**Himabindu Pati**

Chapel Hill, NC

Phone: *919-480-1434* | Email: *patibindu@gmail.com*

LinkedIn: <https://www.linkedin.com/in/bindu-pati/>

GitHub: <https://github.com/patibindu>

* Ambitious **Data Analyst** with a passion for condensing large data into a simple data to understand story.
* Versatile **team player** with proven success in customer service and cross-team collaboration.
* Multidisciplinary, driven Data Analyst with a passion for condensing large data into a simple to understand story.
* Adept at working independently and collaborating with teams across multiple functions to take information from a wide variety of sources and effectively break down concepts in layman terms to diverse audiences. Excels at solving complex problems.

**Technical Skills**

**Tools: Excel, Tableau, Hive, Jira Software, GitHub, Jupyter Notebook, Matplotlib, Scikit-Learn, NumPy, TensorFlow, Pandas, Flask**

**Databases: SQL Server, PostgreSQL, SQLite, Oracle, MongoDB**

**Languages: Python, HTML, JavaScript, SQL, NoSQL**

**Projects**

**Housing Market** | <https://github.com/patibindu> | | <https://github.com/patibindu/Project1> **|**

* *It’s a group project, to showcase how the housing market is analyzed based on the macroeconomic factors we have taken from Kaggle datasets.*
* *By using the Datasets of Monthly, Annual, Housing macroeconomic factors:*
  1. *We checked if there is a correlation between house prices and mortgage rate- for that we created a scatter plot, calculated the R-value, and developed a linear regression line.*
  2. *We checked how the factors are associated with each other by verifying the Unemployment Rate, Real Disposable Income, House Price Index, Mortgage Rate, Inflation, Consumer Price Index and GDP, to achieve that we created multiple graphs.*
  3. *Finally, to compare how household income is related to the macroeconomic factors, While the dataset did include disposable income as a macroeconomic factor, household income tends to be a more accurate indicator of purchasing power, especially when it comes to housing. The dataset did not contain information on household income, but we were able to obtain it from the US Census Bureau’s website.*
* ***Responsibilities: To check the relation between House Price Index and Mortgage Rate and help with the presentation.***
* ***Tools Used: Pandas, NumPy, SciPy***

**Crowdfunding ETL Mini Project: |** [**https://github.com/patibindu/Crowdfunding\_ETL**](https://github.com/patibindu/Crowdfunding_ETL) **|**

* *In this group project, an ETL pipeline was created. We used Python, Pandas and regular expressions while working with the data. We first cleaned the data to improve the readability and later transformed the data into four dataframes (Category/Subcategory, Campaign, Contacts and Crowdfunding). These dataframes were then converted into CSV files. A Postgres database and an entity relationship diagram was created and finally the CSV files were loaded.*
* *Responsibilities: Converted the raw data that’s in excel to Dataframes using Pandas. Created an ERD diagram using PostgreSQL.*
* *Tools/languages Used: Python, PostgreSQL, Pandas, NumPy.*

**State Of American Universities:** | <https://github.com/patibindu/Project3-Group3/tree/main> **|**

* *In this group project, we focused on describing the state of American Universities and their student profiles in 2013. We used Kaggle to develop a map using Leaflet and a dashboard using JavaScript/HTML. The final map displays the US map with pins locating different Universities. If you select each pin, the corresponding University name, Graduation Rate and Total enrollment will populate. We developed a dashboard including a dropdown menu to select specific University, which results in a metadata panel which displays the graduation rates for attaining Bachelors’ Degree in 4, 5 and 6 years.*
* *Responsibilities: Converted the excel data into geoJSON data, helped in creating dashboard, Created Mongo Database and developed the Flask App.*
* *Tools/languages Used: HTML, CSS, JavaScript, Flask.*

**Data Visualization and Machine Learning on Diabetes Dataset: |** [**https://github.com/patibindu/Project4-Group3**](https://github.com/patibindu/Project4-Group3) **|**

* *This final group project is focused on the data visualization and machine learning to better understand diabetes prevalence and how diabetes relates to demographic information and medical history. Findings show that there is a positive correlation between HbA1c and blood glucose levels and the risk of diabetes. The Random Forest Algorithm demonstrated the highest level of accuracy compared to all other classification machine learning models. This random forest machine learning model presents a solid foundation to be built upon to help healthcare professionals easily access patients for risk of diabetes.*
* *Responsibilities: Cleaned the data, created SQLite database, addressed the imbalanced data using Machine Learning, trained and tested the machine learning model.*
* *Tools/languages Used: Python, Tableau, SQLite,*

**Professional Experience**

**Data Analyst 2022 – Present**

**NTT Data** Plano, TX

**Roles and Responsibilities:**

1. I have worked on data mapping, analysis and etl work in the project.
2. Data mapping: We get a set of files which are to be loaded into the system, I have worked to map the incoming data into the existing columns and sometimes new fields are added into the file and for that I alter the table in the database. I have done the data mapping from scratch to a new table and to existing tables by adding a single/multiple columns.
3. Loaded the data using ETL tool into database.
4. Once the development is done, I test the data using some sample data.
5. Once the testing is done, later we push the changes to UAT and do validations on data.
6. A few Validations would include the total amount matches between the source and the data loaded in the tables.

**Education**

***Boot Camp Certificate: University Of North Carolina****, Chapel Hill, NC**[Data Analytics & Visualization]*

A 24-week intensive program focused on gaining technical programming skills in Excel, VBA, Python, R, JavaScript, SQL Databases, Tableau, Big Data, and Machine Learning.

***Degree or Certification****: Bachelor of Technology, KLCE, Nagarjuna University, INDIA*