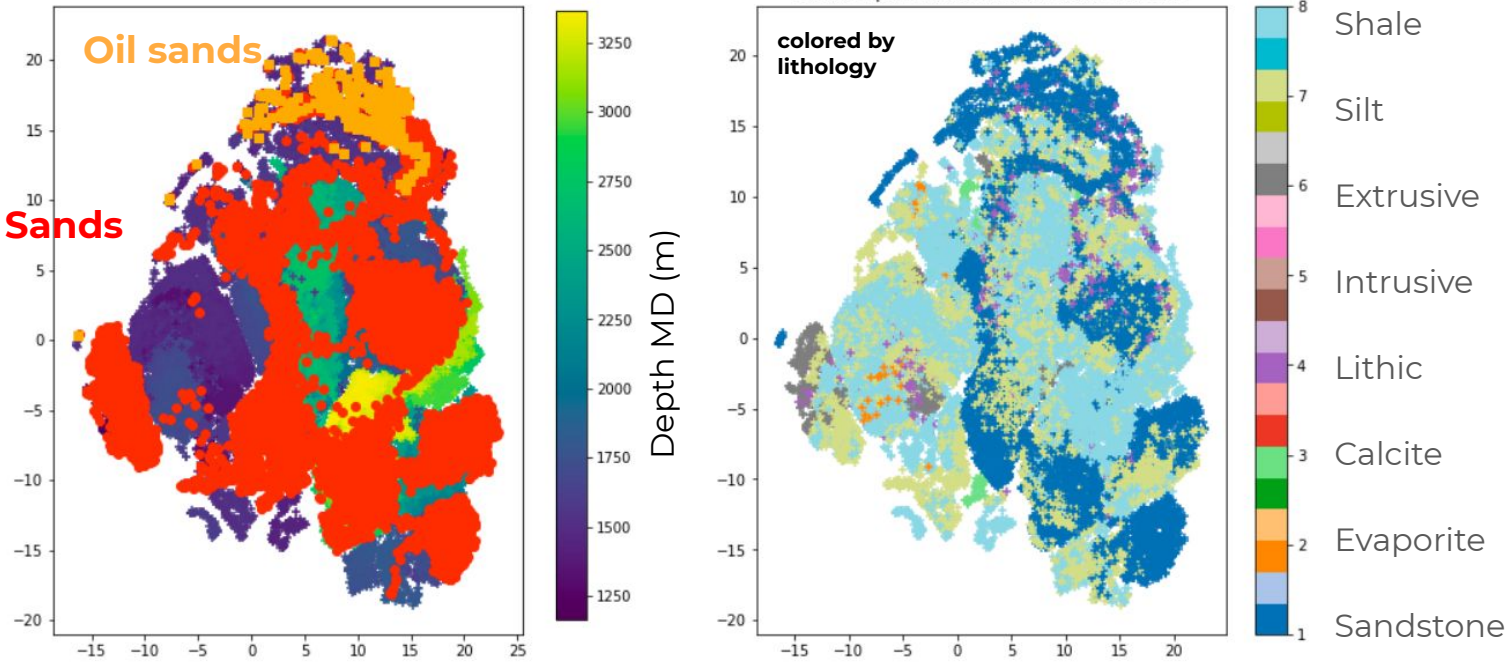


Sands all depths



All depths

UMap of Troll Reservoir Pay and not Pay



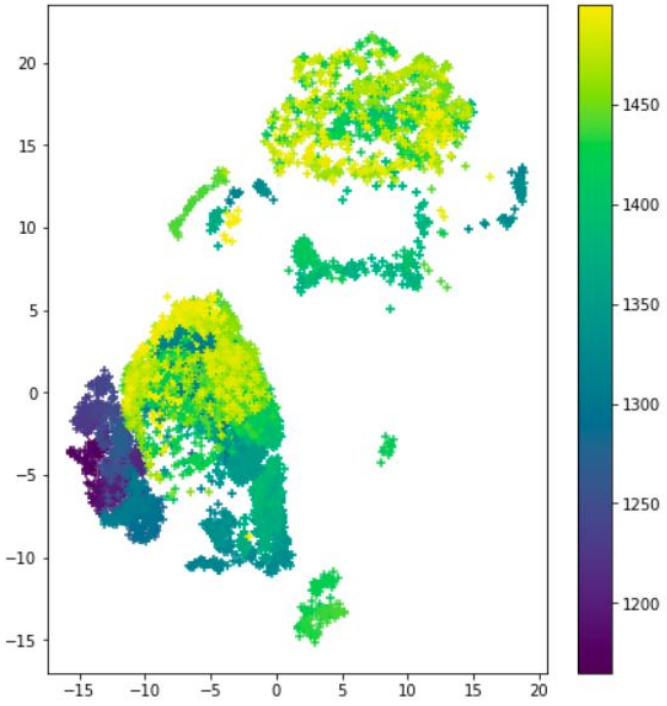
All depths

Conclusions and questions

