Hydraulic Fracturing Optimization:
Predicting Well Production

Katie Peterson
Final Capstone
July 2018

Business Problem

Hydraulic fracturing is used after the drilling of an oil or gas well is complete.

Hydraulic fracturing uses fluid and other materials to create small fractures in a rock formation in order to stimulate the production of oil and gas from the well.

There are no best practices in what fluid systems are used to optimize well production.

Diagram: Greater Lansing Business Monthly

Research Question

What well and frac characteristics are determinant for higher production of the well?

- Fluid Systems:
  - Slickwater
  - Linear gel
  - Cross-linked gel

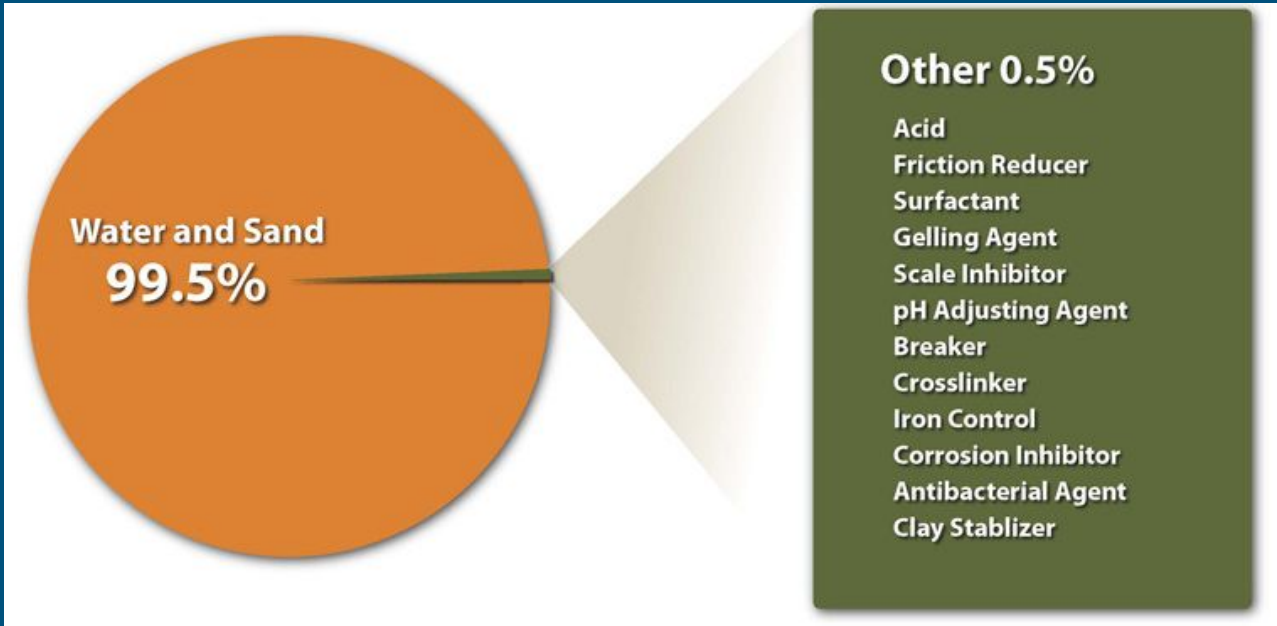


Diagram: EERC

Data Sets

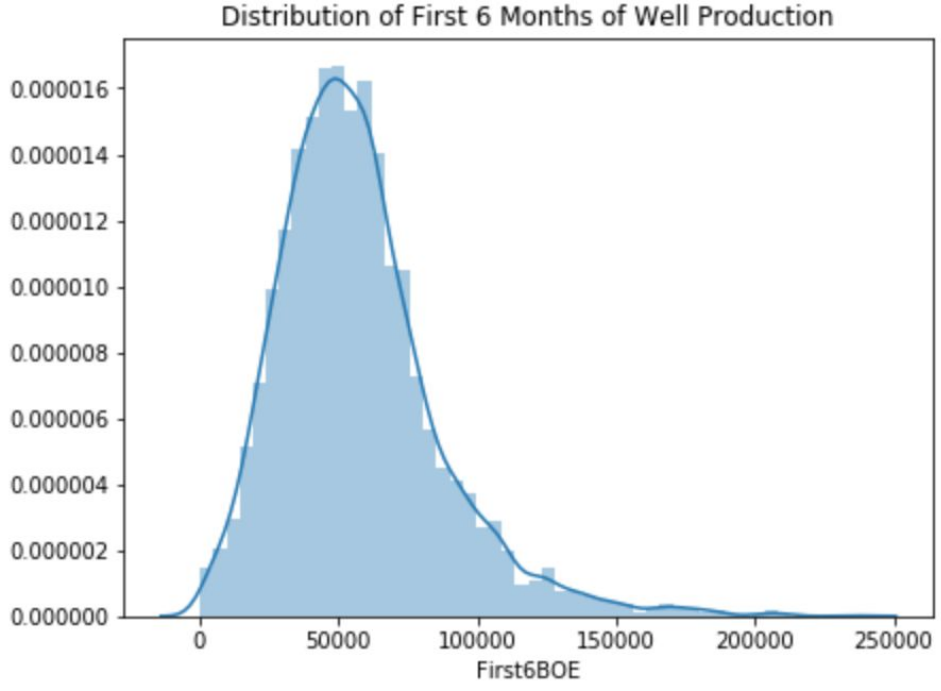
- DrillingInfo: Proprietary site with well and production information on all wells, state by state
  - Colorado: 15K total wells, 5K actually producing
  - 99 Columns
- FracFocus: Public database with all frac fluid additives for all wells across the United States
  - 3.2 Million different ingredients for ~90K wells
  - 39 Columns

