

SHALE ANALYTICS PRO

User Guide

ML · Decline Curves · Economics · A&D · Field Development
Centriv Petrologic Petroleum Engineering Consultants

Demo version · For Python 3.10+ with Matplotlib backend ·

TABLE OF CONTENTS

1.1 What Is Shale Analytics Pro?	4
1.2 What the Workflow Does	4
1.3 Who Should Use This Guide	4
1.4 Software Architecture at a Glance	4
2.1 System Requirements	6
2.2 Installing Dependencies	6
2.3 Running in GUI Mode	6
2.4 Running in Headless Batch Mode	7
3.1 Required CSV Format	8
3.2 Data Quality Rules	8
3.3 Generating the Sample Dataset	8
3.4 Using Your Own Data	9
Tab 0 · Data Loader	10
Tab 1 · Economic Parameters	10
Price Deck (3 sliders)	10
Operating Costs (4 sliders)	10
Capital & Fiscal (4 sliders)	10
Screening Thresholds (3 sliders)	10
Tab 2 · Type Curve Analysis	11
Tab 3 · Decline Curve Analysis	11
Reading the Chart	11
Adjusting Parameters	11
Saving Per-Well Overrides	11
Tab 4 · Machine Learning — EUR Prediction	12
Features Used (13 variables)	12
Early Production Signals (months 1–6)	12
Completion Parameters	12
Reading the Results	12
Tab 5 · Well Economics	13
Metric Definitions	13
The Portfolio Quadrant Map	13
Tab 6 · A&D Portfolio Dashboard	13
A&D Composite Score Methodology	13
Category Definitions	14
Tab 7 · Field Development Planning	14
The Seven Sliders	14
The Four Output Charts	14
Tab 8 · Export	15
5.1 Overview	16

5.2 Slide-by-Slide Description	16
5.3 Customising the Deck	16
6.1 Cash Flow Model	17
6.2 NPV Calculation	17
6.3 IRR Calculation	17
6.4 Arps Hyperbolic Decline.....	17
7.1 All Arguments	19
7.2 Python API (in-session commands).....	19
7.3 Output Files	19
8.1 Economic Screen Benchmarks.....	21
8.2 ML Model Quality Benchmarks	21
8.3 Common Interpretation Pitfalls.....	21
9.1 Common Errors and Fixes	23
9.2 Performance Tips	23
9.3 Getting Help	23
Standard Workflow (5 minutes).....	24
Key Formulas at a Glance	24
Abbreviations	25

1 OVERVIEW

1.1 What Is Shale Analytics Pro?

Shale Analytics Pro is an end-to-end upstream oil and gas analytics platform designed for investment banking, private equity, and engineering teams. It combines production engineering rigor with machine learning and financial economics to produce institutional-grade asset valuations from raw well data.

1.2 What the Workflow Does

Starting from a standard CSV production file, the platform automatically:

- Cleans and validates raw well production data
- Constructs P10 / P50 / P90 type curves by basin, operator, and vintage
- Fits Arps hyperbolic decline curves to every well and estimates EUR
- Trains a Random Forest + Gradient Boosting ML model to predict EUR from early production
- Computes NPV, IRR, payout period, and profitability index for every well
- Ranks the portfolio with a composite A&D score and flags divestment candidates
- Runs a Field Development Planning module to forecast production and revenue from a custom drilling schedule
- Exports a 12-slide PowerPoint executive deck and a master results CSV

1.3 Who Should Use This Guide

This guide is written for four types of users:

User Type	Primary Sections	Typical Use Case
Analyst / Engineer	Sections 2–6	Run the full pipeline; adjust decline parameters
Finance / Investment Professional	Sections 3, 5, 8	Interpret economics; adjust price deck; review A&D ranking
Portfolio Manager	Sections 5, 7, 8	Field development planning; scenario comparison
IT / DevOps	Section 2	Installation, dependencies, headless batch mode

1.4 Software Architecture at a Glance

The application is a single Python file (`shale_analytics_gui.py`) using only standard scientific Python libraries. No database, server, or internet connection is required after installation.

