

# William Humphrey

[whh293@nyu.edu](mailto:whh293@nyu.edu) | +1 (907) 602-7397 | New York, New York

## EDUCATION

---

### NEW YORK UNIVERSITY, TANDON SCHOOL OF ENGINEERING:

2017 - 2021

- Major: Science & Technology Studies, Data Science, B.S.
- Coursework: Causal Inference, Linear Algebra, Data Structures, Machine Learning, Applied Statistics

## EXPERIENCE

---

### Unite Us *Data Scientist*

New York May 2021 – Present

- Implement variational autoencoder aimed towards predicting a given individual's social determinants of health.
- Conduct time series analysis and forecast shifts in community needs based on platform and social media data.
- Perform necessary statistical analyses to derive insights informing leadership, product decisions, and research.
- Maintain ethical AI innovation practices by evaluating model equity and incorporating stakeholder feedback.

### Unite Us *Data Analyst Intern*

New York May 2020 – May 2021

- Experiment with XGBoost and Random Forest classifiers to predict annual customer network retention rates.
- Prototype deep-learning classifiers for predicting social care service types based on client demographics.
- Create and maintain customer facing Tableau dashboards to provide valuable insights across all networks.
- Adopt and implement standard git workflow practices and review pull requests for various team members.

### Vendelux *Data Analyst Intern*

New York Oct. 2019 - Feb. 2020

- Improve overall customer experience by making company website and user workflow more intuitive.
- Design easily understandable and presentable data visualizations in Tableau based on platform ROI calculations.
- Perform HTML scrapes using BeautifulSoup and Selenium in Python to add new data to existing database.
- Actively manage and scale company databases injected from SQL into Django front-end platform website

## RESEARCH / PROJECTS

---

### Predicting Stock Volatility

- Predict annual stock volatility based on SEC 10-K documents and historical volatility using Scikit-Learn.
- Built web scraper using BeautifulSoup4 and Selenium to generate dataset of individual stock 10-K documents.
- Optimized Random Forest Regressor able to achieve an  $R^2$  score of .78 from dataset of 24,000 documents.

### Roll Call Report

- Twitter bot which tweets out roll call voting results from both the House and Senate updated daily.
- Provides followers with an extremely convenient way to stay politically informed from an unbiased source.
- Built using Python, Docker, Selenium, Twitter Developer API, and Google Cloud Platform.

## ADDITIONAL INFORMATION

---

### TECHNICAL EXPERIENCE:

- Python, SQL, R, Java, Excel, Tableau, Scala, Hadoop, TensorFlow, Keras, Scikit-Learn, PyTorch, AWS, GCP

### INVOLVEMENT:

- Volunteer at the Great Strides Event for the Cystic Fibrosis Foundation
- Advisory member at the NYU Data Science and Analytics Club
- Alpha Epsilon Pi Fraternity

---

**Strengths:** Terrific interpersonal and collaborative skills, innately curious, detail-oriented, driven, positive

**Interests:** Cooking, Video Games, Hiking, Soccer, Skiing, Volunteering, Travel