Helen Peng (she/her)

helen.peng04@gmail.com | linkedin.com/in/helenpeng04 | helenpeng04.github.io

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
B.S. in Statistics & Data Science, Expected December 2025
Concentration in Psychology
GPA: 3.44 | Major GPA: 3.51

Relevant Coursework

Statistical Methods in Epidemiology	Causal Inference	Probability Theory
Statistical Machine Learning	Statistical Graphics and Visualization	Meta-Analysis
Advanced Methods in Data Analysis	Statistical Computing	Modern Biology
Modern Regression	Statistical Inference	

Honors

Dean's List, High Honors: Fall 2022, Spring 2025
Dean's List: Spring 2023, Spring 2024

EXPERIENCE

UnitedHealth Group Bridges to Healthcare Technology Research Program, Carnegie Mellon University — Pittsburgh, PA
Research Intern — Summer 2025

- Conducted exploratory data analysis (EDA) and regression modeling to identify predictors of obesity in a team project; presented results to UHG professionals and cohort members to inform potential interventions.
- Performed k-means clustering and EDA on COVID-19 case and death trends across Pennsylvania counties in a team project; presented findings to cohort members.
- Engaged in workshops and mentorship sessions with UHG professionals, gaining hands-on experience in healthcare analytics and data-driven decision-making.

Optimized Algorithms and Knowledge (OAK) Lab, Carnegie Mellon University — Pittsburgh, PA
Research Assistant — Spring 2023 – Present

- Applied text analysis and clustering techniques to extract structured insights from qualitative participant responses.
- Implemented multilevel modeling on data from 100+ participants to evaluate effects of rule matching and interleaved pretraining on learning outcomes.
- Created 10+ visualizations and executed statistical tests on learning outcomes, including analyses of blocked vs. interleaved training and learning support, and prepared reports to communicate findings.
- Designed and analyzed Qualtrics survey comparing learning methods and motivation, synthesizing results and visualizing trends in Excel.
- Anonymized and organized 200+ test papers, ensuring compliance with privacy standards by labeling, scanning, and uploading them to a shared drive.

Research Methods in Cognitive Psychology, Carnegie Mellon University — Pittsburgh, PA
Teaching Assistant — Spring 2025

- Instructed 8 students in debugging R code, conducting analyses, and interpreting statistical results during office hours and lectures, while managing a full academic schedule.

Zhong Ou Asset Management Intl — Hong Kong
Financial Research Intern — Summer 2024

- Produced Excel visualizations to compare investment performance using advanced formulas, pivot tables, and conditional formatting.
- Compiled monthly outlook reports summarizing China's economic indicators and competitor analysis to support strategic planning.

Students Using Data for Social Good, Carnegie Mellon University — Pittsburgh, PA
Data Analyst — Spring 2024 – Spring 2025

- Applied statistical analyses (Fisher's exact test, Kruskal-Wallis test, survival analysis) on healthcare data from 600+ clients with developmental disabilities, identifying patterns in medication errors and risk factors.
- Generated 8+ visualizations and data reports, translating findings into actionable insights that guided nonprofit stakeholders in improving service delivery and resource allocation.

Cognitive Science Student Advisory, Carnegie Mellon University — Pittsburgh, PA

Executive Board Officer Fall 2023 – Present

- Organized and promoted 10+ academic and social events, including guest lectures and networking opportunities that connected undergraduates with graduate students and faculty.
- Launched the Boba Finals Pickup, a recurring end-of-semester event, providing peers with boba as a morale booster and fostering informal community interaction during finals.
- Interviewed and onboarded 5 new board members; mentored 5 students in the statistics/data analytics track on navigating coursework and identifying research opportunities.

PROJECTS

Mental Health Outcomes of COVID-19 Graduates Spring 2025

Class Project – Data Science in Psychology and Neuroscience

- Evaluated 100+ participant survey responses to assess mental health differences between the undergraduate classes of 2020 and 2021; developed and applied statistical modeling (PCA, logistic regression) and sampling (bootstrap) techniques in R to build and validate five predictive models for mental health outcomes.

Gender Stereotype Threat Interventions and Academic Performance Spring 2024

Class Project – Research Methods in Meta-Analysis

- Collaborated in a five-person group to investigate how gender stereotype threats affect academic outcomes through a meta-analysis of 117 studies; articles were screened together, then effect sizes were independently extracted from 13 peer-reviewed articles and analyzed in R to produce an individual research paper summarizing findings.

Gender Stereotype Threat and Memory Performance in a Heading-Recall Task Spring 2024

Class Project – Research Methods in Cognitive Psychology

- Teamed with three classmates to design and conduct a heading-recall task testing memory performance under gender stereotype threat (35 participants, 1,400+ trials); built the experiment in Gorilla, cleaned data in Excel, and performed ANOVA analyses in Jamovi; presented results at the CMU Department of Psychology undergraduate research poster session.

SKILLS

Programming & Data Analysis

- R: tidyverse, data.table, caret, survival, statistical modeling, markdown reporting
- Python: Basic knowledge of pandas and numpy
- SQL: PostgreSQL querying, relational database management

Research Tools & Experimental Design

- Qualtrics, Gorilla: Survey and experimental protocol design, randomization, data collection
- LaTeX: Scientific writing and formatting
- GitHub: Version control and collaborative coding
- Excel: Pivot tables, advanced formulas, conditional formatting, charts
- Jupyter Notebook: Integrated R/Python for reproducible workflows
- Quarto: Markdown reporting, presentations

Languages

- English (Native), Mandarin Chinese (Heritage Proficiency), Spanish (Limited Proficiency)