

## Mehika Patel

(864) 371-2030 | mpatel3105@gmail.com | Portfolio: [mehika.me](https://mehika.me) | Github: [mehiks11](https://github.com/mehiks11)

### EDUCATION

---

**Yale University**, New Haven, CT

Dec 2022

*Sociology with Statistics & Data Science*, GPA: 3.82

**Relevant Coursework:** Machine Learning, Probability Theory, Data Engineering, Information Systems

**Awards:** Mildred Priest Frank Memorial Prize for Thesis work (under projects)

Cohen Summer Fellowship awarded to fund data science research and data system development at nonprofit

### PROFESSIONAL EXPERIENCE

---

**Up & Up**, Data Scientist (Intern), New York, NY

May - Aug 2022

- Designed full data pipeline, ML model, and framework optimized to find high-potential neighborhoods for investing
  - Modular & scalable build integrating several data sources (*Google Maps APIs, Yelp API, BigQuery, etc.*) with automated tuning per market (currently 4) and per stakeholder (4 teams) built using scikit-learn, Clustering, Regression, Time-Series & Spatial analysis (*presentation available at [mehika.upandup.co](https://mehika.upandup.co)*)
  - Expedited successful investing in new real estate markets, and customer acquisition by providing optimal real estate options, with little-to-no local knowledge
  - Feature engineered location “diversity metric” for model using Kullback-Leibler divergence
- Collaborated on building company global dashboard with construction team’s 6 identified key performance metrics
  - Reduced dashboard run costs & time by aiding transition to standard data modeling ETL practices using dbt
- Built route optimization algorithm to cut property managers’ travel time between work sites by 2-8 hours/week

**Freelance Technical Instructor**

Jan 2022-Current

- Provide technical training (*coding, data science, ML, math & statistics*) to professionals at Fortune 500 companies such as Microsoft for individuals and groups of 20+ students of all ages (9-52)

**Detect**, Computer Science Intern, Guilford, CT

Jan - May 2022

- Created data visualization tool for R&D team to facilitate development of at home PCR-quality COVID-19 tests saving researchers ~20 minutes/test and additional time avoiding technical hurdles of manual coding analysis

**Yale Physics Department**, Machine Learning Fellow, New Haven, CT

Jan - May 2022

- Tested & improved Baryon Pasting neural net model (probes nature of dark matter in clusters 25k+ light years away)
- Implemented graphical user interface on GCP VM to host model with increased accessibility and model usage for scientists outside of Yale, saving 30-40 minutes on average in package install-time

**Consumer Credit Counseling Services**, Savannah, GA

*Data Science & Data Systems Architect Fellow*

Jun - Aug 2020, Jun - Dec 2021

- Automated reporting infrastructure inside dashboard with features such as survey management, data analysis, and up-to-date report generation on client feedback and outcomes for 500+ clients (GCP APIs & cloud functions)
- Analyzed efficiency of business services, suggesting improvements on product deficiencies and new marketing decisions in client communities

**Data 2 the People**, Data Science Fellow, New Haven, CT

Jun - Aug 2021

- Delivered data collection solution by identifying quantifiable key performance metrics for stakeholders in 8 weeks used for informed operational decision-making
- Developed web-based interactive data visualization using D3.js used for fundraising within local community

### PROJECTS (ADDITIONAL AVAILABLE ON PORTFOLIO ABOVE)

---

- Natural language processing of breast cancer patient forums** to provide experience-oriented approaches to care
  - Senior thesis project scraping & topic modeling patient forum data (NLTK, GenSim, SpaCy, scikit-learn)

### SKILLS AND INTERESTS

---

*Languages:* Python, SQL, JS, R

*Technologies:* MongoDB, Cloud Computing (Google Cloud), PowerBI, Tableau, Web Scraping

*Stats:* A/B Testing, Significance testing, Supervised & Unsupervised Learning

*Packages/Libraries/Frameworks:* Pandas, Numpy, Scikit-learn, TensorFlow, Keras, dbt, React, Bokeh, Matplotlib

*Interests:* Music (Guitar & Dance), Martial Arts (Shotokan Karate)