

# Jinxin(Jin) Hou

Jinxin.hou@marylandsmith.umd.edu | linkedin.com/in/jinxin-hou-a2708898

(240)610-3226 • College Park, Maryland (Willing to relocate)

## EDUCATION

**University of Maryland, Robert Smith School of Business** (College Park, MD) *Sep.2019-Dec.2020*

M.S. in Business Analytics (GPA: 3.75)

- Database Management System / Big Data and AI / Data Processing & Mining

## TECHNICAL SKILLS

- Python; SQL, RDBMS/NoSQL (Cassandra); MPP(Hadoop/Spark); Tableau; Statistical Modeling
- Snowflake; AWS(S3/Redshift); Databricks

## PROJECT EXPERIENCE

### Data Lake of Music Steaming Service – Udacity Project

*Python, AWS, Spark, SQL*

- Built an ETL pipeline that moves data warehouse with JSON logs to a data lake
- Extracted song and user logs from the data lake, processed data using Spark deployed on an EMR cluster, and loaded data back into S3 as dimensional tables in parquet format

### Anti-COVID19 Recommendations by Twitter Sentiment Analysis - Deloitte Case

*Python, Tableau – 20,000,000 nationwide tweets from Mar to Jun – Project Leader*

- Extracted tweets, preprocessed tweets, applied EDA, and measured overall sentiment(spacy)
- Summarized actions and incidents of best and worst five states based on state-level sentiment ranking, reported recommendations to help control COVID 19

### Customer Propensity to Pay Model – Genpact (NYSE: G) Case

*Python, SQL, Tableau – 45,000 B2B Collection Orders Across 3 Years – Project Leader*

- Led data mining (CRISP-DM); found and created key factors affecting pay date, developed classification models (Decision Tree, Logistic Regression) with 85% accuracy
- Developed executive summary with implementation plan to reduce overdue payments by 60%

## PROFESSIONAL EXPERIENCE

### New Nationwide Service Profitability Analysis - Pilot Freight Services

College Park, MD

*Snowflake, Python, Tableau – 50 Million Rows of record - Analytics Consultant*

*Aug.2020-Nov.2020*

- Led a team to estimate the feasibility and profitability of a nationwide operation plan by conducting ETL, EDA, data mining, and simulation, and proved a profitable operation plan
- Built the ETL pipeline to prepare data and created reports to support C-suite decision making

### University of Maryland, Robert Smith School of Business

College Park, MD

*Graduate Teaching Assistant*

*Aug.2020-Nov.2020*

- Instructed ETL and data analysis using SQL, python (Regex, Pandas, Matplotlib, Beautiful Soup)
- Tutored students on application of sampling and estimation to business cases, decision-making with decision tree, AB and hypothesis test evaluation, and predictive analytics

### New Oriental Education & Technology Group Inc.

Beijing, China

*Business Analyst*

*May.2017-Aug.2019*

- Analyzed conversion rates and user behavior across online marketing channels, such as web, social network, and advertising to design marketing campaigns and increased sales by 7% (> 1Mil)
- Built departmental dashboards using SQL and Tableau to monitor KPIs and support decision-making