Mikiko **Bazeley**

**SF Bay Area, CA |** [**mmbazel@gmail.com**](http://mmbazel@gmail.com) **|** [**linkedin.com/in/mikikobazeley**](http://linkedin.com/in/mikikobazeley) **|** [**github.com/MMBazel**](http://github.com/MMBazel) **|** [**medium.com/@mmbazel**](https://medium.com/@mmbazel)

SUMMARY

Software Engineer with 5+ years shipping ML models & data-driven solutions for high-growth companies in a variety of industries (solar, construction, recruitment-tech, & health-tech). Focused on full-cycle product engineering, including data engineering, ML model development, deployment, serving, monitoring & evaluation. Speaker, consultant, & mentor for initiatives & organizations focused on increasing representation in data science & machine learning.

EXPERIENCE

INTUIT – San Francisco, CA

*Senior SWE, ML Platform & Infrastructure**Nov 2021 – Present*

*SWE, et al. (Mailchimp)**May 2021 – Nov 2021*

**Summary**: SWE focused on developing ML Platform & operations to efficiently scale model pipeline training & deployment, integration, testing & monitoring for the Data Science team using GCP, etc.

* Shipped improvements to DS project templating tools (cookiecutter, etc.), including bug fixes & patches
* Provided initial analysis of project health of ML projects in production
* Lead change management of major upstream templating updates to 50+ downstream repos, as well as managing merge conflicts, performing code reviews, & developing initial repo health checklist
* Helped develop workshops & internal training on software engineering best practices for Data Science team, including testing
* Assisted with interviews for junior & senior software engineers

SIDEWALK.AI – San Francisco, CA

*Machine Learning Engineer**May 2019 – Jan 2021*

**Summary**: First MLE at early-stage start-up focused on building AI-first enablement tools for residential real-state.

* Developed ML roadmap, research & implemented proof-of-concept models from papers (in listing recommendations, user profile clustering) utilizing Python (Scikit-learn, Surprise, Dask, GeoPandas), AWS RDS, Jupyter
* Developed early iteration of ETL pipelines using Python (including SQLAlchemy, paramiko, Dask) with logging, data cataloguing, & job status tracking for processing 1000+ text files (up to 8MB in size)
* Designed data models & schema, created DDL & DML scripts for Postgres (then Amazon RDS)
* Researched, documented, & acquired external datasets through scraping (Python Requests, BeautifulSoup, Scrapy) as well as through APIs (Quandl, Kaggle, etc)

TELADOC – Mountain View, CA

*Data Scientist II, Growth Marketing**Oct 2019 – Nov 2020*

**Summary:** Collaborate closely with senior management to increase enrollment, activation & retention for members living with chronic & behavioral conditions (Diabetes, Hypertension, Pre-Diabetes, Weight Management).

* Provided experimentation support: Designed, set-up, & analyzed experiment results, including A/B tests (Python, Seaborn, SciPy, statsmodels, Metabase, SQLAlchemy)
* Lead projects in developing, documenting, scripting, & visualizing key metrics & analyses around marketing & growth
* Developed machine learning models (classification, regression, time series) & pipelines using Python (Pandas, Scikit-learn, Prophet, statsmodels), Redshift, S3. Leverage Metabase, Tableau & TabPy to engage stakeholders with analyses & machine learning model outputs around forecasting & prediction

WALKME – San Francisco, CA

*Sr BI Analyst, Sales & Revenue Ops Aug 2018 – Oct 2019*

**Summary:** Responsible for leading & executing advanced analytics & data science initiatives for the revenue operations organization. Worked with key stakeholders to scope, develop, & present ML models.