FILE	<code>shale_analytics_gui.py</code> (2,138 lines) · <code>bakken_production_10000.csv</code> (sample data) · <code>generate_dataset.py</code> (dataset builder)
-------------	---

Python	3.10 or later
Key Libraries	numpy, pandas, scipy, scikit-learn, matplotlib, python-pptx
GUI Framework	Matplotlib interactive widgets (no Tkinter or Qt required)
Output formats	PNG charts, PowerPoint .pptx, CSV results, HTML executive summary
Data input	Standard CSV — 11 columns, any row count
Tested on	Windows 10/11, macOS 13+, Ubuntu 22.04

2 INSTALLATION & SETUP

2.1 System Requirements

Python version	3.10, 3.11, or 3.12 (3.12 recommended)
RAM	4 GB minimum; 8 GB recommended for 10,000+ row datasets
Disk space	500 MB for Python environment + model outputs
Display	Required for GUI mode; not required for --no-gui batch mode
Operating system	Windows 10+, macOS 12+, Ubuntu 20.04+

2.2 Installing Dependencies

Open a terminal (Command Prompt / PowerShell on Windows; Terminal on Mac/Linux) and run the following commands in order:

STEP
1

Install Python packages

```
pip install pandas numpy matplotlib seaborn scipy scikit-learn python-pptx
```

→ This installs all required libraries in a single command.
 → If you use conda: `conda install pandas numpy matplotlib seaborn scipy scikit-learn`
 → For python-pptx with conda: `pip install python-pptx`

STEP
2

Verify installation

```
python -c "import pandas, numpy, matplotlib, seaborn, scipy, sklearn, pptx; print('All OK')"
```

→ You should see: All OK
 → If any package is missing, re-run the pip install command for that specific library.

STEP
3

Download the files

Place all files in the same folder (e.g., C:\ShaleAnalytics\ or ~/shale/):

- `shale_analytics_gui.py`
- `bakken_production_10000.csv` (sample dataset — or your own CSV)
- `generate_dataset.py` (optional — regenerate sample data)

TIP

Using a virtual environment is recommended to keep your Python installation clean. Run: `python -m venv shale_env` then `source shale_env/bin/activate` (Mac/Linux) or `shale_env\Scripts\activate` (Windows).

NOTE

On some systems matplotlib will show a warning about the display backend. This is harmless in GUI mode. In headless/server environments, use the `--no-gui` flag (see Section 2.4).

2.3 Running in GUI Mode

This is the standard interactive mode. All nine tabs are available and all charts are displayed in a window.

STEP
1

Navigate to your working folder

```
cd C:\ShaleAnalytics (Windows)
```

```
cd ~/shale (Mac / Linux)
```

STEP

2

Launch the application

```
python shale_analytics_gui.py --csv bakken_production_10000.csv
```

→ Two windows open: a narrow Tab Navigator and a Content window.
→ The application automatically loads the CSV, runs the full pipeline, and shows Tab 0.

STEP

3

Override economic parameters at launch (optional)

```
python shale_analytics_gui.py --csv your_data.csv --oil 80 --gas 3.00 --disc 12
```

→ --oil Oil price in \$/bbl (default: 75)
→ --gas Gas price in \$/mcf (default: 2.50)
→ --disc Annual discount rate in % (default: 10)

2.4 Running in Headless Batch Mode

Use this when running on a server, in CI/CD pipelines, or anywhere a display is not available. All charts and the PowerPoint deck are saved to disk automatically.

```
python shale_analytics_gui.py --csv bakken_production_10000.csv --no-gui
```

Optional overrides:

```
python shale_analytics_gui.py --csv data.csv --no-gui --oil 65 --disc 12
```

Batch mode saves the following files into the shale_pro_output/ folder:

- 02_type_curves.png through 08_operator_benchmarking.png — all analysis charts
- Shale_Analytics_Executive_Deck.pptx — 12-slide PowerPoint presentation
- results.csv — full per-well results (NPV, IRR, EUR, A&D score, category)

3 DATA INPUT FORMAT

3.1 Required CSV Format

The application accepts any CSV file with the 11 columns listed below. Column names must match exactly (case-insensitive, underscores required). Each row represents one month of production for one well.

Column Name	Type	Units	Required	Example
api	String	—	Yes	3310100001
month	Integer	1 = first month	Yes	1
oil_bbl	Float	barrels/month	Yes	5420.5
gas_mcf	Float	mcf/month	Yes	7890.3
water_bbl	Float	barrels/month	Yes	2310.0
ngl_bbl	Float	barrels/month	Yes	124.8
lateral_length_ft	Integer	feet	Yes	9500
proppant_lbs	Integer	pounds total	Yes	17500000
stages	Integer	count	Yes	44
operator	String	—	Yes	EOG Resources
vintage_year	Integer	calendar year	Yes	2021

IMPORTANT

The api column is the unique well identifier. Rows are grouped by api then sorted by month. Each well must have at least 6 months of non-zero production for decline curve fitting. Wells with fewer than 6 active months are still included in type curves and economics.