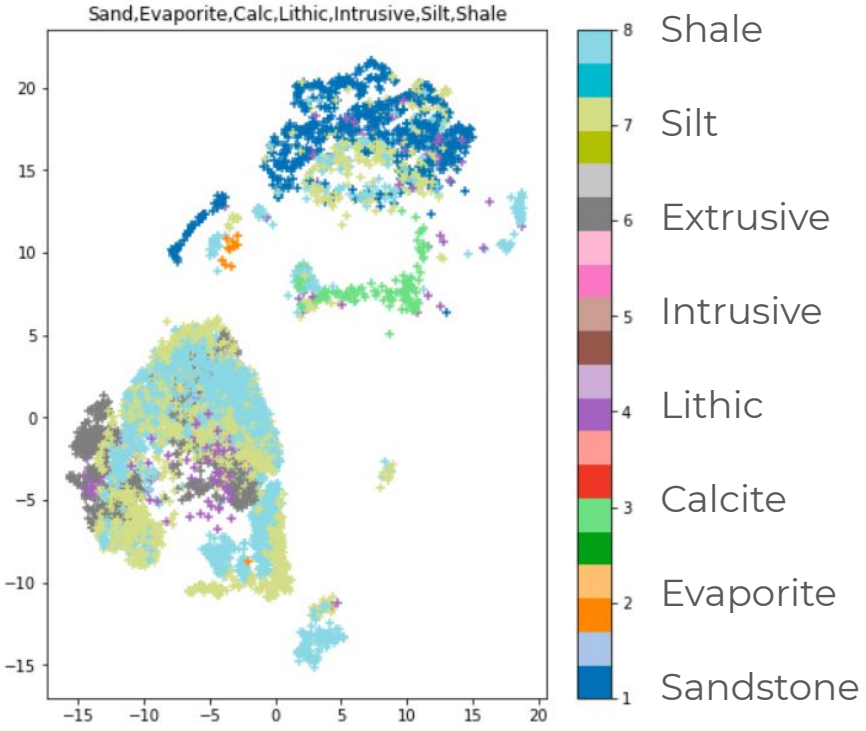
- Clustering is good for:
  - HC detection
  - Unsupervised mapping is effective at QC'ing the labels
  - Depth trends
- Reservoir quality sands form a clear cluster
- Become a troll hunter today!

