

## EDUCATION

## University of Illinois at Urbana-Champaign

Master of Science in Information Science | GPA: 3.86/4

Relative Courses: Data and Statistical Models, Data Mining, Data Visualization, Data Warehousing and BI, Database Systems

Champaign, IL

Aug. 2023 - Dec. 2024

## University of Electronic Science and Technology of China

Bachelor of Engineering in Computer Science and Technology | GPA: 3.73/4

Relative Courses: Probability &amp; Statistics, R for Statistical Analysis, Python Data Analysis, Data Structure &amp; Algorithm

Chengdu, China

Aug. 2019 - Jun. 2023

## SKILLS

**Programming Languages:** Python,R, SQL, Java, Javascript, CSS, HTML**Tech Skills:** Numpy, Pandas, Matplotlib, Sklearn, seaborn, MySQL, Hive, Hadoop, Spark, ggplot, dplyr, sparklyr, shiny, AWS (EC2, S3 Bucket, Lambda), GCP, Azure, VLOOKUP, Pivot Table, Machine Learning Models (Regression, Random Forest, K-Means), EDA, ETL, A/B Test**Platforms:** Excel, Tableau, PowerBI, Matlab, SAS, Looker, Talend, Airflow, Snowsight

## PROFESSIONAL EXPERIENCE

## Apple (through Manpower)

Beijing, China

Data Analyst Intern

May 2024 -Sep. 2024

- Supervised Apple China MONO online stores on the Meituan platform, conducted comprehensive sales data analysis, and formulated promotional strategies, resulting in a year-over-year increase of **17%** in sales and **23%** in GMV.
- Utilized **SQL** to query and manage datasets containing over **2 million** records, ensuring data integrity and accuracy.
- Developed robust **Machine Learning Models** using **Python**, including **Random Forest** models for feature selection and **Multiple Linear Regression** for predicting sales trends, which enhanced the granularity and accuracy of sales forecasts by **15%**.
- Designed and executed an **A/B test** for the 618 (Mid-year) Shopping Festival to evaluate promotional strategies, utilizing **Descriptive Analytics** to analyze the resulting data to identify the most effective approaches, leading to a **17%** YoY improvement in campaign performance.
- Developed and maintained weekly updated data dashboards using **Tableau**, providing dynamic and interactive visualizations that enhanced decision-making processes for senior management and cross-functional teams.

## DiDi Chuxing

Beijing, China

Operational Data Analyst Intern

Jul. 2022 - Jan. 2023

- Performed data analysis on DiDi's driver dataset to support the development of driver retention strategies. Implemented strict data validation and analysis protocols, achieving a **25% reduction** in insurance costs, a **37% decrease** in claim rates, and an increase in driver retention rates from **85.2% to 92.7%**.
- Utilized **SQL** queries for comprehensive data retrieval and built dynamic dashboards with **Tableau Server** for real-time monitoring of driver behavior and insurance claims, with T+1 updates to support immediate operational decisions.
- Designed two **A/B tests** to enhance driver personal insurance enrollment and reduce car insurance claims, using **SQL** for precise data segmentation and extraction.
- Conducted **Inferential and Causal Analysis** using **Python** and **R** to evaluate the impact of operational strategies and A/B tests on reducing insurance costs and improving driver retention rates.

## SusallWave Sustainable Digital

Shenzhen, China

Technology ESGData Analyst Intern

Jan. 2022 - Jul. 2022

- Participated in an ESG data analysis project, performing advanced statistical analysis that improved the efficiency of ESG reporting by **20%** and improved client satisfaction.
- Utilized **Pandas** to quantify unstructured data from the annual and responsibility reports of CSI 300 companies, and developed an ESG data analysis system, using **Python** and **NLP** to automatically extract and quantify ESG performance indicators from textual data.
- Created an ESG scoring framework to assess and rank companies based on their ESG practices, integrating the results into a dashboard for data visualization and analysis using **Pivot Tables** in **Excel**.
- Performed advanced **Predictive Analysis** to identify trends in corporate ESG behavior and correlations between various ESG indicators, providing data-driven insights for investors and stakeholders evaluating the company.

## China Mingsheng Banking Corp., Ltd

Chengdu, China

Financial Data Analyst Intern

Jun. 2021 - Nov. 2021

- Developed a data analysis system for the credit card department, integrating **ETL** processes with interactive dashboards, and conducted **Predictive Analytics** that achieves a **95%** accuracy rate in assessing creditworthiness and identifying risk factors.
- Established automated **ETL pipelines** using **Python**, facilitating real-time data extraction, transformation, and loading, which enabled continuous monitoring and analysis of loan applicant behaviors and credit risk factors.
- Constructed an interactive visualization dashboard utilizing **Power BI**, automating the monitoring of user lifecycle stages, from application to loan repayment, to identify patterns and trends that impact credit decisions.
- Performed in-depth **predictive modeling** and forecasting using **R**, leveraging machine learning algorithms such as **logistic regression** and **ARIMA** models to provide accurate insights into loan demand and default rates.

## PROJECT EXPERIENCE

## EU Unemployment Visualization

Aug. 2023 - Dec. 2023

- Developed an EU unemployment trends dashboard, focusing on age and gender disparities as well as regional trends, utilizing **PowerBI** and **Tableau** for **data visualization**.
- Processed extensive EU unemployment data in **Excel**, performing segmentation by age and gender, and applied advanced data cleaning across 50+ sheets for integration into the dashboard.
- Created the EU Countries Unemployed Map in **PowerBI**, using a color gradient to illustrate unemployment rates, allowing for interactive exploration of regional and temporal data.
- Developed interactive dashboards in **Tableau** for detailed unemployment analysis, including an Unemployment Population Table with interactive filters, emphasizing age-related unemployment trends.

## Spotify Artists Analysis

Jan. 2023 - Apr. 2023

- Conducted data analysis to unveil the distinct musical styles and preferences across a diverse range of genres by analyzing the musical attributes of 50 mainstream artists on Spotify, leveraging the Spotify Web **API** to extract detailed audio features.
- Used **R** packages like to handle API requests and data manipulation.
- Applied **K-means Clustering** in **R** to categorize artists into distinct groups, facilitating the understanding of genre-specific characteristics and artist similarities.
- Utilized **D3.js** for dynamic **data visualization**, creating interactive radar charts and **PCA** scatterplots to illustrate the relationships between various audio features and to display the artist clustering.
- Documented the entire analytical process in an **R Markdown** file, integrating code, visual outputs, and narrative analysis to compile a comprehensive report.

## Statistical Analysis in Phonetics

Mar. 2022 - Jun. 2022

- Conducted a **data-driven** phonetic study on the prosodic features of deferential speech in Mandarin among 26 participants, demonstrating the ability to predict deferential speech with over **95% accuracy** using pitch span, and nearly **100% accuracy** for other acoustic features.
- Utilized **R** to perform **Linear Mixed-Effects Modeling**, analyzing the relationship between speech features and variables such as gender and social attitudes. Refined the models systematically, focusing on statistically significant predictors and their effect sizes.
- Used the **ggplot2** package along with custom scripting to create insightful graphs and charts, illustrating the differential patterns in acoustic features across various speech conditions.