3.2 Data Quality Rules

The application applies automatic data cleaning before any analysis:

- Null / missing production values are filled with 0 (treated as shut-in months)
- Negative production values are clipped to 0
- Production spikes greater than 5× the 3-month rolling median are smoothed
- BOE is calculated as: $\text{oil_bbl} + \text{gas_mcf} / 6 + \text{ngl_bbl}$
- $\text{GOR} = \text{gas_mcf} \times 1,000 / \text{oil_bbl}$ (zero oil months excluded from ratio)
- Water-oil ratio and oil cut are derived from monthly volumes

3.3 Generating the Sample Dataset

If you do not have real well data, use the included generator to create a realistic Bakken synthetic dataset:

COMMAND `python generate_dataset.py`

This creates bakken_production_10000.csv with exactly 10,000 rows (167 wells × 60 months). The synthetic data uses:

- Log-normal IP distribution (median ~420 boe/d, $\sigma = 0.45$)

- Arps hyperbolic decline: $b = 1.1\text{--}1.6$, $D_i = 55\text{--}80\%/year$
- Six operators with different reservoir quality multipliers
- Five vintage years (2018–2022) with improving completion performance

3.4 Using Your Own Data

To use real field data, export your production database to CSV with the 11 required columns. Common data sources include:

- Enverus / DrillingInfo — export well production report
- IHS Markit — well performance export
- Quorum / WellView — production history extract
- Internal company production database (SQL or Excel export)

After exporting, rename columns to match the required names exactly. Example mapping for a DrillingInfo export:

Source Column Name	Required Column Name
API_UWI	api
Prod_Month	month
Oil_BBL	oil_bbl
Gas_MCF	gas_mcf
Water_BBL	water_bbl
NGL_BBL	ngl_bbl
Lateral_Length	lateral_length_ft
Total_Proppant	proppant_lbs
Frac_Stages	stages
Operator_Name	operator
Spud_Year	vintage_year

4 THE NINE TABS — DETAILED GUIDE

The application is organized into nine tabs accessible from the Tab Navigator window at the top of the screen. Each tab is described in detail below.

Tab 0 · Data Loader

This is the first tab you see after launch. It displays a six-panel summary of the loaded dataset.

Panel 1	IP Rate Distribution — histogram of month-1 BOE production with P10/P50/P90 markers
Panel 2	Active Well Count — how many wells are producing each month
Panel 3	GOR Evolution — median gas-oil ratio over time with P25–P75 confidence band
Panel 4	Water-Oil Ratio — median WOR trend showing water breakthrough timing
Panel 5	Production Heat Map — 40-well × 48-month matrix colored by monthly BOE
Panel 6	Wells by Operator — bar chart showing well count per operator

To load a different CSV file while the application is running, call:

```
COMMAND app.load_csv('path/to/your_data.csv')
```

Tab 1 · Economic Parameters

This tab contains 14 interactive sliders organized into four groups. Adjust any slider and click the Apply Parameters button at the bottom to update all downstream economics calculations.

Price Deck (3 sliders)

- Oil Price (\$/bbl) — WTI flat deck; range \$30–\$150; default \$75
- Gas Price (\$/mcf) — Henry Hub flat; range \$1–\$10; default \$2.50
- NGL Price (\$/bbl) — NGL realization; range \$5–\$60; default \$22

Operating Costs (4 sliders)

- LOE (\$/boe) — Lease operating expense; range \$3–\$25; default \$8.00
- G&A (\$/boe) — General & administrative allocation; range \$0.50–\$8; default \$1.50
- Gathering (\$/boe) — Gathering and transport; range \$0.50–\$8; default \$2.00
- Severance Tax (%) — State production tax as % of gross revenue; default 4.6%

Capital & Fiscal (4 sliders)

- D&C Capex (\$MM) — Drilling and completion cost per well (100% WI); range \$3–\$20; default \$7.50MM
- Working Interest (%) — WI fractional share; range 10–100%; default 75%
- NRI (%) — Net Revenue Interest; range 60–100%; default 81.25%
- Discount Rate (%) — Annual WACC / hurdle rate; range 5–30%; default 10%

Screening Thresholds (3 sliders)

- Min NPV Screen (\$MM) — Wells below this are flagged; default \$1.0MM
- Min IRR Screen (%) — Wells below this IRR fail the screen; default 15%
- Max Payout (months) — Wells beyond this payout period fail; default 36 months

WORKFLOW

Always adjust economic parameters in Tab 1 before switching to the Economics (Tab 5) or A&D (Tab 6) tabs. The Apply Parameters button must be clicked for changes to take effect.

