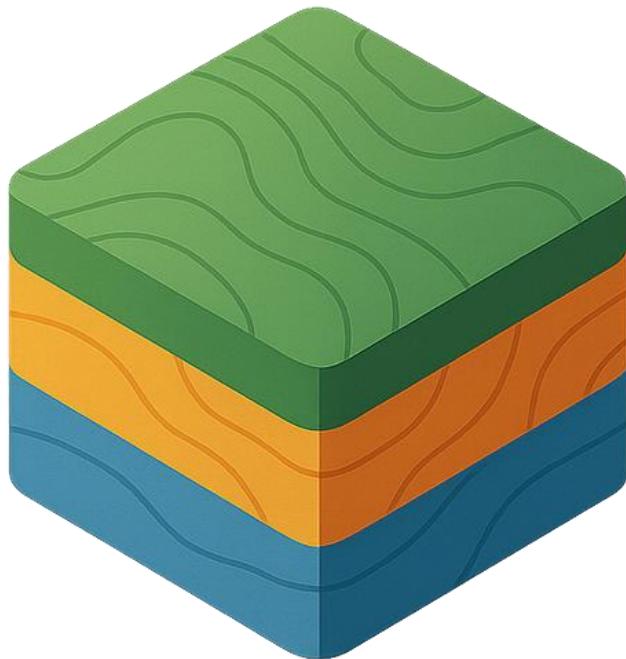


GeoInsights 3D User Guide

Rapid Drilling Data Visualisation and Analysis for Geologists with AI



GeoInsights
3D

Introduction

GeoInsights 3D is a purpose-built geological analysis tool that transforms drilling data into meaningful visualisations in minutes, not hours. Built with Python and Streamlit, it streamlines the process of taking raw drilling data and turning it into actionable insights.

Key Benefits:

- **Rapid Analysis:** Get from raw data to insights in minutes
- **AI Integration:** Large Language Model capabilities for automated geological interpretation
- **Standard Format Support:** Handles typical CSV files and Australian State Geological Survey formats
- **No Formatting Required:** Direct analysis of standard drilling data formats
- **Exploratory Focus:** Perfect first port of call for initial data exploration

Where GeoInsights 3D Fits:

Think of GeoInsights 3D as your first port of call—a powerful tool for quick visualisation and initial analysis. Once you've identified interesting patterns, you can move your data into specialised software like loGAS for multivariate statistical analysis or Leapfrog for detailed 3D modelling.

Try GeoInsights 3D now: <https://geoinsights3d.koyeb.app/>

Getting Started

System Requirements:

- Web browser (Chrome, Firefox, Safari, Edge)
- Internet connection
- Google API key (for AI features - free from aistudio.google.com)

Installation Options:

Option 1: Web Application Simply navigate to: <https://geoinsights3d.koyeb.app/>

Option 2: Local Installation

Using pip, you can install GeoInsights 3D directly from the Python Package Index (PyPI).

1. Create and activate a virtual environment (Strongly Recommended):

It is best practice to install packages in a dedicated virtual environment to avoid conflicts with other projects.

For Windows:

- python -m venv venv
- .\venv\Scripts\activate
- **For macOS/Linux:**
- python3 -m venv venv
- source venv/bin/activate

2. Install GeoInsights 3D:

With your virtual environment active, install the package using pip:

```
pip install geoinsights-3d
```

3. Set up API Key (Optional for AI features):

The generative AI features require a Google Gemini API key.

Create a file named `.streamlit/secrets.toml` in the root of the project directory.

Add your API key to this file as follows:

```
.streamlit/secrets.toml GOOGLE_API_KEY = "YOUR_API_KEY_HERE"
```

Running the Application

Once the installation is complete, run the following command from the project's root directory:

```
geoinsights-3d
```

Data Requirements:

GeoInsights 3D accepts:

- **CSV files** with headers in the first row
- **Australian State Geological Survey formats** with headers in the H1000 row
- Three main data types:
 - **Collar Data:** Drillhole locations and orientations
 - **Assay Data:** Sample elemental concentrations
 - **Lithology Data:** Downhole logging and logging code dictionary

Interface Overview

GeoInsights 3D features an intuitive tab-based interface that guides you through the data analysis workflow:

Navigation Tabs:

- **Data Loading:** Upload and prepare your drilling data
- **3D Visualisations and Cross-Sections:** Explore data in interactive 3D space
- **Automated 3D Grade Shell Generator:** Create anisotropic grade shells
- **3D Solid Model Generation:** Generate 3D solids for categorical data
- **Statistics:** Analyse distributions, correlations, and summary statistics
- **Clustering:** Apply geochemical clustering techniques
- **Geochemical Anomaly Detection:** Identify geochemically distinct samples
- **SHAP Analysis:** Unveil drivers of mineralisation with explainable AI
- **AI Geo Summary:** Automated geological interpretation using LLMs
- **Data Analysis Playground:** Ask questions in plain English
- **Export Data:** Save processed data for other applications

This simple navigation structure allows you to move freely between different analysis types while maintaining the context of your dataset.

The screenshot displays the GeoInsights 3D application interface. At the top center is a stylized 3D cube composed of four horizontal layers in green, yellow, orange, and blue. Below the cube, the text "GeoInsights 3D" is prominently displayed in a bold, sans-serif font. Underneath the main title, the subtitle "Exploratory Data Analysis and Visualisation" is centered. A horizontal navigation bar at the bottom of the main content area contains several tabs: "Data Loading" (which is highlighted in red), "3D Visualisations", "Statistics", "Clustering", "ML Explain", "AI GEO Summary", "Data Analysis Playground", and "Export Data". The "Data Loading" tab's content area is visible, featuring a large input field with the placeholder text "Upload your data files to begin analysis or load demo data to quickly explore the app's features." Below this input field is a small button labeled "Load Demo Data". To the left of the main content area, there is a sidebar titled "LLM Integration" containing two input fields: "Enter your Google API Key" and "Enter the Google AI Model". In the top right corner of the main content area, there are two small icons: "Deploy" and a three-dot menu icon.

1. Getting Started — Data Upload & Compositing

First things first — getting your data in. GeoInsights 3D directly handles the data files in typical csv format with the headers in the first row or the formats specified by the Australian State Geological Surveys with the headers in the H1000 row. Once you have your files, the upload interface is straightforward and the with three main sections:

- Collar Data: Contains your drillhole locations and orientations
- Assay Data: Your sample elemental concentrations
- Lithology Data: Downhole logging and logging code dictionary

Don't have data to hand? Try the Load Demo Data button to begin exploring with a real-world dataset

Assay Data Preview:

	Hole ID	Sample ID	From	To	AuPpm	CuPpm	PbPpm	ZnPpm	AgPpm	AsPpm	BiPpm	CdPpm	CoPpm	FePer	MnPpm	MoPpm	NiPpm	SbPpm	WPpm	Drill Code
0	RC11RE0002	4018216	0	4	0.005	27	21	66	0.1	15	1	0.25	6	4.58	99	0.5	24	1	5	RC
1	RC11RE0002	4018217	4	8	0.038	45	17	111	0.1	18	1	0.25	4	4.74	122	0.5	25	1	5	RC
2	RC11RE0002	4018218	8	12	0.01	32	29	134	0.1	8	1	0.25	6	3.23	83	1	24	1	5	RC
3	RC11RE0002	4018219	12	16	0.006	34	30	222	0.1	18	1	0.25	9	3.9	140	0.5	27	1	5	RC
4	RC11RE0002	4018220	16	20	0.006	42	39	194	0.1	22	2	0.25	20	4.21	669	1	36	2	5	RC

Select Assay Columns

Select HOLE_ID column (Assay)

Hole ID

Select FROM column

From

Select TO column

To

Select all

Select element columns

AuPpm × CuPpm × PbPpm × ZnPpm × AgPpm × AsPpm × BiPpm × CdPpm × CoPpm × FePer ×
MnPpm × MoPpm × NiPpm × SbPpm × WPpm ×

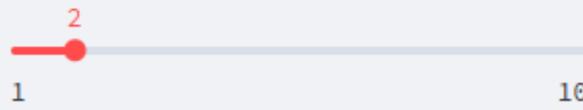
Column mapping and element selection.

Once your data is loaded, GeoInsight 3D's compositing tool offers flexible compositing options that let you standardise your sampling intervals using length-weighted, averaging methods. This flexibility allows you to view and analyse your data at whatever resolution best suits your needs. Compositing standardises your data, smoothing out noise while preserving important trends, and can make cross-hole comparisons more reliable. The tool's immediate visual feedback helps you quickly optimise your composite parameters to match your deposit's characteristics and your analytical objectives.