* Developed sales & marketing pipeline demand model to set sales targets, evaluate pipeline funding options, & monitor key metrics (lead generation, lead coverage, opportunity conversion). Leveraged Python, Google sheets, Redshift, Tableau. Made strategic recommendations using model results to prevent potential risk of $4M+ global revenue target miss.
* Developed pipeline allocation model & pipeline (with Python, Numpy) allocating lead generation targets from Demand Model to teams & individuals based on quota, start date & quota ramp. Model output ingested into external BI tool & CRM (Clari, Tableau) for real-time tracking of sales & actionable insight into opportunity velocity, leading to improvements in sales forecasting.
* Created propensity-to-purchase model using Salesforce data (opportunity, lead, & call records) as well as Python (simple-salesforce, NLTK, Spacy, Scikit-learn). Using features created around lead job titles, prospect account pain points, & historic demo call data, developed random forest model to predict propensity-to-first-purchase of account. Provided model interpretability using pdpbox, SHAP, & Scikit-learn.

AUTODESK – San Francisco, CA

*Data Scientist, Customer Adoption**Jul 2017 – Aug 2018*

**Summary:** Sole data science asset focused on understanding & driving user adoption of core products of construction software portfolio.

* Created customer scoring model in R utilizing account demographics, subscription & product usage data to flag customers struggling with product adoption. Model outputs scheduled for upload to Salesforce & Domo via scheduler, RForcecom, etc. with status reports automatically emailed to account owners via mailR.

VARIOUS COMPANIES – SF Bay Area, CA

*Various Roles, Business Operations**Dec 2013 – Jul 2017*

**Summary:**  Held various roles for a number of early- to mid-stage start-ups in bizops growth, and customer service.

**Companies:**  Sunrun (Analyst), Digimarc Guardian (BizOps Associate), RecruitLoop (Growth Hacker), eSalon (Manager)

ADDITIONAL – TEACHING , MEDIA, HONORS, COMPETITIONS

AMAZON SPONSORED: GHC2020 WOMEN IN BOOTCAMP– AnitaB.Org

*Scholarship:* [*https://bit.ly/2OLNgju*](https://bit.ly/2OLNgju) *Jun 2020*

(NOT IRL) PRIDE SUMMIT– Lesbians Who Tech

*Selected Speaker: https://bit.ly/3fgJY0d Jun 2020*

**Description:** Led 30min Ted talk for over 100+ attendees on data visualization topic, “Visualizing History with Python: Reproducing WEB Du Bois’ Paris Exposition Exhibit, 35 years”.

CAREER PANEL: WOMEN IN DATA SCIENCE – General Assembly

*Invited Speaker:*[*https://bit.ly/2BOZL86*](https://bit.ly/2BOZL86) *Apr 2020*

**Description:** Two-hour panel that featured “women who are successfully navigating the tech industry and developing exceptional work while motivating change.”

**Competition Standings:** Selected Competitor & Team Finals (*MIT Covid Datathon*), Top 8% [*of 800 Teams*] (WiDs GOSSIS 2020 Kaggle Competition)

**Teaching:** Data Science Mentor (*Data Science Dream Job*), Data Science & Analytics Mentor (*Springboard*)

**Additional Speaking Appearance include:** Artist of Data Science Podcast, Ken Jee’s Nearest Neighbors, Career Karma, etc. Can be found here: [*https://github.com/MMBazel#-also-the-medias-*](https://github.com/MMBazel#-also-the-medias-)

**EDUCATION**

UNIVESITY OF CALIFORNIA, SAN DIEGO – La Jolla, CA

*Economics & Anthropology, B.A. 2013*

**SKILLS**

**Languages**: Python *[PyTest, Great Expectations, scikit-learn, pandas, numpy, genism, spacy, OpenCV ],* SQL *[Postgres, MySQL, Oracle],* C++, Solidity, Rust

**Applications**: Docker, Kubernetes, Airflow, GCP *[BigQuery, AI Platform, Dataflow]*, Streamlit, Flask, Databricks/PySpark, FastAPI, Jenkins, Github Actions

**Certificates:** Full Stack Deep Learning, Deep Learning Specialization (*Coursera – deeplearning.ai*), Supply Chain Management (*Micromasters – MITx*), Database Programming: MySQL (*CCSF*), Data Science Bootcamp (*Springboard*), Machine Learning Engineering (*Springboard*)