

Ngoc (Alice) Hua

DATA SCIENTIST

626.677.1998



alicehua11@berkeley.edu



linkedin.com/in/alicehua/



github.com/alicehua11



alicehuacal.com



SKILLS

SQL / Python / R / Matlab / Git / GIS /
ArcGIS / Hadoop / GCP / Docker / Spark /
Tableau / HTML / CSS / Carto / Mapbox

EDUCATION

Master of Information & Data Science
UC Berkeley
Anticipated 2021

Bachelor of Arts, Geography
Magna Cum Laude, Phi Beta Kappa
UC Berkeley
2020

Associate of Arts,
Social Sciences & Social Behavior
Certificate, Automotive Technology
Rio Hondo College
2018

SELECTED COURSEWORK

Applied Machine Learning
Statistics for Data Science
Fundamental of Data Engineering
Research Design

FUTURE COURSEWORK

Machine Learning at Scale
Privacy Engineering
Experiments and Causal Inference

LANGUAGES

English & Vietnamese

PROJECTS

Regression Study of the Spread of Covid-19 2020

- Used classical linear model of OLS to estimate the causal relationship between adults from 19 to 34 and the number of Covid-19 cases in 50 US States
- Assessed the assumptions, estimated coefficients and confidence intervals. Established magnitude and direction that states with more young adults indeed result in higher Covid-19 cases

Machine Learning for Online News Prediction 2020

- Scaled out Python webscraping for raw data of 2020 news articles from Forbes and reproduced features
- Ran multiple ML algorithms and assessed the percentage errors in predicting a continuous outcome which can be improved with more data and feature engineering

PySpark for Business Intelligence 2020

- Used Hadoop and Spark to build a data pipeline, from Docker cluster, consume messages from Kafka to Spark for data transformation (i.e. using flatMap() to make a nested json friendlier for querying) and loaded to HDFS.
- Reported business findings of this data to customers

Descriptive Analysis: Bird Strikes & Vertipport Analysis 2019

- Analyzed bird strikes trend from 1990 – 2018 against airline flights, temperature & migration datasets using R and LIDAR data to understand the relationship between these factors
- Produced suitability analysis for vertipport locations in using ArcGIS for the Civil and Environment Department at UC Berkeley

HIGHLIGHTED EXPERIENCE

UC Berkeley School of Chemistry Berkeley, CA | 2020

Prospect Research & Data Collection Assistant

- Discover and collect large amounts of employment data of alumni using LinkedIn Sales Navigator to better engage UC Berkeley alumni and identify new donor prospects
- Perform query-based requests and analyze publicly available data on candidates to assist supervisor on engagement

UC Berkeley Urban Displacement Project Berkeley, CA | 2019

Undergraduate Research Assistant

- Used GIS and data science skills to produce rigorous analysis on the nature of gentrification and displacement
- Used Python to wrangle and analyze urban and social media data, used ArcGIS to visualize results
- Developed interactive, online maps from the results of data analysis and develop website content
- Created Python course materials for Urban Data Analytics course to teach other students spatial and network analysis