<END/>

Looking at the Troll selected depth range



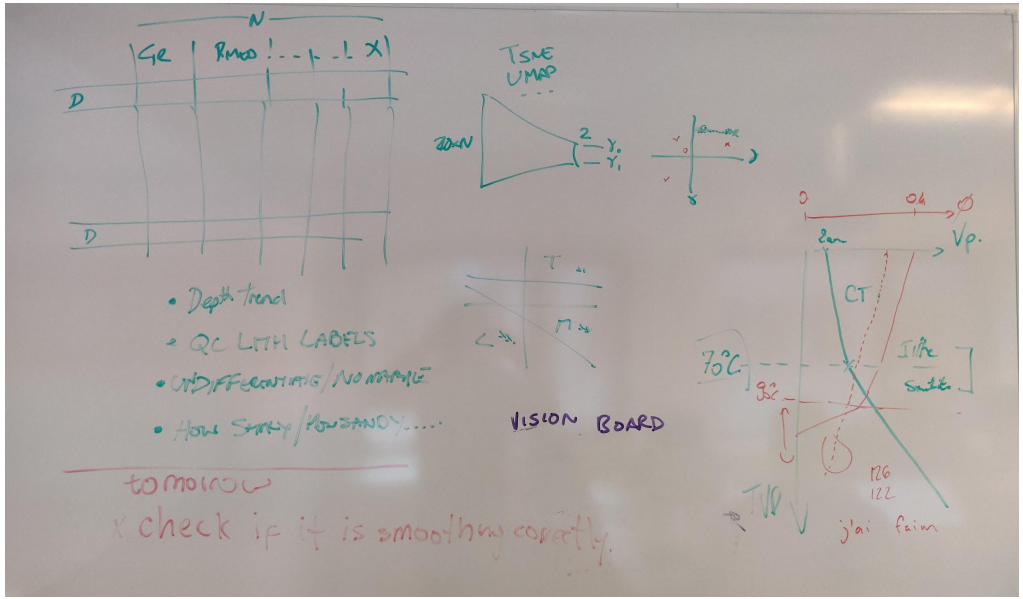
Colored by depth



Colored by lithology

We’re Going on a (data-driven) Adventure!

- Do Something cool with well data!
- Past hacks have messed around with dimensionality reduction and latent space stuff so let's try that...
- UMAP, TSNE, VAE- which will work and what can it solve
- Using GR and others to reduce the set and zooming in on particular intervals and lithotypes