Compositing Options

Composite geochemical data

Composite Interval (m)

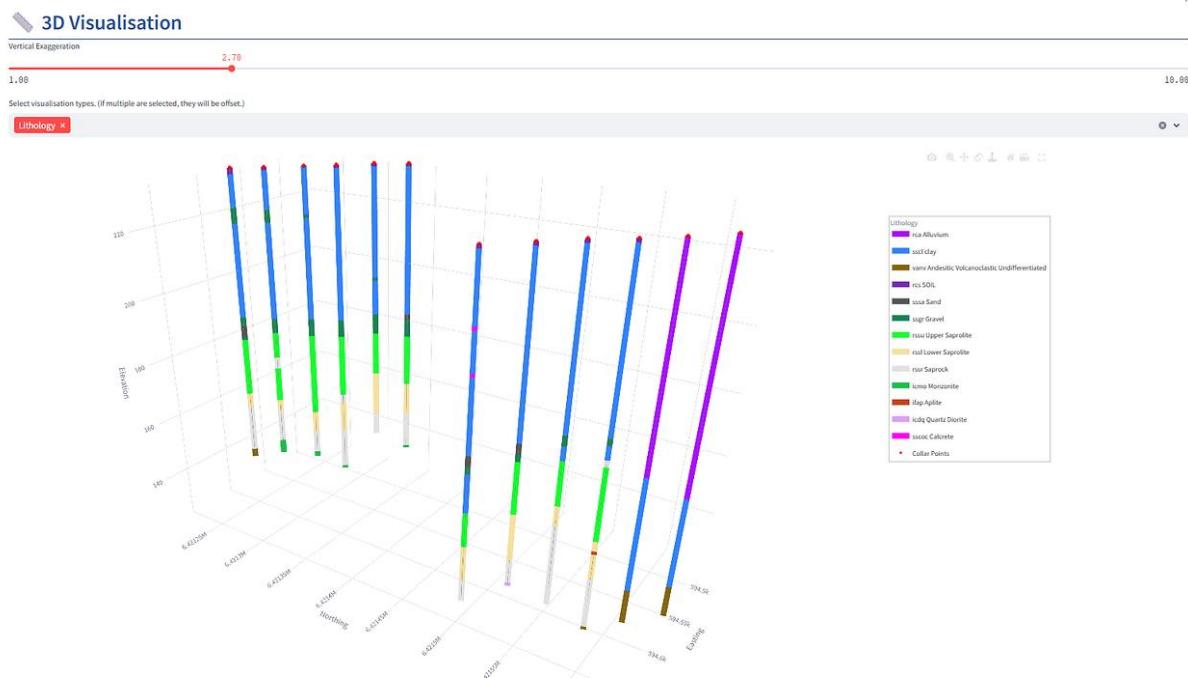


Compositing Options

2. 3D Visualisations

At the heart of GeoInsights 3D is its interactive 3D viewer, designed specifically for drilling data. Once loaded, your data springs to life in a fully interactive 3D space where you can:

- Examine drillholes in true 3D space
- Offset grade and lithology traces for clearer viewing of both datasets simultaneously
- Rotate, zoom, and pan with intuitive mouse controls
- Adjust vertical exaggeration to enhance subtle features in shallow drilling



Interactive 3d Visualisation

Powerful Filtering Options:

- Select specific drillholes or drilling fences
- Filter intervals by lithology or alteration codes

- Set grade thresholds
- Filters can be applied to the visualisations or to the subsequent analysis and calculations
- Toggle between grade views, lithology views, or combined displays

Filter Options

Apply filters to all analyses (not just visualisation)

Select holes to display (empty for all)

Choose an option ▾

Use log scale

Select element for analysis:

Fe ▾

Fe cutoff range

0.51 16.45

0.51 16.45

Select lithologies to display (empty for all)

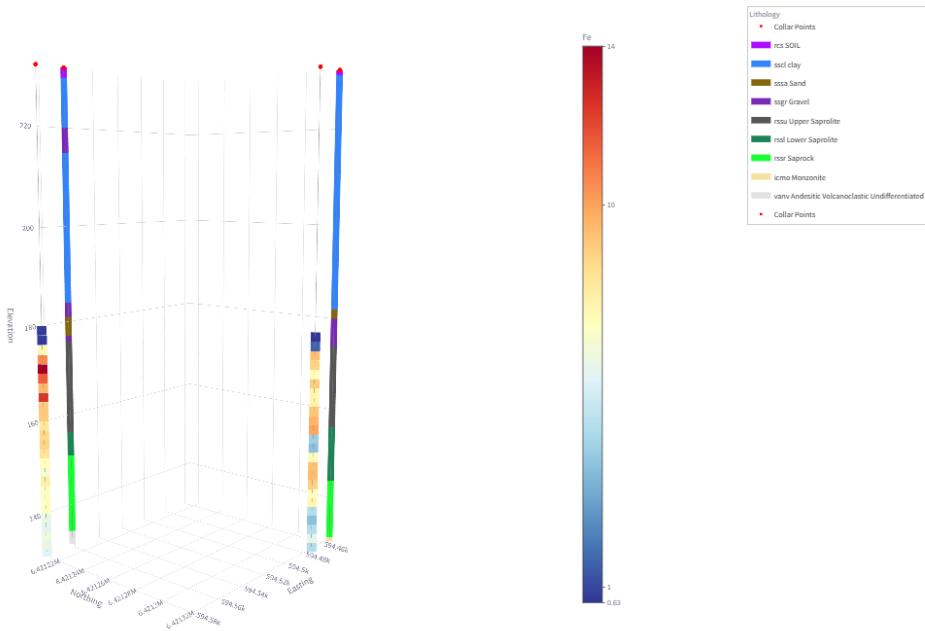
Choose an option ▾

Filtering options

Select visualisation types. (If multiple are selected, they will be offset.)

Grade Lithology

Download 3d Visualisation (HTML)



Grade and Lithology for selected drillholes

Navigation is straightforward, left-click and drag to rotate, right-click and drag to pan, and use your mouse wheel to zoom. The grade filter slider lets you dynamically adjust minimum and maximum values, making it easy to identify and investigate high-grade zones and their relationships to specific rock types. You can download the interactive 3d plot by hitting the Download 3d Visualisation button. You can share this with anyone, it will run natively in any web browser.

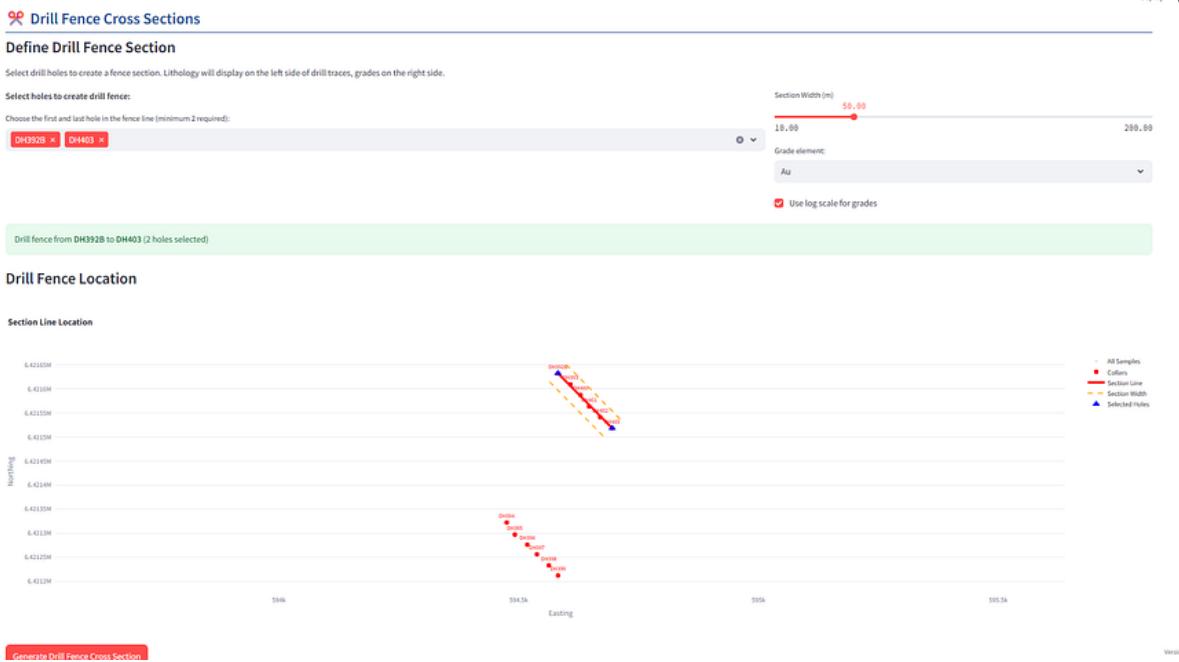
 Download 3D Visualisation (HTML)

Pro Tips:

- Start by filtering to a single drill fence to get familiar with your geology
- Use the combined grade-lithology view to spot relationships between mineralisation and rock types
- Experiment with vertical exaggeration to better visualise shallow or deep features

Cross-Section Analysis — Drill Fence Visualisation

Understanding geological relationships often requires examining data along specific orientations or drill fence lines. GeoInsights 3D includes cross-section tools that create drill fence visualisations, allowing you to examine your data along user-defined section lines.

