

Ruby Cheung

(612)481-6071 | r.cheung763@gmail.com | [/in/rubycheung](https://in/rubycheung)

Summary

Data professional with 4+ years of experience in building data driven solutions. Skilled in statistical modeling, econometrics, research, and cross-functional collaboration with stakeholders ranging from executives to technical teams. Strong foundation in R, Python, and SQL with proven ability to develop dashboards, maintain analytical pipelines, build predictive and causal models, and communicate insights to both technical and non-technical audiences.

Technical Competencies

- **Programming and Querying:** Python (NumPy, Pandas, Matplotlib, SciKit-Learn), R (dplyr, tidyr, ggplot2, Keras), SQL
- **Data Tools:** Tableau, Excel, Git, PostgreSQL, Hadoop, Spark, R-Shiny, Airflow, PowerBI, Visual Studio Code, QGIS, Selenium
- **Data Methods:** Machine Learning, Deep Learning, NLP, Statistical Analysis Econometrics, Time Series Analysis, Forecasting, Sentiment Analysis, Text Classification, Big Data Analytics, Web Scraping, Data Visualization, A/B Testing

Relevant Work Experience

Wabash College

Nov 2022 - Present

Institutional Research Analyst

Crawfordsville, IN

- Conducted in-depth analysis of institutional priorities, including a longitudinal study of the Math department to identify points of attrition in the program sequence and improved tracking of major completion and cohort progression.
- Conducted targeted analyses of internal survey data to uncover trends in employee engagement and satisfaction, and student perceptions on campus climate, identifying key areas of concern and organizational tension that informed strategic planning, policy development, and programming initiatives.
- Developed institution-wide dashboards and reports using R to monitor KPIs including student retention, enrollment trends, demographics, and financial aid metrics comparison against peer institutions to support executive decision-making.
- Produced frequent ad hoc reports by leveraging advanced Excel and R scripting to meet urgent or evolving administrative requirements.
- Standardized Institutional Research protocols through comprehensive documentation, achieving over 50% reduction in reporting time and improved data quality.

Seattle University

Jun 2024 - Sep 2024

Graduate Research Assistant

Seattle, WA

- Collaborated with the Math and Biology departments to develop and deploy an NLP model using Python for sentiment analysis of comments submitted on carnivorespotter.com.
- Processed high volumes of unstructured data by applying feature engineering and vectorization techniques in Python to enhance model performance for downstream analytics.
- Designed a robust evaluation process to test model performance under class imbalance conditions, achieving an accuracy of 81%.

St. Catherine University

Jun 2020 - Aug 2022

Economics Research Specialist

St. Paul, MN

- For the Sudan Labor Market Panel Survey, developed a new census map for Sudan using GIS software, managed data collection, data wrangling, conducted analysis, including discrete-time hazard models and employing stratified random cluster sampling, and co-authored the resulting research paper.
- Led economics research teams on projects focused on development economics and labor markets in North Africa and the Middle East to successfully complete multiple research studies.
- Contributed to research design and survey development by creating Android-based survey instruments using ODK-X with integrated SQLite, streamlining data management and enhancing data collection efficiency.

Education

Seattle University

June 2025

M.S., Data Science

St. Catherine University

May 2020

B.S., Economics

Select Projects

Predicting Operational Vitality for the Office for Catholic Schools Seattle

- Conducted segmentation analysis and ANOVA tests to uncover internal and external factors associated with school financial health for the Archdiocese of Seattle's Office of Catholic Schools.
- Built predictive models using random forest and regularized regression to identify key indicators of school solvency and guide cost optimization strategies.
- Developed a scalable, data-driven PowerBI dashboard to track key performance metrics, integrating live data via API to support real-time decision-making by school administrators.

Evaluating Forecasting Methods for Gross Gaming Revenue Prediction

- Evaluated ARIMA, ETS, GARCH, and Prophet models to forecast Louisiana gross gaming revenue.
- Assessed time series assumptions, and performed residual diagnostics using Ljung-Box and Shapiro-Wilk tests.
- Addressed COVID-era and hurricane shocks by incorporating Prophet features and creating exogenous variables, resulting in a model with low RMSE and a MAPE of 12% (ARIMA) and MAPE of 5.6% (Prophet) for 6-month forecasts, performing competitively with industry standards.

Automated Bayesian A/B Testing Framework Using Airflow & PostgreSQL

- Built end-to-end data pipeline using Apache Airflow to process 112K+ e-commerce transactions, transforming operational data into optimized star schema for analytics.
- Designed and implemented Bayesian A/B testing framework with automated customer randomization, statistical analysis, and business recommendations for promotional campaigns.
- Created customer segmentation analytics identifying VIP customers' average lifetime value and analyzing geographic distribution.

Identifying Key Indicators for Recidivism Among Georgia Parolees

- Leveraged Georgia Department of Community Supervision data to model recidivism risk factors, producing actionable insights for risk mitigation. Developed and evaluated Gradient Boosting and SVM prediction models with emphasis on false positive rate reduction.
- Identified significant correlation between continuous employment and reduced recidivism through variable importance analysis.

Select Publications

- Caroline Krafft, Ragui Assaad, Ruby Cheung. Introducing the Sudan Labor Market Panel Survey. Cairo, Egypt: Economic Research Forum.
- Caroline Krafft, Ragui Assaad, Mohamed Ali Marouani, Ruby Cheung. The Impact of COVID-19 on Middle Eastern and North African Labor Markets: Vulnerable Workers, Small Entrepreneurs, and Farmers Bear the Brunt of the Pandemic in Morocco and Tunisia.. Cairo, Egypt: ERF.
- Ragui Assaad, Caroline Krafft, Maia Sieverding Layla Al-Hajj, Ruby Cheung. Barriers to Economic Participation for Young People in Jordan: A Gendered Analysis. Amman, Jordan: UNICEF.