Vivek Pawar

+91 7414946348 Ravet Pune 412101, Maharashtra - vivekpawar51@gmail.com -
https://www.linkedin.com/in/vivek-pawar-302965207 - https://github.com/Vivek-pawar1411

PROFESSIONAL SUMMARY

Dedicated to exploring emerging technologies and staying updated with the latest trends in data engineering, committed to continuous learning and applying innovative approaches. Proficient in Data Visualization, Python Programming, and Database Management (SQL, MongoDB), with a strong foundation in big data technologies. Thriving in collaborative environments, eager to build efficient data pipelines and drive data-driven solutions

EDUCATION

Institute For Advanced Computing and Software Development - CDAC — **Pune Maharashtra**
PG - Diploma in Big Data Analytics — *September 2023 - February, 2024*

Sinhgad College Of Engineering — **Pune Maharashtra**
BE Production Engineering — *July 2018 - August 2022*

Kendriya Vidyalaya Sarni — **Sarni Madhya Pradesh**
12th Science — *April 2017 - May 2018*

Kendriya Vidyalaya Sarni — **Sarni Madhya Pradesh**
10th — *April 2015 - May 2016*

SKILLS

Programming Languages: C, C++, Python
DataBase: SQL, MongoDb, Cassandra
Libraries and Tools: Apache Spark, Pandas,BeautifulSoup
Data Analytics: Big Data Technologies, ETL Processes
Data Visualisation: Tableau, PowerBI, Defining KPI, Creating Dashboard

PROJECTS

- **Predict the fare amount of future rides using regression analysis:-**(***Technology use in :Machine learning,Python, Pandas, Matplotlib***) Regression analysis facilitates the prediction of fare amounts for future rides by identifying patterns and relationships between factors such as distance traveled, time of day, and location, and the corresponding fare amounts. By analyzing past ride information, we train regression models to comprehend how these factors influence fare prices. Subsequently, the model actively predicts fare amounts for future rides by inputting relevant trip details. GitHub

- **IPL Data Analysis:-*(Technology use in :Data Visualisation Tableau,Python, Pandas***) The IPL (Indian Premier League) Analysis Dashboard is a web-based data visualization tool designed to provide insights into the Indian Premier League matches from 2008 to 2020. It includes details about the winning teams, Purple Cap holders (leading wicket-taker), and Orange Cap holders (leading run-scorer) for each season. GitHub

- **Flipkart Web Scraping:-*(Technology use in :BeautifulSoup, Python, Pandas***) Web scraping Flipkart using BeautifulSoup and Pandas entails extracting data from Flipkart's website, parsing it with BeautifulSoup, structuring it into Pandas dataframes, and then saving it as a CSV file. This process allows for easy manipulation, analysis, and storage of product information retrieved from Flipkart's web pages. GitHub

- **Spotify Data Engineering Project:-*(Technology use in :Amazon Web Service (AWS) ,Extract Transform and Load ,Python ,Pandas***) Preprocessed Spotify data with Pandas for accuracy, then deployed on AWS S3. Automated processing via Glue for seamless extraction and transformation. Established an S3 warehouse, linked with Glue Crawler for cataloging. Integrated with Athena for SQL-based querying, enabling streamlined analysis. This integration streamlined processes, empowering efficient decision-making. Leveraged Glue and Athena for scalable and flexible data processing. GitHub

- **Power BI Job Simulation: Client-Focused Data Insights :-** ***(Technology use in : PowerBI, Defining KPIS, Creating Dashboard, Calculating Measure***) This simulation strengthened my Power BI skills to deliver impactful client visualizations. I designed dashboards effectively conveying KPIs, responding to client requests. Clear email communication with partners ensured valuable insights and actionable suggestions based on data analysis, including identifying HR data trends and root causes for gender balance issues within leadership, demonstrating my commitment to data-driven solutions. GitHub

- **Design and development Road Spike System:-** The Road Spike System is a safety mechanism designed to control traffic flow by deploying retractable spikes on roads. It consists of durable mechanical components, electronic control systems, and safety features like motion sensors and manual overrides.The system enhances road safety and security by deterring unauthorized access and controlling vehicle movement effectively. GitHub

CERTIFICATION

Disha Computer Institute
Got certification for completing C,C++,Python course from Disha Computer institute

MentorMind
Has successfully completed the project Predict the fare amount of future rides using regression analysis.

Forage
Successfully completed Power BI Job Simulation and get certificate (Forage, June 13, 2024)

EXTRACURRICULAR ACTIVITIES

Sports Captain
Had selected for the Sports Captain of Subash House at School

Volunter
Contributed to the planning and execution of Techtonic 2020, a college tech event