

SOPHIA ZHANG

Data analysis and visualization expert seeking Summer 2021 internship

EDUCATION

2020–
Current

● **Tufts University**
MS Candidate in Data Analytics

📍 Medford, MA

2007–
2012

● **University of Alberta**
B.Sc. in Chemical Engineering

📍 Edmonton, AB, Canada

WORK EXPERIENCE

Feb
2017–
Sep 2020

● **Process Engineer**

Bantrel Co.

📍 Edmonton, AB, Canada

- Effectively used empirical operations data to forecast the magnitude and consequence of various risk scenarios in pipeline projects
- Expert in modeling to determine cost effective, compliant, and practicable solutions to mitigate risk
- Experienced in analyzing large and small capital projects with different layers of complexity from single equipment to large connected networks
- Processed and presented trends from large volumes of data (collected from monitoring systems)
- Optimized production by reducing redundancy in plant work processes and design
- Actively communicated and provided timely support to clients for project implementation
- Recognized by client for presenting innovative and cost saving solutions
- Participated in full project cycle from research and information collection, proposal and project bid, cost estimation, presentation and delivery of viable solutions to client and implementation support

Sep 2016
– Feb
2017

● **Project Engineer**

CEPro Energy & Environmental Services

📍 Edmonton, AB, Canada

- Evaluated feasibility of expansion into the strand natural gas delivery market by investigating established and upcoming technologies and their market trends
- Developed business, siting, procurement, and regulatory plans and conceptual design for a clean energy technology, natural gas power generation plant
- Engaged in business development by presenting additional services to current clients and researching and cold calling potential clients via industry contacts
- Coordinated site assessment projects and ensured regulatory compliance between stakeholders
- Applied directives from the federal, provincial and municipal regulations and standards regarding activities for the production, transmission, and distribution of energy



CONTACT INFO

✉ Hao.Zhang634280@tufts.edu

🐙 github.com/sophiahzhang/

in

www.linkedin.com/in/sophiaHzhang

SKILLS

Proficient in R (RStudio), Python (Jupyter, Spyder), Git, QGIS, AFT, MATLAB

Excels at **data analysis**, with a focus on business, health and bio-informatics, economics, and engineering models

Experienced in **data communication**; interpreting data and building models

Machine Learning on cloud platforms (AWS)

8 years of **engineering experience** in industrial applications

Expert in process optimization, risk management, consulting and client relations

Last updated on 2020-11-27.

May
2012 –
May
2015

● **Process Engineer-In-Training**

Edmonton, AB, Canada

📍 Stantec Inc.

- Designed multifunctional models for complex plants and processes by incorporating multiple information sources including system and component drawings, equipment datasheets, technical reports, 3D models, and databases
- Developed department work standards, technical and business reports, scope of work, cost estimates, project workflows, and process implementation guidelines
- Estimated capital cost and operating cost, compared tenders and bid proposals from vendors for various technologies
- Eliminated \$10 million of infrastructure costs by identifying risks and developing solutions using data analysis and simulation in a pipeline system
- Conducted systematic health, safety & risk management assessments
- Managed multiple projects' scope, budget, schedule and performed quality reviews and audits
- Increased productivity by implementing team matrices to align project needs with members' skills



SELECTED PROJECT EXPERIENCE

2020–
Current

● **DATA-200 Foundations of Data Analytics**

Tufts University

📍 Medford, MA

- Developed a model in R which predicted student test scores based on features such as student enrollment and school expenditure
- Iterated the model to select the best approach, given OLS, LASSO, and ridge regression and random forest.

2020–
Current

● **DATA-201 Machine Learning in Python**

Tufts University

📍 Medford, MA

- Constructed a predictive model in Python for US market avocado prices
- Applied linear and logistic regression, principle component analysis (PCA), and KNN algorithms to predict price, analyze time series data, and classify avocado types.
- Created clear and insightful visualizations to communicate the findings.

2020 –
Current

● **COMP-134 Computational Models in Cognitive Science**

Tufts University

📍 Medford, MA

- Created a Bayesian predictive model in R with JAGS for natural language processing
- Predicted human reaction time and classification of conversational pauses by testing various models and evaluating their accuracy on laboratory data.



SELECTED PRESENTATIONS

No-
vember
2020

● **DATA-200 Foundations of Data Analytics**

Tufts University

📍 Medford, MA

- Developed a model for predicting Airbnb listing price based on reviews, property characteristics and location in Cambridge, Massachusetts
- Presented findings of the model and evaluation metrics to my cohort.

October
2020

● **DATA-201 Machine Learning in Python**

Tufts University

📍 Medford, MA

- Presented an exploratory data analysis for heart disease hospital data analyzing patient demographics, cholesterol, blood pressure, and medical ECG readings.
- Created clear and insightful visualizations to communicate the findings.