

# JINGJING LIN (She/Her)

+1 202.460.4927 | [isjingjing.lin@gmail.com](mailto:isjingjing.lin@gmail.com) | Bethesda, MD 20814 | [github.com/jlin-data](https://github.com/jlin-data) | [LinkedIn-imjingjinglin](https://www.linkedin.com/in-imjingjinglin)

## EDUCATION

<b>Georgetown University</b> <ul style="list-style-type: none"><li>Master of Science: Data Science and Analytics, GPA 3.5/4.0</li><li>Peer lead mentor, advised mentors and ~70 graduate students</li></ul>	08/2018 – 05/2020 Washington D.C.
<b>University of Manchester</b> <ul style="list-style-type: none"><li>Master of Science: Management and Information Systems, GPA 3.3/4.0</li></ul>	09/2015 – 12/2016 Manchester, UK
<b>Tianjin Polytechnic University</b> <ul style="list-style-type: none"><li>Bachelor of Engineering: Software Engineering, GPA 85/100   Minor: Finance, GPA 88/100</li><li>2014 Presidential First-Class (top 3%), 2013 Second-Class (top 5%) Scholarship; Outstanding Graduate (top 2%)</li></ul>	09/2011 – 06/2015 Tianjin, China

## SKILLSET

**Programming:** Python (scikit-learn, pandas, numpy), R (dplyr, glmnet), SQL, VBA – [Demo](#), Java (JDBC), HTML, CSS, Shell  
**Machine Learning:** Regression (Linear/Logistic, Lasso), Clustering (K-Means, Hierarchical), Classification (Boosting, SVM, Random Forest), Deep Learning (CNN, RNN), Natural Language Processing (NLP), Dimension Reduction (PCA)  
**Statistics:** Probability, Distribution, Sampling, Hypothesis Testing, Bayes Theorem, Correlation, A/B testing  
**Cloud Computing:** AWS (EMR, S3, Hadoop, MapReduce, Spark, Git); Google Cloud (BigQuery, storage buckets)  
**Visualization & Tools:** Tableau, Power BI, Plotly, Matplotlib, ggplot2; MySQL, Jupyter, Google Analytics (Certificated)

## EXPERIENCE

<b>Whitaker Brothers Business Machines</b>   <i>Retail &amp; Consumer Goods</i> <b>Marketing Data Analyst</b> <ul style="list-style-type: none"><li>Led market and pricing analysis; tracked key improvement metrics and provided analytical breakdowns of pricing structures and sales funnel of 2000+ products</li><li>Designed and implemented dynamic tiered pricing strategies across multiple platforms, increasing the number of orders by 52%, total revenue by 49%, and session conversion rate by 280% (check-out rate by 179%)</li><li>Developed automated competitor pricing analysis processes and a pricing adjustment model to monitor, deliver and update pricing decisions regularly through Shopify using Python, decreasing operation time by 90%</li><li>Created periodical in-depth analysis reports of sales, market trends and products prediction via Tableau and Power BI to provide recommendations for pricing and marketing campaign strategies</li><li>Solved database inconsistency among vendor, legacy and current data by matching and unifying main variables</li></ul>	06/2021 – Present Rockville, MD
<b>Georgetown University Graduate School</b>   <i>Higher Education</i> <b>Data Science Research Assistant</b> <ul style="list-style-type: none"><li>Tracked 100+ vaccines development progresses from scientific reports, merged all data using pandas and NLTK</li><li>Built a map-based Tableau dashboard to quantify and visualize textual processes with geological information</li><li>Generated weekly reports and cleaned datasets, embedded results in the website using HTML and CSS</li></ul>	08/2020 – 06/2021 Washington D.C.
<b>Center for Security &amp; Emerging Technology, Georgetown University</b>   <i>Higher Education</i> <b>Data Science Research Assistant</b> <ul style="list-style-type: none"><li>Accessed data from 3 AI publication databases (130+ million rows of 14+ GB) using BigQuery in GCP</li><li>Merged database A, B by DOIs and metadata (e.g., keywords), built a dictionary from metadata by bag-of-words</li><li>Matched the top 50 articles between database A and B using TF-IDF distance and the year of publication, created the distribution of the longest substring sizes in each match</li><li>Increased matching rates to ~20% in articles without DOIs by running text similarity algorithms using the highest similarity score, based on the threshold (95%) from the scores distribution in matched articles with DOIs</li></ul>	09/2019 – 12/2019 Washington D.C.
<b>Unilever - Dollar Shave Club</b>   <i>Retail &amp; Consumer Goods</i> <b>Marketing Technology Analyst Intern</b> <ul style="list-style-type: none"><li>Sole analyst responsible for analyzing and optimizing current manual-operated 20+ spreadsheets with 200+ tags of Urchin Tracking Module (UTM) management information</li><li>Reduced parameters setting time by 90% for marketing team, and implementing time by 33% for data systems team by developing an automated tags generator and re-designing the dashboard using macros, pivots and SQL</li><li>Delivered a new UTM solution to teams of ~20 colleagues independently, designed a long-term operation plan</li><li>Composed and pitched the 'DSC x Military' business proposal for breaking into the military market with launch plans</li></ul>	06/2019 – 08/2019 Los Angeles, CA
<b>WST Consulting (Startup)</b>   <i>Education Consulting</i> <b>Research Analyst</b> <ul style="list-style-type: none"><li>Gathered info on target firms' finance and development strategy to generate biweekly reports, wrote 4 chapters of the guidebook on recruitment programs of finance, consulting, data and tech fields across global markets</li><li>Led resource management efforts to create and restructure marketing materials that yielded 50% increase – from 3000 to 4500, in average view count of over 15 supported articles on WeChat platform</li><li>Supervised an intern and 2 junior colleagues on document research methods and writing materials revision</li></ul>	09/2017 – 04/2018 Shanghai, China

## PROJECT (More projects accessible on [GitHub](#))

**Data Science: Where Should You Live for Your Health** ([Demo](#), [Code](#))

- Acquired ~20k rows of data through API after scoping the project and translating it into an analytical problem
- Assessed data by performing data wrangling and normalization with pandas, posed hypotheses from the evaluation
- Architected multiple ML models of clustering, association rule mining and regression on factors (air and water quality) to validate hypotheses, investigated the cancer trend based on results in Tableau and network analysis