Tab 2 · Type Curve Analysis

Four production profile charts built from the full well dataset:

- Basin-wide P10 / P50 / P90 (linear scale) — shaded confidence band with annotated milestone production at 12 and 48 months
- Log-scale type curve — reveals late-time decline behavior and terminal production rates
- Operator P50 comparison — side-by-side P50 type curves for each operator with at least 3 wells
- Vintage improvement trend — P50 curves by spud year showing completion technology advancement

Type curves use a 60-month evaluation window and apply a 3-month rolling average to smooth month-to-month measurement noise. Only wells with at least 6 active months are included.

Tab 3 · Decline Curve Analysis

This is the most interactive technical tab. It displays a per-well Arps hyperbolic fit alongside three sliders for adjusting the decline parameters.

Reading the Chart

- Blue dots — actual monthly BOE production for the selected well
- Red line — fitted Arps hyperbolic curve
- Subtitle bar — current EUR, annual decline rate (D_i), and b-exponent for the fitted curve
- Stats box (right side) — q_i , D_i , b, R^2 , and EUR in one glance

Adjusting Parameters

Use the three sliders below the chart:

Slider	Parameter	Range	Typical Shale Value
q_i (boe/mo)	Initial rate at $t=0$	100 – 60,000 boe/mo	5,000 – 25,000
D_i (1/mo)	Monthly nominal decline	0.001 – 0.200/mo	0.04 – 0.08
b-exponent	Hyperbolic exponent	0.10 – 2.00	1.1 – 1.6 (tight oil)

Saving Per-Well Overrides

After adjusting sliders to your satisfaction, click the Save Override button. The adjusted parameters are stored in memory for this session and will be used instead of the automated fit whenever that well appears in economics or portfolio calculations.

TIP

To change which well is displayed, call: `app.show_tab(3, selected_api='3310100001')` — replacing the API with your well identifier. The bottom EUR histogram always shows the full portfolio for context.

Tab 4 · Machine Learning — EUR Prediction

Two ML models are trained: Random Forest and Gradient Boosting. Both use the same feature set and are evaluated identically.

Features Used (13 variables)

Early Production Signals (months 1–6)

- `ip_boe` — month-1 BOE (the single strongest predictor)
- `avg_3mo` — average BOE over the first 3 months
- `avg_6mo` — average BOE over the first 6 months
- `cum_6mo` — 6-month cumulative production
- `peak_boe` — maximum monthly BOE in months 1–6
- `gor_early` — median GOR in months 1–6 (gas quality indicator)
- `wor_early` — median WOR in months 1–6 (water handling early signal)
- `oil_cut` — average oil fraction of total liquids

Completion Parameters

- `lateral_length_ft` — horizontal wellbore length
- `proppant_lbs` — total proppant placed
- `stages` — number of frac stages
- `proppant_per_ft` — proppant intensity (key efficiency metric)
- `proppant_per_stage` — proppant per stage

Reading the Results

- Actual vs Predicted scatter — dots should cluster tightly around the 1:1 red dashed line. The shaded band shows $\pm 20\%$ error.
- Feature importance bar — longer bars indicate variables with more predictive power. `ip_boe` and `proppant_per_ft` typically dominate.
- CV R^2 comparison — shows 5-fold cross-validation performance for both models. Higher is better; 0.80+ is excellent for EUR prediction.
- Residual histogram — ideally centered near 0%. Skew indicates systematic over- or under-prediction.
- Hex-bin density — shows where predictions are most dense. Good models concentrate mass near the diagonal.

NOTE

The ML model predicts EUR from early production data only. It is most useful for ranking new wells in the first 3–6 months of production and for screening acreage acquisitions based on analog early-time data.

Tab 5 · Well Economics

Economics are computed for every well in the dataset using the parameters from Tab 1. Results update automatically when you click Apply Parameters in Tab 1 and then switch back to this tab.

