

# RACHEL DILLEY

✉ racheldilley11@gmail.com  
🌐 racheldilley.medium.com  
☎ (336) 905-5619  
📍 Raleigh, NC 27518  
in r-dilley  
🔗 racheldilley1

## Skills

### LANGUAGES/DATABASES

Python  
SQL Server  
PostgreSQL  
VBA

### PACKAGES

Pandas  
NumPy  
SciPy  
Scikit-learn  
nltk  
Beautifulsoup  
Selenium  
OpenCV

### DATA VISUALIZATION

Tableau  
Matplotlib  
Seaborn

### MACHINE LEARNING

Regression  
Classification  
Natural Language Processing  
Convolutional Neural Network

### INDUSTRY

Lean Six Sigma Yellow Belt  
Stochastic Processes  
Quality Design & Control  
SIMIO (System Simulation)

## Education

### North Carolina State University

2016 to 2020  
B.S. Industrial & Systems Engineering (ISE)

### Metis

Dec. 2020 to Current  
Data Science Certificate,  
Data Science, Machine Learning, and Programming

## Summary

An ISE graduate with 2+ years of industry experience in data analysis and system optimization. Skilled in coding (**Python**) and query-based languages (**SQL**) with a love for data mining, problem solving, and programming that has led to a pursuit of a full-time career in data science/analytics.

## Experience

### Metis

#### Data Scientist

Dec. 2020 to Mar. 2021

##### US Tourist Attraction Recommender and Image Classifier

- **75,000** images from 1,500 tourist attractions scraped from Tripadvisor (**Selenium, BeautifulSoup**)
- Attractions clustered using **topic modeling** (NMF), images classified using a **CNN** with **transfer learning** (VGG-16), and image compared using color distribution feature vectors (**OpenCV**)
- Created a web app via **Streamlit** for showcasing model

##### A Deep Dive into Cryptocurrency

- Provided time series analysis and investment insights utilizing **NLP** techniques, **topic modeling** (NMF, LDA, LSA) and **sentiment analysis** (VADER), on **17,000+** comments from r/Cryptocurrency subreddit obtained with Reddit API
- NLP analysis and price data of various currencies, gathered via Cryptocompare API, visualized using **Tableau**

##### What Operating System Do You Use?

- Predicted a users operating system utilizing **supervised classification** algorithms (Logistic, Decision, Random Forest, Gradient Boosted Trees)
- Trained models on StackOverflow survey data (stored and accessed using **PostgreSQL**), handling class imbalance, feature engineering, hyperparameter tuning (grid search), and gradient boosting to optimize f1 score

##### Car Sales Predictor

- Built **linear regression** models (simple linear, polynomial, lasso, ridge) and engineered features to predict yearly car sales , optimizing for r2-score
- Webscraped (**BeautifulSoup**) to get car sales and spec data

##### Street Team Planning

- Created a plan of suggested times and locations for street teams in order to maximize non-profit fundraiser attendance and donations using NYC subway turnstile data (Pandas, Matplotlib, Seaborn)

### i2M LLC

Raleigh, NC

#### Engineering Intern

Feb. 2019 to Dec. 2020

- Implemented **5S** and **Lean Manufacturing** principles on new COVID-19 related assembly line
- Collected and analyzed **KPIs** to establish work standards across all **15** manufacturing processes
- Created a new process to reduce cycle time of bottleneck to increase production capacity by **50%**, project key steps include:
  - *Paying special attention to quality control, designed scientific experiments and data collection plan that enabled automated tracking of trials*
  - *Formulated various **data visualization** graphs to present experimental results to supervising directors*

### Haworth Inc

High Point, NC

#### Summer Student Worker

May 2017 to Aug. 2018

Entry-level manufacturing engineer for two summers at a wood and upholstery office seating manufacturer

- Performed make vs. buy analysis of parts using **MS-Access SQL** database queries
- Analyzed I/O of operations to determine scrap rate and **QC parameters** of production output
- Conducted capability studies for various machines that led to **10%** improvements in operational efficiency
- Performed time and labor analysis studies (value/non-value added) labor content for various operations

## Projects

### Capstone Project with Ribbon Communications

Jan. 2020 to May 2020

- Implemented new notification system in Oracle to improve communication across 4 departments and decrease information delay time by **65%**
- Collected KPIs, mapped (**Visio**), and analyzed complex installation process to identify information sharing inefficiencies and potential cost and time savings