HUI WANG

Salt Lake City, UT | (385) 212-0590 | u1303081@utah.edu

LinkedIn: linkedin.com/in/hui-wang-823650320 • GitHub: github.com/HazelnutHui

Data Analyst (Entry-Level) | Python, SQL, Tableau, Excel | Regression, NLP, ML | CPT/OPT Eligible

SUMMARY

Data analyst with experience in end-to-end ML workflows including regression, feature engineering, and ensemble modeling. Proficient in Python (pandas, scikit-learn), SQL, Tableau, and R. Kaggle Top 8.5% in a housing price prediction competition. Strong at turning data into actionable insights.

TECHNICAL SKILLS

• Languages & Tools: Python (pandas, matplotlib, scikit-learn), SQL, R, Tableau, Excel, Power BI, Git

• Techniques: EDA, Feature Engineering, Regression, Clustering (K-Means), TF-IDF, A/B Testing, Time-Series, NLP (tokenization, keyword extraction), Ensemble Modeling, Cross-Validation

• Domain Applications: Job Market & Housing Data Analysis, User Segmentation, Dashboard Communication

EDUCATION & CERTIFICATIONS

University of Utah — BS in Mathematics (Statistics Emphasis), Expected Spring 2026

Google Data Analytics (Foundational & Advanced), Tableau Desktop Specialist, Data Camp SQL Associate

EXPERIENCE

Legal & Data Research Intern, Canyon Guard (Sep–Dec 2024)
- Modeled transportation funding and allocation data to explain a 30% budget variance
- designed reproducible data cleaning pipelines in SQL and R.

Summer Research Assistant, Carnegie Mellon University (Aug 2021)
- Preprocessed data for ML models in gesture recognition study; contributed to accuracy improvements.
- Co-authored a research publication and presented data insights to interdisciplinary teams.

SELECTED PROJECTS   (View code & dashboards: github.com/HazelnutHui)

House Prices Prediction (Kaggle Top 8%)
- Performed data cleaning, feature engineering, and log transformation on the Ames Housing dataset (Kaggle).
- Trained Ridge, Lasso, and XGBoost models; ensembled for final submission ranked top 8% (405/4781) on Kaggle.

Shared Mobility Usage Analysis (Google Capstone)
- Analyzed e-moped usage with SQL and R; segmented users via K-Means and visualized patterns in Tableau.
- Recommended fleet rebalancing and targeted marketing based on behavior insights.

U.S. Data Analyst Job Market Insights (Advanced Capstone)
- Extracted key job skills using TF-IDF from 2,800+ LinkedIn postings; visualized results with ggplot2 and wordcloud2.
- Conducted time-series analysis and proposed training alignment based on demand trends.

ADDITIONAL

Languages: English (Fluent), Mandarin (Native)

Work Authorization: F-1 Visa with CPT/OPT eligibility (STEM major)