Metric Definitions

Metric	Definition	Formula
NPV10	Net Present Value at 10% annual discount	Sum of monthly CF / $(1 + r/12)^t$; index 0 = WI capex outflow
IRR	Internal Rate of Return (annualized)	Annual rate where NPV = 0; solved via Brent's method
Payout	Months until cumulative CF turns positive	First month where cumulative CF ≥ 0
EUR	Estimated Ultimate Recovery (Arps)	Integral of Arps rate to economic limit (5 boe/d)
PI	Profitability Index	$1 + \text{NPV} / \text{WI Capex}$; PI > 1 = value-creating
NPV/boe	Value created per barrel of reserves	NPV / EUR in \$/boe

The Portfolio Quadrant Map

The EUR vs NPV scatter plot divides the portfolio into four quadrants:

- Core (high EUR, high NPV) — top right — primary development targets
- Upside Ops (high EUR, low NPV) — bottom right — cost optimization candidates; may be NPV-positive at higher prices
- Marginal (low EUR, high NPV) — top left — smaller wells still generating value; hold or optimize
- Non-core (low EUR, low NPV) — bottom left — divestment or abandonment candidates

Green dots pass all three screening criteria (NPV, IRR, payout). Red dots fail at least one screen. Adjust the thresholds in Tab 1 to tighten or loosen the criteria.

Tab 6 · A&D Portfolio Dashboard

The A&D tab provides a composite ranking of every well using a weighted scorecard designed for acquisition screening and portfolio management.

A&D Composite Score Methodology

Each well receives a score from 0–100 based on five metrics, each percentile-ranked against the full portfolio:

Metric	Weight	Direction	Rationale
NPV	30%	Higher = better	Primary value measure for an acquirer
IRR	25%	Higher = better	Risk-adjusted return vs. hurdle rate
EUR	20%	Higher = better	Long-term reserve base and production duration
PI	15%	Higher = better	Capital efficiency (value created per dollar spent)
Payout Period	10%	Lower = better	Capital recovery speed (inverted for ranking)

Category Definitions

Category	Score Range	Description	Recommended Action
Core Acquisition	≥ 75	Top-quartile across all metrics	Prioritize development capital; defend in A&D process
Development	50–74	Above-average on most metrics	Continue development; optimize completion design
Marginal	25–49	Below-average; some value	Monitor; reassess at next price deck update
Divestment	$< 25 + \text{NPV} < 0$	Negative NPV and low scores	Evaluate for sale, trade, or abandonment

TIP

The divestment table in the bottom-right panel lists the worst-ranked wells by NPV. Before acting on divestment flags, verify the production data quality for those wells — low scores can result from incomplete data as well as poor reservoir performance.

Tab 7 · Field Development Planning

The Field Development Planning tab answers the question: If we drill N wells per year, what does our production and revenue look like? Seven interactive sliders control the drilling schedule and production assumptions.

The Seven Sliders

Slider	Controls	Default
Wells — Year 1	Wells spudded in the first year of the forecast	4
Wells — Year 2	Incremental wells in year 2	5
Wells — Year 3	Incremental wells in year 3	4
Avg IP (boe/d)	Average initial production rate per new well (linked to P50 type curve)	Dataset P50
Annual Di (%)	Annual nominal decline rate applied to new wells	Dataset median Di
Uptime (%)	Production efficiency factor (accounts for downtime, curtailment)	92%
Oil Price (\$/bbl)	Price used for revenue forecast on this tab only (independent of Tab 1)	75

The Four Output Charts

- Portfolio Production by Vintage — stacked bar showing annual production from each well tranche (Year 1 wells in blue, Year 2 in gold, Year 3 in green)
- Cumulative Production — S-curve of total field production in mmboe over 15 years
- Gross Revenue Forecast — annual gross revenue in \$MM based on the price slider
- NPV Sensitivity to Total Well Count — curve showing estimated undiscounted net cash flow as a function of total wells drilled, with a vertical line at the current well count selection

NOTE

The production forecast uses simple exponential decline (no b-exponent here). For a more precise forecast using Arps hyperbolic parameters, use the Decline Curve tab and run the full economics pipeline. The FDP module is designed for quick scenario comparison, not final reserves booking.

Tab 8 · Export

Two export actions are available:

- **Save All Charts (PNG)** — saves every analysis chart from the cached figures to the shale_pro_output/ folder. Files are named 02_type_curves.png through 08_operator_benchmarking.png.
- **Generate PowerPoint Deck** — builds the 12-slide executive presentation (see Section 5 for full slide descriptions). Saved as Shale_Analytics_Executive_Deck.pptx.

IMPORTANT

You must run `app.run_all()` before exporting. If the application is launched with a CSV file argument (the default), `run_all()` is called automatically. If you load a new file mid-session with `app.load_csv()`, call `app.run_all()` again before exporting.

5 POWERPOINT EXECUTIVE DECK

5.1 Overview

The automatically generated PowerPoint deck is designed for use in investment committee presentations, deal team briefings, and management reviews. It requires no manual editing — all KPI values are pulled directly from the current analysis run.