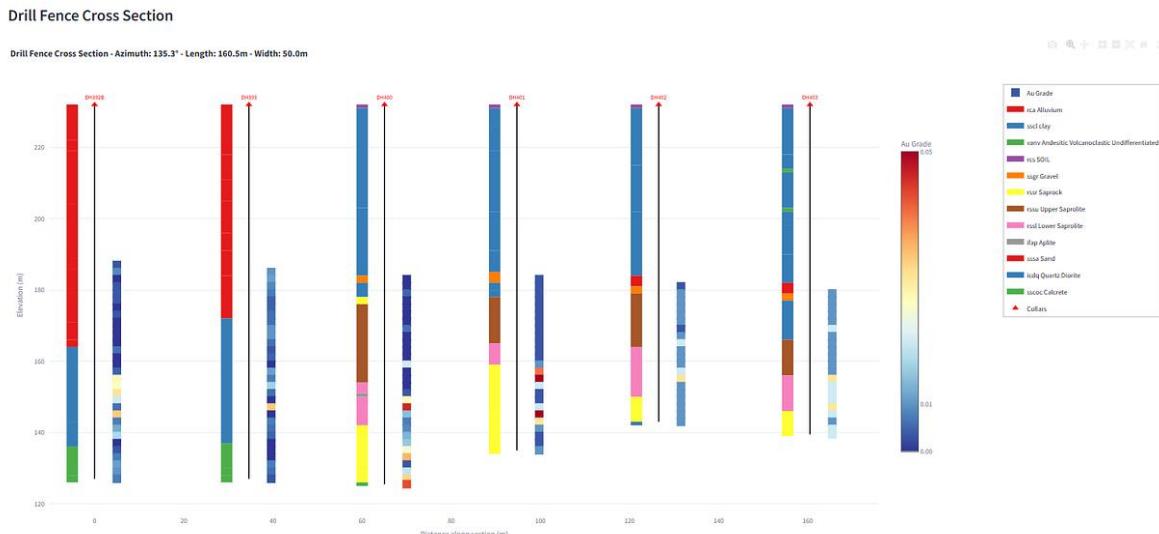


Cross-section location map showing selected drill holes and section line

Creating Cross Sections

The cross-section tool allows you to:

- Select multiple drill holes to create a fence line
- Define section width to capture nearby data
- Display lithology intervals on the left side of drill traces
- Show grade values on the right side of drill traces
- Visualise drill hole traces connecting collar to downhole data



Drill fence cross-section showing lithology bars on left, drill traces in center, and grade points on right

The cross-section automatically calculates the section azimuth and projects all selected drill holes onto the fence line. This provides a clear geological interpretation view where you can:

- Identify geological continuity between holes
- Spot structural controls on mineralisation
- Validate logging consistency across the fence
- Examine grade distribution patterns in 2D space

3. Automated 3D Grade Shell Generator

This feature allows users to quickly and objectively define the shape and orientation of a mineralised body from drillhole data. Traditionally, creating a 3D grade shell requires geologists to select an appropriate cut-off value and then manually interpret and define the orientation of the the strike, dip, and plunge of the mineralisation. This process can be time-consuming, subjective and require specialised software. The GeoInsights 3D Grade Shell Generator automates this critical step. By using Principal Component Analysis (PCA) on the high-grade portion of your data, the tool automatically discovers the natural trend of the mineralisation. This data-driven, automated anisotropic modelling means you can generate objective, repeatable 3D models in a seconds.

How It Works: The “Grade Shell” Tab

1. Load Your Data: As always, start by loading your collar and assay files.
2. Navigate to the Grade Shell Tab: Click on the “ Grade Shell” tab.
3. Set Your Parameters: Select your element of interest, choose your compositing and transform options, and set the grid resolution.
4. Generate! Click the “Generate Grade Shell” button.

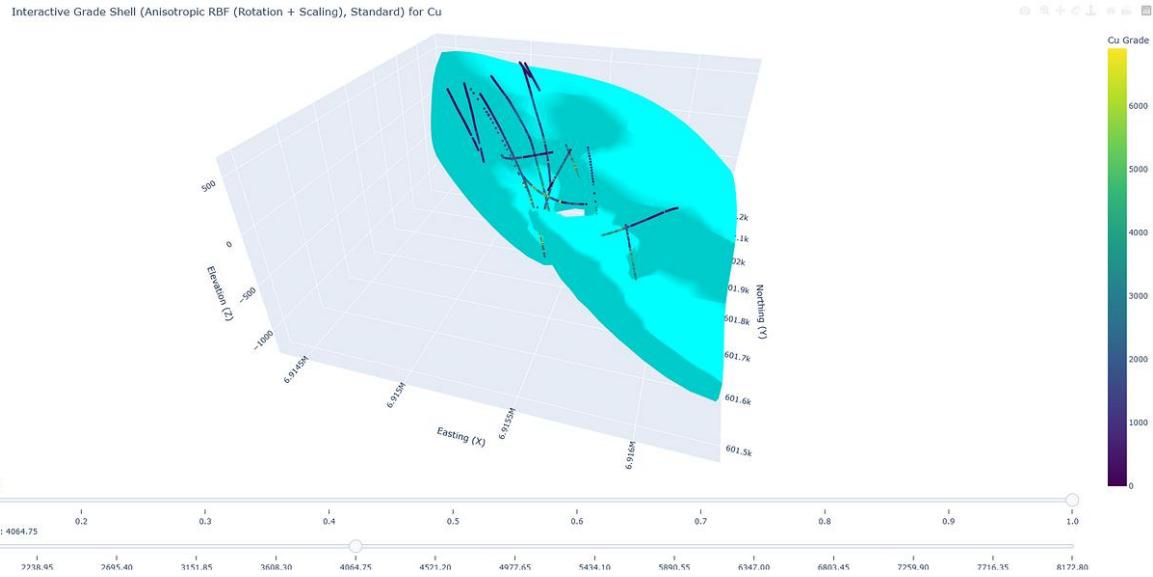
 **Grade Shell Generation**

This tool generates a 3D grade shell using a Radial Basis Function (RBF) interpolator. Adjust the parameters and click the button to begin.

Grade Shell Parameters

Select Element for Shell	Use Log Transform	Grid Resolution
Cu	<input checked="" type="checkbox"/>	20
Select Interpolation Method	<input checked="" type="checkbox"/> Composite Data for Modeling	60
anisotropic	<input type="radio"/> Composite Length (m)	100
	<input type="radio"/> 5.00	1.00
	<input type="radio"/> 2.00	2.00

Generate Grade Shell



The interactive grade shell for copper (Cu) in this example deposit. The slider at the bottom allows for dynamic adjustment of the opacity and grade cutoff, revealing the coherent shape of the high-grade core.

The tool utilises Radial Basis Function (RBF) interpolation. To honor this anisotropy, the model analyses the data to determine the principal direction of mineralisation for rotation and then calculates scaling factors based on data variance to represent the deposit's shape.

4. 3D Solid Model Generation: Visualising Geological Units

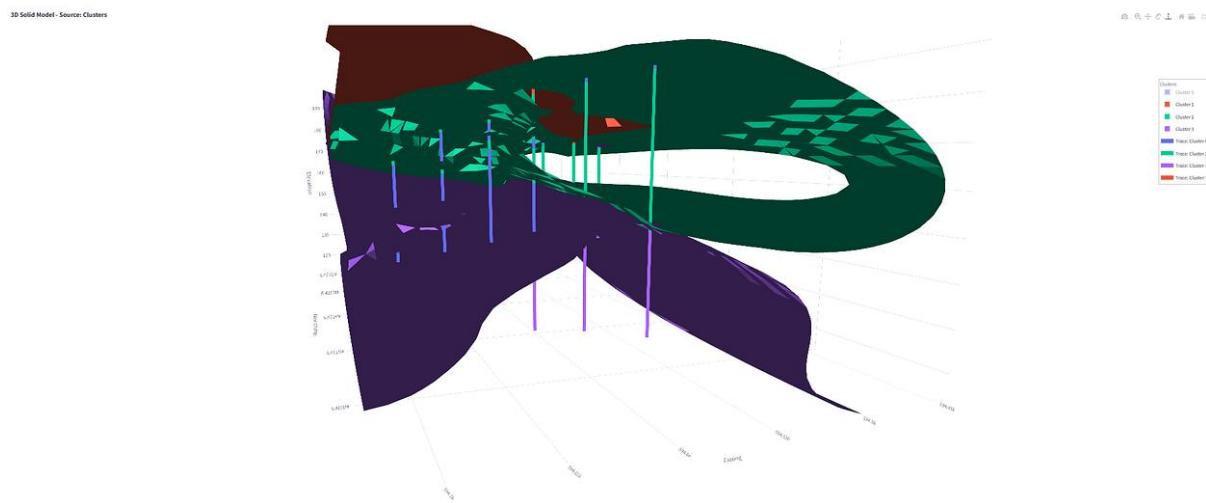
While grade shells are excellent for visualising mineralisation trends, understanding the overall geology requires modelling the rock units themselves. The new 3D Solid Model Generation tab extends the powerful modelling engine to categorical data, allowing you to create 3D solids for logged lithologies or for the domains identified in the geochemical clustering tab.

How It Works:

- Select Source:** Choose whether to model your logged **Lithology** or the calculated **Clusters**.
- Choose Units:** Select the specific rock types or cluster domains you want to visualise.
- Generate:** Click the “Generate 3D Solid Model” button.

To ensure the generated solids are geologically realistic, the tool employs the same data-driven engine that powers the grade shell generator. For each selected rock type or cluster, it automatically discovers the unit’s natural orientation (strike, dip, and plunge) from the drillhole data. This ensures the models reflect the underlying geology, rather than appearing as simple, uniform shapes.

The final interactive plot displays the generated solids, each with a distinct colour. For context and validation, the original drillhole intervals are drawn as thick, colour-coded lines running through the model. You can toggle the visibility of each solid using the plot legend, adjust opacity, and download the entire scene as a standalone HTML file to share with colleagues.



Interactive 3D view showing multiple colored geochemical cluster solids with drillhole traces, demonstrating the spatial relationship between different rock units.

5. Summary Statistics — Understanding Your Data

GeoInsight 3D provides a statistical dashboard that gives you an immediate overview of your dataset:

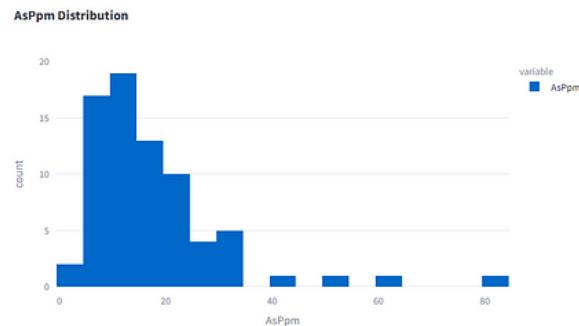
- Statistical summaries showing means, quartiles, and potential outliers
- Histograms for all numeric fields
- One-click transformation between linear and log scales — particularly useful for elements like gold or arsenic that often show strong right-skew distributions.

