# Hongling Lei

honglinglei.github.io | www.github.com/HonglingLei
Apt 614, 6315 Forbes Ave, PA 15217 | +1 (412) 805-8510 | hongling@andrew.cmu.edu

## EDUCATION

**Carnegie Mellon University** U.S.
*Master of Information Systems Management – Business Intelligence and Data Analytics* *Aug. 2021 – Dec. 2022*
- Relevant Courses: Object Oriented Programming in Java, Data Focused Python, Data Structures for Application Programmers, Unstructured Data Analytics (NLP), Advanced Business Analytics, Database Management

**Xiamen University** China
*Bachelor of Economics.* GPA 3.81/4.0 *Sept. 2016 – Jun. 2020*
- Relevant Courses: Data Science, C, Mathematical Analysis, Statistics, Probability Theory, Linear Algebra

**University of California, Berkeley** U.S.
*Semester Exchange, Multidisciplinary Program.* GPA 3.95/4.0 *Jan. 2019 – May 2019*
- Relevant Courses: Data Analysis Using Python, Computing with Data Using R

**Nanyang Technological University** Singapore
*Summer Exchange, Digital Marketing.* GPA 5.0/5.0 *Jun. 2018 – Jul. 2018*

## PROFESSIONAL EXPERIENCE

**Tencent** Shanghai
*Data Scientist Intern, Public Data Science Department* *Apr. 2021 – Jul. 2021*
- Conducted causal inference with algorithms like Causal Bayesian Networks, X-Learner, and Causal Forest; analyzed treatment effects of key drivers for customer consumption time on Tencent's streaming platforms
- Designed and implemented experiments on user experience with Tencent WeSee, TikTok, and Kuaishou
- Developed a data pipeline for automated causal inference, including modules like feature engineering, machine learning model training, and future experiment testing based on analytical results

**DiDi** Beijing
*Data Analyst Intern, Decision Support Department* *Sept. 2020 – Nov. 2020*
- Supported business and strategy decisions by extracting and analyzing billions of data points on Apache Hive; conducted SQL query optimization, speeding up execution processes and boosting computational efficiency
- Designed a market sizing model with SQL and Excel, correctly predicting driver and order growth rates; collected drivers' feedback and wrote analysis reports with directions of improvement and feasible solutions
- Designed and monitored business dashboards that can automatically update data and visualize analytical results

**Bairong Technology** Shenzhen
*Data Scientist Intern, Financial Technology Department* *May 2020 – Aug. 2020*
- Built a semi-supervised learning model using the MixMatch algorithm with PyTorch, which can classify customers into different groups based on credit default risk predictions, achieving 90% accuracy
- Developed a report generation program with over 2,000 lines of Python code, which automatically calculates performance indicators, graphs statistical distributions, produces analysis texts, and creates formatted slides
- Tested over 200K Excel rows with VBA, ensuring data quality before sending them to clients

**LexinFintech** Shenzhen
*Data Management Intern, Big Data Intelligence Department* *Aug. 2019 – Sept. 2019*
- Analyzed over 667K real-time customer profile data using Hive; developed a real-time interactive web report with a search function and visualized results based on Kafka and Spark streaming
- Participated in the Financial Statement Project: extracted 44 testing samples that covered all situations from 170M financial data for the Risk Management Department, shortening report generation time from 3 days to 2.5 hours
- Performed database consistency checks by sorting and fixing abnormal jobs through checkpoint recovery

## PROGRAMMING PROJECTS

**R | Package Development** [GitHub] *Apr. 2019*
- Published an R package called "binomial" that tests the validity of inputs, calculates the probability based on given parameters, graphs the distribution and displays a summary of skewness, kurtosis, etc.

**Python | Angry Projectiles** *(Team Member, with 2 Google Engineers)* [GitHub] *Mar. 2019*
- Applied aerodynamics knowledge and Newton's method using NumPy and SciPy
- Developed a throwing game that simulates the launch, studies the motion, and animates the trajectory of projectiles

**R | Shiny App Development** [GitHub] *Mar. 2019*
- Developed a user-interactive web app that can forecast investment returns and graph the growing trend

## TECHNICAL SKILLS

**Programming/Software:** Python, Java, SQL, R, C, HTML, TensorFlow, PyTorch, Apache Hive, Meltwater, SEMrush
**Languages:** Mandarin Chinese (native), English (fluent, TOEFL 116 with Speaking 30), Korean (basic)