

# Huguang (Luke) Zheng

Tel: 917-891-7460 | Address: 625 W 57<sup>th</sup> Street, Apt 2651, New York, NY 10019  
Email: zheng.huguang@columbia.edu | LinkedIn: <https://www.linkedin.com/in/huguangzheng>

## EDUCATION BACKGROUND

<b>Columbia University in the City of New York</b> Candidate of <i>M.S</i> in Operations Research, specialized in Data Science	<b>12/2017 (Expected)</b> GPA: 3.5
<b>Pennsylvania State University, University Park</b> <i>B.S</i> in Mathematics, <i>B.S</i> in Statistics, and <i>B.S</i> in Economics <b>Honors:</b> Highest Distinction ( <i>Summa Cum Laude</i> ), Schreyer Honors College Scholar, Bates White Fellowship, Dean's List	<b>5/2016</b> GPA: 3.8

## SKILLS

<b>Programming:</b> Python, JavaScript, R/RStudio, MATLAB <b>Database:</b> SQL (MySQL, Postgres), MongoDB <b>Visualization:</b> D3.js, Matplotlib, Shiny, Tableau	<b>OS:</b> Mac OS, Linux (Ubuntu), Windows <b>Hadoop Ecosystem:</b> Apache Spark, Pig, Hive, HBase <b>Cloud Service:</b> AWS (S3, EC2, RDS, DynamoDB, Kinesis Streams)
---	--

## PROFESSIONAL EXPERIENCE

<b>Gogotech</b> <b>Part-time Data Scientist</b> <ul style="list-style-type: none"><li>Conducted data analytics and visualization on E-commerce data using Python NumPy, Pandas and Matplotlib.</li><li>Researched on assessment of customer lifetime value, and acquisition and maintenance strategies to assist decision making.</li></ul>	<b>New York, NY</b> <b>10/2017 – Present</b>
<b>Lux Research Inc</b> <b>Data Science and Artificial Intelligence Intern</b> <ul style="list-style-type: none"><li>Built data ETL (Extraction, Transformation, Loading) pipeline with Python, AWS S3, Postgres SQL database, etc.</li><li>Developed, tested, and deployed proprietary data analytics and visualization Python Flask application in an agile team, used packages like D3.js, Pandas, NumPy, SQLAlchemy and tools including Git, Fabric and Docker.</li><li>Participated in development and integration of article recommendation engine, and performed A/B testing for evaluation.</li><li>Conducted NLP (Natural Language Processing) research project regarding technology innovation using nltk, sklearn, etc.</li><li>Processed streaming data by setting and configuring Amazon Kinesis Streams service.</li></ul>	<b>Boston, MA</b> <b>6/2017 – 8/2017</b>
<b>Local Interactive</b> <b>Data Science Intern</b> <ul style="list-style-type: none"><li>Developed python programs to scrape and store data and displayed them in a web app with Flask framework.</li><li>Applied Machine Learning algorithms in SEO strategy analysis to boost performances.</li></ul>	<b>New York, NY</b> <b>3/2017 – 5/2017</b>

## RESEARCH AND PROJECT

<b>Columbia Data Science Hackathon</b> <ul style="list-style-type: none"><li>Performed NLP and Network Analysis on Enron email corpus in Zeppelin Notebook. Deployed data on AWS EMR clusters and process it with Apache Spark.</li></ul>	<b>9/2017</b>
<b>IBM Watson HealthHacks Hackathon</b> <ul style="list-style-type: none"><li>Designed and developed an AI Psychiatrist coping with working pressure with classification methods, including logistic regression, SVM and random forest algorithms, integrated with Watson Lab IoT Hardware and APIs.</li></ul>	<b>4/2017</b>
<b>Research Assistant in Economics Department, Supervised by: Professor Kala Krishna</b> <ul style="list-style-type: none"><li>Assisted researches on international trade and education in developing countries, responsible for information searching and data cleaning in Stata and R, and got involved in pattern recognition and interpretation.</li></ul>	<b>9/2015 - 5/2016</b>

## LEADERSHIP ACTIVITIES

<b>Program Manager of Enactus</b> <ul style="list-style-type: none"><li>Devised and executed projects of public interest, including straw recycling and reusing, free lunch promotion, etc.</li></ul>	<b>4/2013 - 1/2014</b>
<b>Co-founder of Haisen Health Consulting Group</b> <ul style="list-style-type: none"><li>Co-founded a company and presented business plan to the university fund and got RMB 50,000 grant.</li></ul>	<b>2/2013 - 7/2013</b>