Summary Statistics

Summary Statistics

Statistic	Value
Count	74
Mean	17.635
Median	14
Std Dev	13.137
CV	0.745
Min	1
Q1	9.25
Q3	22
Max	84
Skewness	2.643
Kurtosis	9.816

Histogram



Statistical dashboard showing grade distributions and summary statistics.

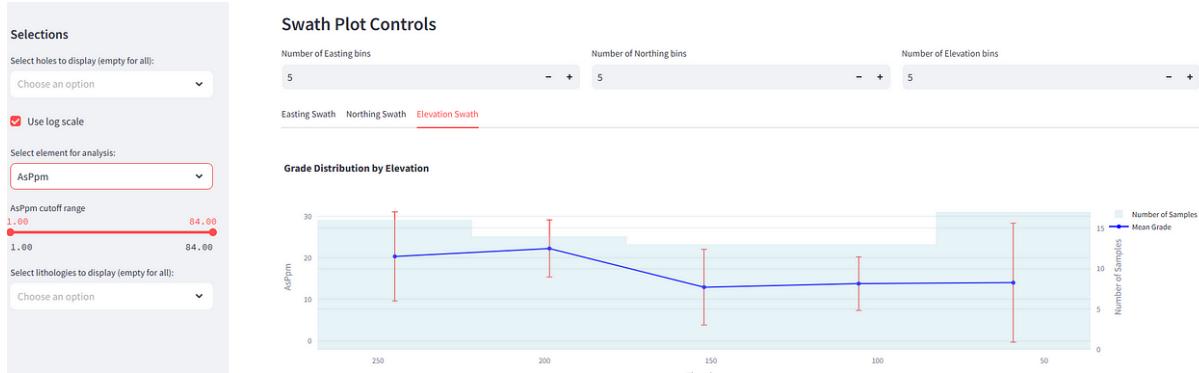
Swath Analysis — The Big Picture

One of the most powerful features is its interactive swath analysis. These plots reveal how grades vary across your deposit through:

- East-West and North-South grade trends along customisable sections

- Vertical grade changes with depth
- Sample density overlays to validate data coverage
- Grade variability bands showing standard deviation — super helpful for understanding grade consistency

The tool automatically calculates and displays these trends, helping you identify grade patterns, structural controls, and potential domains. You can adjust the number of bins and section widths to optimise the analysis for your deposit scale.



Swath plots showing grade trends across the deposit with standard deviation bands (red) and sample density overlay.

Correlation Analysis

Understanding element relationships is crucial for geochemical interpretation. The correlation tools help identify element associations that might indicate mineralisation styles or alteration patterns. The correlation tools help you explore these relationships through:

- Interactive correlation matrices
- Scatter plots for element pairs
- Identify element associations



Global correlation matrix.

Scatter Plots

Select elements for scatter plot (minimum 2)

AuPpm CuPpm PbPpm AlPpm



Correlation Matrix

	AuPpm	CuPpm	MnPPM	AlPpm
AuPpm	1.00	0.16	-0.16	0.32
CuPpm	-0.14	1.00	0.13	-0.03
MnPPM	0.14	0.09	1.00	-0.25
AlPpm	0.32	0.09	-0.26	1.00

Selected scatter plots and correlation matrix

Logged Unit Analysis

Lithologies and logged units play a crucial role in domain definition and deposit understanding. While multivariate geochemistry analysis often helps refine geological domain boundaries, well-logged lithology and alteration can provide initial constraints for domaining. There are several tools for analysing your logged units:

- Box plots for each rock type
- Grade statistics by lithology
- Interactive filtering
- Minimum sample thresholds

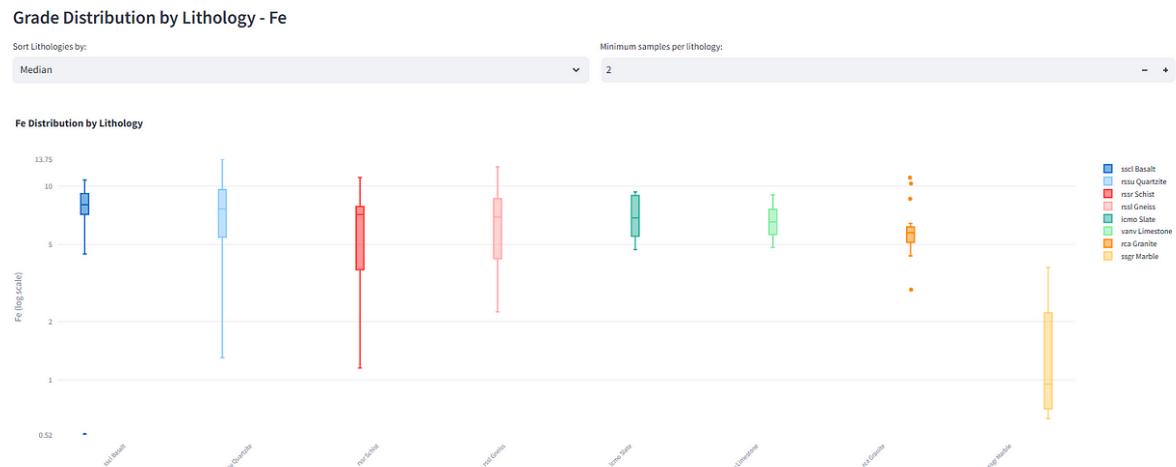
These built-in visualisations help you quickly understand element distributions between rock types. The boxplots effectively display where statistical populations differ, highlighting potential domain boundaries that align with logged units.

Lithology Analysis

Summary Statistics by Lithology - FePer ↗

LITHO	Description	Count	Mean	Median	Std Dev	Min	Q1	Q3	Max	CV	LITHO_LABEL
SS	Siliclastic	42	3.354	3.335	0.519	2.27	2.992	3.848	4.24	0.155	SS Siliclastic
SI	Andesitic	25	3.84	3.84	0.432	2.73	3.59	4.01	4.74	0.113	SI Andesitic
QV	Quartz Vein	6	3.833	4	0.525	2.89	3.675	4.182	4.3	0.137	QV Quartz Vein
CO	Basaltic	1	4.58	4.58	None	4.58	4.58	4.58	4.58	None	CO Basaltic

Summary statistics by logged units



Box plots displaying elemental concentration for each lithology.

Finding the Good Stuff

GeoInsights 3D includes a flexible significant intervals calculator that lets you quickly identify and report your best drill intersections. You can:

- Set custom grade cutoffs for different elements
- Specify minimum intersection lengths
- Adjust maximum internal waste tolerance
- View results with associated lithologies
- Export results as formatted tables for reporting

Significant Intervals

The screenshot shows a user interface for calculating significant intervals. At the top, there are three input fields: 'Minimum Interval Length (m)' set to 1.00, 'Maximum Internal Waste (m)' set to 2.00, and 'Minimum AuPpm Grade' set to 1.00. Below these is a red-bordered button labeled 'Calculate Significant Intervals'. Underneath the button is a table with the following data:

HOLE_ID	FROM	TO	LENGTH	AuPpm_GRADE	INTERNAL_WASTE	LITHOLOGY
0	RC11RE002	38	39	1	1.88	0 SS

Customisable significant intervals calculator with adjustable cutoffs and internal waste parameters

6. ⚙ Geochemical Clustering — Quick Domain Insights

One of the most powerful features in GeoInsights 3D is the geochemical clustering tool. While detailed geological logging is valuable, it often faces challenges:

- Inconsistency between geologists
- Variable logging quality
- Subtle mineralogical changes not visible to the naked eye
- Time-consuming manual review of large datasets

Geochemical clustering provides a rapid, objective first pass at identifying major rock types and alteration domains. GeoInsights 3D offers two approaches: straightforward K-means clustering of your raw data, or a more advanced option that first transforms the data using Principal Component Analysis (PCA) before clustering. Both methods have their merits — direct clustering is simpler and more transparent, while PCA transformation can help reduce noise and highlight subtle patterns.

To demonstrate the core workflow, we'll use the simpler K-means approach on a dataset where the geology consists of a weathered profile of transported cover, saprolite and saprock.

Remember, while clustering is powerful, it should always be integrated with geological understanding and validated against known geology, structure and controls on mineralisation.

Feature Selection: Before clustering, it's important to select appropriate elements. Using biplot and correlation matrix visualisations, we can identify key elements while removing those that are redundant or noisy. As geochemical data typically shows skewed distributions, we apply a natural log transform to normalise the data.

