

Andy Pang

[apang782.github.io](https://github.com/apang782)
[apang00](https://www.linkedin.com/in/apang00)

apang782@gmail.com
[apang782](https://github.com/apang782)

EDUCATION

Georgia Institute of Technology 2022 - Present
M.S. Computer Science
University of California, Santa Barbara 2020
B.S. Statistics & Data Science

PROGRAMMING LANGUAGES

Python (Pandas, Plotly, OpenCV)
C Java
C++ LaTeX
SAS SQL
HTML CSS
R (astsa, qpcR, rpart, glmnet)

TECHNOLOGIES

BigQuery SFTP
AWS (Qubole) Databricks
UNIX
Jupyter Notebook
Git Ver. Control/Collaboration
SAP (Business Objects)
Visual Basics for Applications

TECHNICAL KNOWLEDGE

Computer Networks
Network Security
Knowledge-Based AI
Software Design
Operating Systems
Stochastic Processes
Machine Learning
Web Scraping/Crawling
Feature Engineering

CERTIFICATIONS

BigQuery for Marketing Analysts

Build and Optimize Data Warehouses with BigQuery

WORK EXPERIENCE

TransUnion Remote
Solutions Architect Oct 2021 – Present

- Managed multi-touch attribution product implementation and delivery for 5 clients across various industries, resulting in contract renewals and CSAT scores of 4+.
- Designed and created internal tools for fellow SAs to streamline data validation process, automate report generation, and assist with QA.
- Designed and created an internal tool for visualizing code logic and database relations using the Graphviz package.
- Created and updated Confluence articles to document best practices, troubleshooting, and share key information.

Tawa Services, Inc. Buena Park, CA
Financial Analyst Jan 2021 – Oct 2021

- Designed multilabel NLP ticket classifier to tag over 1000 customer tickets with 12 labels, achieving 85% accuracy.
- Automated the updating process for over 19 reports, saving 30 minutes weekly.
- Developed and maintained over 50 reports (labor, supply, etc) for upper management and all stores nationwide.
- Implemented original ETL pipeline to combine multiple e-commerce data sources for weekly reports on best-selling articles.

Piedmont Racing Ltd. Remote
Analytics Platform Design & Development Intern Dec 2019 – Sep 2020

- Developed psychological index: an indication of a rider's mental state based on factors such as individual recent history and standing/status through career.

PROJECTS

Web Scrape & Visualization of Used Car Prices [apang782.github.io/vroom1](https://github.com/apang782/vroom1)
Python (BeautifulSoup, Pandas, NumPy, Selenium, Seaborn, Matplotlib)

- Designed automated scraper using Selenium and BeautifulSoup to pull data from 1440 dynamic Javascript used car pages with randomized URLs.
- Determined archetypal listed car to be a 2019 Infiniti SUV, with 28,000 miles, 1 previous owner, and a price tag of around either \$19,000 or \$26,000.

Regression Modeling & Prediction of Used Car Prices [apang782.github.io/vroom2](https://github.com/apang782/vroom2)
R (car, MASS, glmnet, rpart, randomForest, gbm)

- Found torque, horsepower, and city MPG as the most important predictors of list price during decision tree, linear modeling, and ensemble method development.
- Predicted 218 car prices with \$2400 margin of error using random forest model.