

# Krissy (Dahl) Gianforte

## Data Scientist & Engineer

krissy.gianforte@gmail.com

484 744 1203

## Career Summary

*Data scientist with 5+ years of engineering work experience, looking for a role focused on data science. An authority on defendable, data-driven decision making after years of working under FDA scrutiny.*

*Strives to bring care and creativity to every problem; known for fast learning, self-motivated research, and unrelenting quality.*

## Education

### UC Berkeley: School of Information

#### Master of Information and Data Science (MIDS), 2019

Coursework included data engineering, statistical methods, experiments and causality, machine learning, natural language processing, data visualization, and data ethics

### California Institute of Technology (Caltech)

#### B.S. in Mechanical Engineering, 2012

Coursework included engineering design, kinematics and robotics, control theory, learning systems, and programming methods

## Portfolio Highlights *(demos at [KrissyG-hub.github.io](https://krissyg-hub.github.io))*

Beating the Game: Predicting Board Game Success Using ML

- This project uses board game data to understand what makes a successful design. Multiple models including kNN, Naive Bayes, and Linear Regression are explored, with insights extracted from each.

*Jupyter notebook, machine learning, python & sklearn*

*Translating model parameters and results into real-world insights*

Political Language Processing: Sentiment Analysis of Political Statements Using Cross-Field Training Corpora

- This project extracts statements regarding healthcare from congress peoples' .gov websites, then performs NLP analysis to predict how those politicians would vote in matters related to the ACA. After employing various known language sets for training, the project compares predictions to actual voting and ultimately uncovers interesting nuances specific to processing political language.

*Jupyter notebook, natural language processing*

Ford GoBike Business Analysis

- This notebook presents an analysis of Ford GoBike use in San Francisco, CA. It frames the business case, pulling and displaying relevant information from a Google BigQuery database. Finally, it makes a recommendation for offering a particular coupon to users.

*SQL, actionable business insights*

NBA Statistics Analysis

- This exploratory data analysis (EDA) uses Tableau to display NBA player statistics in multiple forms, allowing the reader to visually evaluate a few basic hypotheses.

*EDA, Tableau, data interpretation*

Insurance Ethics: The Danger of an Over-Personalized Insurance Industry

- This paper discusses the existing ways in which personal data has entered the insurance industry: demographics, driving metrics, fitness data, etc. It then considers other sources of data that would be useful for outcome prediction (social media, DNA) and explores whether the current legal regulations are sufficient to protect individual privacy.

## Technical Skills

Combining data from multiple sources

Performing statistical analyses (EDA, descriptive statistics, confidence intervals, sample sizing, etc)

Managing data intake, cleansing, filtering, and feature selection

Developing machine learning (ML) models for prediction, categorization, language processing, and image analysis

Understanding the uses and potential of machine learning models, as well as their limitations and pitfalls

Transforming model outputs into terms that are digestible to non-technical audiences

Presenting findings to enable data-based decisions

## Software & Tools

Python programming including:

Jupyter notebooks

Data frames (pandas)

Machine learning (sklearn)

Natural language processing (nltk)

Excel interfacing (openpyxl)

Command line interface (CLI)

Statistical analysis tools in python, R, excel, and matlab

Visualization including:

Tableau                      Html

D3.js                        Excel

Data collection via

SQL, extraction from excel via python, specific site APIs

Version control (via CLI or Github)

***Portfolio & demos at***  
***[KrissyG-hub.github.io](https://krissyg-hub.github.io)***

## Professional Experience

### Ad-hoc Data Science Projects (2019-present)

*During the covid pandemic, I focused on stabilizing and managing my household. However, I did find ways to exercise my data science skills through data manipulation consulting and some fun personal projects.*

Data Manipulation Consultant for a leading Tax Automation Software Company

- Created python script to translate lat/long coordinates into zip codes and US states necessary for tax calculations
- Project demanded 100% accuracy, with error checking around both input and output data
- Included a clear ReadMe and easily readable error logs in order to transition the script to non-technical end users

Machine Learning Creator for [Pixel Perfect Guide](#), the definitive strategy reference for the mobile game Pixel Starships

- Used ML to grade new game characters the day they were introduced, ultimately recommending whether to purchase the character and what roles to apply
  - Pulled game data from api and prepared modeling features by normalizing/scaling numeric categories, converting some features to binary attributes, and creating interaction terms
  - Divided data into randomized train and test sets, using subsampling to account for uneven sample distribution
  - Explored multiple model types, including linear regression, kNN, kernelized SVM, and decision tree
  - Automatically executed the final model on each game update
- Assisted with website readability and design

### Product Creation Studio (2/2018-6/2018)

Quality Engineer

Seattle, WA

- Responsible for all quality engineering project tasks and deliverables at PCS, including requirements development, risk analysis, and verification planning, execution, and reporting

### Philips Healthcare, AED Business (4/2015-12/2017)

Verification Lead, V&V Engineering

Bothell, WA

- Authored hardware verification test plans, sampling plans, protocols, tool packages, and test reports (including compliance testing and certifications)
- Performed statistical analysis on test data to determine test sample sizes and prove product reliability
- Interpreted medical device standards for their application to AED devices and accessories
- Participated in design reviews for input specifications, specifically for testability and consistency
- Developed test tools in C#, including a timing tool for interval testing and an equipment tracking program
- Provided mentorship and guidance to the team as a senior engineer; served as a process expert to ensure that V&V work aligned with the Philips Quality System and applicable standards & regulations
- CAPA owner – performed root cause analysis; created action plan including correction, corrective action, and preventive action; documented progress and effectiveness
- Back-room reference for verification matters during audits; front-room experience as subject matter expert

### Weyerhaeuser (7/2013-12/2014)

Mechanical Engineer, Scale and Development Engineering, Timberlands Technology

Federal Way, WA

- Programmed control logic and interfaces for multiple PLC automated systems using DirectSoft5
- Developed image analysis program in Matlab (including GUI) for inspection of cellulose fibers products
- Designed and installed custom mechanical equipment for use in lumber mills and development labs

### Medtronic Diabetes (7/2012-7/2013)

Associate Mechanical Design Engineer, Research and Development

Northridge, CA

- Wrote design verification test plans and oversaw their execution. Authored corresponding verification reports for product approval and release
- Performed Minitab statistical analyses on test data to quantify design performance
- Constructed Matlab program to standardize a previously subjective visual test method for detecting material stress in injection molded plastic components