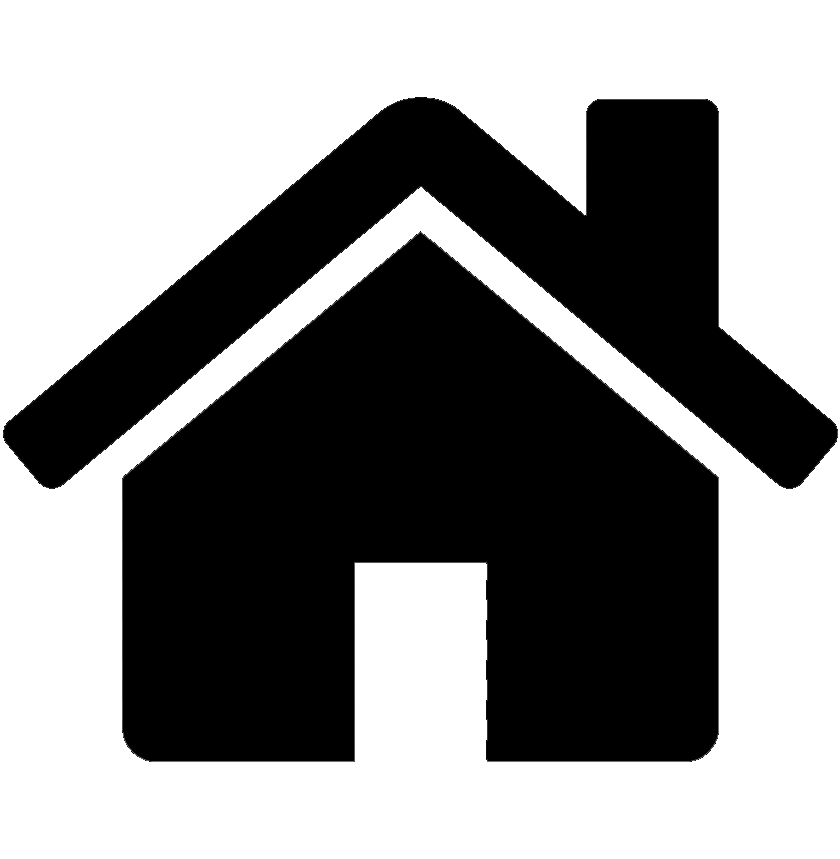
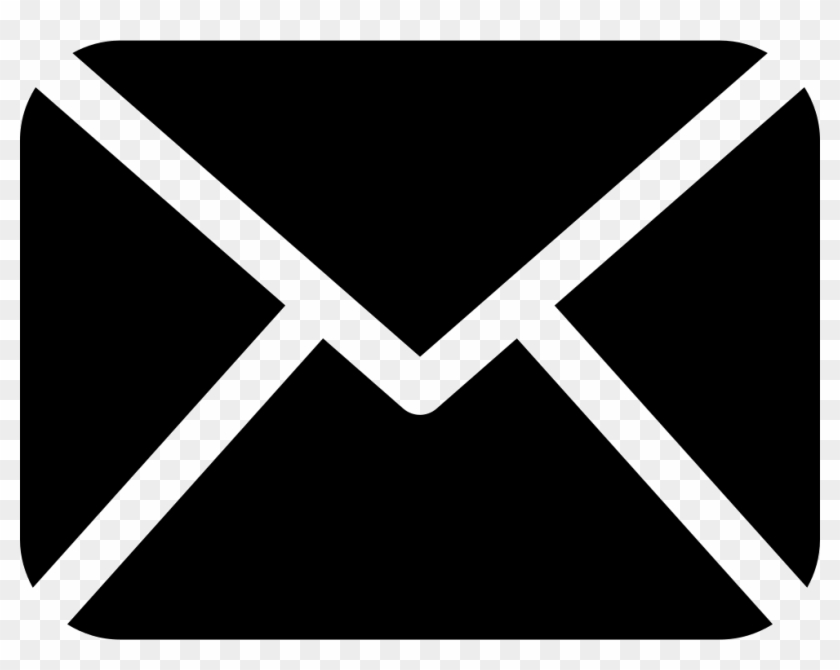
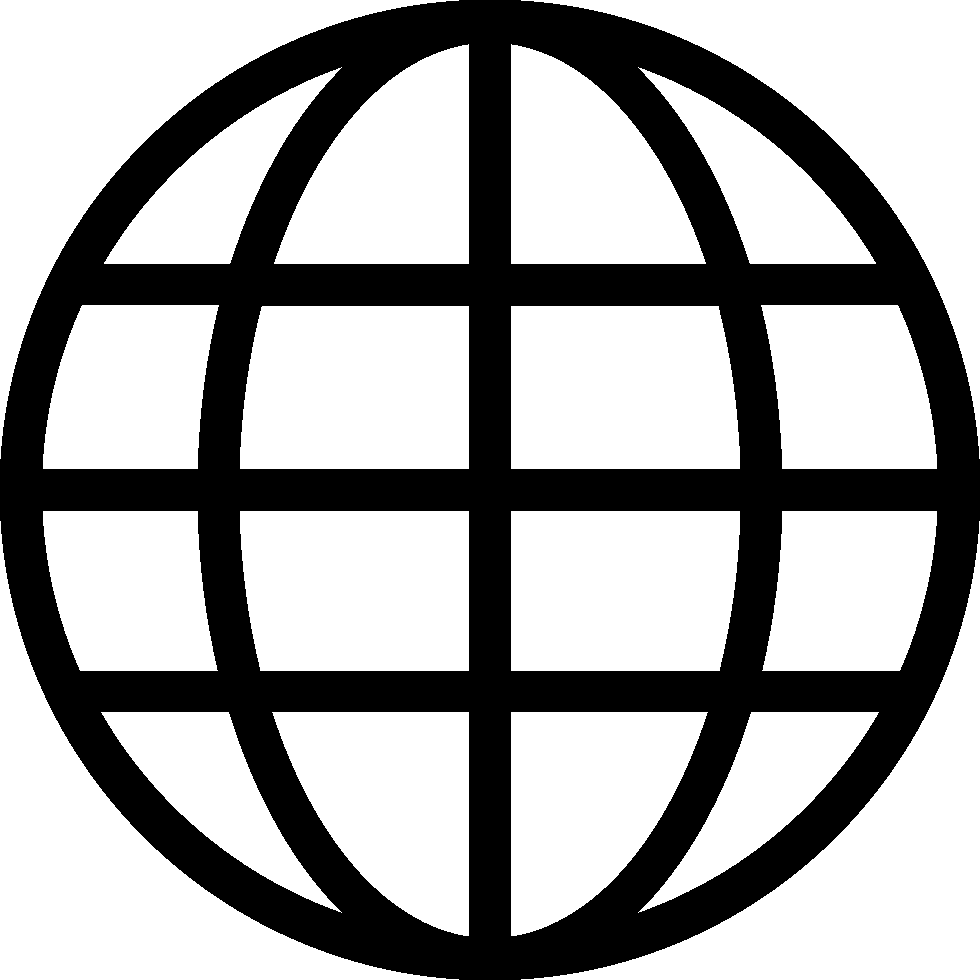
PRISCILLA  CHEN

