Nihal Sidhu

nss94@cornell.edu | 609-613-1978 | Princeton, NJ | GitHub Portfolio

EDUCATION

Cornell University August 2019 - May 2022
B.A. in Data Science, Minor in Business *GPA: 3.64*
- Relevant Coursework: Data Science | Object Oriented Programming & Data Structures | Practical Tools for Machine Learning and Data Science | Statistics | Finance | Business Management

University of Michigan, Ann Arbor August 2018 - May 2019
Transferred to Cornell University in 2019 Fall *GPA: 3.72*

WORK EXPERIENCE

Bank of America July 2022 – Present
Software Engineer – OTC Derivatives Team *Jersey City, NJ*
- Developed an ETL pipeline to aggregate millions of OTC trade data monthly. Created a dashboard summarizing the OTC information, eliminating the 1 hour wait time required to generate the monthly report.
- Automated daily metrics of disk space utilization. The daily metrics helped identify errant jobs, resulting in a 20% decrease in disk space usage.
- Designed a reusable purging system to help remove 10TB of reporting trade data from an overloaded database, which removed the $5,000 cost requirement for a new database server.
- Operated as the main point of contact for ad-hoc requests and live production issues for 2 months, ensuring regulatory requirements were met.
- Led and mentored 3 college interns on their individual projects, timelines, and presentations. Streamlined and documented the on-boarding process to help with future intern in-take.

Johnson & Johnson June 2021 – August 2021
Data Scientist Intern – Pharmacovigilance Team *Edison, NJ*
- Redesigned the FDA Adverse Event Reporting System (FAERS) database, which helped the pharmaceuticals research team with accessing and creating summary reports on over 100k adverse medical events.
- Structured the system as relational database to increase the speed at which relationships could be discovered across different medical events.

PROJECTS

Predicting Touchdowns in the NFL GitHub - Predicting Touchdowns
- Trained a predictive model using regression analysis and probable usage to help identify which players were likely to score more touchdowns in the 2023 NFL season.
- Utilized Pandas and NumPy to analyze over 200k records of NFL play-by-play data and visualized the analysis using Matplotlib and Seaborn libraries.

Optimizing your Fantasy Football Lineups GitHub - FF Lineups
- Created a decision framework to help fantasy football users decide which players to start in their lineups based on positional relationships and underlying player performance data.
- Constructed the optimal strategy by cleaning weekly fantasy football data and mapping player performances across correlation matrices in Python.

New York Housing Market Dashboard Tableau - NY Housing
- Built an interactive dashboard of the New York housing market, allowing the user to filter and uncover key insights in their NY home search.

SKILLS & ACTIVITIES

- **Skills:** Database Modeling, Data Cleaning, ETL Development, Product Analytics, & Project Management
- **Tools & Software:** Python, SQL, Tableau, Git, Jira, Scrum, Pandas, NumPy, Seaborn & Matplotlib
- Member of Cornell Big Red Raas, a competitive traditional Indian dance team. Performed and coached the team to a top 5 placing nationally in 2022.
- Captain of the WW-P High School North Varsity Men's Tennis team in 2017 and 2018. Led the team to back-to-back top 10 NJ state rankings.