The Data

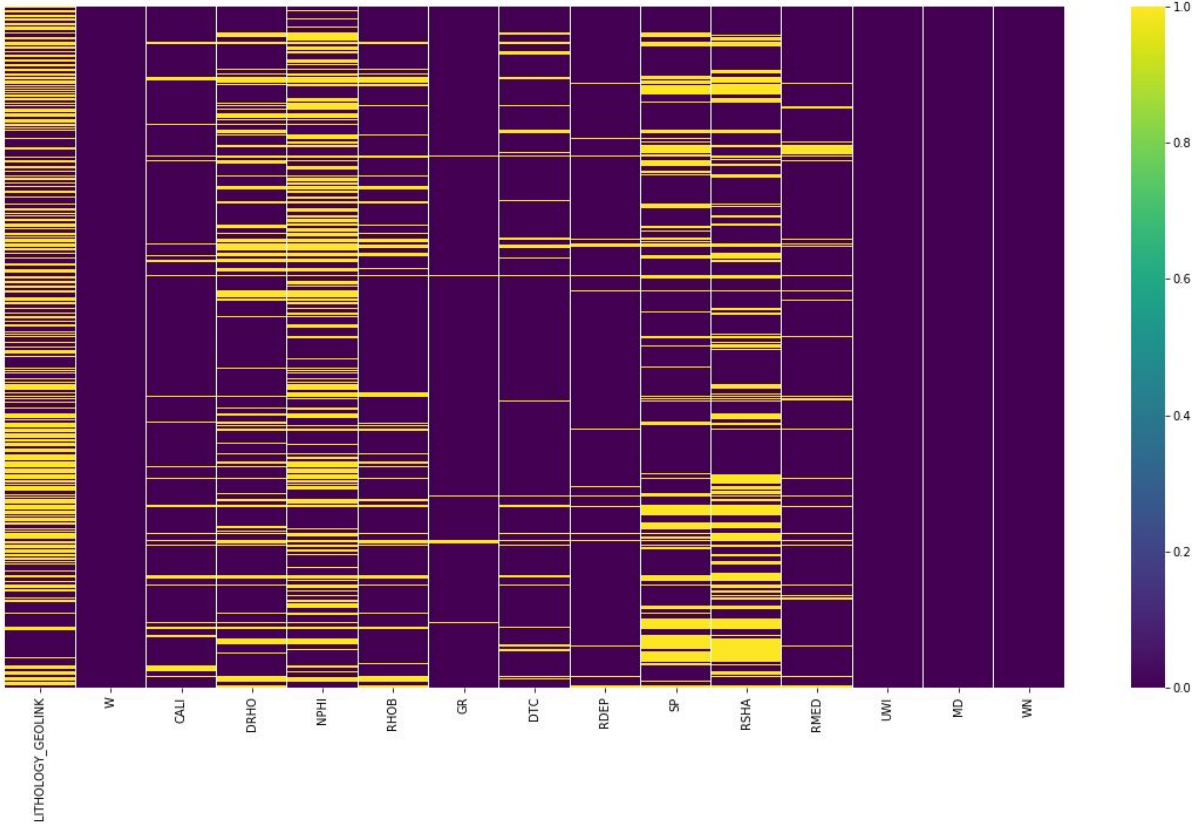
Idx	UWI	Data	Quality	LITHOLOGY_GEOLINK	GR	RDEP	RHOB	RMED	DTC	CALI	NPHI	
			%	223/223 wells	205/223 wells	204/223 wells	200/223 wells	200/223 wells	198/223 wells	196/223 wells	188/223 wells	18
0	15/9-12	13/13 curves	0	LITHOLOGY_GEOLINK 7.90 _	GR 44.66 gAPI	RDEP 3.02 ohm.m	RHOB 2.20 g/cm3	RMED 4.39 ohm.m	DTC 114.04 us/ft	CALI 13.65 in	NPHI 0.32 m3/m3	
1	15/9-13	14/14 curves	0	LITHOLOGY_GEOLINK 7.28 _	GR 63.68 gAPI	RDEP 1.86 ohm.m	RHOB 2.15 g/cm3	RMED 1.86 ohm.m	DTC 125.11 us/ft	CALI 14.77 in	NPHI 0.40 m3/m3	
2	15/9-14	14/14 curves	0	LITHOLOGY_GEOLINK 7.13 _	GR 53.61 gAPI	RDEP 1.87 ohm.m	RHOB 2.16 g/cm3	RMED 3.88 ohm.m	DTC 121.53 us/ft	CALI 14.11 in	NPHI 0.38 m3/m3	
3	15/9-15	16/16 curves	0	LITHOLOGY_GEOLINK 6.25 _	GR 58.82 gAPI	RDEP 1.47 ohm.m	RHOB 2.13 g/cm3	RMED 1.54 ohm.m	DTC 124.09 us/ft	CALI 13.97 in	NPHI 0.38 m3/m3	
4	15/9-17	17/17 curves	0	LITHOLOGY_GEOLINK 6.89 _	GR 57.18 gAPI	RDEP 1.78 ohm.m	RHOB 2.14 g/cm3	RMED 1.73 ohm.m	DTC 126.83 us/ft	CALI 13.56 in	NPHI 0.39 m3/m3	
5	15/9-18	14/14 curves	0	LITHOLOGY_GEOLINK 7.30 _	GR 56.01 gAPI	RDEP 180.95 ohm.m	RHOB 2.21 g/cm3	RMED 112.55 ohm.m	DTC 126.19 us/ft	CALI 13.56 in	NPHI 0.32 m3/m3	
6	15/9-19 A	7/7 curves	0	LITHOLOGY_GEOLINK 6.92 _	GR 47.15 gAPI	RDEP 35.95 ohm.m	RHOB 2.30 g/cm3	RMED 24.89 ohm.m		CALI 9.86 in		

- Approx 250 wells from the North Sea
- Tops and some lithology labels available
- Grungy and we know it!
  - Not all logs in all wells
  - Different log tools and types

Tactics and Strategy

- Collaborative QC
  - Huddle up and eyeball data with welly and other packages
  - Set up Repro and working environments
- Split into specific QC task teams
  - more on this later
- Advance ML party
  - Dumping data into premade tools to test process end to end
  - Pseudocode notebooks for task
- Discuss results and then work out what it all means!

