

# Charlie Guthrie

Data scientist and visualization engineer

[charles.d.guthrie@gmail.com](mailto:charles.d.guthrie@gmail.com)  
[charlieguthrie.com](http://charlieguthrie.com)  
[linkedin.com/in/guthriecharlie](https://www.linkedin.com/in/guthriecharlie)

---

## WORK

### [Udemy](#)

#### **Sr. Data Scientist**

Apr 2019 - Present

- Defined metrics for learning engagement, included in company-level KPIs, built model to study drivers, achieved 11% YoY increase. Led project to expand tracking beyond video consumption to include quiz, labs, and exercises.
- Led effort to establish culture of experimentation & provide tools for easier analysis & faster iteration.
- Built system for identifying and removing bots from experiment results. Up to 60% of traffic had been bots, which was blocking experiment analysis.

### [Facebook](#)

#### **Data Scientist**

Jul 2017 - Mar 2019

- Developed accuracy measurement for Facebook's content moderation systems, which remove porn, hate speech, etc. from 40M posts per month. Reduced error by 40% in a year.
- Determined sampling distribution and headcount requirements for body of expert moderators who judge the accuracy of 15,000 content moderators.
- Developed model to identify ambiguous cases, which realigned team to focus on improving accuracy in more actionable, straightforward cases. Recognition of the limits of Facebook content moderators in ambiguous cases led to creation of the [Oversight Board](#).

### [Amazon \(Audible\)](#)

#### **Data Science Intern**

Spring 2017

- Natural language processing research to cluster user comments and identify trends in customer support needs.
- Keyword extraction and automatic summarization of Audible's library.

### [Cadent](#)

#### **Data Science Intern**

Summer 2016

- Used variational auto-encoding to simulate & preserve sensitive data, critical for our forecasting models.
- Overhauled design of ad forecasting system.

### [New York Times](#)

#### **Data Vis Engineer**

May 2013 - Jun 2015

- Broadly expanded data literacy, access, and decision-making by developing custom interactive visual dashboard in D3.js, used for business intelligence analytics across 10+ departments, including finance & advertising.

---

## EDUCATION

See projects at [charlieguthrie.com](http://charlieguthrie.com)

### **New York University**

Sep 2014 - May 2017

#### **MSc Data Science**

- Coursework included Statistics, Deep Learning, Machine Learning, Big Data, Inference, Natural Language Processing, and Algorithms

### **Dartmouth College**

Sep 2003 - Jun 2007

#### **B.A. Mathematics**

- Coursework included Mathematical Statistics, Modeling, CS

---

## SKILLS

### **Data Science**

- SQL, Hive ETL; Analysis in Python (Pandas, NumPy, SKLearn), R, Excel; deep learning with Tensorflow/Torch.

### **Data Vis / Front-end**

- HTML, CSS, Javascript, specializing in D3 and JQuery.

### **Other**

- Shell scripting; French; Comics at [instagram.com/charlesdguthrie](https://www.instagram.com/charlesdguthrie)