 Grand Rapids, MI           jc253@calvin.edu           https://www.linkedin.com/in/jx-chen

**EDUCATION**

|  |  |
| --- | --- |
|  |  |

**B.S. in Data Science** **Calvin University | Grand Rapids, MI**

**GPA 3.98** with Honors; Junior

**EXPERIENCE**

|  |  |
| --- | --- |
|  |  |

**Research Assistant** Jun 2025 - Present

Supervisor: Prof. Stacy DeRuiter

* Collaborated with 4 research groups across different disciplines
* Provide data consulting service, including **data cleaning**, **statistical modelling**, **data visualization**.
* Assisted in a 3-day technical workshop, teaching over 20 researchers on how to use R for advanced statistical modelling.

**Academic Grader** Dec 2024 - May 2025

Supervisor: prof. Rocky Chang

* graded weekly computer science lab assignments in Python for over 70 students.

**Administration Assistant** Sep 2023 - May 2025

* Served 34 faculty members and over 300 students, facilitated 12 events organization.
* Worked on data cleaning, transformation and organization using Excel, contribute to event-planning, such as admitted student outreach, graduating student celebrations, and faculty retirement celebration.

**PROJECTS**

|  |  |
| --- | --- |
|  |  |

**Statistical Modelling for Rain Garden Contaminants Research**

Faculty Advisor: Stacy DeRuiter; Collaborator: Cassandra Demlow

* To analyze the effects of rain gardens on soil contaminants, we fitted mixed effect generalized linear models (GLM) with contaminants Chloride, Iron, and Zinc, Phosphate as a function of rain garden age, percentage soil organic content, and pH.
* Phosphate and Chloride were modeled using gamma family and log link function while Iron and Zinc were modeled using binomial family with logit link function. Random effect of sample nested in sites is included.
* All models are fitted using R statistics in R studio.

**Stratigraphic Visualization for Pollen Analysis**

* In collaboration with ecology researchers Dr. Melinda Higley and her team, our summer research team created stratigraphic visualizations that align with research standard using R.

**Computational Bayesian Inference of Exchange Rates**

Faculty Advisor: Prof. Stacy DeRuiter

* Developed a Bayesian Inference model to quantify the effects of macroeconomic indicators (interest rate, inflation, GDP growth) on USD/CNY exchange rate
* Preformed extensive **data cleaning and transformation** on historical economic data collected from public sources, such as handling missing values, aligning time indices, calculating first-differences, and standardizing predictor scales.
* Specified **Stan** models and utilized **Markov Chain Monte Carlo (MCMC)** sampling and convergence diagnostics to obtain posterior predictions and assess model fitting.

**Impact of Deer on Forest Ecosystem**

Faculty Advisor: Prof. Stacy DeRuiter; Collaborator: Matthew Dykstra

* Implemented a **generalized linear model (GLM)** pipeline using **R** that examines and predicts the deer impact on local ecosystems. Finding was adopted by collaborator to inform ongoing assessments and future field studies.
* Preformed **data cleaning** and **feature engineering** on three thousand cells of data.
* Conducted **Exploratory Data Analysis (EDA)** to explore patterns and outliers.
* Evaluated multiple model variants for optimal model using AIC and residual diagnostics

**SKILLS**

|  |  |
| --- | --- |
|  |  |

**Programming Languages:**R, Python, C++, C#, MySQL

**Environments:**GitHub, R-studio, Visual Studio Code, Google Colab, Tableau, MS Office

**Methods:**

* Descriptive & Inferential Statistics
* Generalized Linear Models (GLM)
* Bayesian Inference with stan
* Supervised (Regression, Classification) & Unsupervised Learning (Clustering) in Python
* Model Evaluation (cross-validation, AIC/WAIC/PSIS)

**Data Wrangling & Analysis：**

* Data Cleaning & Feature Engineering
* Pipeline Building (dplyr, Scikit-learn)
* Data visualization (ggplot2, seaborn, plotly, vega-lite