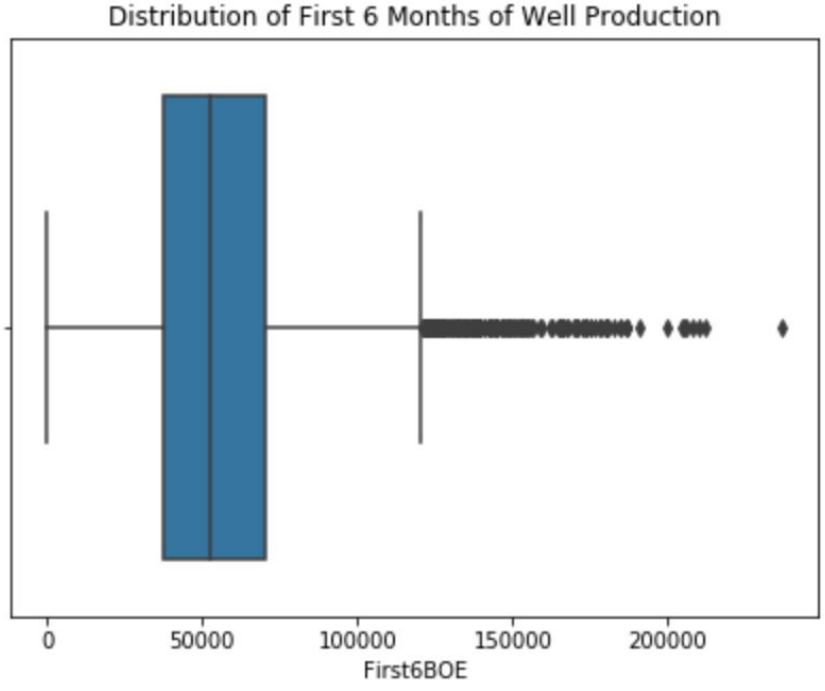
Logos: DrillingInfo.com, Fracfocus.org

Data Cleaning

- Identifying relevant features from well information
  - Imputing missing data (length)
- Cleaning FracFocus Data
  - Categorizing ingredients as frac fluid systems
  - Calculating sand mass from ingredient information
  - Imputing missing data (sand mass)
- Joining two data sets together based on well identifier (API number)

