Mark Susol

Boulder/Denver, CO | 970.680.3194 | marksusol@gmail.com | LinkedIn: linkedin.com/in/marksusol

STAFF BUSINESS DATA SCIENTIST

With over 10 years of experience, I bring a profound expertise in driving technical initiatives within cross-functional environments. My career is marked by a deep involvement in developing and implementing advanced data analysis techniques and machine learning models, leading to significant improvements in business processes and decision-making strategies. Skilled in collaborating with diverse teams, I excel in translating complex data into actionable insights, aligning data science goals with business objectives. My leadership in technical projects is characterized by a hands-on approach and a commitment to mentoring team members, fostering an atmosphere of continuous learning and innovation. ***Areas of expertise include:***

Python • Go • SQL • R • TypeScript	Predictive Analytics	Technical Lead	Program Management
Experimental Design	Machine Learning	Mentorship	Collaboration
Natural Language Processing	TensorFlow	Education	Critical Decision Making

EXPERIENCE

***Google Inc.**, Mountain View, CA* — Sep 2015 - Present

Boulder, CO — ***Staff Business Data Scientist*** — Nov 2018 - Present
Remote — ***SQL Readability Team*** — Mar 2018 - Present
Mountain View, CA — ***Data Engineer*** — Sep 2015 - Oct 2018

gTech Users and Products Analytics plays a vital role in delivering business intelligence and reporting support for a wide range of internal Google products. Key responsibilities encompass:. Data Analysis and Insights, Product Optimization, User Experience Enhancement, and Reporting and Visualization. Additionally, the role requires the mentorship of analysts in data handling, visualization, and advanced analytical techniques.

Highlighted Personas:

Innovation

- Established a Natural Language Processing (NLP) “Center of Excellence” focused on raw user text analysis, enabling symptom clustering, detection of emerging issues in user feedback, and evaluation of contact form efficacy, thereby enhancing responsiveness and insights into user experiences.
- Directed the Personalized, Proactive, and Predictive Support multi year initiative as Lead Data Scientist, playing a pivotal role in the development of advanced online ML models. This leadership significantly increased **symptom match accuracy by 35%**, yielding substantial cost savings for Google. The initiative highlighted not only my technical expertise but also my ability to generate major business improvements through innovative ML solutions. It was a transformative project, upgrading Google’s support routing from outdated rules-based systems and offline models to cutting-edge, real-time serving technologies.

Technical Leadership

- Established and led the **gUP Analytics Tech Working Group**, enhancing the team's methodologies and overall workflow through dedicated expertise. Fostered a collaborative environment, ensuring broad representation and engagement within the group, overcoming challenges in a demanding, often unrecognized context to drive incremental improvements.
- Spearheaded multi-year data warehousing and reporting dashboard initiatives in a program manager capacity, orchestrating sprint iterations and comprehensive project documentation, thereby streamlining data and reporting efficiency.
- Conducted 193 technical interviews for data engineer and analyst roles, playing a key role in the strategic onboarding of new talent to our team and across Google’s Global Business Organization (GBO).

Education and Mentorship

- Developed comprehensive 'how-to' guides and training materials to scale analysts in specialized data analysis methodologies, incorporating custom Python, R and SQL solutions, thereby enhancing team proficiency and analytical capabilities.
- Guided junior analysts through technical mentorship, resulting in the promotion of over 5 peers, highlighting my role in fostering career growth and enhancing team capabilities.
- Provided expert guidance on advanced data analysis techniques through internal forum posts to a community of 70+ analysts, enhancing their skills and improving overall analytical capabilities within gUP Analytics

Adaptability

- Delivered key annotator rating quality metrics for Google's Text to Image Generative AI initiative (not directly supported by gUP Analytics), ensuring high standards of accuracy and reliability in the evaluation of generative models.
- Crafted a comprehensive data engineering pipeline for the 'Queriosity' team (20% project), enabling the creation of offline binary classification models using scikit-learn (python). This pipeline was instrumental in accurately predicting question-answer pairings in forum threads, thereby enhancing the analytical capabilities for user interaction assessment.
- Designed and executed an end-to-end data engineering pipeline complemented by AutoTFX (TensorFlow) modeling, specifically for predicting server defects in manufacturing within the Supply Chain 'QA Threshold' initiative (20% project). This robust system significantly enhanced the precision of defect detection, streamlining the quality assurance process.

Google's **SQL Readability Program** is an entirely voluntary initiative, managed by a diverse group of cross-functional administrators from various technical disciplines.

- Spearheaded a five-person, cross-functional admin team at Google, analyzing new features of GoogleSQL and effectively refining the global SQL style guide, enhancing readability **across the organization**.
- Mentored Googlers in SQL code optimization during code reviews, consistently ranking in the **Top 5** for review completions, demonstrating expertise in SQL and commitment to team development.
- Conducted **GoogleSQL Functions and Modules** training for 30-60 Googlers per session, achieving a high feedback score of 4.7 out of 5, reflecting strong instructional skills and subject expertise.

Cognizant Technology Solutions*, Mountain View, CA* — ***Technical Solutions Engineer*** Sep 2013 - Aug 2015

Contracted to **gTech Engineering Business Intelligence** at Google (Mountain View, CA)

- Developed software in Javascript, Python and Go using Test Driven Development (TDD).
- Conversion of operational ETL pipeline from Python workflow to Pipewrench (PatchPanel) workflow.
- Added, Modified, or Deprecated multiple ETL processes to include Map/Reduce in Python and Go.
- Enhanced Third Party data through Exploratory Data Analysis (EDA) using Google search in ETL pipeline.
- Conversion of operational ETL binaries from Sawzall to Go lang (~10k lines of code).
- Build internal Business Intelligence tools using Google's App Engine integrated with Google Data Services.
- Developed visualization dashboards using Aplos and Google's Dashboard for ETL pipeline quality monitoring.
- Experienced with Google's internal bug tracking, code review, and version control software for each project.
- Worked directly with Third Party data vendors to establish FTP acquisition and ETL pipeline data importation.

REI Systems*, Reston, VA* — ***Software Engineer*** Feb 2012 - Aug 2013

GovDashboard, Software-as-a-Service (SaaS) for Business Intelligence Visualizations.
Contracted to **US Department of Homeland Security (DHS)**.

- Development of open source SaaS for REI's performance analysis tool extending the core Drupal CMS platform.
- Developed OOP software in Python for Predictive Analytics; Logistic & Multivariate Linear Regression Regression.
- Developed interactive data visualizations using AJAX, Drupal's REST API, HighCharts and D3.js graphing libraries.
- Developed custom drupal module for FISMA security policies for REI's federal agency sites.
- Experienced with continuous integration systems (Hudson) and version control/management (Subversion).

EDUCATION

University of Maryland, Baltimore County (UMBC)

MS, Applied Physics - 1996, **BA, Physics -** 1992