JENNIE LE

DATA SCIENCE & ANALYTICS

CONTACT

- tle35@fordham.edu
- (347) 449-2616
- New York, NY
- LinkedIn: linkedin.com/in/jenniele
- GitHub: https://github.com/bigforehead
- Personal website: bigforehead.github.io

PROFILE

Passionate about building solutions through business strategies and data analytics to improve efficiencies and drive revenue. Strong foundation in Web Analytics, E-commerce, and Web Usage Mining.

EDUCATION

SEP 2019 – DEC 2020

FORDHAM UNIVERSITY [NEW YORK, NY]
Master of Science in **Analytics**
GPA: 3.81

JAN 2014 – DEC 2017

ELIZABETHTOWN COLLEGE [ELIZABETHTOWN, PA]
Bachelor of Science in **Finance and Accounting**

TECHNICAL SKILLS

- Data Analysis: Python, R, SQL, Tableau, SPSS
- MS Office: Excel, PowerPoint, Word
- Web Analytics: Google Analytics
- A/B Testing, Applied Regression Analysis
- Pandas, Scikit-learn, NLTK, TensorFlow, Spark, Hive
- Google Cloud Platform (GCP), AWS, Data Bricks, Azure
- Storytelling

AWARD & CERTIFICATIONS

- Dean's List
- Delta Mu Delta
- The National Society of Leadership & Success

EXPERIENCE

GABELLI SCHOOL OF BUSINESS　　　　　NEW YORK, NY
Graduate Assistant　　　　　　　　　JAN 2020 – PRESENT

- Developed and **led over 10 workshops to 120 graduate students** in **Tableau, Python, Databricks, SQL,** and **Azure.**
- **Transformed high-dimensional data sets** to insightful dashboards and **designed reports** using **Tableau** for 5 academic projects.

ALVAREZ & MARSAL　　　　　　　　　NEW YORK, NY
Data Analytics Intern　　　　　　　MAY 2020 – AUG 2020

- **Managed a team of 5** to conduct **ad-hoc data analysis** via machine learning and web scraping to examine the relationship between S&P500 financial performance and organizational culture.
- Collected Glass Door's 60,384 job descriptions and **identified key metrics** to evaluate S&P500 company financial performance.
- Generated **exploratory data analysis report** by creating **Tableau** interactive visual analytics on major organizational culture concepts and Tobin's Q.

HAVER ANALYTICS　　　　　　　　　NEW YORK, NY
Economic Research Analyst　　　　MAR 2018 – FEB 2019

- Maintained and updated extensive time-series database by **tracking economic key metrics** utilizing Data Link Express (DLX).
- **Led a team of 2** by providing technical guidance and engagement opportunities.
- **Streamlined** updating process by **80%** by **building automation program** for consolidation tasks, such as data collection and database updates.
- **Collaborated** closely with database managers to perform database cleansing to ensure validity of updates.

PROJECTS

Fashion Life Cycle Forecasting with Instagram

- **Coordinated** with fashion experts and programmers to **conduct data analysis** and **track** work process on Jira platform.
- **Identified** key fashion trend elements by conducting market research.
- **Collected, aggregated,** and **analyzed** Instagram data using **GCP.**
- **Transformed** 'fashion language' into machine-readable expression.
- **Created presentation** with interactive **Tableau dashboards** to access Prophet model performance.

Lyrics-Based Music Recommendation System

- **Built** a **personalized** music recommendation system **to increase users' engagement and retention** using Spark and MLlib on **Google Cloud Platform**.
- **Queried** 3GB data using **SQL** to visualize and model lyric topics by applying Natural Language Tool Kit and Latent Dirichlet Allocation.
- **Engineered** new lyrics features with TF-IDF, Word2vec, and LDA to **recommend** songs based on the content.

Twitter Malicious Bots Classification

- **Built a Random Forest model** using **SPSS** to classify spam bots, scam bots, and fake followers, resulted with **91.7%** accuracy.
- **Scraped** +200,000 tweets using **Python** and Twitter API, engineered bot temporal behavioral features, and applied **Tweet Semantics analysis**.
- **Maximized** the **organic users' exposure** of social media post on Twitter.

The Art Gallery Tracking System Database

- **Developed a large-scale database** prototype from scratch using Oracle for the museum to **track visitor traffic.**
- Used **SQL** to perform **60** queries from the database to **generate ad-hoc reports.**