

Xavier Genelin

Arcadia, WI | 507-459-0673 | xgenelin@gmail.com | <https://xaviergenelin.github.io/>

EXPERIENCE

Data Scientist – *Freelance project*

July 2022-Present

- Assisted medical team to examine skin cancer data from SEER registry
- Collaborated with medical team in data collection and applying statistical techniques
- Investigated skin cancer data with **R** to assist in research and understand available data
- Created a regression model to determine significant variables in skin cancer incidence
- Publication pending in 2023

Data Scientist - *Mattress Firm*

November 2021-July 2022

- Built customer segments in **python** based on demographic data using **unsupervised models** to analyze customer habits and look for marketing opportunities
- Examined customer demographic data for a variety of marketing teams, working with stakeholders to develop a more data-driven strategy
- Investigated customer survey data in python with **NLP** to analyze feedback from various customer groups and determine opportunities for improvement
- Created an **XGBoost** model to classify customers based on previous transaction habits to aid in customer analysis with **82% accuracy**
- Analyzed the impact of economic stimulus packages on sales, **determining there was an increase**
- Communicated results with stakeholders (managers to C-suite) to both technical and non-technical audiences

Quantitative Analyst - *NC State Baseball*

March 2021-June 2022

- Formed a report in **R** to analyze NC State pitchers to help optimize their performance
- Consulted the coaching staff based on findings from analysis and assist with game strategy
- Advanced to the 2021 College World Series semifinal and 0.632-win percentage in 2022

Business Intelligence Analyst/Data Analyst - *Ashley Furniture Industries*

2019-November 2021

- Automated manual processes writing **SQL** queries, **saving 45 hours per week**
- Designed an app in **python** to optimize the process of diverting shipping containers, **saving 8 hours per week**
- Conducted a statistical analysis in **R** on new product sales and advertisement spending using a linear regression, determined ad spending had no impact on sales, **saving \$300,000**
- Developed and maintained 17 Power BI Dashboards for Supply Chain, Manufacturing, and HR

EDUCATION

Master of Science, Statistics - *North Carolina State University*

Bachelor of Science, Mathematics, concentration in Economics - *Xavier University*

SKILLS

Programming languages: SQL (T-SQL, MySQL), R, Python, PyTorch, PySpark

Statistical Modeling: Classification, Regression, Clustering, Deep Learning (CNN, RNN), NLP

Tools: Google Cloud Platform (GCP), Jupyter Notebook, RStudio, Jupyter Lab, Google Colab, Tableau, BigQuery

PROJECTS

Emotion Detection [\[github repo\]](#)

- Developed deep learning and traditional models with **PyTorch** that detects the emotion behind a conversation

Terrain Identification [\[github repo\]](#)

- Classified the type of terrain using **PyTorch** from a prosthetic limb based on accelerometer and gyroscope data using a CNN with **88% accuracy**

Alzheimer's Risk Factors [\[github repo\]](#)

- Identified risk factors associated with individuals identified with Alzheimer-onset dementia using **PySpark** and used classification models for prediction with **87% accuracy**