Target: First 6 Month's Production

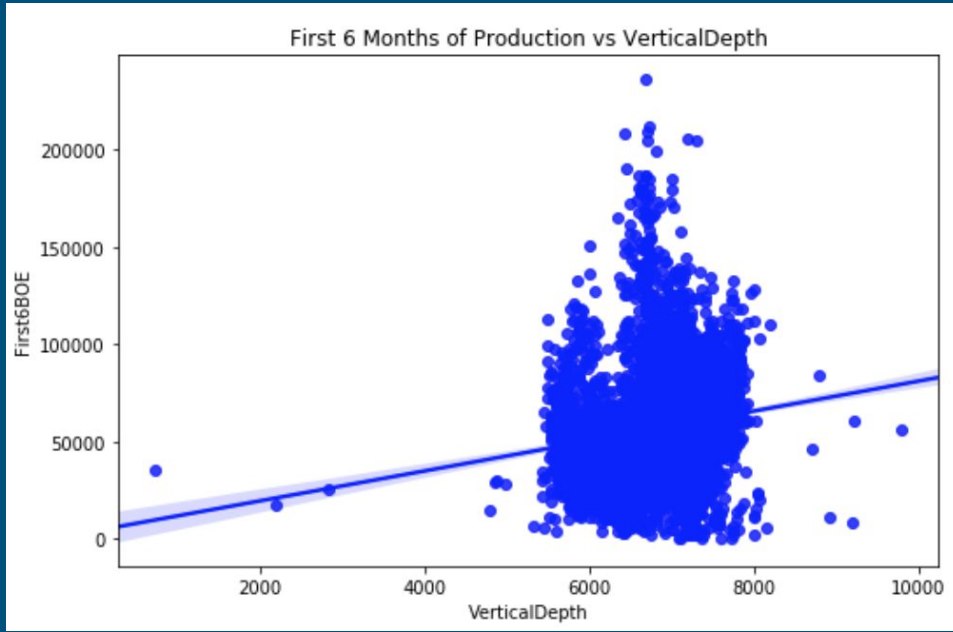




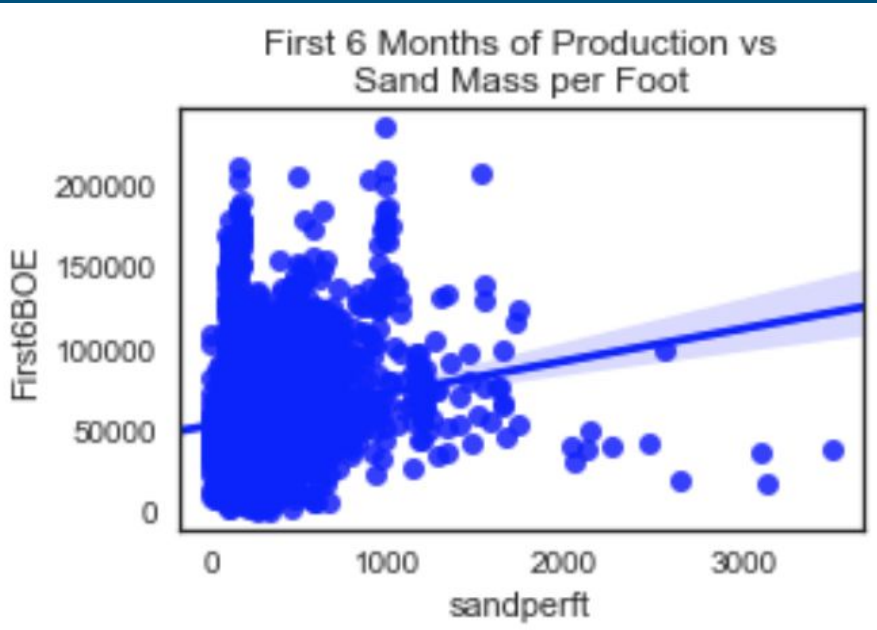
Features

Well Characteristics

Vertical Depth

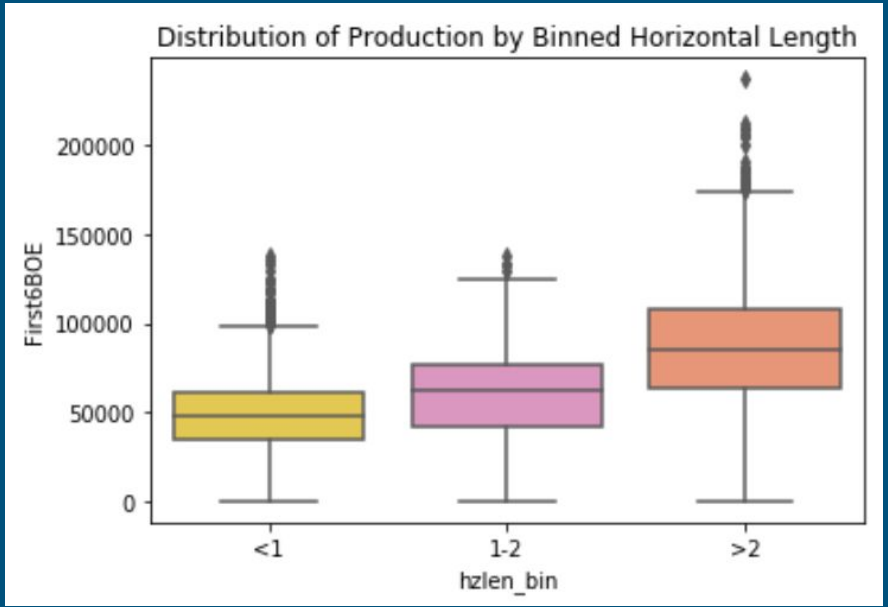


Sand Mass per Foot

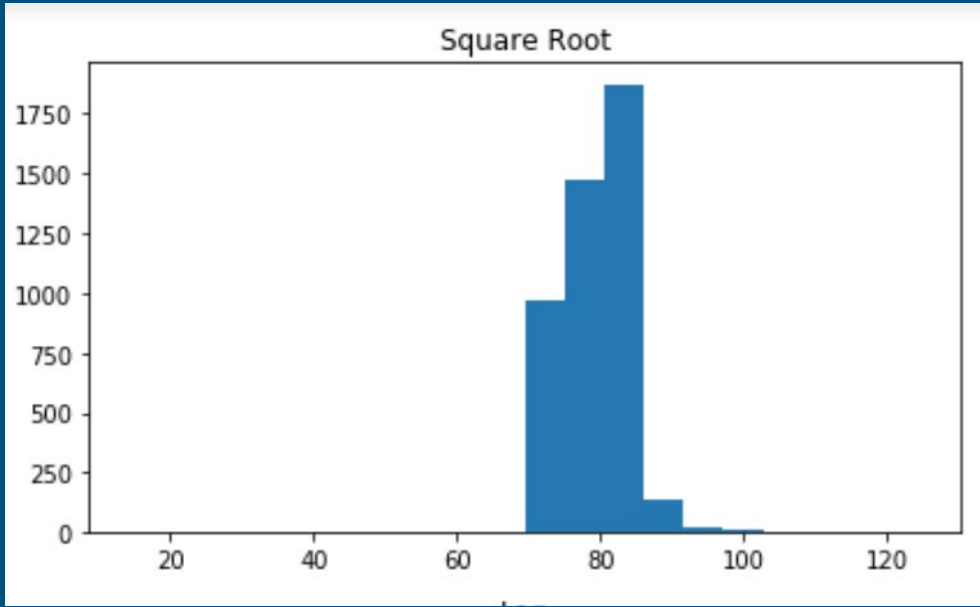


Well Characteristics

Binned Horizontal Length

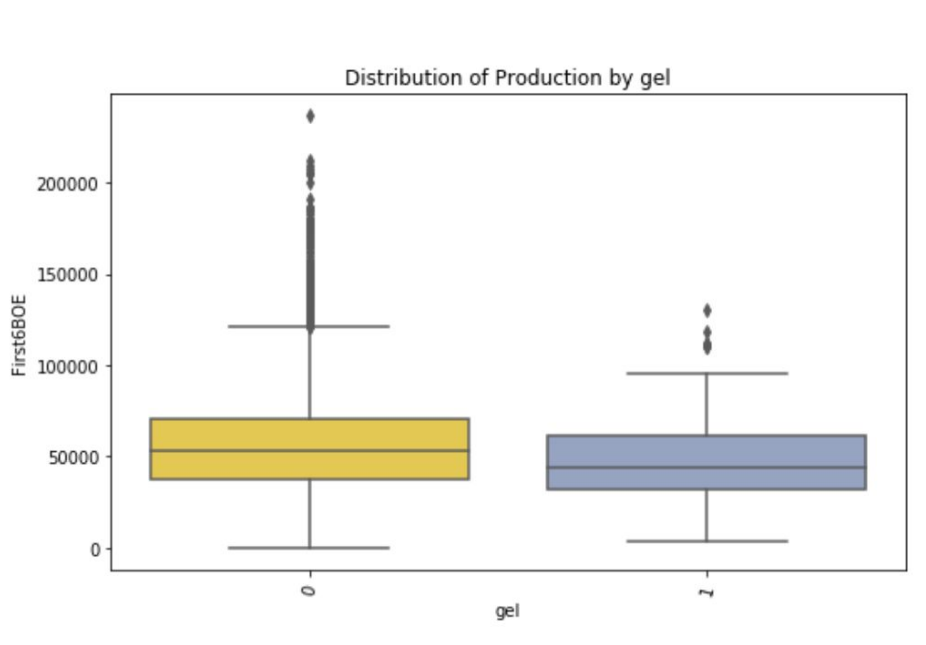


