# Synthesis of Denver-Julesburg Basin Production (1999-2019)
# A Data Science Approach

## Abandonment Rate Prediction
## Modelling Production Likelihood In a Well's Lifecycle

A Capstone Project in Fulfillment of
Springboard's Data Science Career Track Program

By


Laura Elliott

And In Association with

Ajith Patnaik

Springboard Mentor

# Table of Contents

# 1. ABSTRACT / PROBLEM STATEMENT

The Denver-Julesburg (DJ) Basin is a mature oil and gas producing province centered in southeastern Wyoming and northeastern Colorado. The earliest production occurred in 1901 with the drilling of the McKenzie well which led to the discovery of the Boulder Oil Field in Boulder County Colorado (Wikipedia). The Wattenberg Field, discovered in 1970, is one of the top 10 largest oil and gas fields in the United States, according to the Energy Information Agency (Sonnenberg, 2019). It has been revitalized in the last ten years by the success of horizontal production in the Codell and Niobrara Formations. However, weak commodity prices and investment uncertainties over the past few years coupled with increasing costs of collecting and/or subscribing to geological, geochemical, and engineering datasets and software have put operators under increasingly difficult financial pressures.

A great deal of perspective can be gained by examining the plethora of free publicly available data and by using open source software. Such questions as the average well lifecycle span, the most productive target formations by location, the optimal lateral length (in the case of horizontal wells) and spacing, and expected ultimate recovery prediction can be ascertained. In this paper, the author uses a 'top-down' approach based on historic well production behavior to investigate whether it is possible for a given well A and for a given point X in its life cycle, to predict the probability of its permanent abandonment in the next N years. Only publicly available data sources are used, primarily data from the Colorado Oil and Gas Conservation Commission (COGCC). This approach is in contrast to more traditional techniques using subscription and proprietary data.

Python APIs, geospatial and statistical analysis, and machine learning techniques are used to synthesize, classify, and predict horizontal and vertical well performance in the basin. The approach employs the building of segmentation use cases, in which geo-located data points are generated that describe and classify the well's behavior. The segments are created based on economic, geologic, and engineering criteria. These segments can then be used as either ranked categories and/or further used to synthesize, classify, and predict horizontal well performance in the basin.

For those prospectors working the area with proprietary data available to them, insight gathered from this type of analysis can be used as a complement to traditional methods, providing an independent, higher level look at the basin to optimize development strategy and recovery enhancement.

### Data Description

The primary data for this study were obtained from the Colorado Oil and Gas Conservation Commission. These data were downloaded via the following link:

https://cogcc.state.co.us/data2.html#/downloads

Production reports ranging from 1999 to 2020 are available in comma separated format. See listing below:

### Production Data ▾

**Production Summaries ▸**

**All Production Reports Received By Year ▾**

Data Dictionary for Production Files

| | | |
|---|---|---|
| **1999 Production Report:** | Zip file (3.62 mb) | CSV file (32.3 mb) |
| **2000 Production Report:** | Zip file (8.49 mb) | CSV file (66.0 mb) |
| **2001 Production Report:** | Zip file (7.05 mb) | CSV file (55.1 mb) |
| **2002 Production Report:** | Zip file (9.93 mb) | CSV file (70.7 mb) |
| **2003 Production Report:** | Zip file (13.8 mb) | CSV file (95.2 mb) |
| **2004 Production Report:** | Zip file (11.1 mb) | CSV file (82.5 mb) |
| **2005 Production Report:** | Zip file (12.4 mb) | CSV file (86.5 mb) |
| **2006 Production Report:** | Zip file (13.1 mb) | CSV file (91.0 mb) |
| **2007 Production Report:** | Zip file (13.9 mb) | CSV file (97.0 mb) |
| **2008 Production Report:** | Zip file (15.2 mb) | CSV file (106 mb) |
| **2009 Production Report:** | Zip file (19.7 mb) | CSV file (109 mb) |
| **2010 Production Report:** | Zip file (18.8 mb) | CSV file (107 mb) |
| **2011 Production Report:** | Zip file (20.6 mb) | CSV file (120 mb) |
| **2012 Production Report:** | Zip file (25.4 mb) | CSV file (164 mb) |
| **2013 Production Report:** | Zip file (17.2 mb) | CSV file (117 mb) |
| **2014 Production Report:** | Zip file (19.9 mb) | CSV file (136 mb) |
| **2015 Production Report:** | Zip file (18 mb) | CSV file (138 mb) |
| **2016 Production Report:** | Zip file (26.0 mb) | CSV file (156 mb) |
| **2017 Production Report:** | Zip file (18.1 mb) | CSV file (139 mb) |
| **2018 Production Report:** | Zip file (21.0 mb) | CSV file (135 mb) |
| **2019 Production Report:** | Zip file (20.8 mb) | CSV file (136 mb) |
| **2020 Production Report:** | CSV file | |

Production and Sales by County (1999 - Present) - Database Search

Oil and Gas Production by County (1995 - 2000)

These 22 separate production reports comprise over 2 GBytes of time-series data documenting monthly oil, water, and gas production for each actively producing well in Colorado over that time period.

A list of each of the 33 attributes contained in these files is provided in the Appendix 9.1.

In addition, well surface locations, directional bottomhole locations, and directional line shapefiles were downloaded to provide pertinent well attribute, geometry, and location information. Only information for existing wells (no undrilled locations) were included in the study.

**GIS ▾**

**Shapefiles ▾**

**Well Surface Location Data (Updated Daily) ▾**

February 25, 2019 - An additional attribute field was added to the Well Spots (APIs) download: Stat_Date - the date of the latest facility status-code update. Automated download/ingestion scripts may need to be updated. If you have questions, please contact Chris Eisinger or Dennis Ahlstrand.

Well Spots (APIs)(10 Mb) - metadata

Active and plugged wells - including active and expired well permits

Well Permits (300 Kb) - metadata

Active well permits

Pending Well Permits (20 Kb) - metadata

Well permits that are in process and not yet approved

**Directional Well Data (Updated Daily) ▾**

Prior to January 1, 2012, straight-line wellbore traces (from surface hole locations to corresponding bottom hole locations) were automatically generated for directional and horizontal wells. The generated wellbore traces do not represent the true wellbore path. After January 1, 2012, directional lines have been generated using operator-submitted well survey information.

Directional Bottom Hole Locations (1 Mb) - metadata

Actual and planned bottom hole locations

Directional Lines (5 Mb) - metadata

Actual and planned directional wellbore survey surface projections

Pending Directional Lines (1 Mb) - metadata

Pending directional wellbore surface projections are from Forms 2, 4, and 5 that HAVE NOT been been verified or approved

This data should be used with caution!

A list of each of the attributes contained in these files is provided in the Appendix 9.2 - 9.4.

State and county boundaries were obtained from the Database of Global Administrative Areas (GADM.org) website using the following link:

https://gadm.org/download_country_v3.html

PLSS township and section boundaries were obtained from the Bureau of Land Management via the following links:

https://data.doi.gov/dataset/plss-township

https://data.doi.gov/dataset/blm-colorado-plss-first-division

All shapefiles were loaded into QGIS, an open source GIS package, for display and visual analysis. Information and software downloads for QGIS can be obtained from the following link:

https://www.qgis.org/en/site/

The map below shows the distribution of wells contained in the shapefiles and the area of interest boundary chosen for the study. A total of 71116 well surface entities were present within the AOI at the inception of this project on February 17, 2020. All existing horizontal wells within the DJ Basin in Colorado (shown in lavender) were included.

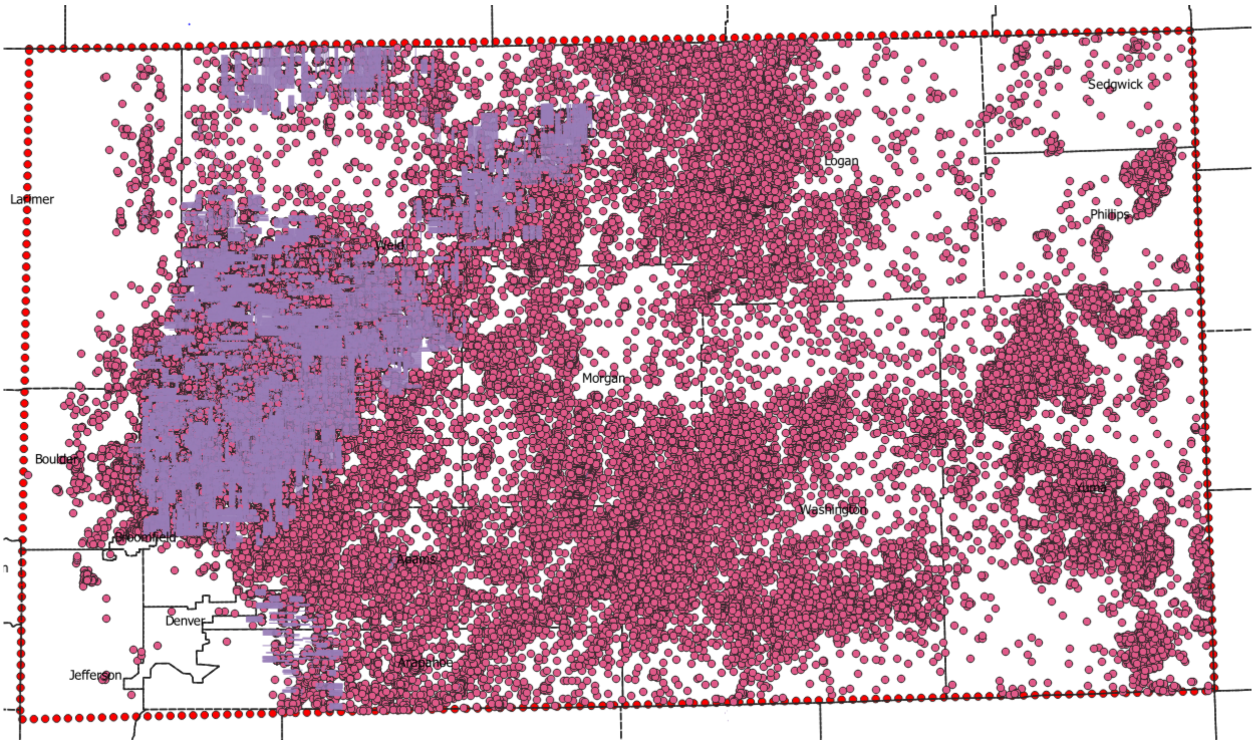

*Fig 1: The study area boundary is shown by the dotted red outline. Red dots show the location of the 71116 wells included in the study. Lavender lines show the subsurface traces of the horizontal wells.*

Structural depth and thickness attributes for the Niobrara and Codell units and regional faults were obtained from georectification and digitization of maps found in published literature (Milne, 2014) (Higley and Weimer, 2003). The maps (saved as .tif files) were loaded into QGIS and georectified into the project's projection system using county line corner point coordinates for registration. The fault traces and contour line features were then digitized in QGIS to create new line features. The fault shapefile was then used as an engineering feature to determine if proximity of a well to a fault affects its performance. The structural and thickness contours were loaded into LandMark's ZMap mapping package (the only licensed software package used in this study) to convert the contours to meshgrids. The resulting grids have a 500 meter grid increment and cover almost all of the area of interest. The grids were then loaded into geopandas for analysis and comparison with the well shapefile information. These grids were sampled to the well locations to be used as engineering features.

All spatial data for the study is defined in CRS EPSG 6342 (NAD83(2011) / UTM Zone 13N / Meters, which is consistent with the original shapefile information downloaded from the COGCC.

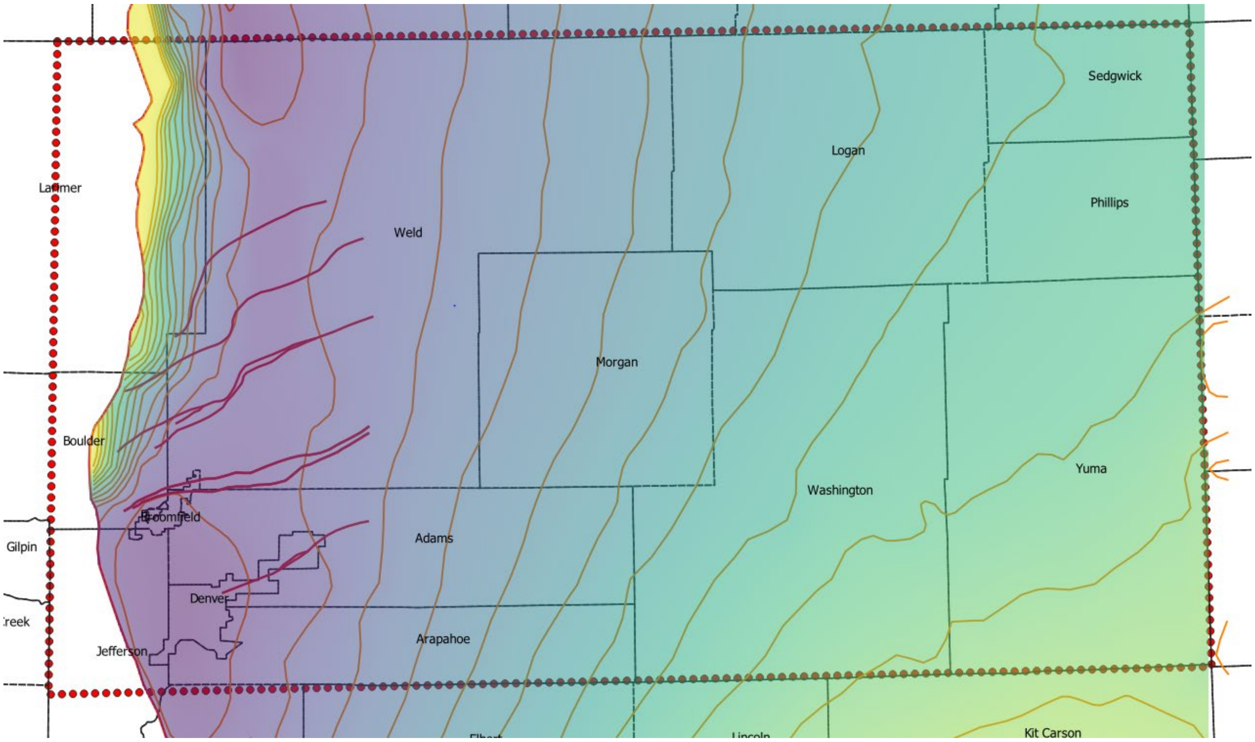


*Fig 2: Digitized fault traces shown by red lines, including the basin-bounding thrust fault to the west. Brown lines are digitized structural contours. Colorfill indicates the structural depth of the resulting meshgrid (Top Niobrara structure).*

Monthly spot prices for WTI (West Texas Intermediate Oil) were downloaded from the Energy Information Agency at this address:

https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=rwtc&f=m.

