Scott McCallum, MS

Pittsburgh, PA (*open to remote & relocation*) • (412) 722-3416 • scott.mccallum.42@gmail.com
linkedin.com/in/scott-mccallum-4a406816/ • www.mccallum-pda.com (*company website*)

DATA ANALYST | DATA SCIENTIST

- ❑ ***Executing Data Project Milestones***: Go-to data services specialist with 15+ years of developing workflows and fit-for-purpose models to support operations, business development, and client needs for capital projects.

- ❑ ***Driving Data-Centric Solutions***: Guarantees product and service quality using supervised / unsupervised data science methods, machine learning, programming languages, applied mathematics, and statistics.

- ❑ ***Disrupting Industry Status Quo***: Innovative data manager at the intersection of technology and energy with millions saved in operational expenses based on cutting-edge analysis and forecasting techniques.

TECHNICAL COMPETENCIES

- ❑ **Version Control Tools**: Git (open-source version control system), GitHub (distributed version control system)
- ❑ **Software**: Petrel (subsurface data analysis), Studio (subsurface evaluation), Kingdom (seismic & geological interpretation), Geographix (structure / stratigraphic analysis & mapping), PowerLog (petrophysical core logging)
- ❑ **Model Development:** supervised (labeled dataset) / unsupervised (unlabeled dataset) machine learning, regression, classification, clustering, association, petrophysical, geocellular, and geologic core calibrated
- ❑ **Programming**: Python (pandas / numpy / scipy / plotly libraries), Excel VBA, R, Knime, Octave SQL, C
- ❑ **Analytical Methods**: applied / computational mathematics, statistics, multivariate analyses, clustering, genetic algorithms, neural networks, random forest, and Monte Carlo methods

PROFESSIONAL WORK HISTORY

MCCALLUM PETROPHYSICS AND DATA ANALYTICS (MPDA)   9/2019 to Present

Founder | Lead Data Analyst / Scientist

Launched an independent company after Rice Energy acquired EQT Corporation and replaced management. Consult for 3+ clients to deliver quantitative solutions with full transparency of workflows and code development using data science, machine learning, applied / computational mathematics, and statistics skill sets.

- Consult under a long-term contract for SpiralGen documenting GAP code due to Python expertise.
- Build Python / Gap scripts replicating signal processing, machine learning, and physics applications to optimize algorithms for high performance computing (i.e., Poisson Algorithm, Hockney Algorithm, Synthetic Aperture Radar, Space Time Adaptive Processing, and Convolutional Neural Network).
- Update grid cells based on Pseudo Spectral Analytical Time Domain using Maxwell's equations.
- Apply only polynomial operations to implement Brakerski, Gentry, Vaikuntanathan (BGV) methods.
- Execute 4 EQT contracts to update petrophysical, Petrel, and mineral (core data) models for reservoir modeling, freshwater casing depth identification, and 2 major business acquisitions (CVX and Alta).

EQT CORPORATION   7/2006 to 9/2019

Geoscience Advisor & Petrophysical Lead (9/2016 to 9/2019)
Senior Geologist, Petrophysicist & Petrophysical Lead (9/2013 to 9/2016)
Senior Exploration Geologist (9/2010 to 9/2013)
Geologist II & III (9/2006 to 9/2010)

Promoted 4+ times over a 13-year tenure with EQT. Built a Petrel database for 2 major asset regions from the ground up and wrote complex Python scripts for a large data centralization project comprised of 27K files and 250M rows of data for an Engineering group. Released after the Rice Energy contingent acquired EQT.

Multivariate Analysis of Gas Production Data Project – Python Language

- Built several multivariate production models using geologic and engineering parameters as inputs via multiple linear regression, k-means, ensemble / neural network, random forest, and Monte Carlo simulation methods.

Geophysical Log Projection Project – Python Language

- Saved hundreds of hours by automating the process for using geosteering data and vertical geophysical log data to project the log data along 1,300 horizontal wells via data scraping / centralization / standardization, interpolation, cross-correlation, and nearest neighbor methods.

Horizontal Well Spacing Project – Python Language

- Optimized the process of survey data applications to calculate north and south distance of a target well to its nearest offset wells via data compilation / standardization, data filtering, and k-nearest neighbor methods.

X-Ray Fluorescence (XRF) Source Signature Project – Python & Knime Languages

- Automated the classification of an unknown XRF sample to its most likely lithology and geologic formation using k-nearest neighbor, fuzzy k-means, and random forest methods.

Centralization and Standardization of Second-By-Second Completions Data Project – Python Language

- Centralized and standardized over 27K files with 250M rows of data varying in vintage, structure, and file type using complex logic and data scraping / centralization / standardization methods.

Real-Time Operating Center Tracker Project – Python Language

- Built an automated workflow for tracking well targeting efficiency and visualization in Spotfire for a client using data scraping / centralization / standardization methods, elementary statistics, and custom metrics.

Petrel Database Management Project – Petrel, Python & Knime Languages

- Built and managed the Petrel database for EQT's Marcellus and Utica operations, including maintaining all well data, structural models, petrophysical models, and 3D geologic models. Used petrophysical modeling methods and developed scripts in Knime and Python for data transfer from non-Petrel software (Geographix & Kingdom).

Shallow Geologic Prognosis Automation Project – Python & Knime Languages

- Automated the prognosis process necessary for building a casing plan in the shallow section of a well using manual entry of scanned documents into a structured format via k-nearest neighbor and statistical methods.

Cross-Correlation for Automated Geologic Top Selection Project – Python Language

- Streamlined the manual geologic process of correlating formation tops using geophysical log data via k-nearest neighbor and cross-correlation methods.

K-Means Lithologic Model Project – Python & Prizm Languages

- Leveraged deep knowledge of Euclidean distance as a method for classifying lithology based on geophysical log signatures, applying statistical methods and k-means with user-defined centroids to deliver accurate results.

Geologic Areas Definition Project – Python Language

- Identified areas of similar geologic properties by applying k-nearest neighbor methods for structural gridding of geologic parameters as inputs to a fuzzy k-means analysis.

EDUCATION

Master of Science (MS) in Computational Math (*pending completion*), Duquesne University

Master of Science (MS) in Geoscience, Southern Illinois University

Bachelor of Science (BS) in Geoscience, Indiana University of Pennsylvania