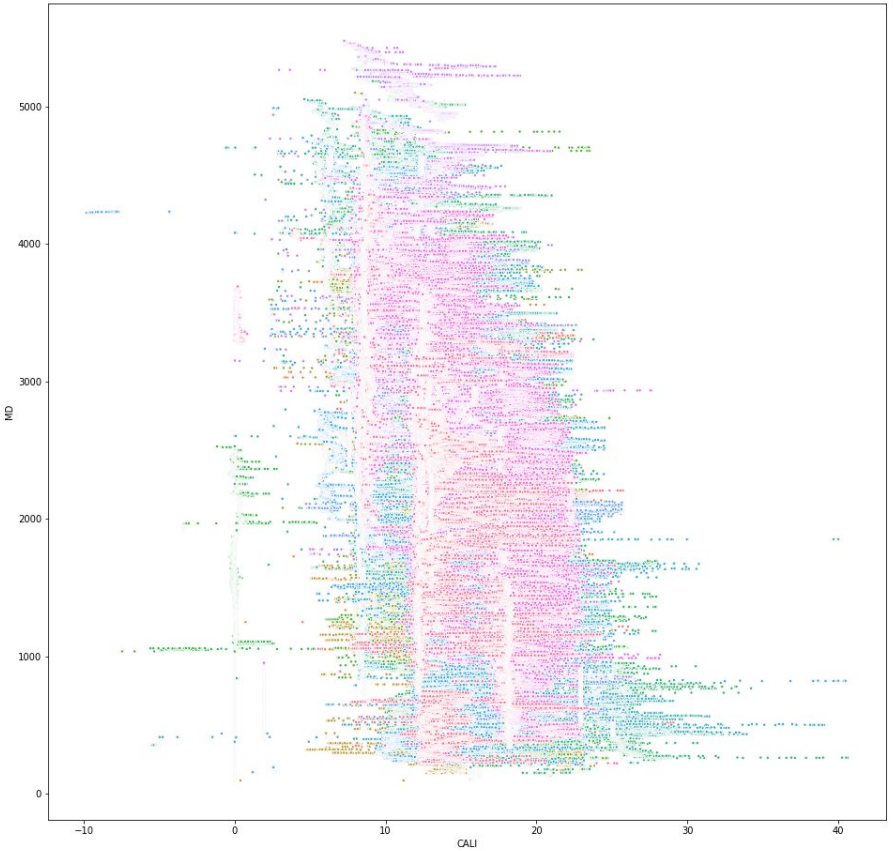
LAS Data with selected logs



- Yellow are NaN values
- Blues are non NaN values
- Note that Lithology labels have lots of NaN

