ALBERT SUN

Durham, North Carolina

albertyusun.com ⋄ albert.sun310@duke.edu

EDUCATION

Duke University *Expected May 2023*
B.S., *Data Science: Sociology and Racism* (Self-Designed Degree) Overall GPA: 4.0/4.0

· Activities: Huang Fellow in Duke Initiative for Science and Society, Resident Assistant, Duke Chronicle Opinion Columnist, Statistical Science Major Union Communications Officer, Chamber Music

WORK EXPERIENCE

Duke Rhodes Information Initiative Summer 2020
Data Scientist

· Investigated class and gender during the Consumer Revolution with Professor Astrid Giugni and a team of Duke graduate students using NLP tools for academic publication and conferences.
· Developed infrastructure to scrape, clean, and visualize topic models and word embedding models on 26,000 old English texts from the EEBO database to answer empirical research questions.

Duke Law Center for Science and Justice May 2020 - Present
Data Scientist

· Advocated for data pipelines for the first state-wide jail database in North Carolina with Law Professor Brandon Garrett and a Bass Connections Research Team.
· Used R and Python to scrape data and create RShiny Dashboards to leverage existing data from California and Texas state jail databases

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.
· Qualified to Opening Round Championships before COVID-19 prevented us from competing.

Duke University Center on Social Equity Aug 2019 - Jan 2020
Research Assistant

· Summarize, transcribe, and discuss academic articles for Professor Adam Hollowell's paper-in-progress, *Why Stratification Economics Includes Disability Justice.*

PERSONAL PROJECTS

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

STA 199L: "Hotel Cancellations"
Trained 96 percent accurate kNN Classification and logistic models to predict hotel cancellations in Portugal.

SKILLS

Languages and Tools	Python, R, Java, Git, Command Line, HTML/CSS
R frameworks	tidyverse, ggplot/ggmap (visualization), dplyr, stm, RShiny
Python frameworks	pandas, numpy, scikit-learn, BeautifulSoup, gensim
Other	project management, Microsoft Office, Adobe Photoshop