

# Preet Modi

Linkedin: <https://www.linkedin.com/in/preetjmodi/>

Website: <https://impreetmodi.github.io/>

Email : [prmodi@iu.edu](mailto:prmodi@iu.edu)

Mobile : +1 8123182011

## EDUCATION

- **Indiana University Bloomington** Bloomington, IN  
• *Masters in Data Science; GPA: 3.82* Aug 2022 - May 2024  
• **Courses:** Advanced Database Concepts, Big Data Management, Introduction to Statistics, Algorithms, Computer Science, Data Mining, Cloud Computing, Machine Learning, Software Engineering, Predictive Analytics (**Kelley School of Business**)
- **Dharmsinh Desai University** Gujarat, India  
• *Bachelor of Information Technology ; GPA: 3.7 (8.29/10.0 )* Aug 2018 - May 2022

## SKILLS SUMMARY

- **Languages:** Python, R, SQL, Java, C, C++, HTML, CSS, JavaScript, C#, Linux, ReactJS, Node JS
- **Database & Tools:** SQL Server, PostgreSQL, Hive, MongoDB, Tableau, PowerBI, Airflow, Kafka, SAP, SAS, TensorFlow, Keras, Excel, VS Code, AWS, GCP, PySpark, Databricks, Snowflake, Github, Azure, EC2, MATLAB
- **Data Science:** ETL, BI, Predictive Modeling, Regression, Classification Trees, Time Series Analysis, Data Warehousing, Natural Language Processing, Hypothesis Testing, Artificial Intelligence, Statistical Analysis, Data Management

## EXPERIENCE

- **Graduate Research & Teaching Assistant** August 2023 - Current  
• *Indiana University* Bloomington, IN
  - **SAS, SQL, MS Power Tools, Data Visualization, Carnegie Classification:** Collaborating with Dr.Victor Borden, I engaged in data metric analysis, processing datasets of over 1 million records, and developed 10 novel interactive visualizations for Carnegie Classification.
  - **Python, Flask, Kubernetes, NoSQL, Cloud Computing, Big Query, Distributed Computing, Docker:** As a Graduate Teaching Assistant for INFO-I 535 Management, Access, And Use of Big And Complex Data, I crafted 13 assignments, graded for 80 students, achieving 95% completion and boosting student engagement by 30%.
- **Data Science Intern** May 2023 - August 2023  
• *Sacoma Specialty Products,LLC* Edinburgh, IN
  - **SQL, Epicor, SAP, Excel:** Integrated Epicor and SAP systems with AWS services, utilizing custom Business Activity Queries (BAQs), resulting in a 20% improvement in supply chain efficiency.
  - **Amazon Redshift, QuickSight, HPC:** Created a centralized data lake, streamlining data extraction, transformation, and loading (ETL) processes, reducing data processing time by 40% and enhancing data accessibility.
- **Data Analyst** Oct 2022 - May 2023  
• *Indiana University* Bloomington, IN
  - **Power BI, Advanced Excel, Dash, R, Tableau & DAX:** Worked with Residential Program and Services to conduct financial data analysis for housing and dining facilities, utilizing DAX (Data Analysis Expressions) for tailored computations, leading to a 15% decrease in operational expenses and a 20% uptick in revenue.

## ACADEMIC PROJECTS

- **HPC Analytics Dashboard Application Development:** Engineered an application for High Performance Computing (HPC) with a theoretical peak performance of 1 Petaflop (PF) and an infrastructure comprising over 10,000 CPU cores and 44 GPU cards. Integrated **Oracle RDMS** for data storage and management. Leveraged **Dash, Matplotlib, and Seaborn Python** libraries and **ReactJS** for data visualization. This project entailed implementing Software Development Life Cycle methods such as **Agile** and **Scrum**, utilizing **Jira** for project management. Harnessing problem-solving, communication abilities, and collaborative teamwork, the analysis delivered actionable business insights.(Nov 2023)
- **Exploratory Data Analysis for Bureau of Transportation Statistics Flight Performance:** Implemented a robust **data pipeline** using **Python** and **ETL tools i.e Alteryx**, devised a storage model in a **MySQL** server, and executed an algorithm using a parallel programming framework, **Hadoop**, resulting in a 30% reduction in data processing time. Conducted statistical analysis using **R** and **Stata**. Proposed and implemented a cleaning improvement solution, explored a big data cloud platform environment, and developed a reliable data management plan. Leveraged the **K-Means Clustering** algorithm to enhance data analysis accuracy, achieving an impressive **90%** precision rate.(May 2023)
- **Epicor-Driven Data Enlightenment: Transforming Business Intelligence:** Implemented a data-driven dashboard to enhance operational efficiency and decision-making. Integrated data from diverse sources and conducted analyses to extract insights, refining data manipulation skills. Utilized **Amazon Redshift** for data pipelines and **Excel Macros** for streamlined processing. Developed interactive dashboards using **Power BI, Tableau & Quicksight**, and in ERP systems like **Epicor** and **SAP**. Provided stakeholders with real-time visibility into key indicators, facilitating informed decisions. Wrote **Business Activity Queries (BAQ)** in Epicor and generated **MRP** dashboards, achieving a **40%** increase in efficiency.(Aug 2023)

## PUBLICATIONS

- **“Insurance Management with Premium Prediction ”, Volume 9, Issue XII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 1222-1238, ISSN: 2321-9653 (Impact Factor: 7.429): - DOI: <https://doi.org/10.22214/ijraset.2021.39416>**
- **”An efficient Artificial Neural Network for Coronary Heart Disease Prediction ”, Volume 9, Issue XII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 1474-1483, ISSN: 2321-9653 (Impact Factor: 7.429): - DOI: <https://doi.org/10.22214/ijraset.2021.39559>**