Geochemical Clustering

Select All Features Select features for clustering

Al As Cu Na K Mg Zr Th Y Cr Ni Fe Ca

Perform PCA before clustering
 Apply natural log transform

Maximum number of clusters to consider:

Analyse Clusters

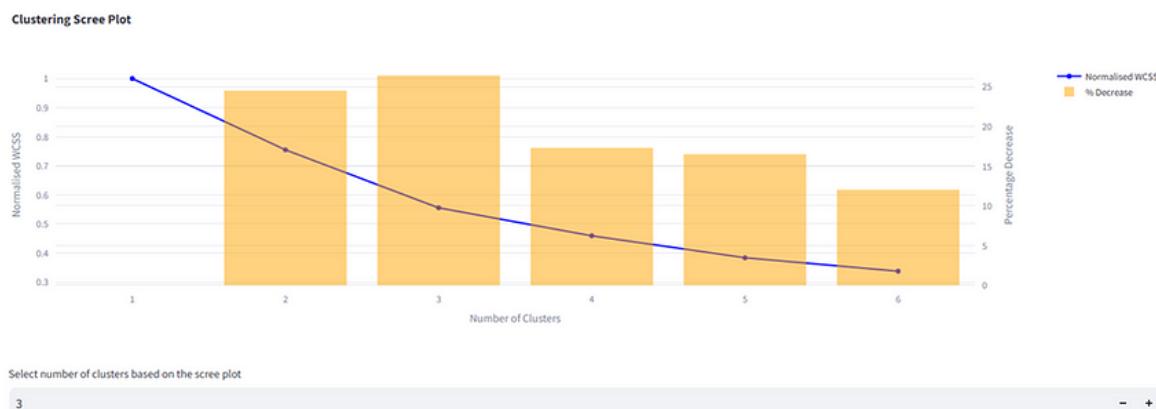
Geochemical clustering feature selection and options

Pro Tips:

1. Start with major elements (Al, Fe, Ca, Mg, Na, K) for broad lithological domains
2. Add pathfinder elements to refine alteration and mineralisation patterns

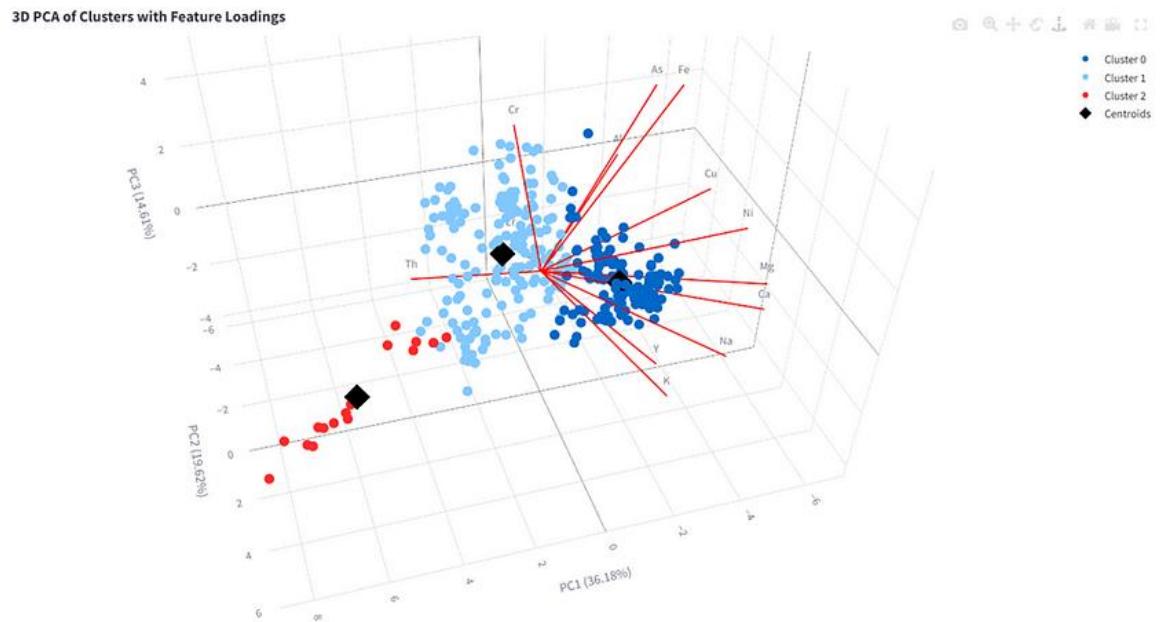
Clustering: Clustering: Applying k-means clustering to the transformed geochemical data revealed three distinct domains. While the weathering profile contains additional complexity, this three-cluster solution provides a simple test case, capturing the main lithological units of interest. The choice is supported by both the scree plot's elbow at k=3 and the coherent spatial groupings that form in the 3D viewer

Clustering Scree Plot ↗



K-means clustering scree plot

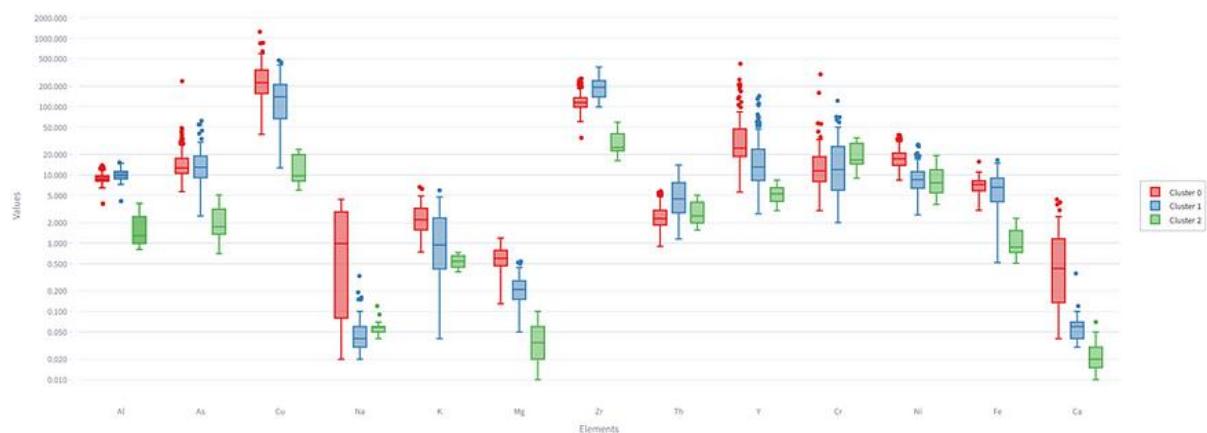
3D PCA of Clusters



3D PCA of Clusters along the first 3 principal components

Statistical Validation: Summary statistics and box plots reveal distinct geochemical fingerprints for each cluster:

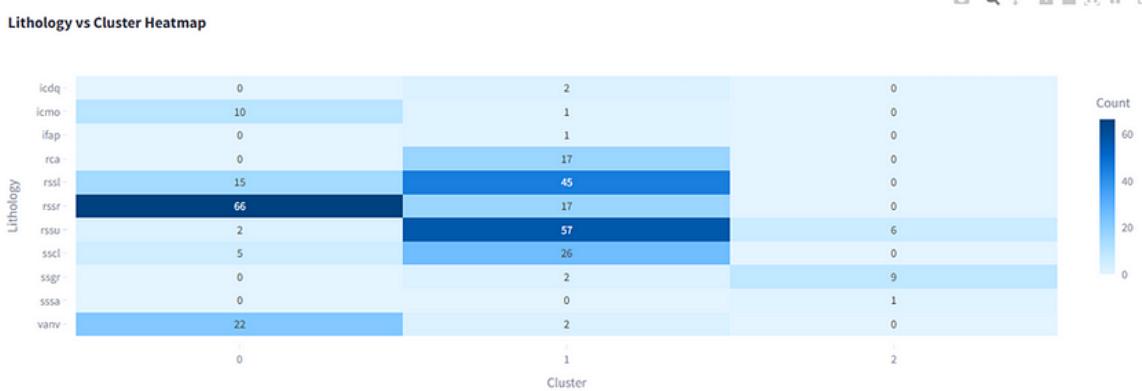
- Cluster 0 (Saprock): Elevated base metals
- Cluster 1(Lower saprolite): Lower concentrations of mobile elements (Na, Mg, and Ca) compared to cluster 0
- Cluster 2 (Cover): Depleted in base metals and mobile elements



Box plots displaying the distributions of the selected features

Comparison with Logged Data: The lithology vs cluster heatmap provides a quick correlate clustering results against traditional logging. For example, the heatmap allows us to cross check our clusters against logged geology: Cluster 0 correlates strongly with the saprock (rssr) as well as the minor andesite (vanv) and monzonite (icmo) units, Cluster 1 correlates with the

upper (rssu) and lower (rssl) saprolite units, and Cluster 2 correlates most strongly with the logged gravel (ssgr) units.

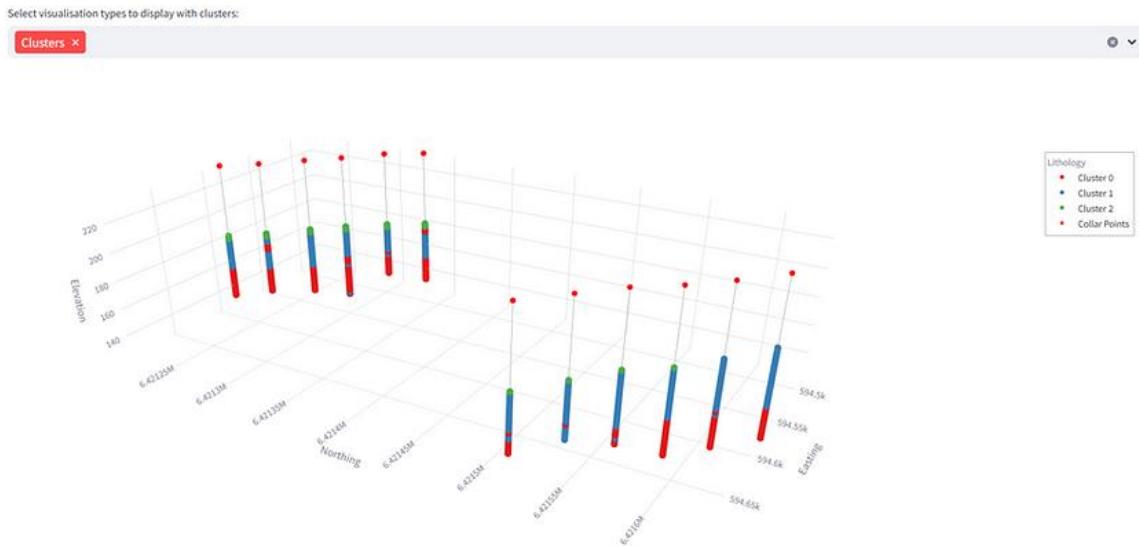


Cluster vs logged units heat map

Spatial Validation — 3D View

The critical final step is checking that clusters form coherent geological shapes in 3D space — random or scattered patterns would indicate poor clustering results. In our case, these distinct geochemical signatures generally align well with the logged weathering boundaries. The clusters display strong horizontal continuity, reflecting the expected geometry of weathering horizons.