Monthly Henry Hub natural gas spot prices were downloaded from the Energy Information Agency at this address:

https://www.eia.gov/dnav/ng/hist/rngwhhdM.htm.

# 2. DATA WRANGLING

A good amount of data wrangling and cleansing were required to prepare the data for analysis. The primary data issues found were: 1) typographical errors; 2) missing values; 3) duplicated records in the production spreadsheets; 4) differences in unique well identifier (API) formatting and difficulties associating the correct sidetrack numbers (API with 10 digits) in the production records with the parent API8 (API with 8 digits) in the surface shapefile; 6) N-COM production assignments in the first few months of a well production which needed reassignment to the correct producing formation; and 5) other inconsistencies between the surface, bottomhole, and production datasets. Many of these issues required a first pass at cleansing, merging, and comparison to identify, so the process was an iterative one. The following section describes the steps taken during the data wrangling and cleansing process.

### Reading data into Pandas DataFrames

With all but a few exceptions, the production report files were formatted and named similarly, so the 'glob' module of python could be used for batch loading using Jupyter Notebook. The 22 separate files were then appended together into one large pandas dataframe, and a separate list was loaded to define the column labels. The resulting raw dataframe (allproddf) consisted of 12,922,083 rows and had a file size of over 3.2 GBytes.

The three shapefile datasets and an area of interest polygon were loaded into a separate Jupyter notebook using the geopandas module. Each of the three resulting geo-dataframes were then clipped to the area of interest polygon. The procedure was to iterate through the x, y coordinates (geometry) for each well. If the well's coordinates fell within the area of interest polygon, it was added to a shortlisted API well list. This list was then used to filter those wells into new subsetted geo-dataframes. These three new filtered geo-dataframes were also written out to new filtered shapefiles for later merging with the production information. The WELL.SHP file resulted in the gdf_surf_aoi dataframe. The DIRECTIONAL_BOTTOMHOLE_LOCATIONS.SHP file resulted in the gdf_bh_aoi dataframe. And finally the DIRECTIONAL_LINES.SHP file resulted in the gdf_dirlines dataframe.

An important feature contained in the gdf_surf_aoi dataframe was the ground elevations of the wells. Many of these values were missing and were required to calculate the correct bottomhole subsea bottomhole depth of the wellbore. This required an additional iteration step (See the Missing Values section below).

### Consistent API Formatting

The API number is the unique well identifier and a key attribute shared between datasets. However, the format of the attribute varied between the datasets, so reformatting was required. For example, the API was parsed into separate attributes of 'api_county_code', 'api_seq_num', and 'sidetrack_num' in the production dataset and these attributes were defined as integers. In the shapefiles, the API was an object padded with leading zeros. In the WELLS shapefile, the value was limited to 8 digits, while the other shapefiles included two additional digits (sidetrack number) appended on the end. The API attribute was converted to integers with the leading zeros in the county code removed to match the production dataframe. The 8 digit API in the WELLS shapefile was renamed to API8, and two API attributes were created in the production dataframe, an API column and an API8 column.

## Production Data Redundancy

There was some redundancy between the separate yearly production reports which resulted in duplicate rows for the same API and date, once merged. These were removed by subsetting on the date, API, and formation_code attributes and dropping the duplicates. The last row was kept with the assumption that the earlier records appended to the file first are superceded by corrected information from a later date. This brought the resulting dataset down to 11,910,725 rows. The formation_code attributes needed some subsequent re-formatting to remove variants, so the de-duplication process was run a second time after to remove additional duplicates due to these formation_code variants (See Categorical Variables section below).

## Handling Datetime

Separate month and year attributes in the production dataset needed to be converted to a datetime series. These separate attributes were converted to strings and the month attribute was padded with a leading zero. Initial qc of the month column indicated one row with an erroneous month value of '0', which was preventing the conversion to a datetime series from completing successfully. This row (116971) was for a well outside the study area boundary, so was dropped. Then the 'Date' attribute could be successfully created to a datetime series and the dataframe indexed using this series. End-month???

The Stat_Date and Spud_Date attributes in the WELLs shapefile also needed conversion to datetime. The replacement of two erroneous Stat_Date values of '3019-12-19' and '2029-09-06' had to be corrected to '2019-12-19' and '2019-09-06' respectively. Erroneous Spud_Date values with years of '2029', '2109' and '2108' were also corrected to 2019 and 2018 respectively in 6 wells.

In addition the prod_days attribute was converted to a timedelta dtype for direct comparison with a calculated feature 'total elapsed production time'.

## Categorical Variables

The formation_code and well_status attributes are important categorical identifiers and had varying formats which had to be corrected. Lowercase and mixed case strings were all converted to uppercase. Trailing spaces in the formation_codes were removed. This brought the number of unique well_status categories down from 15 to 9 and the number of unique formation_code categories down from 155 to 100. Both fields were converted from objects to categoricals.

It was noted during analysis of the formation_code attribute, that the code 'N-COM' is assigned to early production in many of the wells, prior to an actual formation assignment. To get more accurate production volumes by formation_code, the N-COM code was reassigned to the actual formation ('fm_code_realloc') where known. Other wells having a formation_code of N-COM throughout their production history were left intact.

## Missing Data

Approximately 3000 wells were missing important Ground_Ele information in the surface location shapefile. This attribute is crucial to calculating the bottomhole depth (TVDSS) values for the wellbores. The missing ground elevations were imputed using the Shapely nearest_points, Point, and MultiPoint functions for Python. The location of the nearest wells with a ground_elevation was taken to impute the value.

Additionally, some monthly production records for producing wells were missing. These were imputed using a rolling moving average value from the available monthly production data for those wells.

## Outliers

A few outliers were identified in the TVD attribute max and min ranges, which needed correction to calculate accurate TVDSS values. Five wells were found with extraordinarily large values of 73699, 77018, 82090, and 70993250. Fortunately, the correct values could be ascertained by querying the COGCC's COGIS on-line database. Examination of the correct TVD values for those wells indicated that they should be clipped at the first 4 digits. One extraordinarily small value of -4553 was corrected to a positive value after confirmation with the online COGIS database. Other outliers were values of 0,7,150,217, and 558 in directional wells with MDs of over 7000 ft and were revised to the MD of the well.

There are far too many vertical wells with Max_MD = 0 to correct properly. These were excluded from theTVDSS calculation.

19 horizontal wells were mistakenly identified as well_type_cat of 'Vertical'. These were identified by their extraordinarily large Max_MD values. They were re-categorized to well_type_cat' = 'Horizontal'.

44 wells had an incorrect Facil_Stat which did not correspond to the production records and well_status assignment from the production records. For instance, a 'DA' well (drilled and abandoned) that had known production through to the end of the production records was corrected to 'PR'. And other wells assigned as 'DG' (Drilling), 'XX' (Permitted Location), and 'AL' (Abandoned Location) with current known production were also corrected to 'PR'.

38 rows in the production dataframe had prod_days greater than 31 days, which is impossible for one month. These were reset to the maximum number of days in the month.

Monthly reported water volumes for many wells were extraordinarily and unrealistically large. Many of these wells were actually injection wells in which the injected water volume has been systematically entered into the same water volume fields as produced volumes. For this reason, injection wells were removed from the dataset (183 vertical and directional wells). Other unusually large monthly water volumes for 10 producing wells were found to be data entry errors and were replaced with values consistent with the production in that well from past and future months.

## Merging DataSets

The production time-series dataframe was merged with the gdf_surf_aoi geodataframe using an inner join to produce a datetime dataframe subsetted to only producing wells within the AOI boundary.

The dataframe was also joined with the bottomhole location geodataframe (gdf_bh_aoi) to obtain the bottomhole locations and other pertinent attributes of the deviated and horizontal boreholes. This join was performed in a left fashion to maintain the complete list of producing wells.

### Removing Unwanted Columns and Renaming Others

Unnecessary columns were dropped and the remaining columns were re-organized to bring the most pertinent and key attributes to the left side of the dataframes. Additional feature attributes were calculated that were pertinent. These features will be discussed in the Feature Engineering section of this report.

### Final Production TimeSeries DataFrame

The final production time-series dataframe (Prod_DT_Series_Final_WQuantileRank) has a total of 6,393,783 rows and 85 columns and is over 3.5 GBytes in size. A slice of the dataframe is shown below. A full list of the attributes is provided in section 9.5 of the Appendix.

| Date | API | API8 | sidetrack_num | well_type_cat | Oper_Cur_Num | Oper_Cur_Name | Oper_Hist_Num | Oper_Hist_Name | Well_Title | Ground_Ele |
|---|---|---|---|---|---|---|---|---|---|---|
| 2012-12-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10338 | CARRIZO OIL & GAS INC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-01-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10133 | HILCORP ENERGY COMPANY | 4-28-11-3-64 WEP | 5579.0 |
| 2013-02-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10439 | CARRIZO NIOBRARA LLC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-03-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10338 | CARRIZO OIL & GAS INC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-05-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10338 | CARRIZO OIL & GAS INC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-06-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10439 | CARRIZO NIOBRARA LLC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-07-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10439 | CARRIZO NIOBRARA LLC | 4-28-11-3-64 WEP | 5579.0 |
| 2013-08-01 | 10975301 | 109753 | 01 | Horizontal | 10646 | AXIS EXPLORATION LLC | 10338 | CARRIZO OIL & GAS INC | 4-28-11-3-64 WEP | 5579.0 |

*Fig 3: Slice of final production time-series dataframe*

### Final Rollup DataFrame

To gather statistics for each well and perform further exploratory analysis on non-time-series features associated with each well, a rollup dataframe was created. The final rollup dataframe (rollup_prodhead_final) has 49,280 rows (the number of unique wells and producing formation pairs in the final production time-series dataframe) and 66 columns and is 21.8 MBytes in size. A slice of the final rollup dataframe is shown below.

See the Feature Engineering section for more discussion on the attributes within this dataframe. A full list of the attributes is provided in section 9.6 of the Appendix.

| | API | API8 | API_County | well_type_cat | well_type_cat2 | Oper_Cur_Num | Oper_Cur_Name | Oper_Hist_Num | Oper_Hist_Name | Field_Code | Field_Name | UTM_X_SF | UTM_Y_SF | U |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 10502900 | 105029 | 1 | Vertical | Non-Horizontal | 72085 | PETRO-CANADA RESOURCES (USA) INC | 94090 | WALSH PRODUCTION INC ... | 60000.0 | NOONEN RANCH | 590719 | 4403657 | |
| **1** | 10504400 | 105044 | 1 | Vertical | Non-Horizontal | 72085 | PETRO-CANADA RESOURCES (USA) INC | 94090 | WALSH PRODUCTION INC ... | 60000.0 | NOONEN RANCH | 591032 | 4404791 | |
| **2** | 10507000 | 105070 | 1 | Vertical | Non-Horizontal | 95620 | WESTERN OPERATING COMPANY | 96155 | WHITING PETROLEUM CORP ... | 9000.0 | BUSY BEE | 575649 | 4406679 | |
| **3** | 10524200 | 105242 | 1 | Vertical | Non-Horizontal | 10330 | INVESTMENT EQUIPMENT LLC | 39150 | HEARTLAND OIL & GAS COMPANY ... | 5050.0 | BADGER CREEK | 607081 | 4412874 | |
| **4** | 10526300 | 105263 | 1 | Vertical | Non-Horizontal | 59100 | MONAHAN* REX FAMILY TRUST | 59100 | MONAHAN* REX ... | 54800.0 | MIDDLEMIST | 603503 | 4413136 | |
| **5** | 10528900 | 105289 | 1 | Vertical | Non-Horizontal | 10330 | INVESTMENT EQUIPMENT LLC | 39150 | HEARTLAND OIL & GAS COMPANY ... | 5050.0 | BADGER CREEK | 607060 | 4413880 | |
| **6** | 10529900 | 105299 | 1 | Vertical | Non-Horizontal | 10330 | INVESTMENT EQUIPMENT LLC | 39150 | HEARTLAND OIL & GAS COMPANY ... | 5050.0 | BADGER CREEK | 607063 | 4414081 | |
| **7** | 10532100 | 105321 | 1 | Vertical | Non-Horizontal | 46290 | KP KAUFFMAN COMPANY INC | 46290 | K P KAUFFMAN COMPANY INC ... | 39350.0 | IRONDALE | 561199 | 4414081 | |

*Fig 4: Slice of final rollup-dataframe,*

# 3. EXPLORATORY DATA ANALYSIS

The three final dataframes produced during the data wrangling process were used to conduct visual and exploratory analysis. The main purpose of this phase was to: 1) quality check the data, and identify outliers and issues requiring further data cleaning and wrangling; 2) gather general statistical information, identify correlations and trends, and develop insights about the wells and their production characteristics; and 3) begin to build a basis for predicting future well behavior.