5.2 Slide-by-Slide Description

Slide	Title	Contents
1	Cover	Asset name, basin, price deck, WI capex, confidentiality statement
2	Executive Summary	10 KPI tiles (P50/P10/P90 NPV, IRR, EUR, payout, screen pass rate, core targets, divest count, ML R^2) + 6-bullet findings narrative
3	Type Curve Analysis	Basin P10/P50/P90 charts, operator P50 comparison, vintage improvement
4	Decline Curve Analysis	3 example Arps fits, EUR distribution histogram, b-exponent chart, fit quality pie
5	ML EUR Prediction	Actual vs predicted scatter, feature importance, CV comparison, residual distribution
6	Well Economics Engine	NPV/IRR/payout histograms, portfolio quadrant map, operator NPV ranking
7	A&D Portfolio Dashboard	A&D score distribution, NPV vs PI scatter, category donut, top-20 ranking
8	Field Development Plan	Drilling schedule production stacks, cumulative production, revenue forecast
9	Completion Optimisation	Proppant vs EUR/ft scatter, quartile box plots, NPV/EUR dual-axis chart
10	Operator Benchmarking	NPV by operator, EUR/ft ranking, normalised performance heatmap, portfolio mix
11	Strategic Recommendations	Six action items with rationale (numbered, color-coded by priority)
12	Key Risks & Mitigants	Risk matrix: 2 high / 2 medium / 2 low with mitigation strategies

5.3 Customising the Deck

To add a company logo, open the generated .pptx in PowerPoint and insert the image on the Slide Master (View → Slide Master). This will apply it to all slides.

To change the color scheme, modify the color constants at the top of `shale_analytics_gui.py` before running:

EXAMPLE

Change `NAVY = "0D2137"` to your firm's brand color (6-digit hex). All header bars, KPI tiles, and table headers will update automatically on the next run.

6 ECONOMIC METHODOLOGY

6.1 Cash Flow Model

Monthly cash flows are constructed for each well as follows:

Line Item	Formula	Parameters
Gross Revenue	oil × oil_price + gas × gas_price + ngl × ngl_price	Monthly production from dataset
NRI Revenue	Gross Revenue × NRI	NRI from Tab 1 slider (default 81.25%)
Severance Tax	Gross Revenue × severance_pct	State tax % from Tab 1 (default 4.6%)
Operating Costs	(LOE + G&A + Gather) × BOE × WI	All unit costs from Tab 1
Net Operating CF	NRI Revenue – Severance Tax – Operating Costs	Monthly operating income
Capex	WI Capex at month 0 (negative)	D&C capex × Working Interest
Net CF	Net Operating CF – Capex	Combined CF; index 0 = capex outflow

6.2 NPV Calculation

NPV uses monthly discounting with the annual discount rate from Tab 1:

$$NPV = \sum CF(t) / (1 + r/12)^t \quad \text{for } t = 0, 1, 2, \dots, N$$

where r = annual discount rate · t = month index · $CF(0) = -WI \text{ Capex}$

6.3 IRR Calculation

IRR is the annualized rate r such that $NPV = 0$. It is solved numerically using Brent's method on the NPV function. If no sign change is found in the range $[-99.9\%, +5,000\%]$, IRR returns NaN (the well is uneconomic at any positive rate).

6.4 Arps Hyperbolic Decline

The decline curve model follows the standard Arps hyperbolic formulation:

$$q(t) = q_i / (1 + b \cdot D_i \cdot t)^{(1/b)}$$

q_i = initial rate · D_i = monthly nominal decline · b = hyperbolic exponent · t = months

EUR is computed by integrating $q(t)$ from $t=0$ to the economic limit (where $q = 5$ boe/month) using Brent's root-finding to locate the time at which $q = 5$, then applying the trapezoidal rule.

Fitting uses bounded nonlinear least-squares (`scipy.optimize.curve_fit`) with bounds:

- q_i : 1 to 1,000,000 boe/month
- D_i : 0.001 to 0.990 per month
- b : 0.10 to 2.00

Wells with fewer than 6 months of data above 5 boe/month return NaN for all Arps parameters and are excluded from the EUR distribution but retained in economics (EUR is set to NaN).

