

# Sarah Donehower

sarahdonehower@gmail.com

## EXPERIENCE

### Verisk

OCTOBER 2024 - PRESENT — *Senior Data Engineer*

- Leading data modeling efforts for the enterprise data warehouse across all business lines.
- Designing and implementing the data stack and architecture using Snowflake, dbt Cloud, AWS, and Airflow:
  - Implementing CI/CD pipelines in dbt Cloud for automated testing and linting of data transformations.
  - Architected multi-layered data pipeline using medallion architecture (bronze/silver/gold) for incremental transformations and improved data quality.
  - Established Airflow orchestration on AWS MWAA with patterns for S3 file processing and Snowflake loading, including Fargate task deployment for long-running workloads.
- Developed XML file processing solution with Golang for parsing and loading data to Snowflake.
- Lead the POC of SelectStar as our data cataloging solution.

### Freelance

MARCH 2024 - OCTOBER 2024 — *Data Engineer*

- Migration project
  - rebuilt pipelines in new orchestration tool (Ascend.io), including extraction from source (client social media accounts), transformations with SQL, and ingestion to Snowflake,
  - built custom python connectors used by Ascend.io to extract data from social media channels,
  - built custom python client for interacting with Snowflake.

### Cresset Capital

JULY 2022 - DECEMBER 2023 — *Lead Data Engineer*

- Designed and developed Cresset's data platform with Matillion, AWS, dbt, and Snowflake enabling data-driven decision making and third party integrations.
- Developed data integrations:
  - ingesting data from sources systems using both Matillion components and python scripts deployed in AWS/Fargate,
  - delivering data to external systems for compliance, reporting, and client services.
- Redesigned our data warehouse including:
  - modeling data to improve incremental development and ease of use with reporting and BI tooling,
  - establishing role based access hierarchy and managing Snowflake users,
  - created independent development environments to improve the development experience and data reliability.
- Created AWS architecture to use containers for reducing load on our Matillion EC2 instance.
- Created and managed version control in GitLab for all data warehouse and data pipeline code.

## TECHNOLOGIES

python, SQL, Snowflake, Airflow, dbt, Docker, AWS (S3, EC2, ECS, ECR, Lambda, API Gateway, Transfer Family/SFTP, IAM, SQS, SNS), Matillion, git, Linux/BASH, data warehousing, data modeling, DataDog

## EDUCATION

### UCLA Extension

Python Programming  
Relational Database Management  
BASH  
Architecting Cloud Solutions in AWS  
Predictive Analytics  
Java

### General Assembly

Data Science  
Data Analytics

### University of Colorado, Boulder

B.A. Biology  
Authored Grant proposal and undergraduate thesis

## VOLUNTEER

### Food Bank of the Rockies

**Peace Corps**, Solomon Islands *Math & Science Teacher*

Math and Science Secondary School Teacher

Technical Writer and Trainer for World Health Organization (WHO)

## EXPERIENCE (continued)

### Jupiter Intelligence

MARCH 2022 - JUNE 2022 — *Data Engineer*

- Created data quality checks for data loads into Snowflake to ensure data reliability.
- Developed monitoring and alerting dashboards in DataDog for Snowflake and AWS for easier cost management and monitoring.

### Floodlight

JUNE 2021 - JANUARY 2022 — *Data Engineer*

- Built ELT infrastructure using Python, SQL(postgres), and AWS services,
- Partnered closely with data scientist on all engineering needs, as well as data analysis, data source evaluation, and model development,
- Assembled large complex data sets to meet business requirements,,
- Contributed to product strategy, customer discovery, and product development efforts.

### HomeAdvisor/Angi

MARCH 2020 - JUNE 2021 — *Data Engineer*

- Built and maintained HomeAdvisor's data lake and enterprise data warehouse in Snowflake for reporting by all business units,
- Implemented data pipelines from external sources using python jobs (deployed with Docker on Linux) and Snowpipe with S3,,
- Automated internal processes with stored procs (in javascript),
- Contributed to data modeling in data warehouse.

### Independent/Self Study

JANUARY 2018 - FEBRUARY 2020 — *Data Engineer*

- Completed projects in data modeling, data warehousing, data lakes (using python/boto3)
- Created ETL processes with CloudFormation, EC2, and S3
- Proficient in data analysis using pandas, numpy, scipy, SQL, visualizations with seaborn, matlab, bokeh, tableau.

### RED Interactive

JUNE 2017 - DECEMBER 2017 — *Data Analyst*

- Created and maintained data pipelines in Domo for reporting and research on client social media accounts and organic campaign performance;
- Created Tableau dashboards for analyzing app usage (with raw data from Google Analytics) and mobile gaming metrics (with data from Facebook Analytics and raw log data).
- Conducted audience and platform research through social listening.

### HYFN

JUNE 2016 - MAY 2017 — *Data Analyst*

- Provided data analysis, reports, and actionable insights to stakeholders to maximize social media campaign performance;
- Created client presentations and slides on campaign and audience performance;
- Contributed to improving the analytics workflow and evaluating business intelligence tools for use by the analytics team.