All plotting was done within Jupyter notebooks using either matplotlib or the plotnine ggplot package for python.

## Rollup DataFrame EDA

In the analysis of the surface location shapefile, we found that there were 71116 surface locations within the study area. Analysis of the rollup dataframe for all the wells listed in the production spreadsheets indicates that there are a total of 37116 unique wells listed with entries in those spreadsheets over the 21 years analyzed. This suggests that more than half of the wells existing in the study area have been in production over the 1999 - 2019 time period, although as will be shown in the following bar charts, some of these wells are of a status that suggests they have not yet produced.
The bar chart below provides a breakdown of the number of producing wells by county within each of the 16 counties included within the study area. Clearly the bulk of the producing wells are in Weld County, where the giant Wattenberg Field exists.

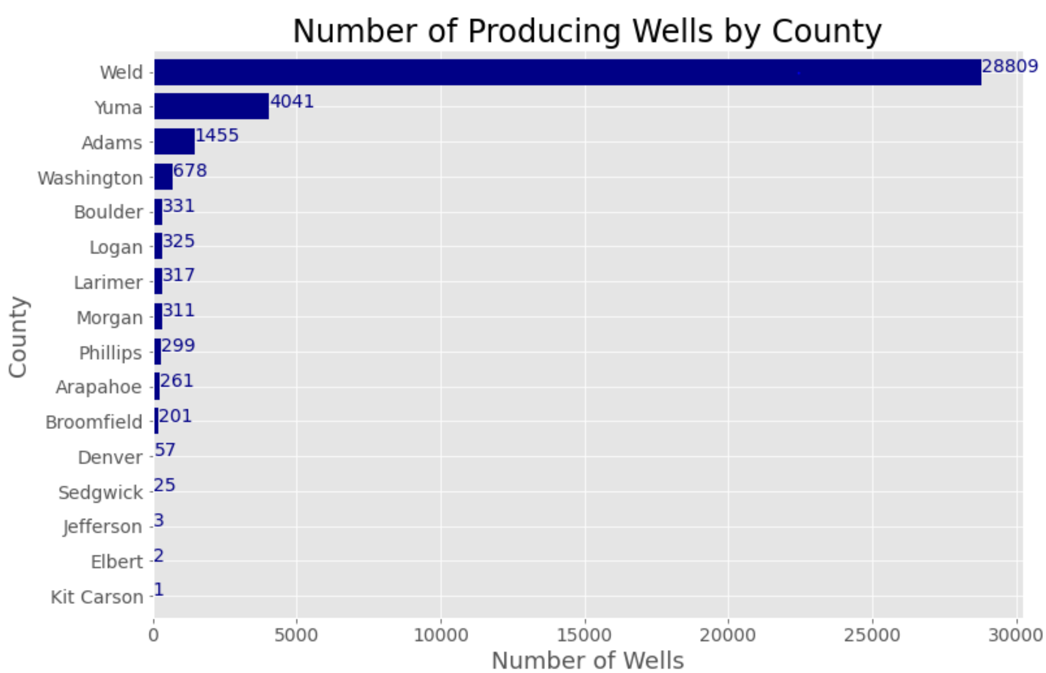


*Fig 5: Horizontal bar chart showing the count of producing wells by county from 1999 through 2019 within the study area.*

The following barchart shows a breakdown of the type of wells producing within the study area. The bulk of wells producing over the 21 year time period are non-horizontal wells.

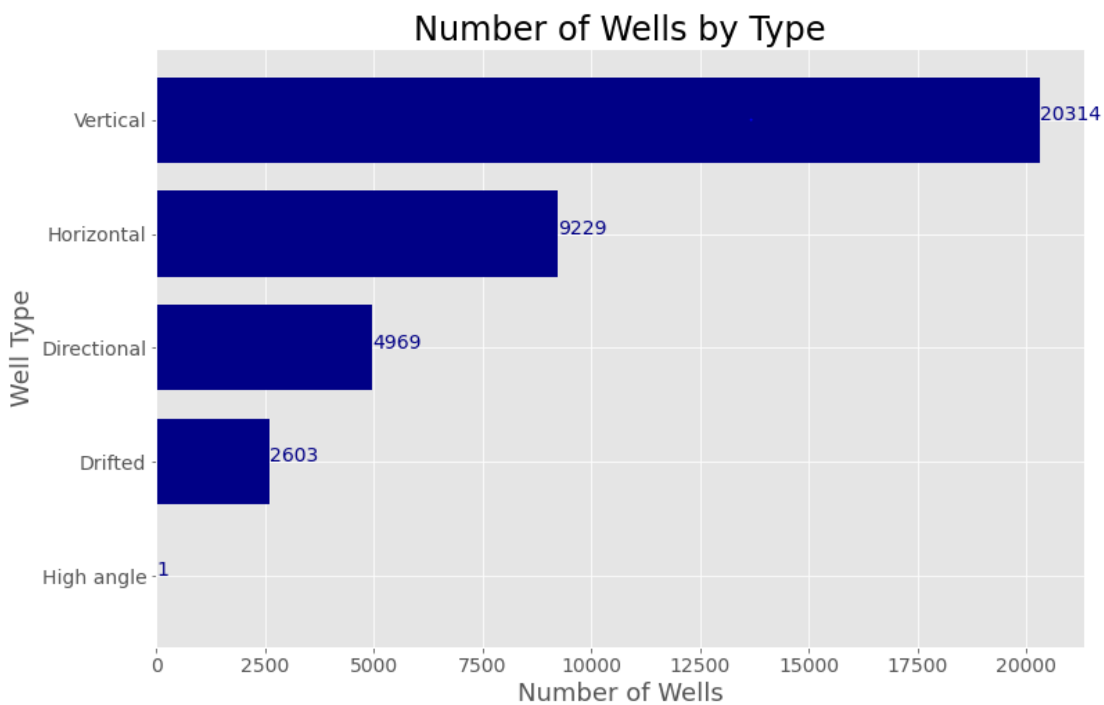


*Fig 6: Horizontal bar chart showing the count of producing wells by well type (horizontal vs vertical and near vertical variants).*

The following barchart shows a breakdown of the type of each well's final status. The bulk of wells are still actively producing.

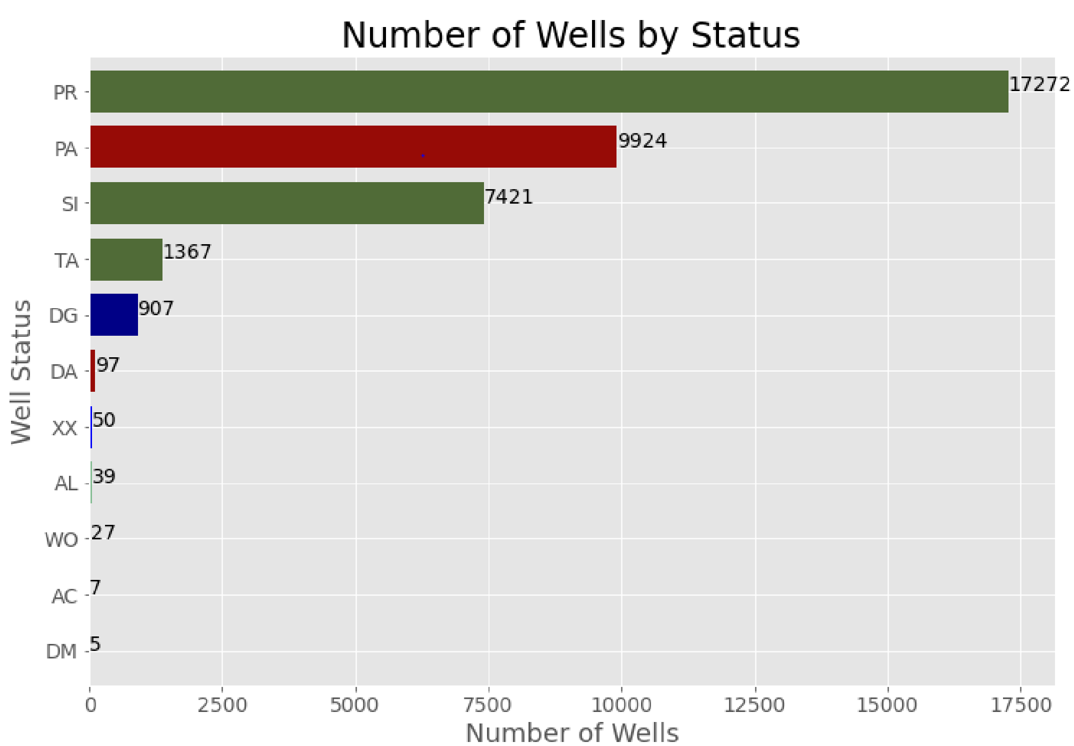


*Fig 7: Horizontal bar chart showing the count of producing wells by their last status. Those shaded in green are considered active. Those shaded in red are considered abandoned and those shaded in blue are in progress (permitted, drilling, active location, etc.)*

Each well was also tagged with an 'Active' or 'Abandoned' flag, given it's status. The breakdown of active versus abandoned wells is shown in Figure 8 below.

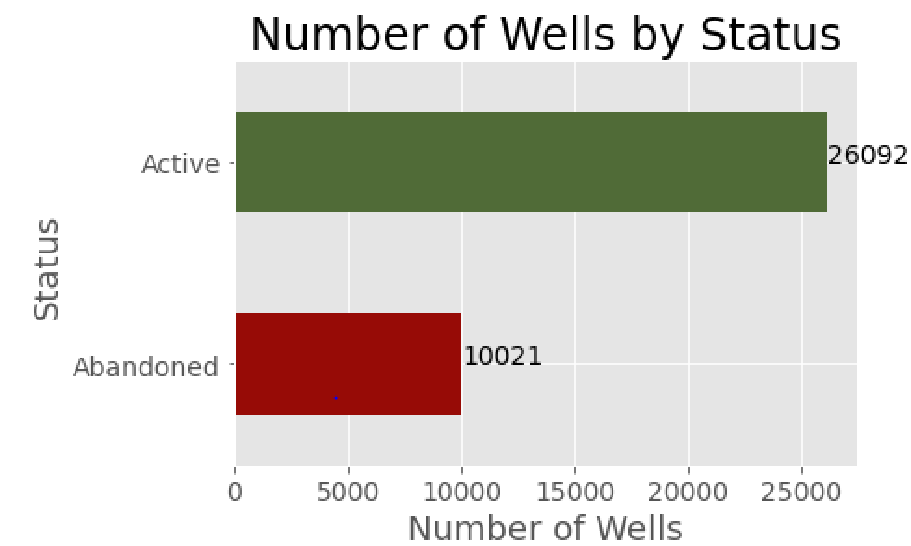


*Fig 8: Horizontal bar chart showing the count of producing wells by active or abandoned category. Wells with final status of 'DA' or 'PA' have gotten an abandoned flag (1). Wells with final status of 'PR', 'SI', 'TA', 'WO', or 'DM' have gotten an active flag (0). Wells with final status of 'DG', 'XX', 'AL', or 'AC' were not included in the count.*

The following stacked horizontal bar chart shows a breakdown of producing wells by well type (horizontal or vertical) and active / abandoned category. There is a much larger percentage of abandoned vertical wells than there are horizontal wells. In fact, only 120 horizontal wells have been abandoned in the study area.

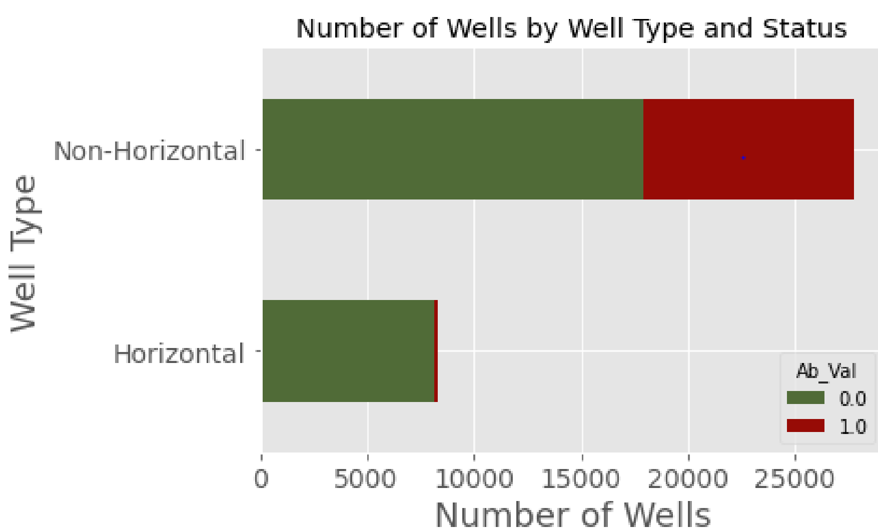


*Fig 9: Stacked horizontal bar chart showing the count of producing wells by well type and active (0) or abandoned (1) category (color-coded).*

The following horizontal bar chart shows a breakdown of operators by producing well count. The top 20 producers are shown.

