

Cici Chen

San Francisco, CA • (254) 315-2642 • cc4291@columbia.edu
• [LinkedIn.com/in/iamcici/](https://www.linkedin.com/in/iamcici/) • [Github.com/iamcici0424](https://github.com/iamcici0424) • [Project Portfolio](#)

SUMMARY

3 YOE motivated, detail-oriented, and responsible data analyst with solid experience in business analytics and data science area, comprehensive understanding of ETL, statistical analysis, dashboard reporting, and machine learning techs. Proven ability to work with cross-functional teams, and communicate insights and recommendations to stakeholders effectively

SKILLS

- Programming/Tools: SQL, Tableau, R (ggplot2, dplyr, glmnet, shiny, etc.), Excel (Pivot Table, Vlookup), SPSS, Python (Matplotlib, NumPy, pandas, etc.), SAS, RedShift, Asana
- Certifications: Google Analytics, CITI Researcher, SAS Advanced/Statistical Business Analyst, Hootsuite Platform
- Machine Learning: Decision Tree, Random Forest, PCA, SVM, Logistic Regression, KNN, Clustering, Regularization

EXPERIENCE

Amazon.com, Inc. and its affiliates and subsidiaries

San Francisco, CA

Business Intelligence Engineer, AWS

01/2022 – Present

- Created and maintained 6+ metrics to track the performance of critical values in each report, captured abnormality in data pipelines by identifying appropriate thresholds using SQL, visualized overtime trends in Tableau, and deeply dived root causes to improve services under high peak usage of Data Center
- Validated service compliance by analyzing large amounts of datasets in DataGrip and Airflow, sliced complicated domain concepts into several layers of components, and designed clear diagrams for customers and teammates to understand data strategies and logic behind
- Delivered ~4 individual projects and cooperated 8+ dashboards yearly with high-quality outcomes, prepared clean documentation, and the standardized procedure now is widely used within the team after demonstrating its efficiency

SchoolCare, formerly CareDox

New York, NY

Data Analyst, Product Insights

06/2020 – 12/2021

- Attracted ~15% more clients and reduced ~33% operating cost by analyzing ~350+ tables with ~2B+ data and producing data-driven recommendations, maintained Covid-19 dataset to diagnose potential risk and trends, and researched social and mental health reasons behind chronic conditions in nurse notes
- Improved SQL queries efficiency by ~18%, tracked 10+ KPIs from Electronic Health Record platform, conducted the statistical analysis, prioritized data goals across teams, and built dashboards to maximize product performance
- Obtained ~26% more deals by examining sales velocity from HubSpot, quantifying survey results to build long-term school partnerships in our marketing funnels through acquisition, engagement, and retention

PROJECTS

Marketing Mix/Multi-Touch Models: Advertising Response Measurement in AdTech

- Evaluated the effectiveness of different advertising channels such as emails, social media, display, and direct mails by building 6+ Marketing Mix Models (MMM)/Multi-Touch Attribution Modeling using 500,000+ data along with advertising effect to identify the cause-and-effect relationships between sales and customers' reactions
- Increased the interactive application of the report by designing R Shiny App to visualize the results of advertising testing methods such as hold-out test and last-touch analysis

Applied Data Science: Image Recognition With 10 Machine Learning Techniques

- Proceeded 70,000+ images with 'data.table', reduced the dimension from 28x28 to 7x7, and classified the images by 10 models such as Neural Network, Logistic Regression with Lasso, KNN and Random Forest with 80+% accuracy
- Defined a score function to evaluate classification models based on the proportion of training set, computing time, and test error, then summarized model performances under different circumstances in dashboards

Data Visualization: Workforce HR Evaluation with R, Tableau

- Extracted 100,000+ data from 10+ tables, visualized ER Diagram with 'datamodelr' to display the relationships among employees' information, and animated the story with dashboards for better interpretation and presentation
- Explored the association between gender and their average rate, observed the discipline of the hiring season, predicted the new quota for sales by linear regression, and deliciated those results by Tableau and 'ggplot2' graphs

EDUCATION

Columbia University – M.A. in Statistics | GPA: 3.7/4.0 – New York, NY

2/2020

Baylor University – B.B.A. in Marketing | Dean's Honor List | *Cum Laude* | GPA: 3.88/4.00 – Waco, TX

12/2017