Square Root of Non-Perforated Feet

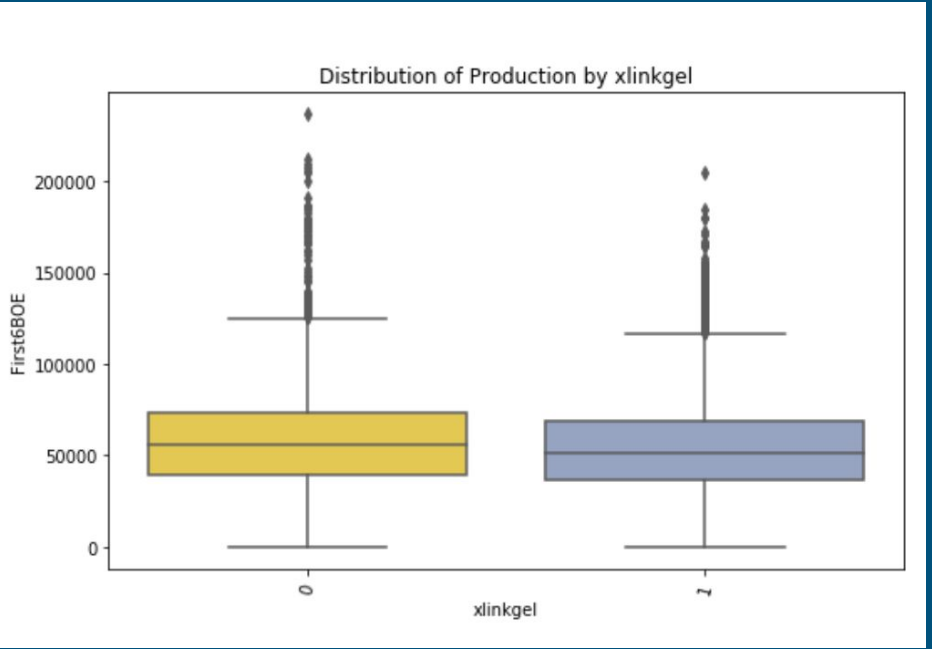


Frac Fluid System

Linear Gel

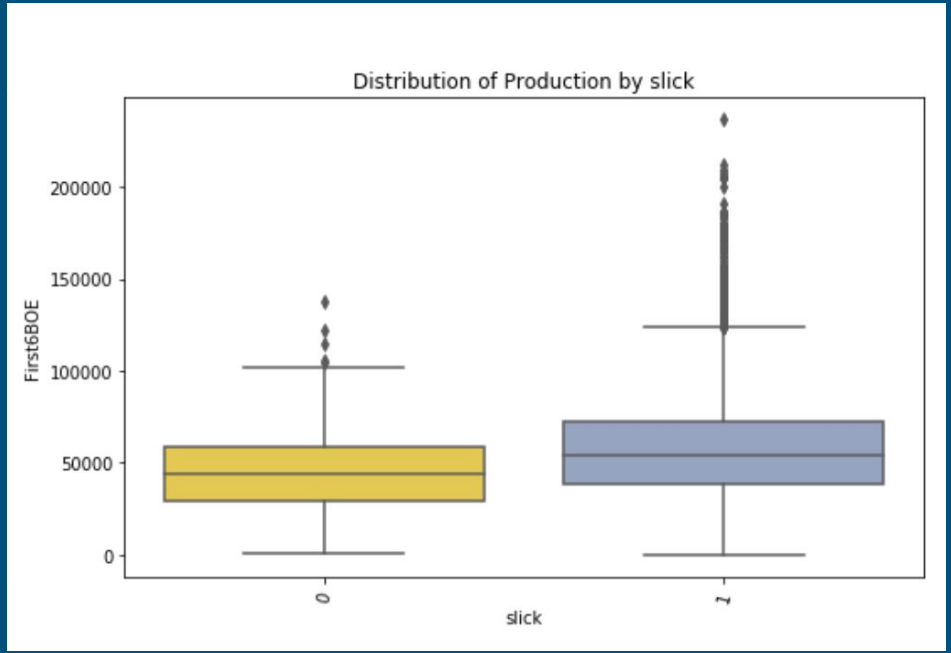


Cross-linked Gel

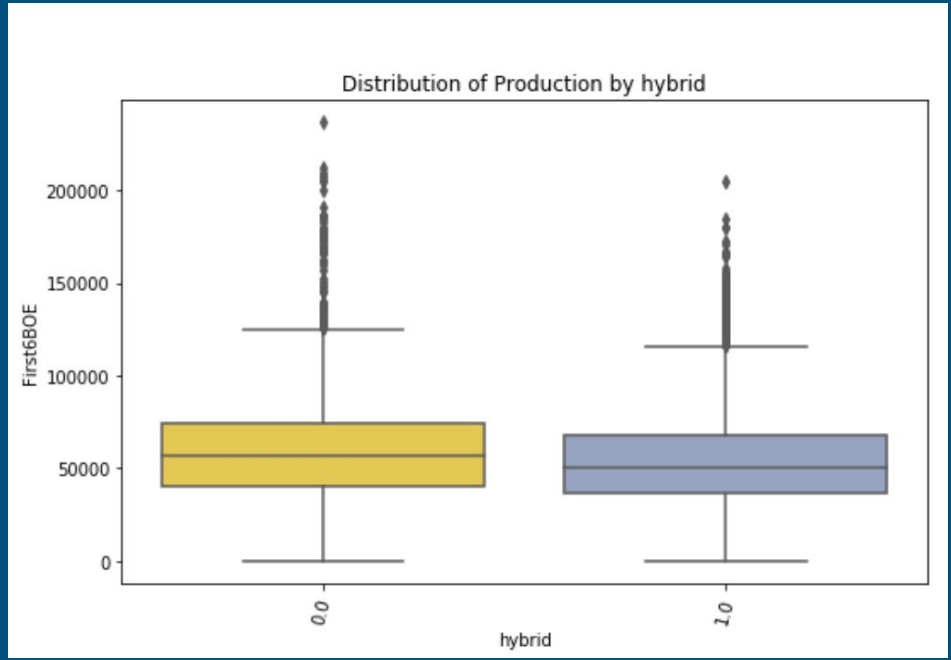


Frac Fluid System

Slickwater

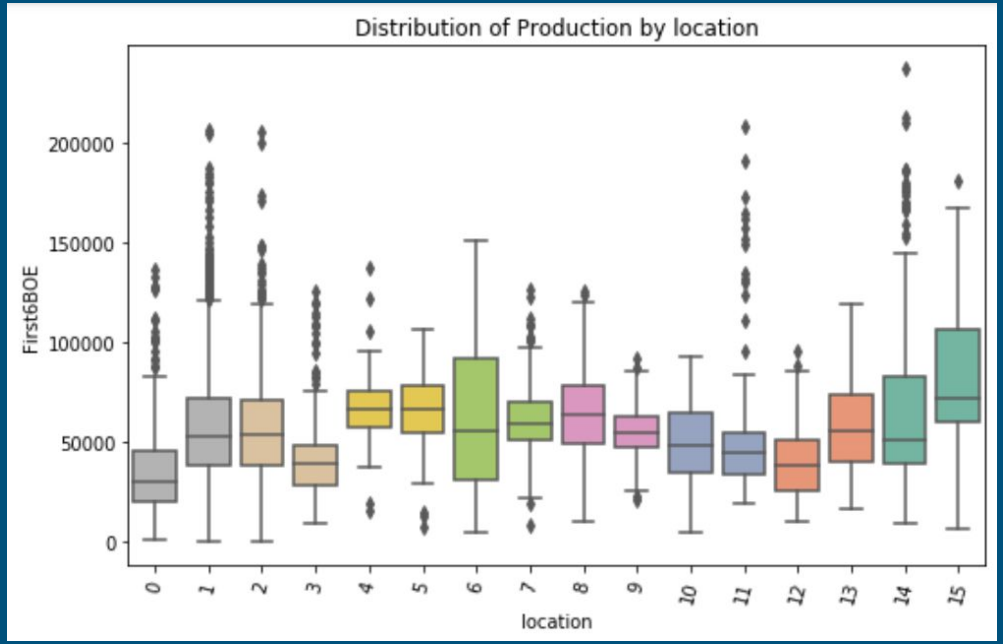


Hybrid Approach



Norming for Location

Clustering using KMeans to normalize for County, Township, and Range



Modeling

R-squared values

Random Forest Regression	0.5353 (+/- 0.1267)
Random Forest Quantile Regression	0.5306 (+/- 0.1336)
Gradient Boosting Regression	0.5040 (+/- 0.0959)
Linear Regression	0.3236 (+/- 0.0785)

Overall Feature Importance

Most important features in predicting First 6 month's production:

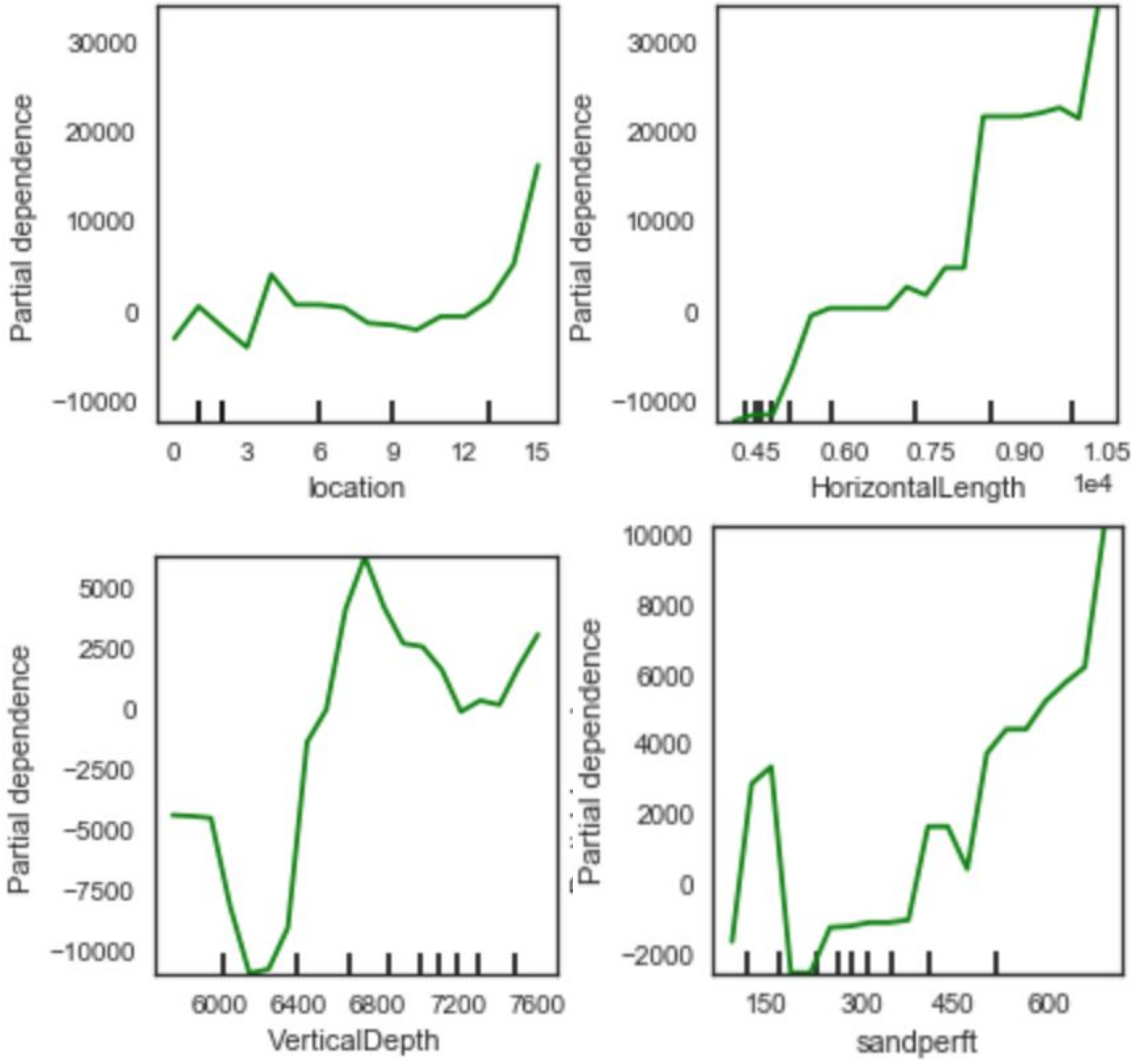
- Vertical depth
- Horizontal length
- Sand mass per foot
- Perforated interval
- Location
- Frac fluid system*

	importance
VerticalDepth	0.207645
hzlen_bin_>2	0.191659
sandperft	0.167481
hzlen_bin_<1	0.115931
nphf_sqrt	0.115197
location	0.105148
hybrid	0.035178
hzlen_bin_1-2	0.028181
xlinkgel	0.026348
gel	0.005019
slick	0.002213

Model Interpretation

Higher odds of production increases with:

- Increased horizontal length
- Vertical depth around 6600 ft
- 150 or above 400 pounds of sand per foot of casing