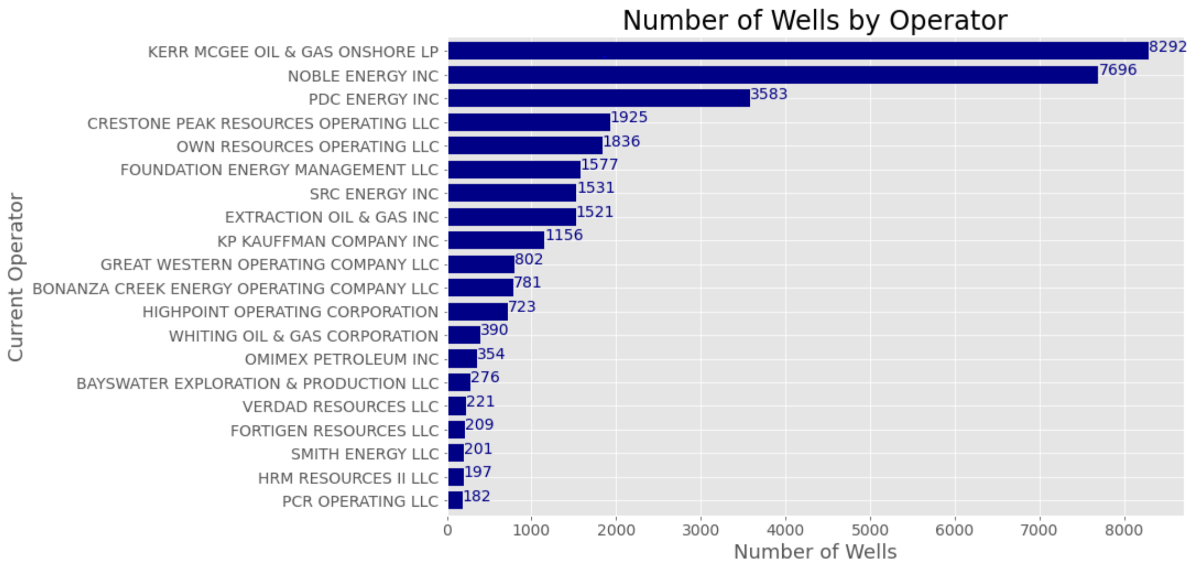


*Fig 9: Horizontal bar chart showing the count of producing wells by current operator, with the top 20 most common operators shown.*

Need to add in additional stacked bar charts showing production by unit and well type and scatter plots of production by unit.

## QGIS Map-Based EDA

Need to work on this section.

## Time Series DataFrame EDA

The time series provides an excellent data source for reviewing the production information over time. Figure 5 shows the total production of gas, oil, and water over the time period provided in the production spreadsheets. Production of all entities has increased over time. A noticeable steep increase in oil and gas production can be seen in 2014. The oil and water rates then decline over the 2015 to 2017 time period with gas production flattening. Since 2017, production of all fluids have risen sharply.

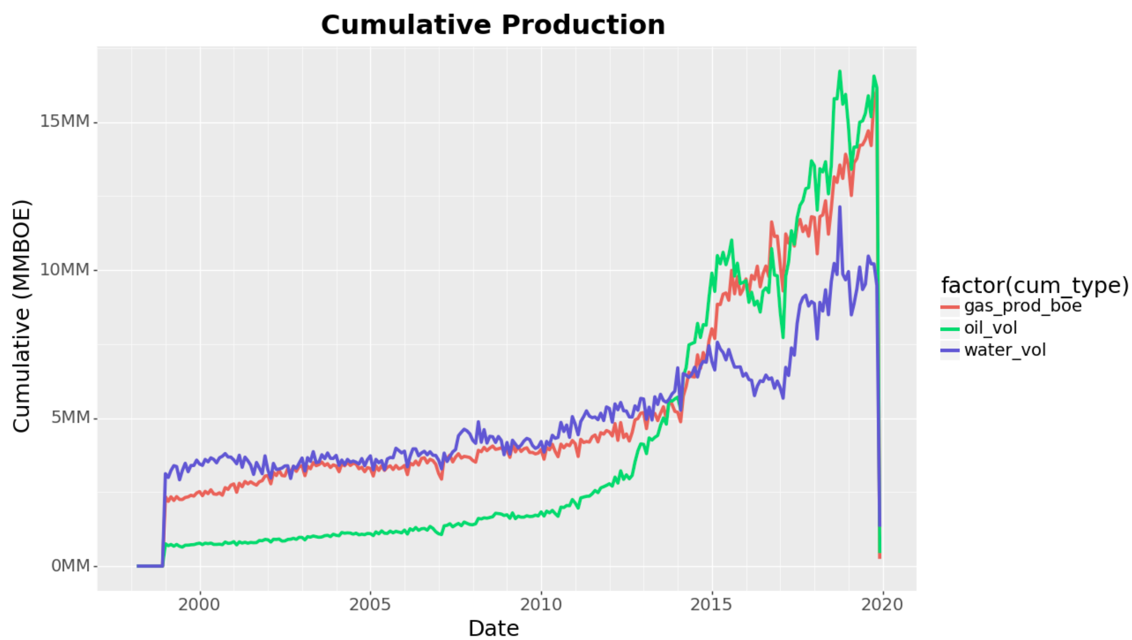


*Fig 10: Cumulative production of gas, oil and water (MMBOE) over the production file time period (1999-2019).*

Figure 11 shows the same information, but with the production entities stacked.

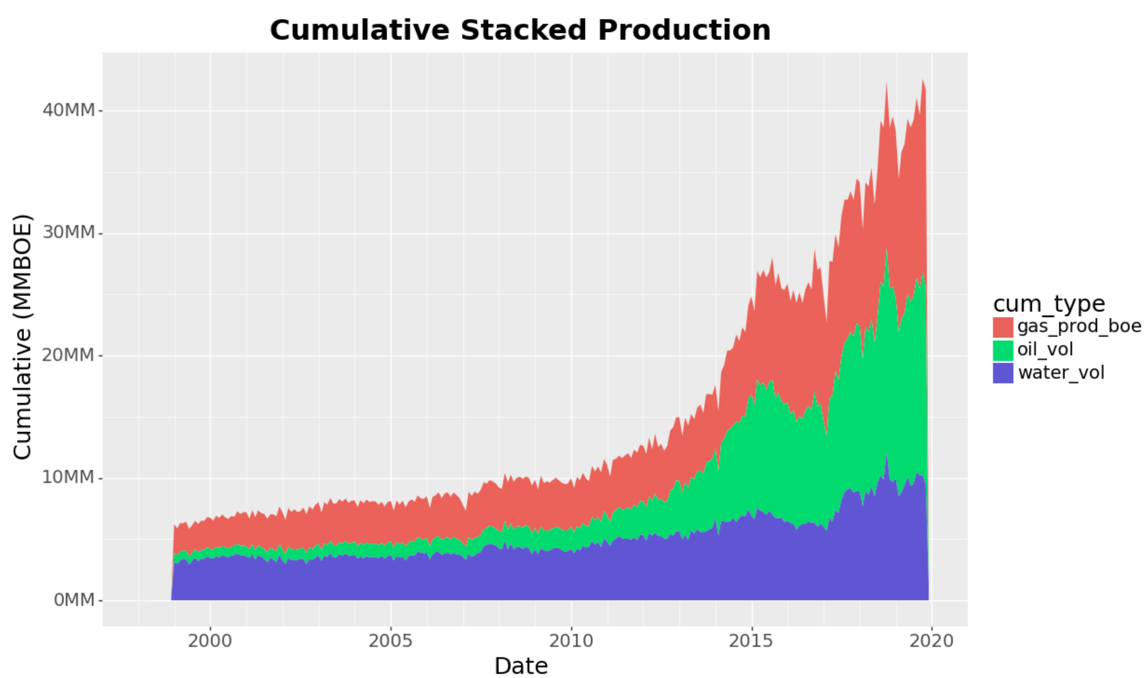


*Fig 11: Cumulative Stacked Production of gas, oil and water (MMBOE) over the production file time period (1999-2019).*

The following two figures show the total production over time for all non-Horizontal and Horizontal wells broken out separately. The same y-axis scale has been used as in Figure 6 above to understand the relative production from these 2 well type categories as a part of the total production.

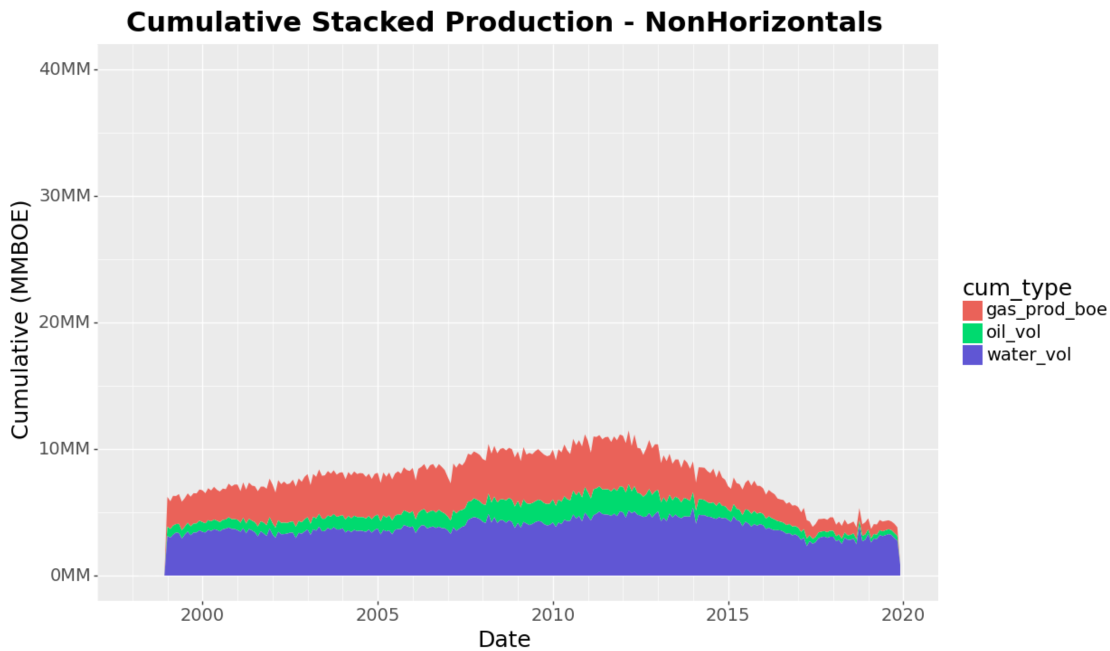


*Fig 12: Cumulative Stacked Production of gas, oil and water (MMBOE) for non-Horizontal wells over the production file time period (1999-2019).*

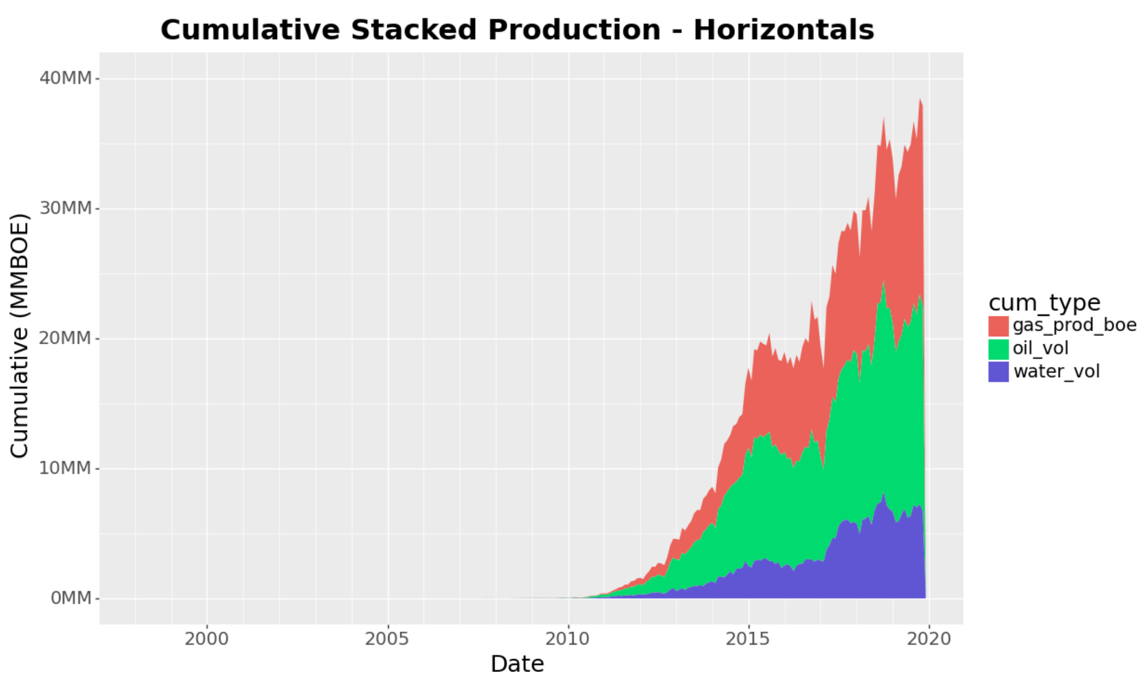


*Fig 13: Cumulative Stacked Production of gas, oil and water (MMBOE) for Horizontal wells over the production file time period (1999-2019).*

These plots indicate that non-Horizontal production peaked toward the end of 2011 and has continued to decline since that time, while horizontal production, since its inception in 2010, has quickly grown to account for the majority of the fluid produced over the last five years.

The same cumulative stacked production plots were also created for the non-Horizontal and Horizontal well population by producing formation. Figure 14 below indicates that the JSND and

NB-CD units have accounted for the majority of the production from non-horizontal wells in the study area. CODL production peaked in 2007 and has been declining since that time, but this has been compensated by the increasingly common practice of commingling Niobrara and Codell production together (NB-CD). The latter peaked near the end of 2012 and has continued to decline. Note also the substantial amount of water production from the JSND and DSND, and the very low oil production (mostly gas) in the NBRR non-horizontals.

