

# Benedict Neo

☎ 515-388-0996 ✉ [benthecoder07@gmail.com](mailto:benthecoder07@gmail.com) [linkedin.com/in/benedictneo](https://www.linkedin.com/in/benedictneo) [github.com/benthecoder](https://github.com/benthecoder)

## Technical Skills

**Programming Languages:** Python, R, SQL, Java, JavaScript, HTML/CSS, Bash

**Machine Learning:** Pandas, NumPy, Matplotlib, scikit-learn, plotly, Tidyverse

**Data Science:** Data Science pipeline (cleaning, wrangling, visualization, modeling, interpretation), hypothesis testing

**Tools:** SPSS, Excel, Tableau, AWS, Google Cloud, Google BigQuery, Git

## Professional Experience

### Data Science Intern

Apr. 2021 – Present

*bitgrit Inc.*

*Tokyo, Japan*

- Collaborated with data scientists to develop and formulate solutions for Data Science competitions
- Wrote technical data science articles and tutorials with over **400,000 views**
- Taught 50+ people data cleaning in Python in a data science workshop

### Undergraduate Research Assistant

Jan. 2022 – Present

*Iowa State University*

*Ames, IA*

- Undergraduate statistics research with Dr. Jarad Niemi, Department of Statistics
- Coded functions to transform, analyze, and visualize soil and land data
- Dealt with over 2TB of untidy data in a Linux server

## Projects

### LinkedIn Connections Insights | Python, Streamlit, pandas, thefuzz, plotly, pyvis, networkx

Jan. 2022

- Built a web app with **500+ users** that allows them to gain insights into their LinkedIn connections
- Utilized fuzzy matching to clean and manipulate user text data for more accurate insights
- Visualized user connections with interactive bar charts, time series plots, and network graphs

### Next Word Prediction | R, Shiny, Tidytext, Tidyverse

Aug. 2021

- Built a Shiny web app for users to enter text and obtain multiple next word predictions
- Analyzed and cleaned over 4 million lines of text corpus data sourced from news, tweets, and blogs
- Utilized the Katz Back-Off (KBO) language model and Markov Chains to generate next word predictions

### COVID-19 Dashboard | R, Flexdashboard, Tidyverse, Plotly, Docker

Feb. 2021

- Informed users of key metrics and trends of COVID-19 worldwide with interactive maps and charts
- Implemented automatic daily updates to dashboard using Github Actions and Docker

## Education

### Iowa State University

Ames, IA

*Bachelors of Science in Statistics, Minor in Computer Science*

*May 2023*

- CGPA: 4.0/4.0
- Relevant Coursework: Probability and Statistics, Object-Oriented Programming, Data Structures, Discrete Mathematics, Design and Analysis of Algorithms
- Statistics Department Undergraduate Scholarship (\$1,000 awarded)

## Certification

### Data Science Specialization

Coursera

*10 course specialization on Data Science by Johns Hopkins University*

- R Programming, Data Analysis, Data Cleansing, Data Manipulation, Statistical Inference, Regression Models, Machine Learning

## Leadership

### Vice President of American University Program Society

Kuala Lumpur, MY

*INTI Subang College*

*Aug. 2019 - Sep. 2020*

- Organizing chairperson of a social event with over 100 participants
- Chaired meetings and delegated tasks to 10 committee members