3D Visualisation with Clusters

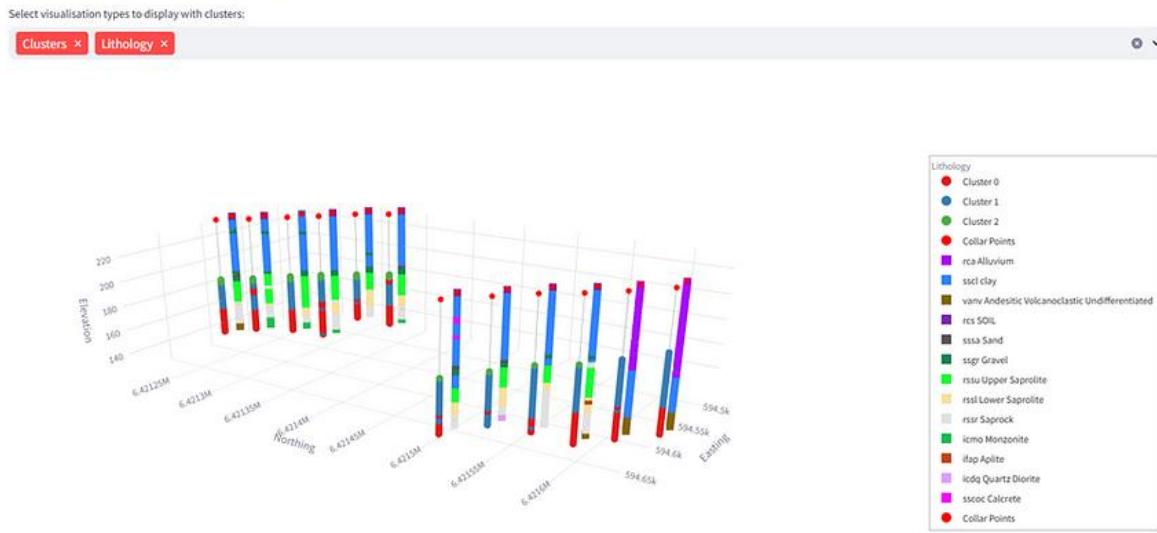


3D validation of created geochemical clusters

When viewed in 3D alongside the lithological logging, the geochemical domains show better hole-to-hole correlation than the original logging, particularly in the southeastern portion of the project. Here, two holes show significant logging inconsistencies compared to the broader

dataset — a common challenge when multiple geologists are involved in logging weathered material.

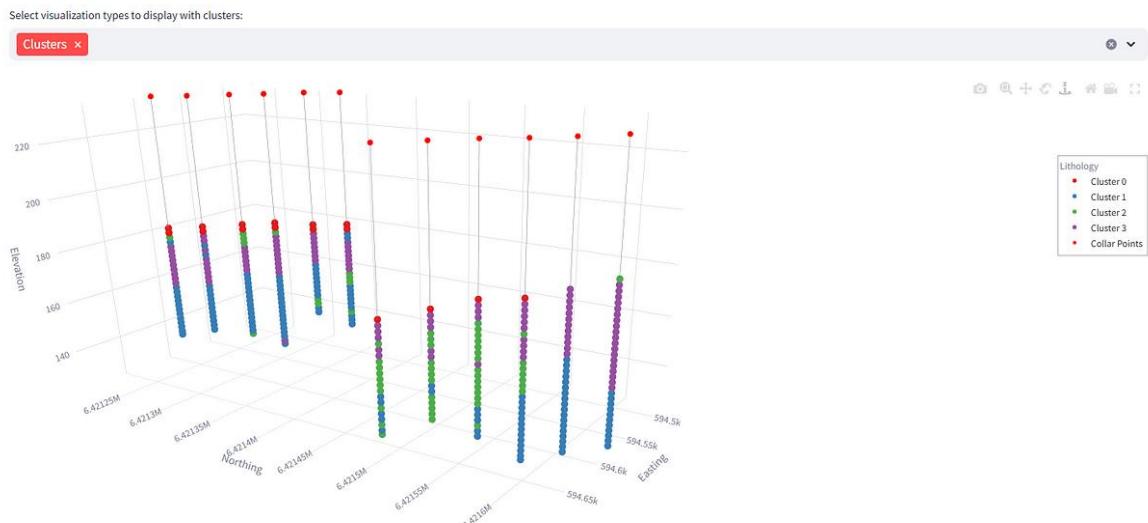
3D Visualisation with Clusters



3D validation of geochemical clusters offset against logged units

While visual logging can effectively identify these weathering horizons, determining the exact boundaries between lower saprolite, upper saprolite, and saprock units often remains subjective. The objective nature of geochemical clustering helps standardise these boundary determinations. If we wanted to delve into further detail, applying k-means with k=4 would reveal distinct geochemical differences between the northern and southern saprolite units, reflecting variations in weathering intensity across the deposit.

3D visualisation with Clusters

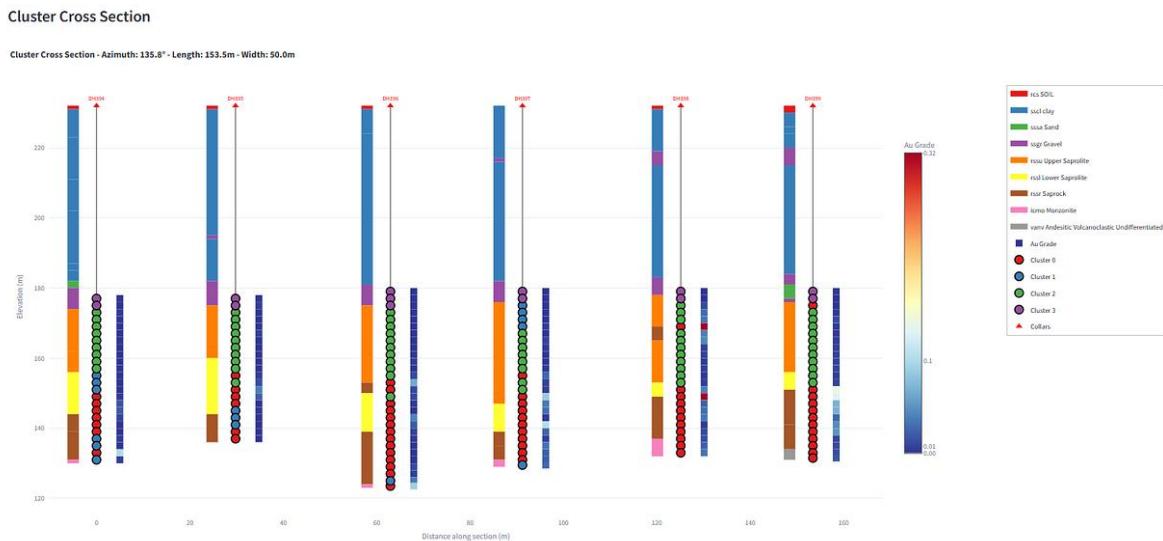


3D visualisation of geochemical clusters with increased resolution (k=4)

Spatial Validation — Cross-Section

The cluster cross-section tool provides:

- Clusters on drill traces: Colored circles directly on the drill hole traces showing cluster assignments
- Optional lithology display: Lithology intervals on the left side for geological context
- Optional grade overlay: Grade values on the right side for mineralisation context
- Spatial continuity validation: Visual confirmation that clusters form coherent geological domains



Cluster cross-section with both lithology bars and grade points, showing the relationship between geochemical domains, rock types, and mineralisation

7. 🔎 Geochemical Anomaly Detection

The Geochemical Anomaly Detection tool is designed to identify samples that are geochemically distinct from a defined baseline. Its primary uses are for data quality control, such as highlighting logging inconsistencies, and for identifying distinct geochemical populations that may represent different alteration styles or rock units. The tool uses an Autoencoder neural network to learn a “normal” geochemical signature and then flags any samples that deviate from this pattern.

The Method: How Autoencoders Work

An Autoencoder is a type of neural network designed for unsupervised learning. It operates in two stages. First, it compresses the input geochemical data into a lower-dimensional representation, forcing it to learn the most important underlying patterns. Second, it attempts to reconstruct the original data from this compressed representation.