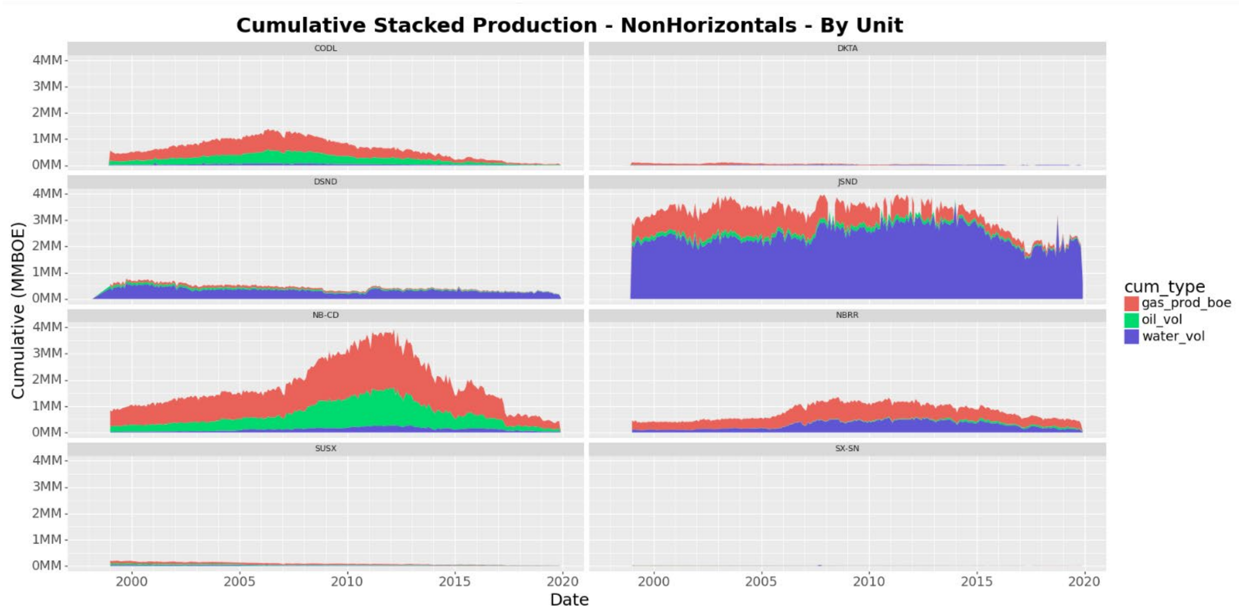


*Fig 14: Cumulative Stacked Production of gas, oil and water (MMBOE) for non-Horizontal wells over the production file time period (1999-2019) grouped by producing formation.*

Figure 15 below indicates that the NBRR unit is by far the largest producing unit from the horizontal well population. The CODL production is quite small in contrast.

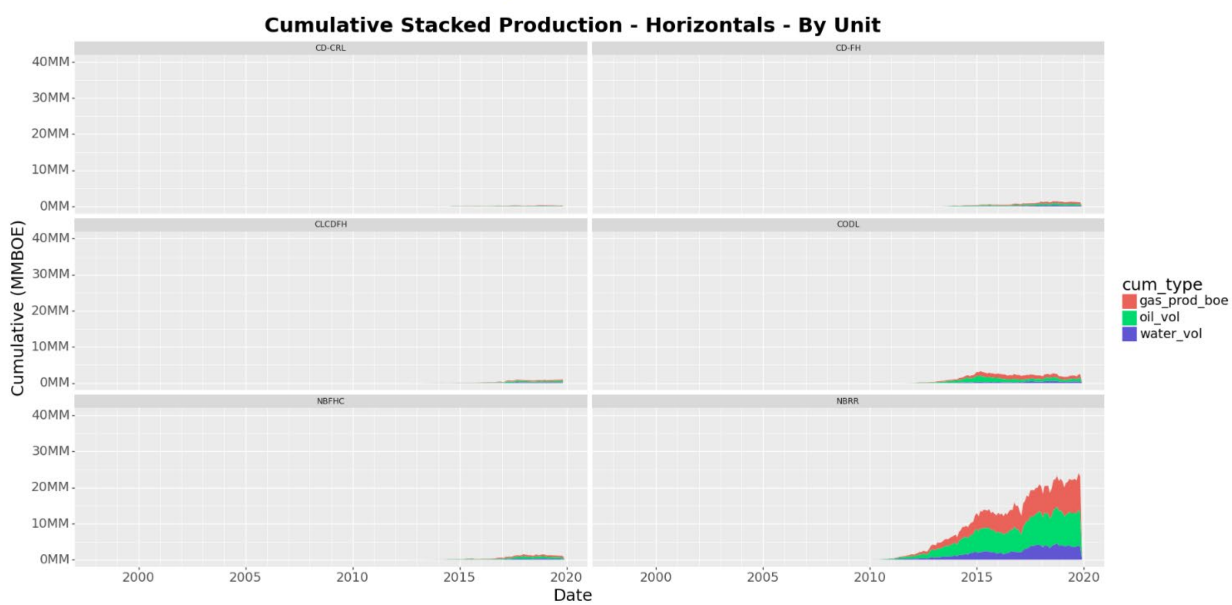


*Fig 15: Cumulative Stacked Production of gas, oil and water (MMBOE) for Horizontal wells over the production file time period (1999-2019) grouped by producing formation.*

The stacked plots below contrast further the NBRR, CODL and NB-CD production broken out by well type.

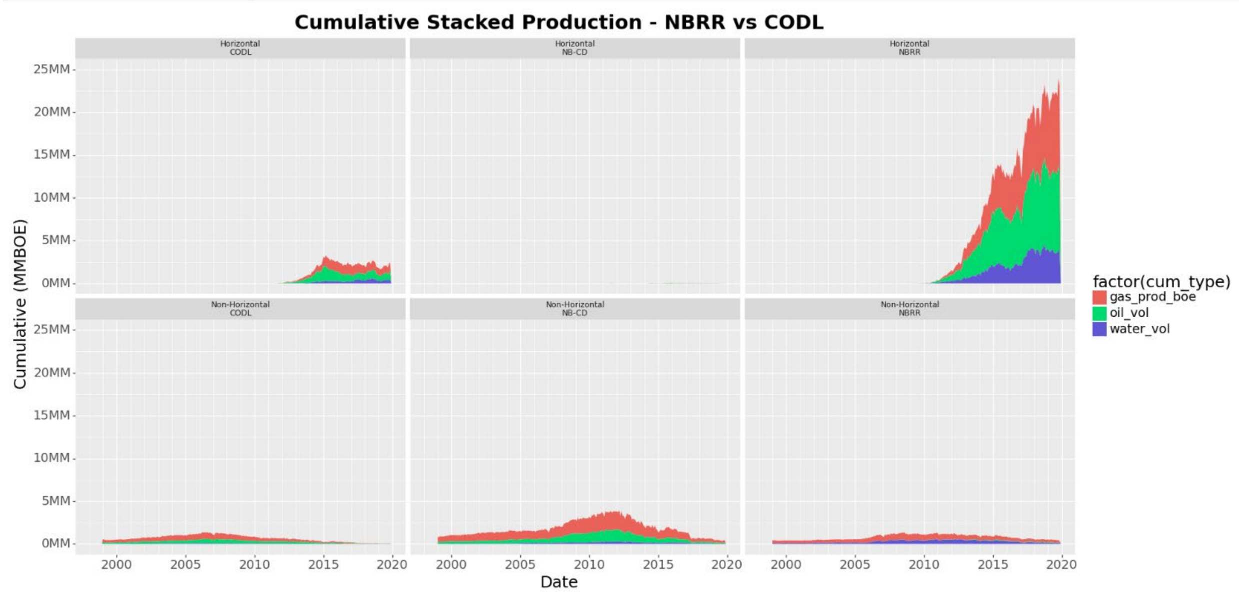


*Fig 16: Cumulative Stacked Production of gas, oil and water (MMBOE) for all NBRR and CODL producing wells over the production file time period (1999-2019) grouped by producing formation and well type (Horizontal and Non-Horizontal).*

To understand the production decline trends in both the NBRR and CODL by individual horizontal wells, the production was broken out by quintiles, first using the first 6 months initial production (with first 3 months production removed). For each quintile, the production was then plotted for each well over time starting with the first month of production. The values were also averaged for each quintile. (Need the count of wells in each quintile)

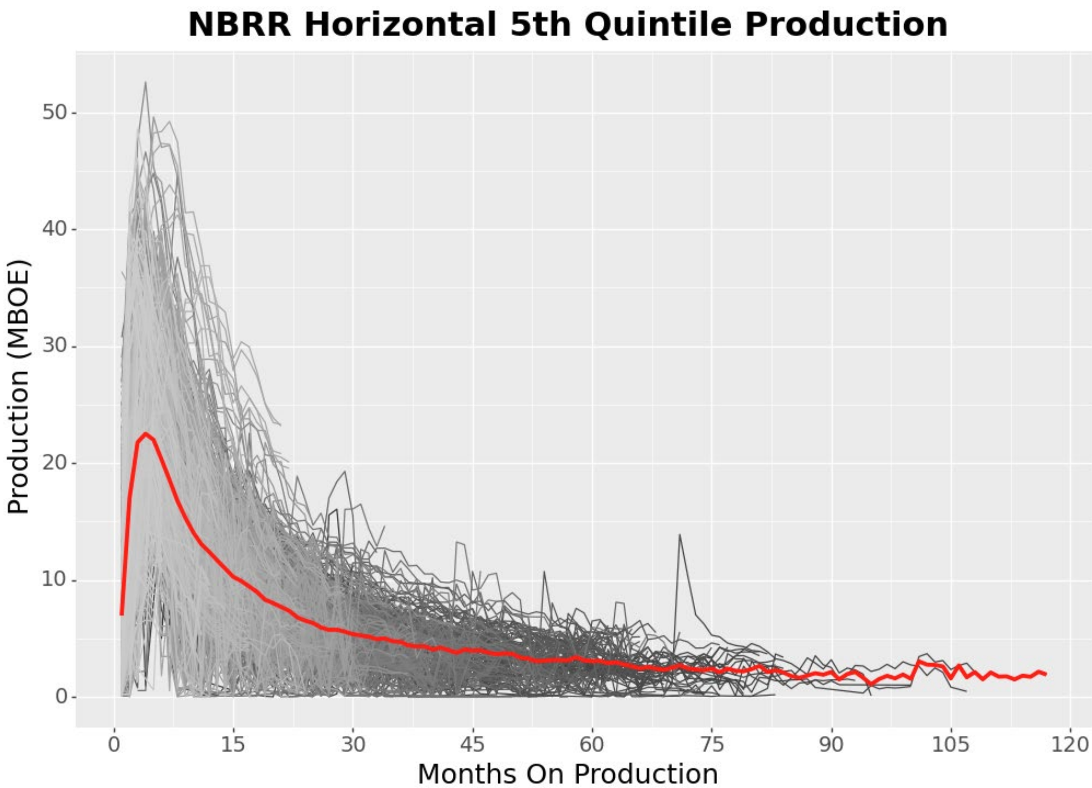


*Fig 17: NBRR 5th Quintile Production by Well Since First Month. Red line is average of all wells by month.*

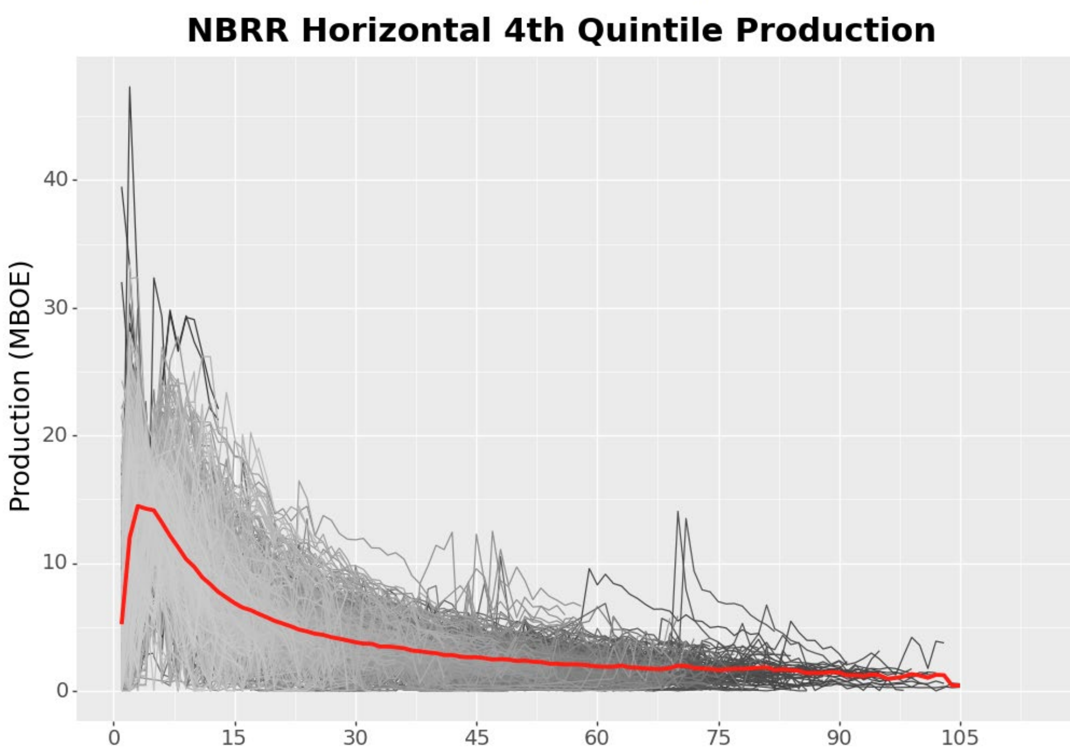


*Fig 18: NBRR 4th Quintile Production by Well Since First Month. Red line is average of all wells by month.*