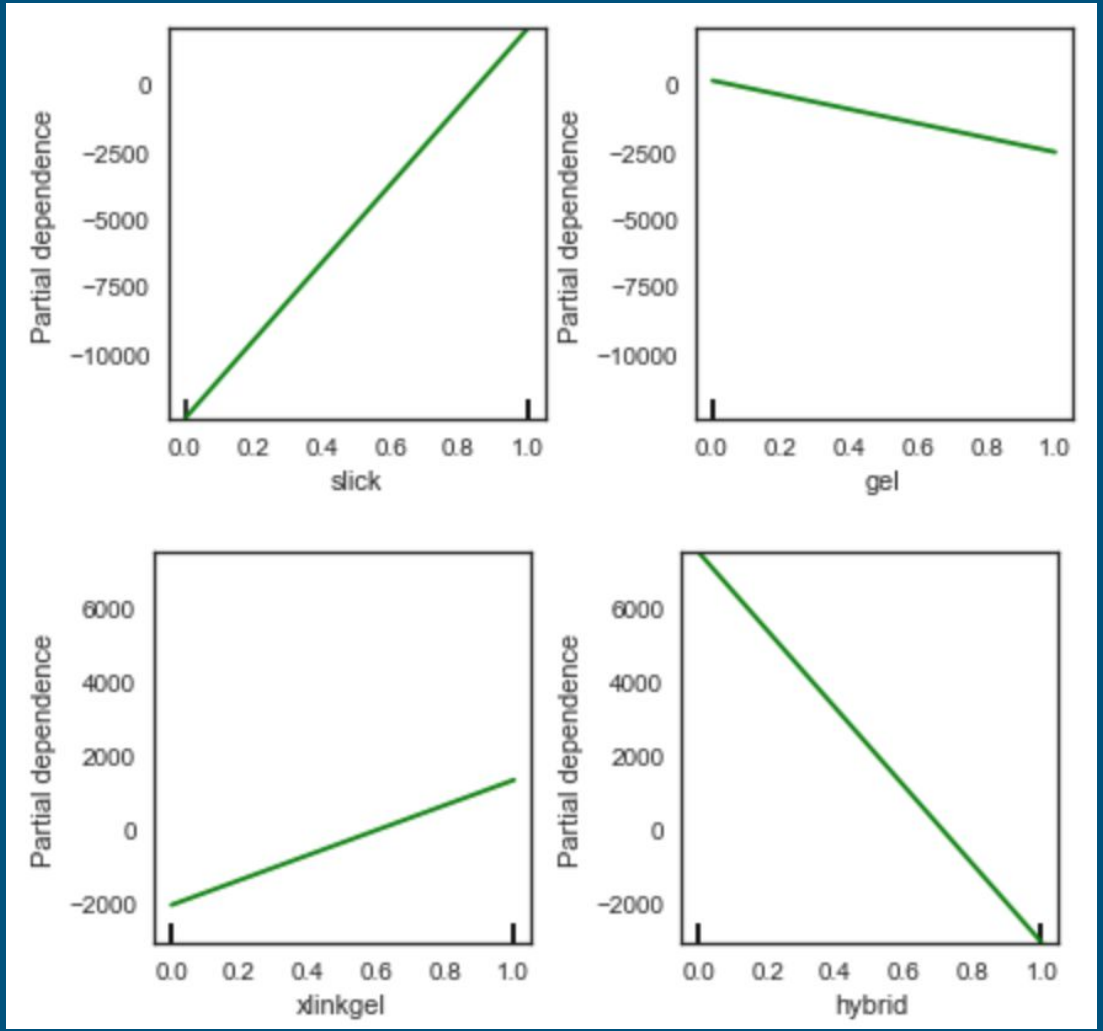
Model Interpretation

Higher odds of production increases by using:

- Slickwater system
- Crosslinked gel

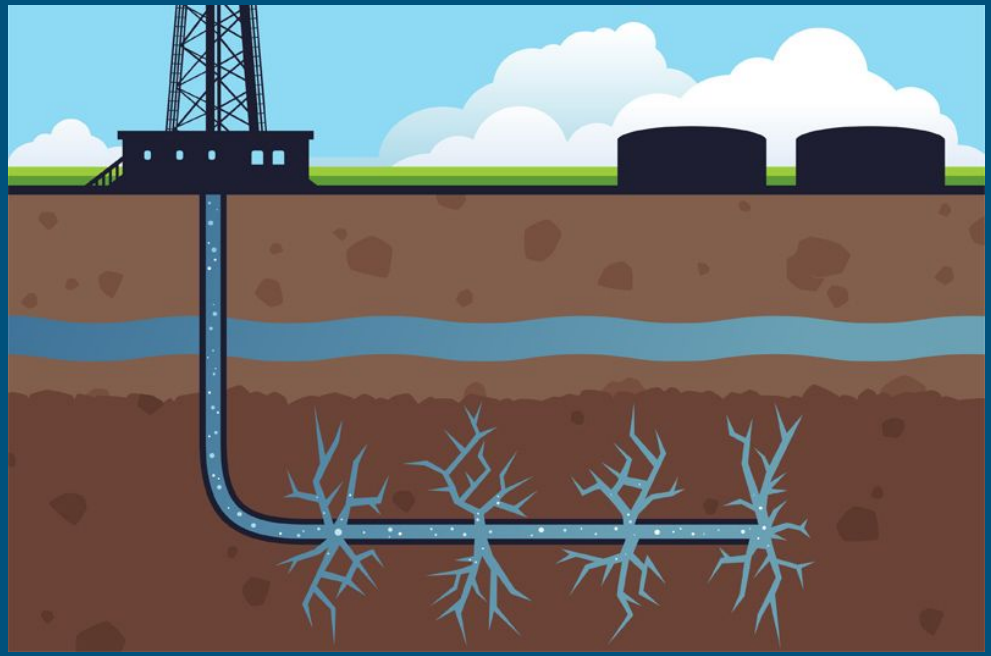
Lower odds of high production by using:

