

SHAUN ZHENG

Philadelphia, PA 19104 | (267) 324-1399 | shaun.zhong@gmail.com | shaunzhong.github.io/portfolio/

As a life-long learner and motivated data scientist with years of learning experience, I strive to pose and answer questions with data-driven insights. I am passionate about taking challenges and building models that fix real world problems.

SKILLS

- **Programing:** Python (Scikit-learn, Pandas, NumPy, Pyspark), R (Tidyr, Caret), SQL
- **Data Visualization:** JavaScript, HTML, CSS, Python (Matplotlib, Seaborn, Altair, HoloViz)
- **Hands-on experience:** A/B Testing, Machine Learning/Deep Learning, Time Series Analysis, Spatial Analysis, Dashboard, Computer Vision, NLP, Convolutional Neuro Network, Open-source Hardware & Programming
- **Others:** Git, Spark, Statistics, Google Cloud Platform, AWS, System Thinking & Problem Solving

EDUCATION

UNIVERSITY OF PENNSYLVANIA

Master of Urban Spatial Analytics – GPA: 3.94/4

Philadelphia

2021 - Present

TONGJI UNIVERSITY

B.S. Industrial Design – GPA: 3.90/4

Shanghai

2017 - 2021

PROJECTS

FOOT TRAFFIC PATTERN DATA ANALYSIS & HUFF MODEL VALIDATION [Link](#)

- Collaborated with **Philadelphia Parks & Recs Department** to better program resources among **500+ facilities**
- Utilized **external foot traffic pattern data** to cluster with attendance data using **K-Means** to find correspondence and conflicts
- Created & published R library, '**Socially Aware Huff Model**', and **established operating dashboard**

STREET VIEW IMAGE SEGMENT EXTRACTION, ANALYSIS [Link](#)

- Scrapped 12,000+ Bing Street View images in Philadelphia Road Net through **Bing API**
- Trained with **SegNet (Semantic Segmentation)** to extract street components and proportions from image collections
- **Spatially operated** with community demographic data, and crime data to explore the society equity topic

DELAY PREDICTION WITH TEMPORAL & SPATIAL FIXED AND LAG EFFECT [Link](#)

- Explored **ground truth meta data**, created time series panels, and constructed **time fixed and lag indicators** in different scales
- Constructed a **segmental model system** using logit regression to classify firstly and apply OLS regression respectively

SVM & KNN CLASSIFICATION EFFECT COMPARISONS IN COMPUTER VISION PRACTICE [Link](#)

- **Decreased image dimensions** using binarization and outline extraction (find first foreground pixel of each pixel row & column)
- Looped KNN and SVM models using different values of K and **matrix of different SVM parameters**
- Calculated & compared responding prediction accuracies in each category, and used **PCA** to visualize prediction outcomes

WORK & EXPERIENCE

PENN DATA SCIENCE GROUP

Kaggle Competition Team Member

Philadelphia

Jan. 2022 – Present

- Participated in 5 Kaggle Data Science Competitions to solve business problems with machine learning methods
- Learned new techniques, and chatted with data science professions from different industries at campus

JLL

GIS Analyst Intern

Shanghai

Feb. 2021 – Jul. 2021

- Built database pipelines and interactive dashboards to track the price and transactions for logistics properties in Shanghai MSA
- Conducted macro and micro geospatial and market analyses in 10+ consulting reports for Top500 clients from FMCG, Medicine, Food, Chemical, and Automotive Manufacturing industries etc.

INTERBRAND

Strategy Intern of President Office

Shanghai

Jan. 2021 – Feb. 2021

- Supported president with analysis and research in supervising project strategies for more than 5+ international pitches
- Digitalized database for completed projects and provided query and wrangling assistance to ongoing pitches

CREATER

Project Management Intern

Shanghai

Apr. 2019 – Dec. 2019

- Proposed business-optimized solution using data analysis to predict cash flow and balance cost in construction and operation
- Assisted with general manager in managing cooperation among investors, governments, designers, and constructors