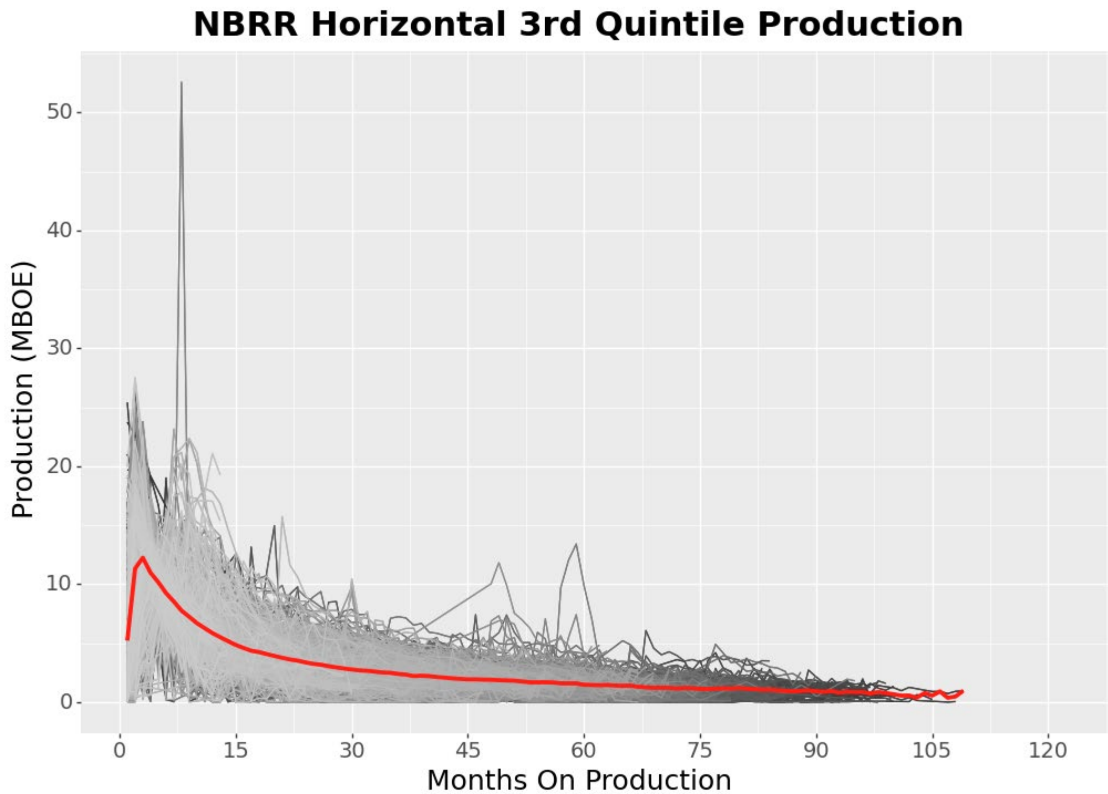


*Fig 19: NBRR 3rd Quintile Production by Well Since First Month.Red line is average of all wells by month.*

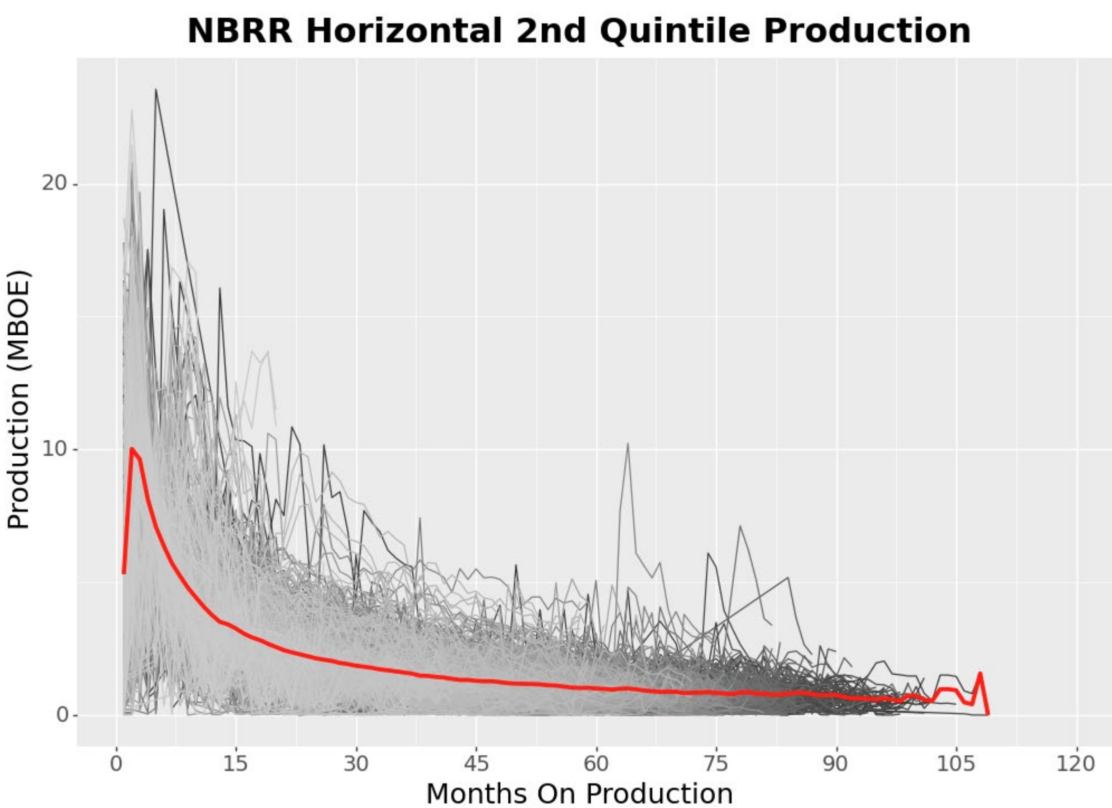


*Fig 20: NBRR 2nd Quintile Production by Well Since First Month. Red line is average of all wells by month.*

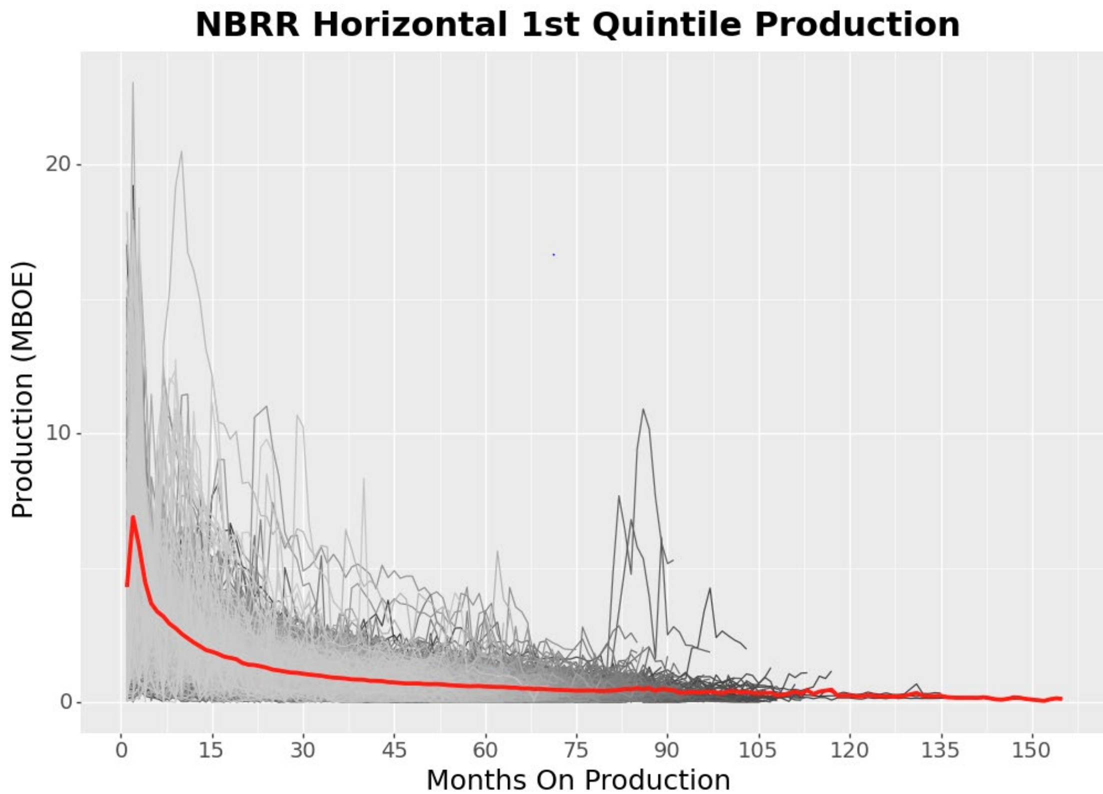


*Fig 21: NBRR 1st Quintile Production by Well Since First Month.Red line is average of all wells by month.*

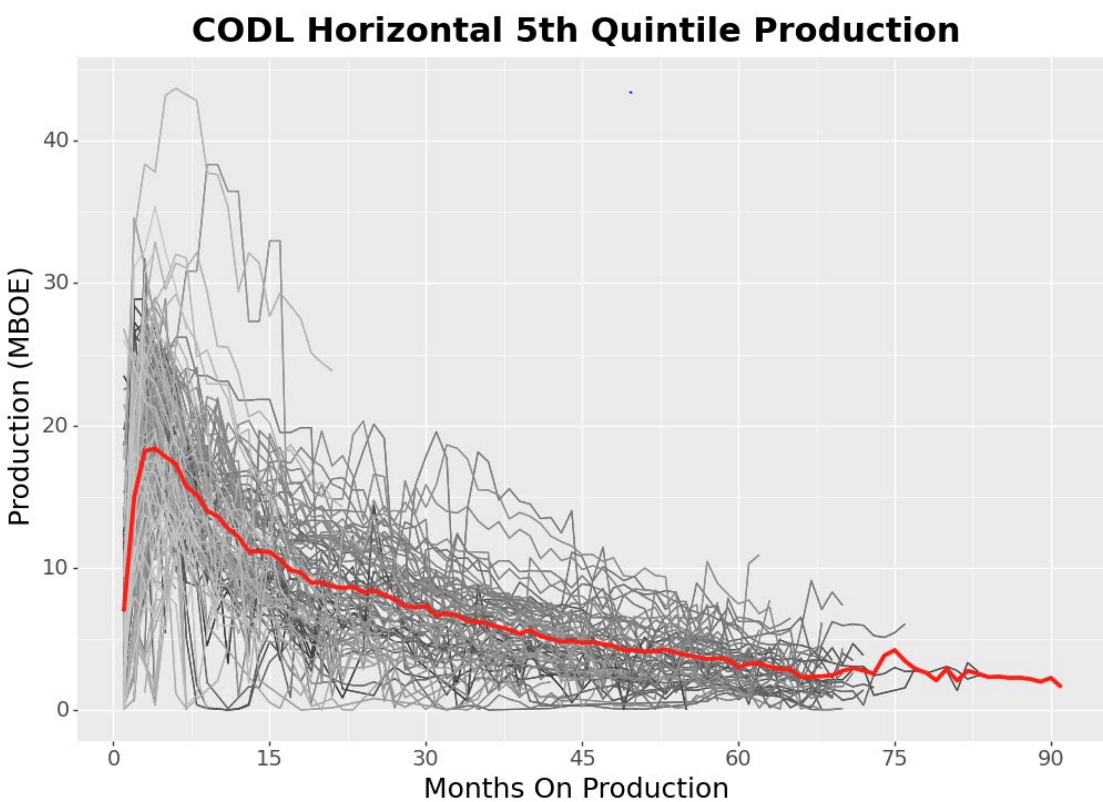


*Fig 22: CODL 5th Quintile Production by Well Since First Month. Red line is average of all wells by month.*

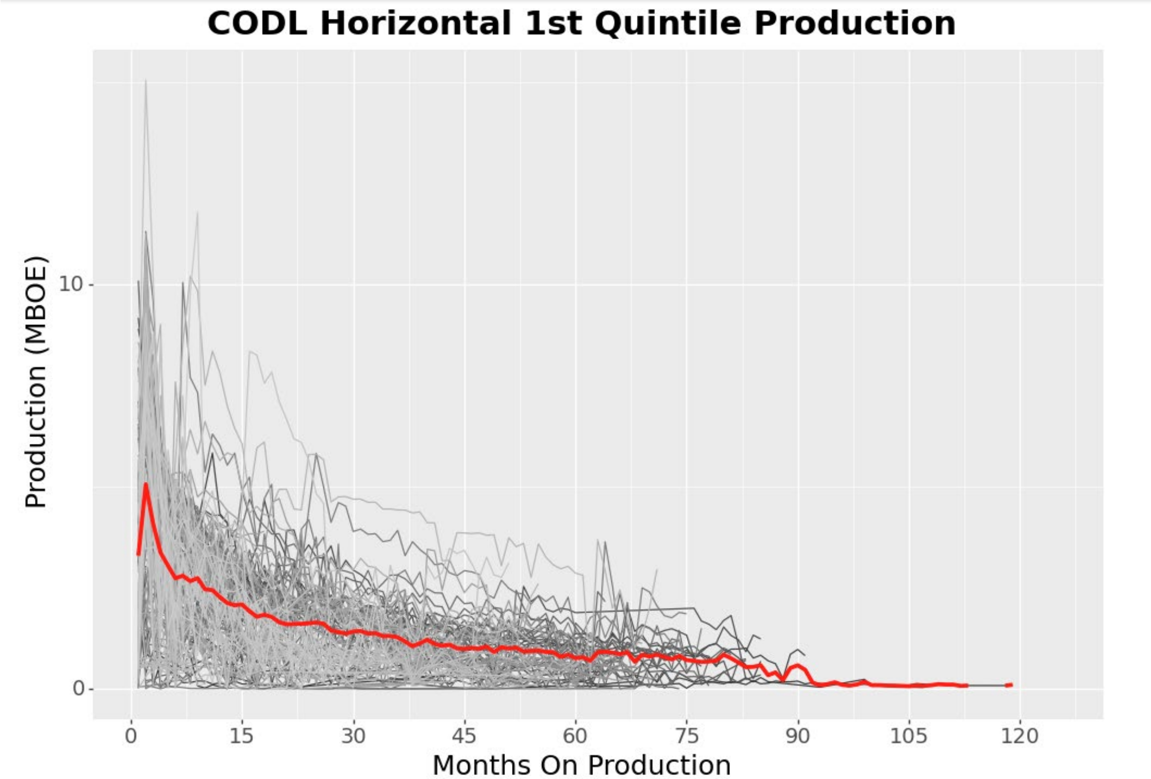


*Fig 23: CODL 1st Quintile Production by Well Since First Month. Red line is average of all wells by month.*

Even though it is clear that each NBRR quintile shows sequentially less initial production, the longevity of the wells across quintiles is similar. Comparison between the NBRR and CODL indicates that even though there are fewer CODL wells and less production, the best 5th quintile CODL wells are almost on par with the best 5th quintile NBRR wells in terms of initial production rates. The 5th quintile CODL wells, however, have not produced as long.