- Linear gel
- Hybrid system



Product: Frac Fluid System Optimization

Product Explanation



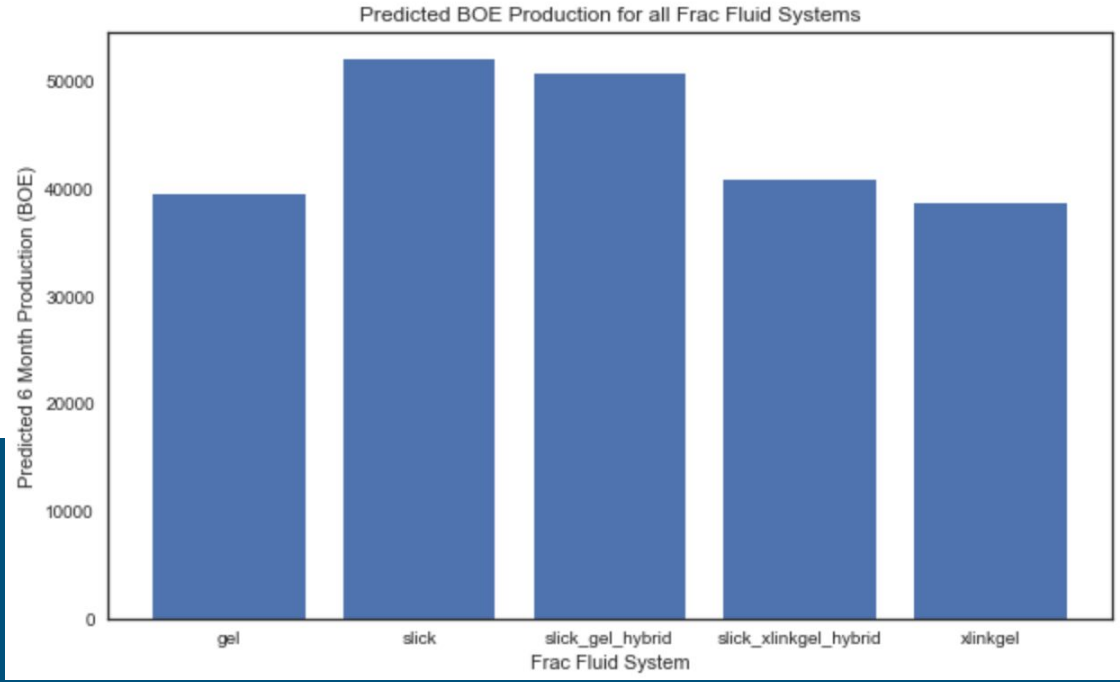
Utility for Well Operators

- After drilling a well:
  - in a given location,
  - with a known vertical depth, and
  - horizontal length
- Select:
  - an amount of length to perforate, and
  - a mass of sand to use per linear foot.
- Finally:
  - model the predicted production based on all frac fluid systems to decide which system to use in hydraulic fracturing

Diagram: World Wide Metric

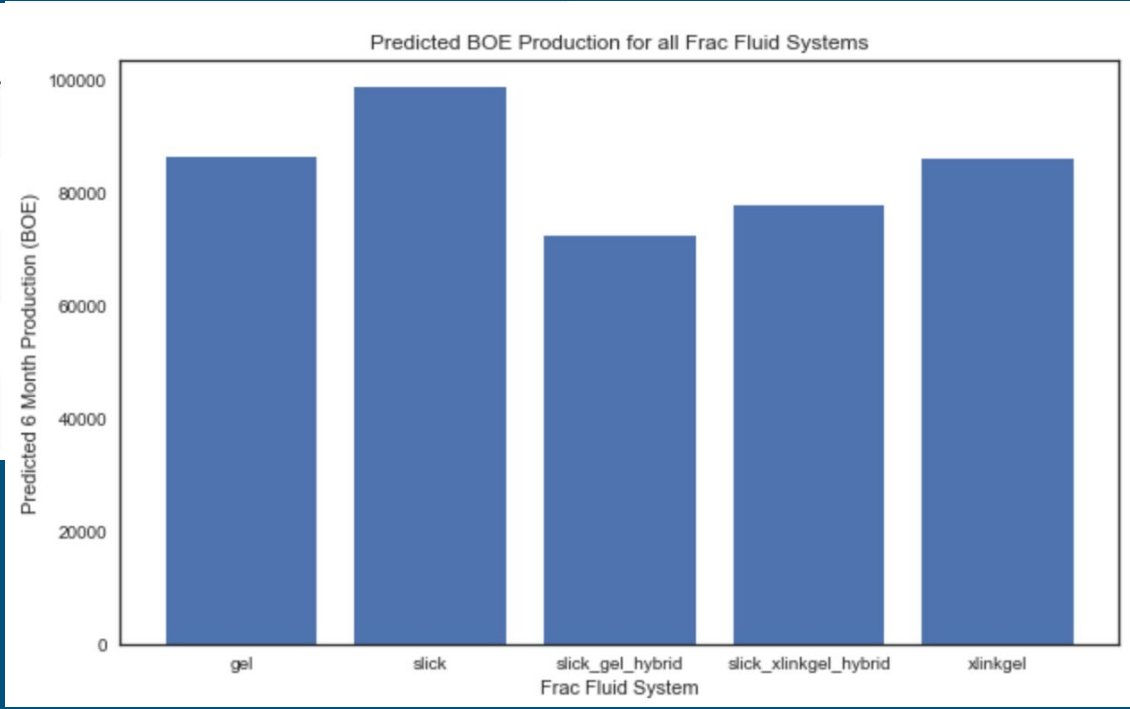
Product Demonstration - Well 1

Fluid_System	Predicted_Production(BOE)
slick	51951
slick_gel_hybrid	50683
slick_xlinkgel_hybrid	40718
gel	39395
xlinkgel	38596



Product Demonstration - Well 2

Fluid_System	Predicted_Production(BOE)
slick	98773
gel	86501
xlinkgel	85995
slick_xlinkgel_hybrid	77810
slick_gel_hybrid	72537



Conclusions

Model Trends

- Drilling
  - Optimal vertical depth around 6600 ft
  - Increase horizontal length*
- Frac Conditions
  - Optimal sand use around 150 pounds per foot and above 400 pounds per foot
- Frac Fluid System
  - Improved predicted production using slickwater and hybrid systems

Final Thoughts

- Model upkeep
  - Update the model with most recent production information and ingredients
- Opportunities for further exploration:
  - Add in costs for each feature, including current BOE rates to determine cost/benefit of each decision
  - Add in a more detailed ingredient analysis for each fluid system and type of sand
  - Streamline data collection and reporting for greater data consistency

Thank you!

Thank you to my mentor, **Max Sop**, for being a supportive thought partner and providing valuable insight into the real world of data science. Special thank yous to **Jeffrey Beunier** for providing access to the data, and to **Ryan Dornbos** for his insight on COGCC scout cards. Finally, I have extra special appreciations to **Andrew Peterson** for his passion, support, and innovative thinking that brought this project to fruition, to my mother, **Suzanne Peterson**, for enduring 21 hours of labor until being graced with my presence, and to **Miles Shuck** for his encouragement, support, and understanding through this whole program.