

# Po-Yu Hsieh

pykenny@gmail.com | GitHub: pykenny | LinkedIn: po-yu-hsieh | Website: pyhsieh.net

---

## EDUCATION

**Boston University**, Boston, Massachusetts, United States  
Master of Science in Computer Science, GPA: 3.66

Sep 2017 — May 2019

**National Taiwan University (NTU)**, Taipei, Taiwan  
Bachelor of Business Administration in Finance

Sep 2009 — Jun 2014

## WORK EXPERIENCE

**CYBERBIZ Co.**, Taipei, Taiwan

Apr 2021 — Present

Data Engineer

- Develop and maintain data transfer pipeline between service database (MySQL, PostgreSQL), data stream (Kafka), and external services (Google Analytics) and BigQuery data warehouse; design database schema and SQLAlchemy data model to facilitate development
- Develop and maintain Data-AI Team's web API service, Data Science/AI/Analytics pipelines, and integration with other products (e-Commerce, e-POS, plugin applications); set up Celery queue service to support background job functionality
- Migrate legacy batch pipelines into Apache Airflow for centralized management; create modularize pipeline utilities, set up staging/testing environment, documentation, monitoring, and error report mechanism for the new system; lead Python runtime system migration process to deprecate environments reaching end-of-life
- Evaluate application scenarios that utilize commercial LLM products (ChatGPT, Gemini, etc.) and develop product prototype to collect feedback, including Slack Bot application and automated data labeling pipeline
- Handle analytical requests from product and business departments and manage internal business intelligence dashboard

**Klaviyo**, Boston, Massachusetts, United States

Sep 2019 — Sep 2020

Data Science Software Engineer

- Supported team data scientists to gather service data for analysis and potential product ideas
- Sampled and compared service data from different sources to ensure data consistency of feature under development with other services
- Developed product prototype with React, Django, Flask as an internal analytics tool to collect feedback from customer success and product design teams
- Maintained automated data collecting process for COVID-19 insight analysis

**Research Center for Digital Humanities, NTU**, Taipei, Taiwan

Jan 2016 – Jul 2017

Research Assistant

- Developed 5 database building and text analysis webpage tools for DocuSky, a database platform to store and analyze text corpus
- Introduced DocuSky and research findings using this platform at the 2016 Conference of Digital Archiving and Digital Humanities

## TECHNICAL SKILLS

### **Programming Language**

Python, Ruby, JavaScript (ECMAScript, TypeScript)

### **Databases, Web Applications, Software Development**

- **Database & Data Warehouse:** MySQL, PostgreSQL, SQLite, BigQuery, SQLAlchemy
- **Web Applications:** Flask, Django, Celery, Redis, React, jQuery, CSS/SCSS, HTML
- **Data Science:** Pandas, NumPy, scikit-learn, Jupyter Lab
- **Data Engineering:** Apache Airflow, Kafka, Apache Avro, Apache Arrow, Protocol Buffers
- **Cloud Service:** Amazon Web Services (AWS), Google Cloud Platform (GCP)
- **Containerization:** Docker, Kubernetes
- **Other Development Tools:** Git/Github, Poetry

## LANGUAGE SKILLS

- English (TOEFL Total 106 (2016); GRE V153/Q167/AW3.0 (2016))
- Mandarin (First Language)