

DARSHANKUMAR PATEL

Skills

- SQL (SQL Server, MySQL, SQLite)
- Python (Pandas, NumPy, SciPy, Matplotlib, BeautifulSoup, seaborn)
- Tableau
- Excel (VLookup, HLookup, Conditional Formatting, Pivot Tables & Charts)
- Machine Learning & Algorithms
- PySpark
- R (ggplot2, dplyr, tidyr, rematch, tidyverse)
- Microsoft Power BI

Projects

[COVID-19 DEATHS EDA](#) – [PERSONAL PROJECT](#) – TORONTO, ON

March 2024

- Utilized **Excel** to clean the raw dataset and delete unnecessary columns and filter null values.
- Extracted and interpreted vaccination and death data using **SQL** to provide actionable insights for public health interventions, leading to a 25% improvement in vaccination coverage and a 10% decrease in fatalities.
- Implemented **CTE Blocks, Temp Tables, queries, and aggregate functions** through the dataset and identified the key judgement.
- Aggregated and visualized the data by using **Tableau** to create dashboard.

[CUSTOMER SEGMENTATION USING MACHINE LEARNING](#) – [PERSONAL PROJECT](#) – BARDOLI, IN

March 2023

- Cleaned and preprocessed the dataset, successfully dealing with missing values and outliers using **Pandas & NumPy** libraries.
- Perform exploratory data analysis to identify patterns using **RFM (Recency, Frequency, and Monetary)** analysis and insights into consumer purchasing behavior using **Matplotlib, Seaborn, scikit-learn** Libraries.
- To cluster the customers and determine the optimal number of clusters, we applied **Clustering algorithms (AFM and K-Means)** and evaluated the clustering performance using silhouette score.
- The clusters were visualised in 3D space to aid understanding and interpretation.

[NETFLIX DATA ANALYSIS](#) – [PERSONAL PROJECT](#) – TORONTO, ON

July 2023

- Engineered **Python** scripts to automate data preprocessing tasks, significantly enhancing data quality and ensuring accurate insights for decision-making processes.
- Proficient in large-scale data processing and analysis using **Big Data** technologies such as **Hadoop, HDFS, Spark, and Zeppelin**.
- Data analysis and visualization using **R** Language using **dplyr, ggplot2, tidyverse** Libraries with strong background in statistical analysis and machine learning.
- Skilled in using **Tableau desktop** to create dynamic and insightful data visualizations.

[PREDICT INSURANCE PREMIUMS USING ML](#) – [ACADEMIC PROJECT](#)- TORONTO, ON

March 2023

- Utilize Python for data preprocessing and cleaning tasks to assure data integrity and quality.
- The team conducted exploratory data analysis (EDA) by utilizing visualizations to better comprehend patterns and linkages in the data.
- To estimate insurance rates for new consumers, we used machine learning algorithms such as Linear Regression and Random Forest Regression.
- Used pandas, seaborn, and scikit-learn libraries for data manipulation, visualization, and modelling.
- We employed cross-validation techniques to assess model performance and ensure generalizability.
- Interpreted model results and offered stakeholders with insights to help them make educated decisions.

Work Experience

SOFTWARE DEVELOPER (C# WINDOWS FORMS, SQL SERVER, TCP) CO-OP – IIHT – Navsari, IN May 2016 – July 2016

- Created and developed two software systems: a Library administration System and a Medical Store System, using C# Windows Forms, SQL Server for database administration, and TCP/IP for communication protocols.
- Created user-friendly interfaces for efficient management of library resources and medical inventories.
- Enforced strong database systems for tracking books, borrowers, medical supplies, and sales transactions.
- For smooth communication between client and server applications, integrated TCP/IP communication.
- Collaborated with team members to analyze requirements, develop system architecture, and implement solutions.

WORDPRESS DEVELOPER CO-OP – ASL Consulting – Etobicoke, ON May 2023 – August 2023

- Crafting a visually captivating and user-centric website. Collaborated with designers to ensure responsive design, while integrating interactive forms, data visualization, and secure authentication to enrich user experience. This facilitated efficient data collection, empowering users to monitor their journey and make informed choices.

Education

GRADUATE CERTIFICATE IN INFORMATION TECHNOLOGY SOLUTIONS – Humber College – Toronto, ON August 2023

Majors: Business Intelligence, Data Analytics, Machine Learning, Deep Learning, SQL, Big Data.

Dean's List Every Semester

BACHELOR DEGREE IN COMPUTER SCIENCE & ENGINEERING – R.N.G.P.I.T– Bardoli, IN August 2020

Majors: Numerical & Statistical Method of Computer Engineering, Analysis & Design of Algorithms, Theory of Computation, Data Mining & Business Intelligence, Artificial Intelligence, Python.

[Certificates](#)

DATA ANALYTICS BOOTCAMP – ALEXANDER FREBERG – MAY 2024

DATA ANALYST 101: EXCEL FORMULAS & FUNCTIONS –SIMPLILEARN– MARCH 2024

SQL ESSENTIAL TRAINING – WALTER SHIELDS – SEPTEMBER 2023

SQL BASIC & INTERMEDIATE – HACKERRANK – MARCH 2023

PYTHON BASICS – UNIVERSITY OF MICHIGAN – AUGUST 2020