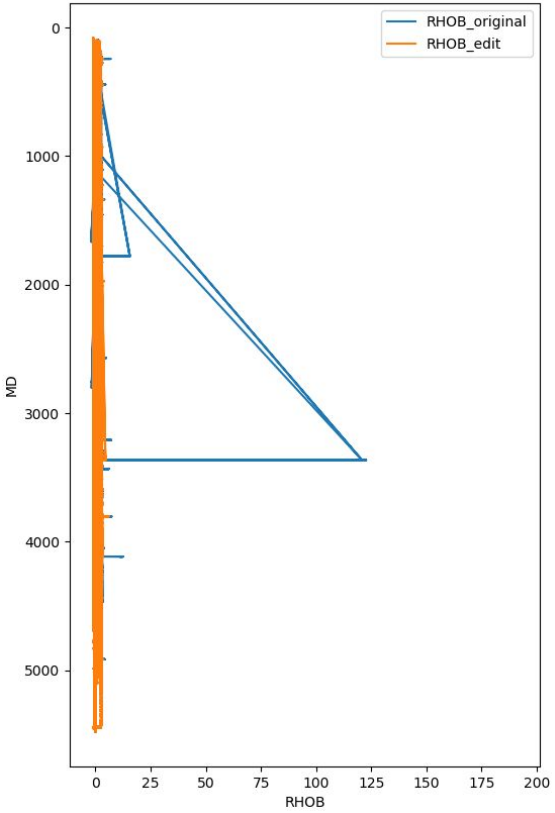
Caliper QC

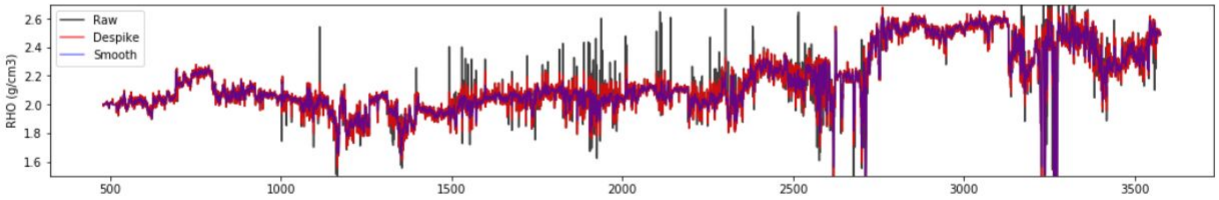
- Caliper in the LAS dataset plotted against MD and colored by well.
- Note the massive range in Caliper values that fall outside valid caliper values



Despiking QC

- Procedure: smooth (rolling window) → despike (Z score based)
- Density and DT despiking for all wells using a rolling window and Z-score.
- Problem with PHIE: fraction vs. percentage

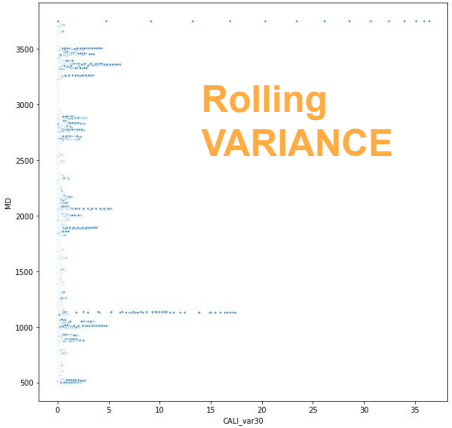


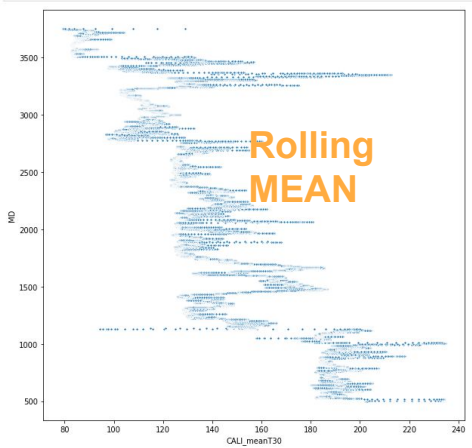


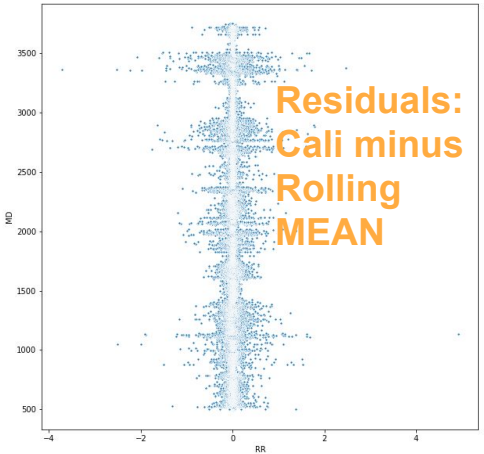
Trying to clean the Caliper

1. Tried to clean the Caliper data using the ***Variance of a rolling window*** and using ***residuals of the deviation to the mean*** on a rolling window
2. Failed! After testing different window sizes, we could not get a sensible threshold to filter the caliper based on variance or residuals

