

# Charlie Guthrie

Data scientist and visualization engineer

[charles.d.guthrie@gmail.com](mailto:charles.d.guthrie@gmail.com)

[charlieguthrie.net](http://charlieguthrie.net)

[linkedin.com/in/guthriecharlie](https://www.linkedin.com/in/guthriecharlie)

---

## WORK

### Facebook

#### *Data Scientist*

Jul 2017 - Present

- On Community Integrity team, which enforced content policies, removing porn, hate speech, and other bad behavior.
- Designed, prototyped, and worked with engineer partners to develop system and dashboard for measuring accuracy of our content reviewers.
- Estimated sample and headcount requirements for stable metrics with narrow confidence interval, designed system to test reviewer expertise to ensure reliable metrics, tracked accuracy in dashboard.
- Developed model to identify ambiguous cases, which realigned team to focus on improving accuracy in more actionable, straightforward cases.

### Amazon (Audible)

#### *Data Science Intern*

Spring 2017

- Natural language processing research at leading audiobook provider.
- Clustered user comments and identified representative samples.
- Used textrank to automatically tag books with keywords for search.

### Cadent

#### *Data Science Intern*

Summer 2016

- Data science consulting position for TV & digital advertising network.
- Used variational auto-encoding to simulate sensitive data.
- Planned upgrades to ad forecasting system to be built after my departure.

### The New York Times

#### *Data Vis Engineer*

May 2013 - Jun 2015

- Developed interactive visual dashboard used for business intelligence analytics company-wide, and live-updating dashboard measuring how people use nytimes.com, including most popular articles and referrers.

### Dstillery

#### *Data Science Analyst*

Mar 2011 - Apr 2013

- Data analysis and testing on ad targeting models for a digital advertiser.
- Queries in Hive, modeling in R, visualization in D3, shell scripting.
- Managed intern and educated employees on our bidding process.

### Carfax

#### *Marketing Analyst*

Sep 2009 - Feb 2011

- Researcher in company that produces used vehicle history reports.
- Developed statistical models for predicting annual driver mileage and for identifying best client prospects.

---

## EDUCATION

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

### 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

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

### Data Vis / Front-end

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

### Other

- Shell scripting, Microsoft Office Suite; Conversant in French.