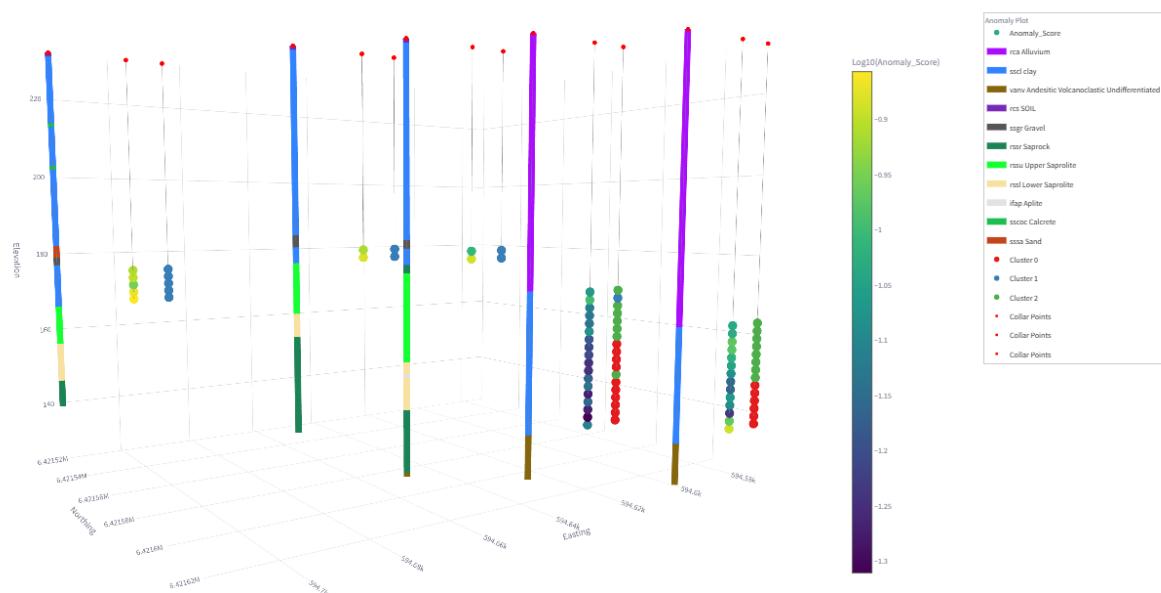
The model is trained on a dataset you define as “normal”. It becomes proficient at accurately reconstructing samples that fit this expected geochemical signature. When the model is given a sample that is geochemically different, its reconstruction will be poor. The degree of this

inaccuracy is quantified as the Anomaly Score. A high score indicates a significant deviation from the learned norm, flagging the sample for investigation.

Example Workflow: Identifying logging errors

In this example, we will look at the example company logging data. As suggested in the Geochemical Clustering section, the logging in the two southeastern-most drillholes appeared inconsistent with the rest of the deposit. We can use the anomaly detector to quantify this inconsistency.

The process involves filtering the analysis to train the autoencoder only on the intervals logged as “clay” within those two suspect southeastern holes. This forces the model to learn their specific geochemical signature as the definition of “clay”. When this specific model is applied the intervals logged as “clay” across the rest of the deposit are all flagged with high Anomaly Scores. This demonstrates quantitatively that the material logged as “clay” in the two southeastern holes is geochemically distinct from the clay found elsewhere, as was indicated in the example geochemical clustering workflow. This provides additional evidence of a logging inconsistency.



A 3D view showing all drillholes. The intervals logged as clay (blue, left) in the two southeastern holes show a low Anomaly Score (centre), in contrast, the clay intervals in all other holes display high Anomaly Scores, and their cluster (right) assignment confirms they belong to a different geochemical populations.

8. ⚡ SHAP Analysis: Unveiling Drivers of Mineralisation with Explainable AI

GeoInsights 3D now incorporates SHAP (SHapley Additive exPlanations) analysis to provide deeper insights into what drives your mineralisation. SHAP is a cutting-edge method from the field of explainable AI that helps you understand the output of machine learning models by quantifying the contribution of each geological feature to a prediction, for instance, the predicted grade of a target element.

Exploratory Data Analysis and Visualisation

Data Loading Visualisations Statistics Clustering SHAP AI GEO Summary Data Analysis Playground Export Data

SHAP Analysis

SHAP (SHapley Additive exPlanations) is a method to explain the output of machine learning models.

It provides insights into the contribution of each feature to the model's predictions.

Select target element for SHAP analysis

Cu

Select data subset for SHAP analysis

- All Data
- Specific Cluster
- Specific Lithology

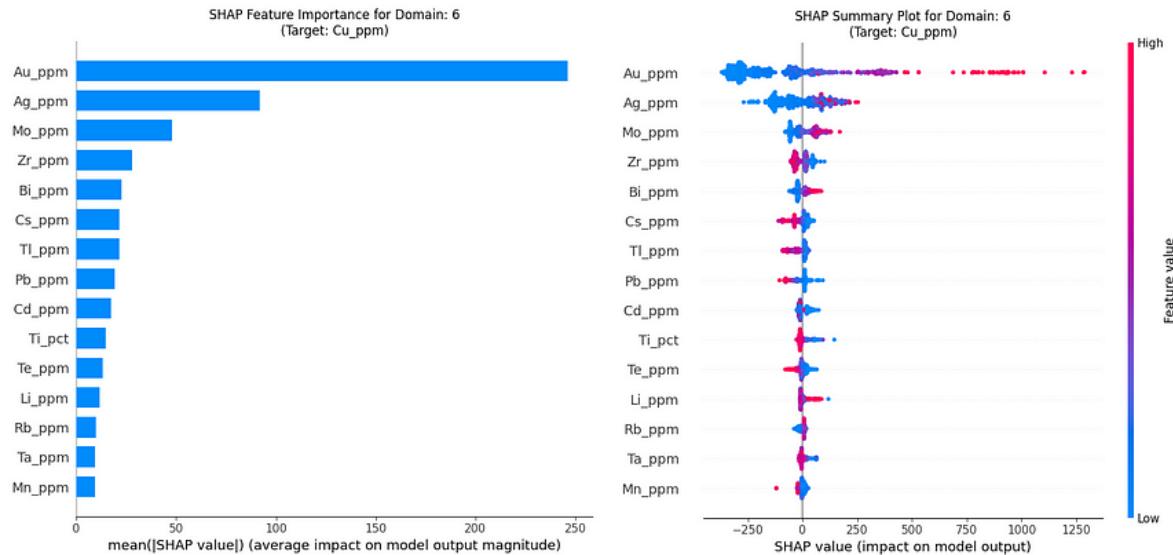
Training Model for SHAP Analysis

Target: Cu

Model trained successfully.

How GeoInsights 3D Implements SHAP Analysis:

- **Target Element Focus:** You can select a specific target element (e.g., Gold, Copper) for which you want to understand the influencing factors
- **Flexible Data Subsetting:** Run SHAP analysis on your entire dataset, or focus on specific geological domains by selecting particular lithologies or geochemical clusters identified in GeoInsights 3D
- **Automated Model Training:** GeoInsights 3D automatically trains a RandomForestRegressor model in the background using your selected data subset and multielement geochemistry to predict the target element's values
- **Calculating Feature Contributions:** SHAP values are calculated for each feature for every individual data point. These values represent how much each feature has pushed the model's prediction for the target element away from the average prediction
- **Summary Bar Plot:** Clearly shows the overall importance of each feature
- **Beeswarm Plot:** Provides more detail, showing not only the importance of each feature but also how the value of that feature affects the prediction



This powerful feature allows geologists to:

- Identify key multivariate elemental associations and geological factors controlling mineralisation
- Validate or challenge existing geological hypotheses with data-driven evidence
- Gain a more nuanced understanding of the complex relationships within their geochemical data, leading to more informed exploration decisions

New AI-Powered Features

9. AI Geo Summary: Automated Geological Interpretation

The new LLM Summary tab leverages the power of Google's Gemini models to provide comprehensive geological interpretations of your data. After processing your drilling information through GeoInsights 3D's visualisation and analysis tools, you can now:

- Generate detailed geological summaries with a single click
- Add your own geological context to guide the interpretation
- Receive insights about element associations, spatial trends, and domain characteristics
- Ask follow-up questions to explore specific aspects of the interpretation

This feature essentially gives you an AI geological assistant that can help identify patterns, suggest interpretations, and highlight areas that might warrant further investigation. GeoInsights 3D utilises automatic prompt engineering that builds context-rich prompts from your data and analysis. The system works by analysing all available information from your dataset, including:

- Statistical distributions and correlations between elements
- Spatial trends identified in swath plots

- Geochemical cluster analysis results and their relationship to lithology
- Significant intervals and their geological context

You can then have a natural conversation with the AI about your data, asking follow-up questions or providing additional information to refine the interpretation.

Data Loading Visualizations Statistics Clustering Export Data AI GEO Summary Data Analysis Playground

AI GEO Analysis

You: No additional context provided.

AI: Okay, here is a geological interpretation based on the provided geochemical analysis summary:

Geological Interpretation Summary

This interpretation synthesizes the provided geochemical data from 12 drillholes (306 samples) focusing on Au within a weathered profile setting.