Showing data for one well only





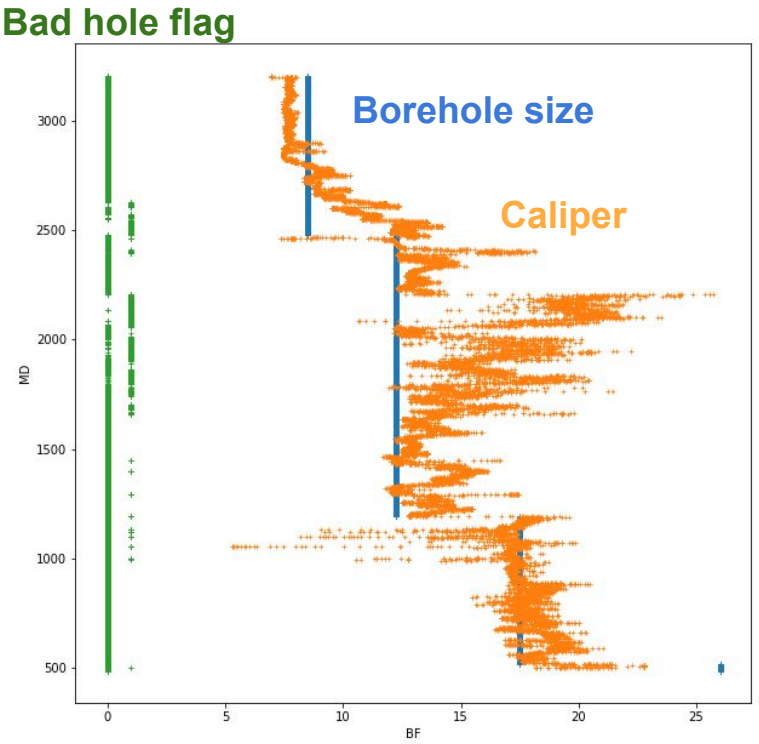


Data mining for casing data - clip the Caliper (finally!)