7 COMMAND LINE REFERENCE

7.1 All Arguments

Argument	Type	Default	Description
--csv PATH	string	bakken_production_10000.csv	Path to the input CSV file
--oil PRICE	float	75.0	Oil price in \$/bbl
--gas PRICE	float	2.50	Gas price in \$/mcf
--disc RATE	float	10.0	Annual discount rate in %
--no-gui	flag	off	Headless batch mode — no display required

7.2 Python API (in-session commands)

When running in GUI mode, the following commands can be called from a Python console or by modifying the script directly:

Command	Description
app.load_csv('path.csv')	Load a new CSV file and trigger data cleaning
app.run_all()	Run DCA, ML, economics, and portfolio in sequence
app.show_tab(N)	Switch to tab N (0–8)
app.export_pptx()	Generate and save the PowerPoint deck
app.save_all_charts()	Save all cached charts as PNG files
ECON.oil_price = 85	Override oil price programmatically (then run app.run_all())
ARPS_OVERRIDES['api'] = (qi, Di, b)	Set manual Arps parameters for a specific well API

7.3 Output Files

All output files are saved to the shale_pro_output/ folder, created automatically in the current working directory.

File	Description	When Created
02_type_curves.png	P10/P50/P90 type curve analysis chart	After run_all() or Save All Charts
03_decline_curves.png	DCA example wells + EUR distribution	After run_all()
04_ml_results.png	ML model performance charts	After run_all()
05_economics.png	Economics distribution + quadrant map	After run_all()
06_ad_dashboard.png	A&D portfolio ranking + category mix	After run_all()
07_completion_optimisation.png	Proppant loading efficiency charts	After run_all()
08_operator_benchmarking.png	Cross-operator comparison dashboard	After run_all()
Shale_Analytics_Executive_Deck.pptx	12-slide PowerPoint presentation	After export_pptx()

File	Description	When Created
results.csv	Per-well NPV, IRR, EUR, A&D score, category	Batch mode only

8 INTERPRETING RESULTS

8.1 Economic Screen Benchmarks

The following ranges are typical for Bakken and Eagle Ford tight-oil plays at \$70–\$80/bbl WTI. Adjust expectations for Haynesville gas or other basin types.

Metric	Below Average	Average	Above Average	Exceptional
NPV10 per well	< \$0.5MM	\$0.5–\$2MM	\$2–\$5MM	> \$5MM
IRR	< 10%	10–20%	20–35%	> 35%
Payout period	> 48 months	36–48 months	24–36 months	< 24 months
EUR (mboe)	< 200	200–400	400–700	> 700
EUR/ft (boe/ft)	< 30	30–50	50–80	> 80
PI	< 1.0x	1.0–1.3x	1.3–2.0x	> 2.0x

8.2 ML Model Quality Benchmarks

R ² Score	Interpretation	Action
> 0.85	Excellent — model captures most EUR variance	Use confidently for A&D screening
0.70–0.85	Good — typical for real-world production data	Use with moderate confidence; flag outlier predictions
0.50–0.70	Fair — meaningful but significant unexplained variance	Use for ranking only; not for absolute EUR estimates
< 0.50	Poor — model is not reliable	Review data quality; check for outliers or data gaps

8.3 Common Interpretation Pitfalls

PITFALL 1

High A&D score ≠ high absolute NPV. The score is a percentile rank within the portfolio — a well can score 80/100 because it is better than 80% of peers, even if its absolute NPV is small. Always check the raw NPV column too.

PITFALL 2

EUR from Arps fitting on early production data (months 1–12) can overestimate EUR if the b-exponent is too high. Bakken b-exponents above 1.6 should be reviewed carefully. Use the Tab 3 sliders to test sensitivity.

PITFALL 3

The ML model is trained on your loaded dataset. If you use it to screen acreage in a different basin or with different completion designs, feature importance and prediction accuracy may be materially different.

PITFALL 4

NPV uses a flat price deck by default. In a volatile commodity environment, run sensitivity analyses (adjust the oil price slider in Tab 1 to \$50, \$60, \$80, \$100) and compare which wells remain economic across the full range.

9 TROUBLESHOOTING

9.1 Common Errors and Fixes

Error / Symptom	Likely Cause	Fix
ModuleNotFoundError: No module named 'pplx'	python-pplx not installed	Run: pip install python-pplx
No module named 'sklearn'	scikit-learn not installed	Run: pip install scikit-learn
No display name and no \$DISPLAY	No display in headless environment	Add --no-gui flag to the command
All IRR values are NaN	All wells are uneconomic at the current price deck	Lower capex or raise oil price in Tab 1 and click Apply
DCA converged: 0/N wells	Wells have fewer than 6 active months	Use longer production history; minimum 6 months required
KeyError: 'ngl_bbl'	Column missing from CSV	Add an ngl_bbl column; fill with 0 if NGL not produced
PowerPoint generation fails with AttributeError	run_all() not completed before export	Call app.run_all() first; then retry export
Charts look blank or very small	matplotlib display scaling issue	Resize the content window manually; or use --no-gui and view saved PNGs
Very slow DCA fitting (> 5 minutes)	Large dataset with many wells	This is normal for 500+ wells; reduce --n-wells in batch or filter CSV first

