

Jibin Madayil John

(647) 868-3741 | Toronto Canada

jibinmadayil/github.io | <https://www.linkedin.com/in/jibinmadayil/> | jibinmadayil@gmail.com

SUMMARY

Detail-oriented and results-driven graduate with a robust foundation in data analytics. Eager to leverage expertise in transforming raw data into actionable insights to drive organizational success. Adept at employing advanced statistical techniques and machine learning algorithms. Seeking entry-level opportunities to contribute technical acumen, collaborate with seasoned professionals and deliver impactful solutions.

EDUCATION

Post-Graduate Diploma in Big Data Analytics

GPA: 3.57 / 4.0

Lambton College

April 2024

Coursework: Big Data Tools, NoSQL Database, Python Programming, Big Data Algorithm & Statistics, Application Design for Big Data, Big Data Visualization, Advanced Python Artificial and Machine Learning Tools

Bachelor's in Mathematics, Statistics and Computer Science

GPA: 3.48 / 4.0

Osmania University, India

August 2021

SKILLS

Programming Languages: Python

Database Technologies: SQL (SQL Server, MySQL), NoSQL (MongoDB)

Big Data Technologies: Hadoop, Hive

Data Visualization Tools: AWS Quicksight, Microsoft PowerBI, tableau, Microsoft Excel

Version Control: Git

PROJECTS

Breast Cancer Prediction

- Developed a machine learning model to classify breast cancer tumors as malignant or benign based on features extracted from biopsy images
- Employed logistic regression, decision tree, and random forest algorithms to train and evaluate the model
- Achieved accuracy rates of 95.61% with logistic regression, 92.11% with decision tree, and 94.74% with random forest on the test set
- Leveraged Python libraries including pandas, scikit-learn, matplotlib, and seaborn for data preprocessing, model building, evaluation, and visualization

Auto Theft Analysis

- Conducted comprehensive data cleaning and preprocessing, reducing duplicates by 9% utilizing Python libraries such as pandas, seaborn, matplotlib for efficient data manipulation and quality assurance
- Analyzed trends in theft incidents, revealing a 220.15% increase from 2018 to 2023
- Visualized theft incident distribution by police division, identifying hotspots and areas of concern
- Generated interactive heatmaps to visualize theft incident density across neighborhoods

Covid Data Analysis

- Applied SQL expertise to analyze COVID-19 data, generating vital insights such as a global death percentage of 2.11% out of 150,574,977 reported cases
- Developed views and temporary tables to streamline data analysis processes, facilitating efficient calculation of infection rates, death percentages, and population vaccination rates
- Visualized data trends using Tableau, presenting continental death counts and forecast indicators for key countries like China, India, Mexico, United Kingdom, and United States
- Enhanced database performance and ensured data integrity through SQL query optimization and thorough database maintenance tasks

Amazon Sales Dashboard

- Created an interactive Amazon Dashboard in Power BI highlighting vital metrics and data visualizations
- Highlighted key sales metrics including \$89.08M in revenue, 120K units sold, 2M returns, and 4M reviews
- Presented sales trends through a 'Units Sold by Month' graph, featuring figures like 35K for March and 46K for April
- Enabled tooltips for deeper data comprehension and analysis, allowing users to hover over data points and gain additional insights into sales metrics, trends, and patterns

PUBLICATIONS

Deciphering Data Storage Solutions	April 2024
The Significance of Data Cleaning	February 2024
The Art and Science of Data Visualization	February 2024
Decoding the Null	January 2024