- Key Geochemical Characteristics & Element Associations:**
 - Low Gold Tenor: The overall Au grade is very low (Mean 0.02 ppm, Median 0.01 ppm), suggesting the sampled area represents background levels or only weakly mineralized material within the weathered zone. The skewed distribution (mean > median) indicates the presence of a few higher values influencing the mean.
 - Weathered Profile Dominance: The dominant lithologies (Upper Saprolite, Saprock, Lower Saprolite) confirm sampling primarily within a deeply weathered regolith profile.
 - Subtle Lithological Control on Au: Slightly higher median Au grades are noted in Saprock (rsr) and Lower Saprolite (rsll) compared to Upper Saprolite (rssu) and especially Gravel (ssgr). This may suggest either better preservation of primary mineralization deeper in the profile or depletion in the more intensely weathered upper zones. However, the absolute difference in median grades (all 0.01 ppm except gravel) is minimal.
 - Multi-Element Variability: Despite low Au, the cluster analysis reveals significant variability in major and trace elements, driven primarily by weathering intensity and likely protolith variations.
- Interpretation of Geochemical Clusters:** The four clusters represent distinct geochemical environments within the sampled area:
 - Cluster 3 (Gravel - ssgr Dominated): Characterized by strong depletion in nearly all analyzed elements (Al, Fe, K, Mg, Mn, base metals, lithophiles). This clearly represents highly weathered, leached, and potentially transported material (Gravel), or a residual quartz-rich lag. It is essentially barren of Au (Median 0.00 ppm).
 - Cluster 0 (Saprock - rssr Dominated): Enriched in elements often associated with less intense weathering or mafic/ultramafic precursors (Mn, Co, Ni, Mg, Ca, Sr, P) and base metals (Cu, Zn). The Saprock dominance supports less weathering. This may represent the geochemical signature of the less-weathered parent rock, potentially with some mafic affinity or Mn-oxide accumulation. Au median remains low (0.01 ppm).
 - Cluster 1 (Mixed/Upper Saprolite - rsll Dominated): Characterized by high Al, K, Rb, Cs, Ga, and immobile/resistate elements (Th, U, REEs, Zr, Hf, Nb, Ta, Ti, Sn). This signature suggests weathering of a more felsic protolith and/or concentration of resistate minerals and clays (high Al) potentially associated with potassio-alteration (K, Rb). Au median is low (0.01 ppm).
 - Cluster 2 (Upper Saprolite - rssu Dominated): Distinguished by high Al, Fe, Sc, V, Ti, Ga, Ge, and notably elevated As and Sb, along with moderate Cu. This points to intense weathering (Upper Saprolite dominance) with significant clay (Al) and Fe-oxide/hydroxide development (Fe). The elevated As, Sb, and potentially Cu may indicate either scavenging by Fe-oxides or, more significantly, reflect an underlying primary mineralization signature (e.g., oxidized sulfides) where these elements act as pathfinders. Au median is low (0.01 ppm).
- Spatial Distribution Patterns:**
 - Easting Trend: The scatter plot reveals a subtle E-W trend in Au grades, with a peak (mean 0.03 ppm) centered around 594541 E. This suggests a potential N-S oriented control on mineralization or fluid flow (e.g., structure, lithological contact) influencing even the low background Au levels. Grades decrease east and west of this central zone.
 - Cluster Distribution (Inferred): Although explicit spatial plotting isn't provided, the lithological associations suggest potential spatial patterns. Cluster 3 (Gravel) might represent surficial or channel deposits. Clusters 1 and 2 (Upper Saprolite) likely dominate shallower zones, while Cluster 0 (Saprock) would be deeper. Spatially correlating cluster membership with the Existing trend could reveal if the peak Au zone corresponds to a specific geochemical environment (e.g., enrichment within Cluster 2).
- Significance of High-Grade Intervals:**
 - Two significant intervals exceeding an unspecified cutoff grade were detected.
 - Their occurrence within Lower Saprolite (rsll) and Upper Saprolite (rssu), despite the low median grades for these lithologies, highlights that localized zones of higher Au concentration exist.
 - These intervals likely represent narrow structures, veins, or specific horizons that are not broadly pervasive but are important indicators of mineralization potential.
 - Their geological context is crucial: determining if they align with the E-W grade trend peak (~594541 E) and whether they fall within a specific geochemical cluster (e.g., Cluster 2 with high As/Sb) would be key to understanding their origin and targeting similar zones.

Overall Geological Synthesis & Implications:

The geochemical data depicts a deeply weathered landscape with low background gold concentrations. However, distinct geochemical variations exist, successfully captured by the cluster analysis, reflecting differences in protolith, weathering intensity, and potentially alteration styles (e.g., K-rich Cluster 1 vs. Fe-As-Sb-rich Cluster 2).

While widespread economic mineralization is not evident in the bulk statistics, the presence of a subtle spatial Au trend (E-W) and discrete higher-grade intervals within the saprolite profile confirms the presence of gold-bearing processes. The association of Cluster 2 with elevated pathfinder elements (As, Sb) in intensely weathered zones (Upper Saprolite) might be particularly significant, potentially pointing towards oxidized primary mineralization below.

Recommendations for Further Work:

- Spatially plot the cluster distributions and significant intervals relative to the Existing trend and known geology/structures.
- Investigate the specific element associations within the significant intervals.
- Consider if the pathfinder elements (As, Sb in Cluster 2) provide a better vector towards potential primary mineralization than low-level Au itself in this weathered environment.
- Integrate these findings with geological mapping and geophysical data to refine targets.

Ack: A follow-up question or provide additional information:

How does this work?

- First, provide your Google Gemini API key in the sidebar and optional geological context.
- Click 'Generate Initial Summary' to analyze your drill hole data.
- The AI will create an initial interpretation based on your data.
- You can then ask follow-up questions or provide additional information.
- The system maintains the conversation context throughout your session.
- Click 'Clear Conversation' to start over with a fresh analysis.

10. Data Analysis Playground: Ask Questions of your data in Plain English

The Data Analysis Playground tab introduces a dynamic way to interact with your drilling data. Instead of writing custom scripts or performing manual filtering or calculations, you can now simply ask questions in plain English:

- “What’s the average gold grade in holes that intersect the andesite unit?”
- “Plot a histogram of arsenic values in the saprolite zone”
- “Show me the correlation between copper and gold in cluster 2”

- “Create a box plot comparing silver grades across all lithologies”

Behind the scenes, GeoInsights 3D:

- Interprets your question using the Gemini model
- Generates appropriate Python code to answer your query
- Executes the python code against your loaded dataset in a sandboxed environment; visualisations use matplotlib or plotly
- Returns the results as text, or visualisations

Data Loading Visualisations Statistics Clustering Export Data AI GEO Summary Data Analysis Playground

Data Analysis Playground

Ask a question in plain English about your drillhole data (e.g. What's the average Au grade for Hole B1? or Plot a histogram of gold values)

which hole has the highest au intercept and what is it?

[Get Answer](#)

Generated Code

```
max_au_idx = df['Au'].idxmax()
hole_id_max_au = df.loc[max_au_idx, 'HOLE_ID']
max_au_value = df.loc[max_au_idx, 'Au']
answer = f"The hole with the highest Au intercept is {hole_id_max_au} with a value of {max_au_value:.2f}."
print(answer)
```

Text Output

The hole with the highest Au intercept is MYAC398 with a value of 0.32.

Data Analysis Playground

Ask a question in plain English about your drillhole data (e.g. What's the average Au grade for Hole B1? or Plot a histogram of gold values)

create a scatter plot of log10 au vs log10 as

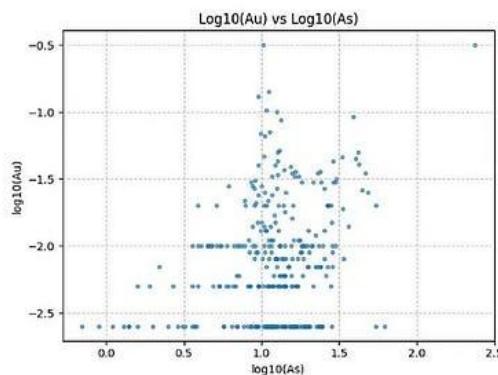
[Get Answer](#)

Generated Code

```
log10_Au = np.log10(df['Au'].apply(lambda x: x if x > 0 else np.nan))
log10_As = np.log10(df['As'].apply(lambda x: x if x > 0 else np.nan))

plt.figure()
plt.scatter(log10_As, log10_Au, alpha=0.6, s=10) # Added alpha and size for better visualization if points overlap
plt.title("Log10(Au) vs Log10(As)")
plt.xlabel("log10(As)")
plt.ylabel("log10(Au)")
plt.grid(True, linestyle='--', alpha=0.7)
plt.tight_layout()
```

visualisation



[Example Visualisation Questions](#)

This feature dramatically speeds up the exploratory data analysis process, allowing you to explore and follow your geological intuition.

How to Use the New AI Features

To access these new capabilities:

1. Enter your Google API key in the sidebar (your key remains on your device and is never stored). Why Google? You can get started with a free account from aistudio.google.com/
2. Process your drilling data through the standard GeoInsights 3D workflow
3. Navigate to either the “SHAP”, “AI Geo Summary” or “Data Analysis Playground” tabs
4. For summaries: Add any additional geological context and click “Generate Initial Summary”. You can then provide further feedback to the LLM or ask follow up questions
5. For Q&A: Type your question and click “Get Answer”

The system works best when you've already explored your data using GeoInsights 3D's visualisation and analysis tools, and you can provide the system with as much geological context as possible for interpretation.

11. Save Processed Data:

Once you're happy with your clusters, you can export the processed data for use in other applications — for example, importing into Leapfrog to create 3D geological models guided by the geochemical domains. These clusters will also serve as our “ground truth” domains in the next article, where we'll explore how machine learning can use these geochemical patterns to predict domains in new drilling data.

Getting Started with GeoInsights 3D

Want to give GeoInsights 3D a try?

Jump over to the web-hosted app at: <https://geoinsights3d.koyeb.app/>

Or install locally

Using pip, you can install GeoInsights 3D directly from the Python Package Index (PyPI).

1. Create and activate a virtual environment (Strongly Recommended):

It is best practice to install packages in a dedicated virtual environment to avoid conflicts with other projects.

For Windows:

- python -m venv venv
- .\venv\Scripts\activate

For macOS/Linux:

- python3 -m venv venv
- source venv/bin/activate

2. Install GeoInsights 3D:

With your virtual environment active, install the package using pip:

```
pip install geoinsights-3d
```

3. Set up API Key (Optional for AI features):

The generative AI features require a Google Gemini API key.

Create a file named `.streamlit/secrets.toml` in the root of the project directory.

Add your API key to this file as follows:

```
.streamlit/secrets.toml GOOGLE_API_KEY = "YOUR_API_KEY_HERE"
```

Running the Application

Once the installation is complete, run the following command from the project's root directory:

```
geoinsights-3d
```