

# Net Zhang

614-815-8511 | zhang.11091@osu.edu  
Columbus, Ohio  
linkedin.com/in/net-zhang/

## EDUCATION

---

### The Ohio State University

May 2022

Computational Data Analytics - Bachelor of Science

Columbus, OH

- Magna Cum Laude
- Cumulative GPA : 3.81 / 4.0

## SKILLS LIST

---

- Probability Theory, Statistical Modeling Fundamentals, Optimization Techniques (MILP)
- Data Wrangling (Python & R & SQL), Operating System (Linux, C), Version Control (Git)
- Visualization, Web-Application Development (R Shiny, Docker, JavaScript, HTML, CSS)
- An R package developer and active member of the Columbus R community

## PROFESSIONAL EXPERIENCE

---

### The Ohio State University Translational Data Analytics Institute

Aug 2022 - Present

Data Scientist

Columbus, OH

- Equity Mapping Tool for Community Health Workers (CHWs)
  - Developed a Modularized R Shiny Web Application to support the day-to-day work of the Community Health Workers (CHW) across Ohio. The application uses new technologies such as bslib to generate the front-end UI and Apache Arrow to reduce backend memory usage and querying time dramatically. It was made to be publicly available, so please explore the full utilities at [emthub.osc.edu](https://emthub.osc.edu).
  - Designed file structure and Git version control workflow to orchestrate data ingestion and archives. Used Python package BeautifulSoup to scrape static web pages to collect food pantries and homeless shelter information. Partitioned large dataset into parquet files to accommodate the frequent query requests from Shiny app users.
  - Converted textual, graphical, and tabular information into ggplot objects and created a summarized Rmarkdown report template in multi-column layouts. Scheduled cron jobs to render PDF reports weekly for each jurisdiction and distribute them to respective recipients.
- The Franklin County Opioid Crisis Activity Levels (FOCAL) Project
  - Led developers in producing a Modularized R Shiny Application that helps Ohio State researchers and community partners identify patterns of opioid overdose events in space and time to understand better the social and environmental determinants associated with these patterns.
  - Initiated the process with the DevOps team to deploy Nominatim from OpenStreetMap to address the secured geocoding challenges. Developed a framework utilizing the future and furr packages to leverage the multi-core processing power of the OSC server for efficient geocoding and reverse-geocoding of a high volume of sensitive overdose incident addresses. Packaged this framework into an R utility package and shared it with Nationwide Children's Hospital to assist in geocoding drug-related crime incidents.
  - Hosted bi-weekly training sessions for users from FCPH, CPH, Safe Point, Columbus Fire Department, and ADAMH for product demonstration and collecting user feedback.

### Progressive Insurance

May 2022 - Aug 2022

Commercial Line Pricing Analyst

Cleveland, OH

- Liability Coverage Costs Prediction
  - Trained regularized linear models for predicting monthly liability costs for commercial line policies from the state of Texas.
  - Developed an internal R package that put the model in production by integrating the data pipeline

from the BI team and delivering the prediction plumber API for the R&D team.

## **Worthington Industries**

May 2021 - May 2022

Corporate Advanced Analytics Intern

Worthington, OH

- Steel Processing Analytics Tool - PHYLO
  - Collaborated closely with Quality Engineers to model complex coil lineage data from Oracle with binary search trees that capture 5,000,000 unique operations and more to come.
  - Implemented interactive coil lineage tree with R Shiny that allows plant workers to mark accurate defect locations, trace the origin of defects, and expose unidentified defects.
  - Enabled supply chain managers, plant workers, and quality engineers to quickly identify defections from received raw materials to sitting finished goods, significantly improving operational processes and suppliers relationships.
- Self-Serve Time Series Analytics Tool - TSA
  - Developed R packages and R Shiny application that enables supply chain managers and sales representatives to assess, clean, and explore time series, at last, make a future forecast that leads to specific decision making.
  - Applied step-wise strategy to search for the optimal lags for leading external indicators, which improved the cross-validation accuracy by 30% on average.
  - Saved \$60,000/year in reduced cycle time for generating statistical forecasts that create opportunities for inventory reduction, on-time delivery, and operational efficiencies

## **The Ohio State University College of Public Health**

Apr 2020 - May 2022

Student Research Associate

Columbus, OH

- COVID-19 Analytics and Targeted Surveillance System (CATS)
  - Led design and development of the CATS app that specialized in informing school district leaders in the Columbus area about emerging trends in COVID-19 cases in their districts.
  - Automated the reproducible weekly report for superintendents across 30 departments to assist their decision-making process in alternating their districts' learning modes.
  - Informed 100,000+ students' families in Columbus about the school-aged diagnosis population in their communities.

## **Battelle Center for Science, Engineering and Public Policy**

Apr 2020 - Sep 2021

Data Science Research Developer

Columbus, OH

- OSU & Yale Collaboration Project: Targeting Testing with Mobile Units in Ohio ([Bandit](#))
  - Developed an information system using R Shiny and Tableau to run active surveillance for SARS-CoV-2 and target the best places to find new cases on an ongoing basis with mobile outreach and testing units.
  - Applied Multi-Armed Bandit Reinforcement Learning Algorithm with batch Thompson Sampling strategy to construct a recommendation system for workload assignments