

# Robert Thorstad, Data Scientist

thorstads@gmail.com | (610) 348-6680 | Atlanta, Georgia, US

<https://www.linkedin.com/in/robert-thorstad/> | <https://github.com/rthorst>

---

## EXPERIENCE**Emory University, PhD Researcher**

2014-08-01 — 2019-12-20

- Trained convolutional neural network to distinguish 100,000s of pictures of objects and scenes (Cheng, Thorstad, & Dilks, in prep.).
- Used L2-regression, tSNE dimensionality reduction and DBSCAN clustering to show that Reddit posts predict whether people have a mental illness ([tinyurl.com/y3lhhec5](https://tinyurl.com/y3lhhec5)).
- Deployed sentiment analysis and dynamic time warping to show that most personal narratives follow one of six common culturally derived arcs. ([tinyurl.com/narr-arcs](https://tinyurl.com/narr-arcs)).
- Won Emory Research Catalyst Award, mentored 2 undergraduates, performed 15 hours statistical consulting, published 8 peer-reviewed journal and conference papers.

## **Emory University, Data Services Fellow**

2018-08-01 — 2019-12-20

- Used R and SPSS to mentor students in acquiring and cleaning research data.
- Led workshops on Python and social media mining ([tinyurl.com/tc4wvfr](https://tinyurl.com/tc4wvfr)).
- Used R to analyze over 10 million households of US census microdata.

## **Insight Data Science, Fellow**

2020-01-13 — Present

---

## EDUCATION**Emory University**

2014-08-19 — 2018-12-20

PhD - Cognitive Psychology

- Machine Learning
- Data Mining the Mind
- Multilevel Modeling
- Regression
- ANOVA

---

**SKILLS****Machine Learning:** Natural language processing (word embeddings, topic modeling, sentiment analysis), Deep learning (CNN, RNN), Dimensionality reduction (tSNE, pPCA), Clustering (DBSCAN, kmeans++)

**Programming Languages:** Python (Tensorflow, scikit-learn, gensim, pandas, seaborn, numpy, bokeh, pysqlite), SQL, R (ggplot, tidyr, dplyr), SPSS, HTML, Java

**Analytic Skills:** A/B testing, Experimental design, Multilevel modeling, Regression, ANOVA, Time Series Analysis

---