

# NANDITHA SHANKAR

[nandithashankar98@gmail.com](mailto:nandithashankar98@gmail.com) | 862-810-0233 | [linkedin.com/in/nanditha-shankar](https://www.linkedin.com/in/nanditha-shankar) | Harrison

## EDUCATION

**New Jersey Institute of Technology, Newark, USA**

Sep 2021 – Dec 2022(Expected)

Master of Science in Data Science and certified in Data Mining

GPA: 4.00/4.00(21 credits)

Courses: Operation System, Probability and Statistics, Data Management and Database Design, Data structures and Algorithms, Data Mining, Machine Learning, Data Integration, Data Visualization, Web Application Development.

**New Horizon College of Engineering, Bangalore, India**

Jun 2015 - May 2019

Bachelor of Technology in Electronics and Communication Engineering

GPA: 8.49/10.00

## PROFESSIONAL EXPERIENCE

**New Jersey Institute of Technology, Newark, USA**

Mar 2022 – Present

Data Visualization Programmer

- Performed a **data cleaning** and carried out **web scrapping** by writing a python script to fetch the required data.
- Formulating reports, developing SQL queries, and creating advanced **Tableau dashboards** for college high schools outreach projects to study the college trends regarding applications, admissions, and enrollment.
- Accomplishing all tasks related to advance analysis actions, calculations, parameters, background images, maps, trend lines, statistics, groups, hierarchies, and data sets to build summary levels and detailed reports.

**EXL Services, Bangalore, India**

Aug 2019 – Aug 2021

Reporting Analyst

- Automated ETL** processes make it easier to wrangle data and **reduced** time by **40%**.
- Increased efficiency** of data fetching by approximately **30%** utilizing query optimization and indexing.
- Constructed data-rich, actionable and interactive dashboards and operated statistical modeling with python to predict High/Low Potential claims to minimize manual intervention and save operation resources.
- Saved 30% cost** of operation resources by stopping manual work on low/non-potential claims.
- Worked closely with the Agile team on weekly meetings to constantly enhance features and functionality.

**BEML Limited, Bangalore, India**

Feb 2019 – Apr 2019

Intern

- Pre-processed passenger data of Bangalore metro for the year 2018 and built **scd-type2** mapping to maintain historic data of passengers using **SQL Server**.
- Generated advanced Tableau dashboards with quick/context/global filters, parameters, and calculated fields that allowed to track and improve customer units' **KPI by 12%** within a month.

**New Horizon College of Engineering, Bangalore, India**

Jan 2017 – May 2019

President, Technology sharing Club

- Trained students' concepts of **Data Visualization**; explained data wrangling techniques and dashboard creation in Python.
- Conducted seminars** and organized **sessions** in batches of 20 students, to train and strengthen their skills through hands-on project-based learning by **coordinating** with the other club members.

## TECHNICAL SKILLS

**Programming Languages:** Python, SQL, C, HTML, MATLAB, SAS, R

**Libraries & Frameworks:** Pandas, NumPy, Django, Seaborn, Scikit-Learn

**Cloud Technologies:** AWS Resources (EC2, S3, Lambda)

**Databases:** SQL Server, Oracle, IBM DB2

**Software:** SSIS, SSRS, SSMS, MS Excel, GitHub

**Tools:** Talend, Informatica, Tableau, Power BI

## PROJECTS

**Interactive Visualization Dashboard for GitHub Analysis (SQL, Google Cloud Platform, Tableau)**

Apr 2020 - Aug 2020

- Extracted aggregate tables using optimized SQL queries using open-sourced 3TB of GitHub data on Google Big Query database
- Architected multi-tab dashboard with dropdowns and sliders for interactive visualization in Tableau to enhance UI
- Designed Docker image by containerizing dashboard code, deployed it on Kubernetes in GCP, and exposed IP address

**Hospital Database Management System (RStudio, MS SQL Server, Power BI, Mockaroo)**

Jan 2020 - Mar 2020

- Architected ERD and database in star schema for hospitals, achieved normalization between entities to solve redundancy issues
- Generated data of 50,000 rows in R and Mockaroo and performed DML, DDL, TCL scripts; created SQL queries as views for easy accessibility, revised data processing efficiency
- Integrated SQL Server with Power BI and visualized data for analysis and further predictions

**Churn Classification using Machine Learning (Python, Scikit-Learn, Pandas, NumPy)**

Sep 2021- Dec 2021

- Performed exhaustive feature engineering and mitigated target class imbalance using SMOTE on IBM sample dataset
- Evaluated Performance of classification algorithms such as Logistic regression, Decision Trees, Random Forest, and XG Boost with Cohen's Kappa, AUC metrics, and ROC curves
- Achieved accuracy of 87% with XG Boost, cross-validated model using k-Fold and analyzed for best fit