9.2 Performance Tips

- For datasets above 500 wells, DCA fitting takes 2–5 minutes. This is expected — fitting scipy curve_fit to hundreds of nonlinear problems is compute-intensive.
- ML training is fast (under 30 seconds) for datasets up to 50,000 rows.
- If you are re-running economics only (not DCA), skip re-fitting by using ARPS_OVERRIDES to cache the previous fit results.
- On macOS, the matplotlib window may require you to click on it once to bring it into focus after launch.

9.3 Getting Help

If you encounter an issue not covered here:

1. Check that all dependencies are installed with the correct versions
2. Verify that your CSV file has all 11 required columns with correct names
3. Try running in --no-gui batch mode to isolate display-related issues from calculation issues
4. Inspect the Python traceback — the last line always identifies the file, line number, and error type

10 QUICK REFERENCE CARD

Standard Workflow (5 minutes)

STEP 1	Launch with your data <code>python shale_analytics_gui.py --csv your_data.csv</code> <i>→ Application opens, loads data, runs full pipeline automatically.</i>
STEP 2	Review data quality (Tab 0) Check the IP distribution and active well count chart. <i>→ Flag any wells with anomalous early spikes before proceeding.</i>
STEP 3	Set your price deck (Tab 1) Adjust oil price, discount rate, and capex sliders. <i>→ Click Apply Parameters — all economics update immediately.</i>
STEP 4	Validate key DCA fits (Tab 3) Review the top 3 wells shown. Adjust sliders if needed; click Save Override. <i>→ Repeat for any wells that are critical to your analysis.</i>
STEP 5	Review economics & A&D ranking (Tabs 5–6) Note the screen pass rate and identify Core / Divestment candidates. <i>→ Use the quadrant map to spot outliers for further review.</i>
STEP 6	Build the drilling schedule (Tab 7) Adjust well count per year, IP, and decline sliders. <i>→ Check the NPV sensitivity curve to optimize well count.</i>
STEP 7	Export (Tab 8) Click Generate PowerPoint Deck. <i>→ File saved to shale_pro_output/Shale_Analytics_Executive_Deck.pptx</i>

Key Formulas at a Glance

Formula	Expression
BOE conversion	$1 \text{ mcf} = 1/6 \text{ boe} \cdot 1 \text{ bbl NGL} = 1 \text{ boe}$
Arps hyperbolic rate	$q(t) = q_i / (1 + b \cdot D_i \cdot t)^{(1/b)}$
Monthly NPV discount	$DF(t) = 1 / (1 + r_{\text{annual}}/12)^t$
Net revenue	$\text{Gross Rev} \times \text{NRI} - \text{Gross Rev} \times \text{severance_pct}$
Net operating CF	$\text{NRI Revenue} - \text{Severance Tax} - (\text{LOE} + \text{G\&A} + \text{Gather}) \times \text{BOE} \times \text{WI}$

Formula	Expression
A&D score	$0.30 \times \text{NPV_pct} + 0.25 \times \text{IRR_pct} + 0.20 \times \text{EUR_pct} + 0.15 \times \text{PI_pct} + 0.10 \times (1 - \text{Payout_pct})$
Proppant intensity	Total proppant (lbs) / Lateral length (ft)
EUR/ft efficiency	EUR (boe) / Lateral length (ft)

Abbreviations

Abbrev.	Meaning
A&D	Acquisitions & Divestitures
Arps	Arps decline curve (named after J.J. Arps, 1945)
BOE	Barrel of oil equivalent
Di	Initial decline rate (nominal monthly)
DCA	Decline Curve Analysis
EUR	Estimated Ultimate Recovery
GOR	Gas-Oil Ratio (scf/bbl)
IP	Initial Production rate
IRR	Internal Rate of Return
LOE	Lease Operating Expense
ML	Machine Learning
NGL	Natural Gas Liquids
NPV	Net Present Value
NRI	Net Revenue Interest
P10	10th percentile (top 10% of wells)
P50	50th percentile (median)
P90	90th percentile (bottom 10% of wells)
PI	Profitability Index
WI	Working Interest
WOR	Water-Oil Ratio

SHALE ANALYTICS PRO · Demo Version

For questions about methodology, data requirements, or customization, refer to the inline code comments in `shale_analytics_gui.py`. Every function includes a full docstring explaining inputs, outputs, and assumptions.