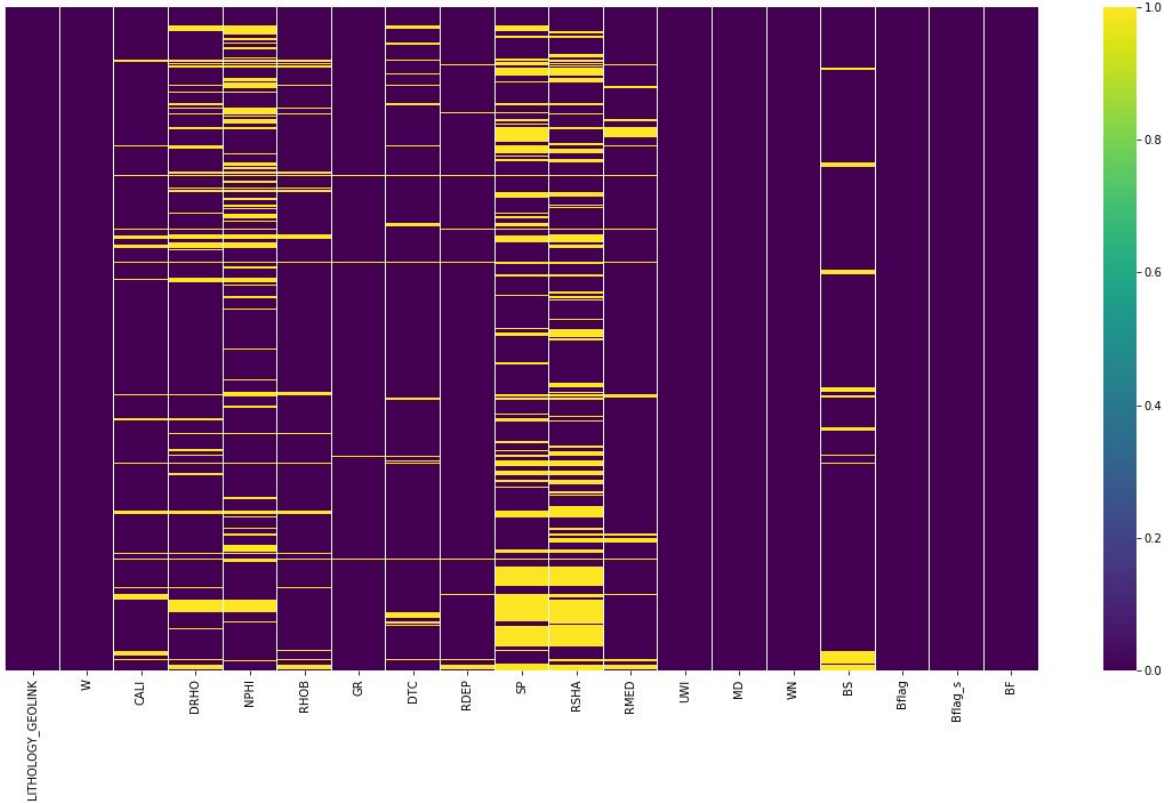
1. Had to find borehole size and borehole size depth online.

2. Created a Bad hole flag, defined by the Caliper being 30% bigger or smaller than the borehole size

3. Then, clip the Caliper where there is bad hole



Subset of LAS data only for the reservoir with no bad hole

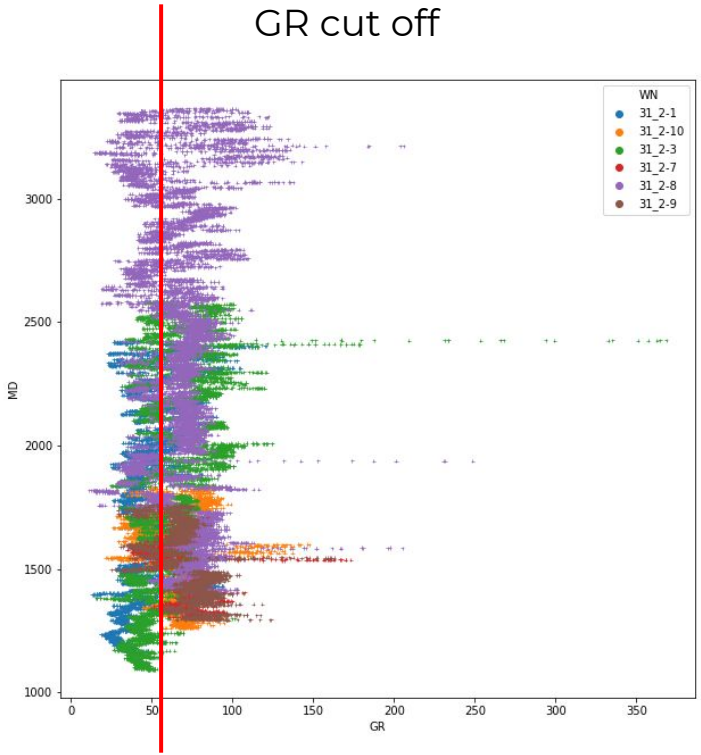


1. Yellow are nan values
2. Blues are non nan values
3. Note that the Lithology log exists for every row

USED AS TRAINING SET

Our awesome sand

GR cut off



Density cut off