The following figures show the production for only the few horizontal wells that have been abandoned, for both the NBRR and CODL. Only 3 CODL wells have been abandoned as of the beginning of 2020.

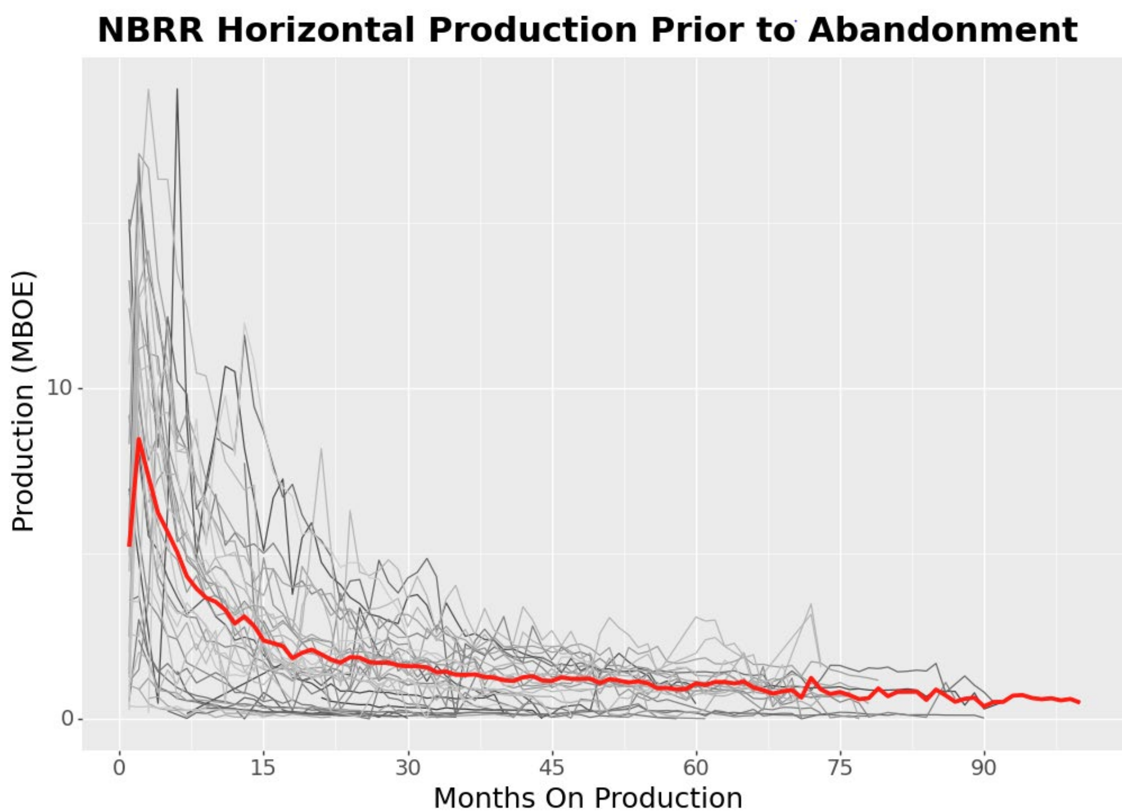

Fig 24: NBRR Production by Well Since First Month only for wells that have since been abandoned. Red line is average of all wells by month.

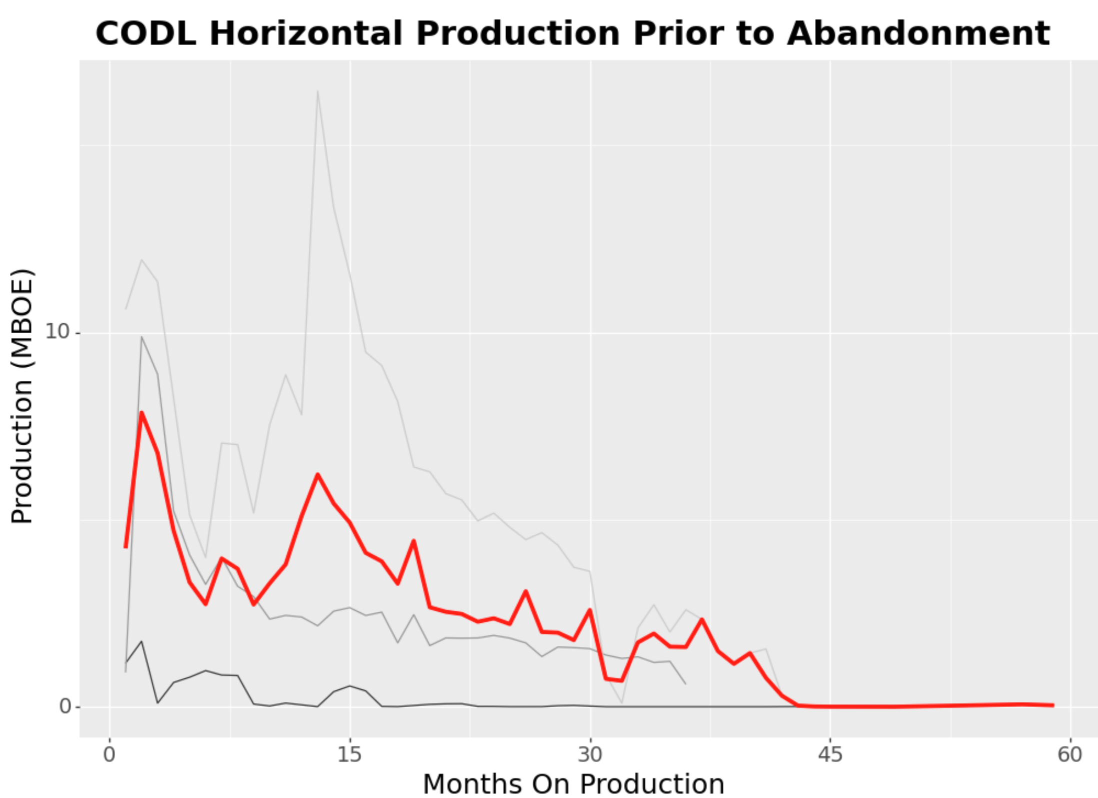

Fig 25: CODL Production by Well Since First Month only for wells that have since been abandoned. Red line is average of all wells by month.

Further Work - Questions Remaining to Be Answered

*Remove outliers?; compare and contrast other quintilized population - which is better.*

# 4. FEATURE ENGINEERING

A number of additional features were calculated or extrapolated from the original datasets to aid in the analysis.

A total of ?? features are indexed by API

And a total of ?? features are indexed by date.

The following is a list of features and their description.

Production Time-Series

TVDSS from ground_ele - Max_tvd if vertical or tvd if deviated
BOE of gas
BOE - gas + oil

Rollup - Header attributes
well_type_cat
well_type_cat2
OpCurNum1,2
OpCurName1,2
OpHistNum1,2
OpHistName1,2
Field_Code
Field_Name
UTM_X_SF
UTM_Y_SF
UTM_X_BH
UTM_Y_BH
well_status
Ab_Val
TVDSS
Spud_Date

Rollup - Production attributes

Start
End
FORM1
FORM2
oil_cum
gas_boe_cum
gas_mcf_cum
boe_cum
wtr_cum
prod_days
norm_oil_cum

norm_wtr_cum
norm_gas_mcf_cum
norm_gas_boe_cum
norm_boe_cum
gor
wor
oil_first_mo
gas_boe_first_mo
prod_days_first_mo
IP_one_mo
oil_two_mo
gas_boe_two_mo
prod_days_two_mo
IP_two_mo
oil_three_mo
gas_boe_three_mo
prod_days_three_mo
IP_three_mo
TotalElapsedProdTime
ProdDayRatio
Total Days on Prod

Needed
Lateral length (for horizontals)
Unit Thickness (for deviated / verticals) Niobrara and Codell only?
Distance to nearest offset
Horizontal orientation

| *Features* | *Description* |
|---|---|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

# 5. MACHINE LEARNING / STATS DATA ANALYSIS

# 6. MACHINE LEARNING: TIME SERIES FORECASTING

## 7. CITATIONS


- Barlacchi, G. et al. A multi-source dataset of urban life in the city of Milan and the Province of Trentino. Sci. Data2:150055 doi: 10.1038/sdata.2015.55 (2015).
- Telecom Italia, 2015, "Telecommunications - SMS, Call, Internet - MI", https://doi.org/10.7910/DVN/EGZHFV, Harvard Dataverse, V1
- Telecom Italia, 2015, "Milano Grid", https://doi.org/10.7910/DVN/QJWLFU, Harvard Dataverse, V1


## 8. REFERENCES


- Colorado Oil and Gas Conservation Commission https://cogcc.state.co.us/data2.html#/downloads

- Milne, 2014

- Wikipedia https://en.wikipedia.org/wiki/Denver_Basin

# 9. APPENDIX

## 9.1 COGCC monthly production reports - attributes

### Production Record Data Dictionary

CSV Header File

| | Field | Description of field | Data Type |
|---|---|---|---|
| 1) | report_month | Month reported by operator | Char(2) |
| 2) | report_year | Year reported by operator | Smallint |
| 3) | St | API state code "05" | Char(2) |
| 4) | api_county_code | API county code | Char(3) |
| 5) | api_seq_num | API sequence number | Char(5) |
| 6) | sidetrack_num | API sidetrack number | Char(2) |
| 7) | formation_code | COGCC formation code | Char(6) |
| 8) | well_status | Well status | Char(2) |
| 9) | prod_days | Number of production days for month | Int |
| 10) | water_disp_code | Water disposal code | Char(1) |
| 11) | water_vol | Water volume in Bbls | Int |
| 12) | water_press_tbg | Water tubing pressure in psia | Int |
| 13) | water_press_csg | Water casing pressure in psia | Int |
| 14) | bom_invent | BOM inventory in Bbls | Int |
| 15) | oil_vol | Volume of oil produced in Bbls | Int |
| 16) | oil_sales | Oil sales volume in Bbls | Int |
| 17) | adjustment | Adjusted oil volume in Bbls | Int |
| 18) | eom_invent | EOM inventory in Bbls | Int |
| 19) | gravity_sales | Gravity sales | Real |
| 20) | gas_sales | Gas sales volume In MCF | Int |
| 21) | flared | Gas flared in MCF | Int |
| 22) | gas_vol | Volume of gas used on lease in MCF | Int |
| 23) | shrink | Gas shrinkage in MCF | Int |
| 24) | gas_prod | Gas produced in MCF | INt |
| 25) | btu_sales | BTU sales | Int |
| 26) | gas_press_tbg | Gas tubing pressure | Int |
| 27) | gas_press_csg | Gas casing pressure | Int |
| 28) | operator_num | Operator number | Int |
| 29) | name | Operator name | Char(50) |
| 30) | facility_name | Well name | Char(35) |
| 31) | facility_num | Well number | Char(15) |
| 32) | accepted_date | Date processed by COGCC | Datetime |
| 33) | revised | This is a revised record | Char(1) |

