

# ALBERT SUN

Durham, North Carolina

408-348-1091 ◇ albert.sun310@duke.edu ◇ albertyusun.com

## EDUCATION

---

### Duke University

*Expected May 2023*

B.S., Data Science: Sociology and Racism (Self-Designed Degree)

Overall GPA: 4.0/4.0

- Huang Fellow, Resident Assistant, Duke Chronicle Opinion Columnist, Duke Statistical Science Major Union Communications Officer, Duke Chamber Music

## WORK EXPERIENCE

---

### Duke Law Center for Science and Justice

May 2020 - Present

*Data Scientist*

- Leveraging data conclusions from CA and TX jail databases to advocate for data infrastructure and pipelines for the first local jail database in North Carolina.
- Using R and Python to scrape data and create RShiny Dashboards. Working on a Duke Bass Connections Research Team with Duke Law Professor Brandon Garrett.

### Duke Rhodes Information Initiative Data+

Summer 2020

*Data Scientist*

- Investigating early histories of the Consumer Revolution with Professor Astrid Giugni and a team of Duke graduate students using Natural Language Processing (NLP).
- Successfully scraped, cleaned, and created structural topic models and word embedding models on 26,000 old English texts from the EEBO TCP database.

### Duke Mock Trial

Aug 2019 - Present

*Travel Team Competitor*

- Received 3rd Most-Outstanding Witness Award out of 240 competitors at Vanderbilt University's National Mock Trial Invitational in November 2019.
- Competed on Duke's Mock Trial Travel Team in 5 national invitationals. Qualified to Opening Round Championships before COVID-19 ended hopes of winning the championship.

## PERSONAL PROJECTS

---

### Louisiana Jail Data - Duke COVID-19 DataFest 2020

Fit multivariate linear models to research the likelihood of overall releases after COVID-19 in Louisiana Jails using data from the Vera Institute of Justice. Found significant correlation for predictors like percent Black, private jails, and political profile.

### Hotel Cancellations - Duke STA 199L

Trained 96 percent accurate kNN Classification and logistic models to predict hotel cancellations in Portugal.

## SKILLS AND INTERESTS

---

### Languages and Tools

#### R frameworks

#### Python frameworks

#### Other

Python, R, Java, Git, Command Line, HTML/CSS  
tidyverse, ggplot, tidytext, stm, RShiny (still learning)  
pandas, numpy, scikit-learn, BeautifulSoup, gensim  
Marathons, Historical Fiction, Jazz