HAI NAM VU

✉ namhaivu97@gmail.com • ☎ 857-326-7638 • linkedin.com/in/hai-vu • github.com/namhaivu173

RELATED SKILLS

Data Tools and Programs: R, Python (Pandas, NumPy, Matplotlib, Seaborn, Plotly, Sklearn, PySpark), SQL (SQL Server, MySQL), Microsoft Office (Excel, PowerPoint, Word), Dashboarding (Tableau, Power BI, Shiny)

Proficiencies: Data Analysis, Data Science, Financial Analysis, Dataset Optimization, Data Visualization, Data Processing, Database Programming, Operational Analysis, Quantitative Analysis, Data Warehousing, Data Modeling

PROFESSIONAL EXPERIENCE

Case Team Assistant Data Analyst Ho Chi Minh, Vietnam **Sep 2020 – Dec 2020, Jun 2021 – Aug 2021**

Boston Consulting Group (BCG)

- Collaborated with 5 cross-functional teams at the client site to define long & short-term KPIs and redesign pricing & brand activation schemes accordingly at 4 distribution channels, resulting in a **5% gain of market shares** from local competitors
- Developed generalized **Python scripts** that extracted and transformed data from periodic sales reports; designed interactive **Excel dashboards** from the processed data to measure performance, monitor compliance, progress & impacts of newly launched schemes, and present to non-technical stakeholders on a monthly basis
- Cooperated closely with 3 client-side teams to handover 4 custom-designed KPI tracking tools & Excel dashboards; performed maintenance as well as bug fixes on an on-demand basis to ensure clients obtained accurate and interactive market reports
- Analyzed data from multiple work streams and visualized key findings in order to recommend next-step actions for bi-weekly company performance reviews with the management board

Market Analysis Consulting Intern Ho Chi Minh, Vietnam **Jun 2019 – Aug 2019**

Dream Incubator Southeast Asia (DI SEA)

- Preprocessed and analyzed multiple quantitative datasets from 3 clients and trusted online resources to extract key insights and provide relevant statistics to 4 business proposals that supported foreign partners in exploring new business opportunities within Southeast Asia
- Served 3 senior associates and 1 managing director by conducting comprehensive **market research** on the competitive landscape and industry value chains that facilitated the collaboration between a major digital solution provider and gas station business partners in Thailand
- Collaborated with 3 senior associates in designing and revising **PowerPoint** presentations to ensure the accuracy and transparency of 5 project deliverables, resulting in data analysis and predictions being presented in a timely manner

DATA PROJECTS

Uber & Lyft ride-share prices in Boston area analysis & predictions (Python)

- Performed data cleaning and conducted exploratory analysis along with hypothesis tests on roughly **700k rows** of data to explore time, distance, and fair price patterns of **13 Uber & Lyft products** in Boston
- Utilized **Google Map API** to extract foot traffic information of popular pickup/drop-off locations in Boston
- Developed machine learning pipeline using 3 regression models (Lasso, Ridge, Random Forest) while controlling for multicollinearity to estimate fare prices based on relevant predictors such as time, weather, pickup/drop-off locations, etc.
- Constructed 4 classification models (Naive Bayes, Logistic Regression, Decision Tree, Random Forest) to forecast whether ride-share price would surge. The best model was evaluated based on F1, accuracy, and AUC scores

Housing prices in Nashville, Tennessee analysis & predictions (R)

- Imputed missing/irrelevant data, removed outliers and regrouped categorical features to allow for more precise analysis on 56k rows of data. Created relevant visualizations to extract insights on the main factors affecting house prices
- Trained 4 predictive models (Linear Regression, K-Nearest Neighbors, Random Forest, Gradient Boosting) on partitioned data set to estimate housing prices using 15 predictors including land value, property age & type, room quantity & quality, year of sales, etc. The best model was evaluated using RMSE, MAPE, R-squared and adjusted R-squared from the testing set

NYC Open Data – 2015 Street Tree Census data dashboarding (R Shiny)

- Performed appropriate data cleaning and feature engineering steps on **600k rows** of data that contain information of trees planted in New York City to improve data quality and avoid redundancy
- Designed interactive **Shiny dashboards** following effective visualization practices that allow users to explore the relationships between different characteristics of trees such as species, trunk diameter, health perception, planted location, etc.
- Optimized R script and used random sampling to successfully deploy the app to Shinapps.io cloud server for public access without exceeding the memory size limit of 1GB

EDUCATION

Master of Professional Studies in Analytics and Statistical Modeling (GPA 3.6) **Dec 2022**
Northeastern University, Boston, MA

Master of Finance in Alternative Investments (GPA 3.88) **Sep 2020**

Bachelor of Business Administration in Operations & Information Management (GPA 3.67) **May 2019**
University of Massachusetts - Amherst, Amherst, MA