## 9.2 WELLS Shapefile - Attributes

| | Attribute | Attribute_Description | Attribute_Definition_Source |
|---|---|---|---|
| 1 | FID | Internal feature number;Sequential unique whole numbers that are automatically generated | Esri |
| 2 | Shape | Feature geometry; Coordinates defining the features | Esri |
| 3 | API | 8-digit API Number (County Code and Sequence Code) | COGCC COGIS Database |
| 4 | API_County | API County Code | COGCC COGIS Database |
| 5 | Api_Seq | API Sequence Code | COGCC COGIS Database |
| 6 | API_Label | 10-digit API Number (State Code - County Code - Sequence Code) | COGCC COGIS Database |
| 7 | Operat_Num | Operator Number | COGCC COGIS Database |
| 8 | Operator | Operator Name | COGCC COGIS Database |
| 9 | Well_Num | Well Number | COGCC COGIS Database |
| 10 | Well_Name | Well Name | COGCC COGIS Database |
| 11 | Well_Title | Well Number and Well Name | COGCC COGIS Database |
| 12 | Citing_Typ | Citing Type (Actual or Planned) | COGCC COGIS Database |
| 13 | Spud_Date | Best Spud Date Available | COGCC COGIS Database |
| 14 | Ground_Ele | Elevation of Ground Level at Well Location in Feet AMSL | COGCC COGIS Database |
| 15 | Max_MD | Maximum Measured Depth of All Sidetracks Associated with API in Feet | COGCC COGIS Database |
| 16 | Max_TVD | Maximum Total Vertical Depth of All Sidetracks Associated with API in Feet | COGCC COGIS Database |
| 17 | Field_Code | Field Code | COGCC COGIS Database |
| 18 | Field_Name | Field Name | COGCC COGIS Database |
| 19 | Facil_Id | Facility ID in COGIS Database | COGCC COGIS Database |
| 20 | Facil_Type | Facility Type in COGIS Database (Well) | COGCC COGIS Database |
| 21 | Facil_Stat | Facility Status in COGIS Database | COGCC COGIS Database |
| 22 | Stat_Date | The date of the latest facility status-code update. | COGCC COGIS Database |
| 23 | Loc_Qual | Location Qualifer | COGCC COGIS Database |
| 24 | Loc_ID | Location ID in COGIS Database | COGCC COGIS Database |
| 25 | Loc_Name | Location Name | COGCC COGIS Database |
| 26 | Dist_N_S | Location in Feet from the N or S Section Line | COGCC COGIS Database |
| 27 | Dir_N_S | Section Line from which the Location is Measured | COGCC COGIS Database |
| 28 | Dist_E_W | Location in Feet from the E or W Section Line | COGCC COGIS Database |
| 29 | Dir_E_W | Section Line from which the Location is Measured | COGCC COGIS Database |
| 30 | Qtr_Qtr | Quarter Quarter Location | COGCC COGIS Database |
| 31 | Section | Section | COGCC COGIS Database |
| 32 | Township | Township | COGCC COGIS Database |
| 33 | Range | Range | COGCC COGIS Database |
| 34 | Meridian | Meridian | COGCC COGIS Database |

| | | | |
|---|---|---|---|
| 35 | Latitude | Latitude | COGCC COGIS Database |
| 36 | Longitude | Longitude | COGCC COGIS Database |
| 37 | Utm_X | Utm X Coordinate | COGCC COGIS Database |
| 38 | Utm_Y | Utm Y Coordinate | COGCC COGIS Database |

## 9.3 DIRECTIONAL_BOTTOMHOLE_LOCATIONS Shapefile - Attributes

| 1 | Attribute | Attribute Description | Attribute Definition_Source |
|---|---|---|---|
| 2 | Shape | Feature geometry | Esri |
| 3 | API | API Number | COGCC |
| 4 | API_Label | 12-digit API Number (State Code - County Code - Sequence Code - Directional Sidetrack Code) | COGCC COGIS Database |
| 5 | Operator | Operator Name | COGCC |
| 6 | Well_Name | Well Name | COGCC |
| 7 | BH_Status | Bottomhole Status; Actual or Planned | COGCC |
| 8 | MD | Measured Depth (Feet) | COGCC |
| 9 | TVD | Total Vertical Depth (Feet) | COGCC |
| 10 | Deviation | Directional or Horizontal | COGCC |
| 11 | Field_Code | COGCC Field Code | COGCC |
| 12 | Field_Name | COGCC Field Name | COGCC |
| 13 | Lat | Bottomhole Latitude Coordinate (NAD83) | COGCC |
| 14 | Long | Bottomhole Longitude Coordinate (NAD83) | COGCC |
| 15 | Utm_X | Bottomhole UTM X Coordinate (NAD83, Zone 13N) | COGCC |
| 16 | Utm_Y | Bottomhole UTM Y Coordinate (NAD83, Zone 13N) | COGCC |

## 9.4 DIRECTIONAL_LINES Shapefile - Attributes

| | Attribute | Attribute_Definition | Attribute_Definition_Source |
|---|---|---|---|
| 1 | FID | Internal feature number | Esri |
| 2 | Shape | Feature geometry | Esri |
| 3 | API | API Number | COGCC |
| 4 | API_Label | 12-digit API Number (State Code - County Code - Sequence Code - Directional Sidetrack Code) | COGCC COGIS Database |
| 5 | Operator | Operator Name | COGCC |
| 6 | Well_Name | Well Name | COGCC |
| 7 | Dir_Status | Directional Line Status: Actual or Planned | COGCC |
| 8 | MD | Measured Depth (Feet) | COGCC |
| 9 | TVD | Total Vertical Depth (Feet) | COGCC |
| 10 | Deviation | Directional or Horizontal | COGCC |
| 11 | Field_Code | COGCC Field Code | COGCC |
| 12 | Field_Name | COGCC Field Name | COGCC |

## 9.5 Final Production Time-Series Dataframe - Attributes

| Column Index | Attribute | Dtype |
|---|---|---|
| 0 | Date | datetime64[ns] |
| 1 | API | int64 |
| 2 | API8 | int64 |

| | | |
|---|---|---|
| 3 | sidetrack_num | object |
| 4 | well_type_cat | category |
| 5 | Oper_Cur_Num | int64 |
| 6 | Oper_Cur_Name | object |
| 7 | Oper_Hist_Num | int64 |
| 8 | Oper_Hist_Name | object |
| 9 | Well_Title | object |
| 10 | Ground_Ele | float64 |
| 11 | Max_MD | float64 |
| 12 | MD | float64 |
| 13 | Max_TVD | float64 |
| 14 | TVD | float64 |
| 15 | TVDSS | float64 |
| 16 | Field_Code | float64 |
| 17 | Field_Name | object |
| 18 | Spud_Date | datetime64[ns] |
| 19 | Stat_Date | datetime64[ns] |
| 20 | well_status | category |
| 21 | Facil_Stat | category |
| 22 | API_Form | object |
| 23 | formation_code | category |
| 24 | fm_code_realloc | object |
| 25 | prod_days | float64 |
| 26 | water_vol | float64 |
| 27 | oil_vol | float64 |
| 28 | gas_prod | float64 |
| 29 | gas_prod_boe | float64 |
| 30 | LAT_SF | float64 |
| 31 | LONG_SF | float64 |
| 32 | LAT_BH | float64 |
| 33 | LONG_BH | float64 |
| 34 | UTM_X_SF | int64 |
| 35 | UTM_Y_SF | int64 |
| 36 | UTM_X_BH | float64 |
| 37 | UTM_Y_BH | float64 |
| 38 | Township | object |
| 39 | Range | object |
| 40 | Section | object |
| 41 | water_disp_code | object |
| 42 | water_press_tbg | float64 |
| 43 | water_press_csg | float64 |
| 44 | bom_invent | float64 |
| 45 | adjustment | float64 |
| 46 | eom_invent | float64 |

| | | |
|---|---|---|
| 47 | gravity_sale | float64 |
| 48 | gas_vol | float64 |
| 49 | shrink | float64 |
| 50 | gas_press_tbg | float64 |
| 51 | gas_press_csg | float64 |
| 52 | facility_name | object |
| 53 | facility_num | object |
| 54 | accepted_date | object |
| 55 | revised | object |
| 56 | year | object |
| 57 | month | object |
| 58 | api_seq_num | object |
| 59 | API_Label_x | object |
| 60 | Well_Num | object |
| 61 | Well_Name | object |
| 62 | Citing_Typ | object |
| 63 | Facil_Id | int64 |
| 64 | Facil_Type | object |
| 65 | Loc_Qual | object |
| 66 | Loc_ID | float64 |
| 67 | Loc_Name | object |
| 68 | Dist_N_S | float64 |
| 69 | Dir_N_S | object |
| 70 | Dist_E_W | float64 |
| 71 | Dir_E_W | object |
| 72 | Qtr_Qtr | object |
| 73 | Meridian | object |
| 74 | BH_Status | object |
| 75 | geometry_SF | geometry |
| 76 | geometry_BH | geometry |
| 77 | ProdHist | category |
| 78 | NBRR_Hor_IP_Quintile | category |
| 79 | NBRR_Hor_NormBoeCum_Quintile | category |
| 80 | CODL_Hor_IP_Quintile | category |
| 81 | CODL_Hor_NormBoeCum_Quintile | category |
| 82 | prod_month_by_API_Form | int64 |
| 83 | well_type_cat2 | category |
| 84 | oil_gas_vol_boe | float64 |
| 85 | Ab_Val | int64 |

## 9.6 Final Rollup Dataframe - Attributes

| Column Index | Attribute | Non-Null Count | Dtype |
|---|---|---|---|

| | | | |
|---|---|---|---|
| 0 | | 49280 | int64 |
| 1 | API8 | 49280 | int64 |
| 2 | API_County | 49280 | int64 |
| 3 | well_type_cat | 49280 | category |
| 4 | well_type_cat2 | 49280 | category |
| 5 | Oper_Cur_Num | 49280 | int64 |
| 6 | Oper_Cur_Name | 49280 | category |
| 7 | Oper_Hist_Num | 49280 | int64 |
| 8 | Oper_Hist_Name | 49280 | category |
| 9 | Field_Code | 49280 | float64 |
| 10 | Field_Name | 49280 | category |
| 11 | UTM_X_SF | 49280 | int64 |
| 12 | UTM_Y_SF | 49280 | int64 |
| 13 | UTM_X_BH | 22269 | float64 |
| 14 | UTM_Y_BH | 22269 | float64 |
| 15 | well_status | 49280 | category |
| 16 | Facil_Stat | 49280 | category |
| 17 | Stat_Date | 49269 | datetime64[ns] |
| 18 | Ab_Val | 48277 | float64 |
| 19 | TVDSS | 48734 | float64 |
| 20 | Spud_Date | 44969 | datetime64[ns] |
| 21 | API_Form | 47158 | object |
| 22 | fm_code_realloc | 47158 | object |
| 23 | Start | 47158 | datetime64[ns] |
| 24 | End | 47158 | datetime64[ns] |
| 25 | oil_cum | 47158 | float64 |
| 26 | gas_boe_cum | 47158 | float64 |
| 27 | gas_mcf_cum | 47158 | float64 |
| 28 | wtr_cum | 47158 | float64 |
| 29 | boe_cum | 47158 | float64 |
| 30 | prod_days | 47158 | timedelta64[ns] |
| 31 | norm_oil_cum | 47158 | float64 |
| 32 | norm_wtr_cum | 47158 | float64 |
| 33 | norm_gas_mcf_cum | 47158 | float64 |
| 34 | norm_gas_boe_cum | 47158 | float64 |
| 35 | norm_boe_cum | 47158 | float64 |
| 36 | gor | 42532 | float64 |
| 37 | wor | 42532 | float64 |
| 38 | 30Day_Oil | 47158 | float64 |
| 39 | 30Day_GasBoe | 47158 | float64 |
| 40 | 30Day_ProdDays | 47158 | float64 |
| 41 | 30Day_IP | 47158 | float64 |
| 42 | 90Day_Oil | 47158 | float64 |
| 43 | 90Day_GasBoe | 47158 | float64 |
| 44 | 90Day_ProdDays | 47158 | float64 |
| 45 | 90Day_IP | 47158 | float64 |
| 46 | 180Day_Oil | 47158 | float64 |

| | | | |
|---|---|---|---|
| 47 | 180Day_GasBoe | 47158 | float64 |
| 48 | 180Day_ProdDays | 47158 | float64 |
| 49 | 180Day_IP | 47158 | float64 |
| 50 | 270Day_Oil | 47158 | float64 |
| 51 | 270Day_GasBoe | 47158 | float64 |
| 52 | 270Day_ProdDays | 47158 | float64 |
| 53 | 270Day_IP | 47158 | float64 |
| 54 | 180Day_IP_Corr | 44868 | float64 |
| 55 | 180Day_Oil_Corr | 47158 | float64 |
| 56 | 180Day_GasBoe_Corr | 47158 | float64 |
| 57 | 180Day_ProdDays_Corr | 47158 | float64 |
| 58 | GrossProdTime | 47158 | object |
| 59 | ProdDayRatio | 47158 | object |
| 60 | ProdHist | 44969 | object |
| 61 | GrossProdTimeRev | 42999 | timedelta64[ns] |
| 62 | NBRR_Hor_IP_Quintile | 5357 | category |
| 63 | NBRR_Hor_NormBoeCum_Quintile | 5500 | category |
| 64 | CODL_Hor_IP_Quintile | 750 | category |
| 65 | CODL_Hor_NormBoeCum_Quintile | 824 | category |

## 9.7 List of Jupyter Notebooks

ShapeFiles_Loading_Conditioning Begin File(s): WELL.SHP
DIRECTIONAL_BOTTOMHOLE_LOCATIONS.SHP
DIRECTIONAL_LINES.SHP
End File(s): gdf_surf_aoi (Wells_filtered.shp)
gdf_dirlines (Dirlines_filtered.shp)
gdf_bh_aoi (DirBH_filtered.shp)
surf_bh_mrg_FINAL_CLEAN.pickle

ProductionDataImportMerge Begin File: COGCC Production Reports.csv
Intermediate Raw File: allproddf
End File: allprodaoi_wbh.pickle

ProdDataClean Begin File: allprodaoi_wbh.pickle
End File: allprodaoi_wbh_dt.pickle

ReassignNCOM Begin File: allprod_wbh_dt
Intermediate Raw File: allprodaoi_wbh_srtd_Form2.pickle
End File: allprodaoi_dt_Final_Clean.pickle

ProdDataCleanPass2 Begin File(s): allprodaoi_dt_Final_Clean.pickle
Intermediate Raw File: allprodaoi_dt_Final_Clean_No_Inj3.pickle
End File(s): Prod_DT_Series_Final_WQuantileRank.pickle

GenerateRollup Begin Files(s): allprodaoi_dt_Final_Clean_No_Inj3.pickle
Intermediate Raw File:
Prod_DT_Series_Final_WQuantileRank.pickle

| | | |
|---|---|---|
| | End_Files(s): | allprodaoi_final_rollup.pickle |
| RollupExplAnal | Begin Files: | allprodaoi_final_rollup.pickle |
| TimeSeriesEDA | Begin Files: | Prod_DT_Series_Final_WQuantileRank.pickle |
| WOE | Begin Files: | allprodaoi_final_rollup.pickle |