YASH BHAIRAO

DATA CONSULTANT

📞 +918888683070

✉ bhairaoyash@gmail.com

📍 Pune, india

in [LinkedIn](#)

⌗ [GitHub](#)

PROFILE

Dedicated and detail-oriented Data Analyst with 3+ years of experience. My expertise lies in leveraging data to drive decision-making processes and foster organizational development. Whether it's through data visualization or predictive modeling using Machine Learning techniques, I thrive on extracting value from data to fuel business success.

WORK EXPERIENCE

FEB 2023 - PRESENT
Capco Technologies | Pune, India
Data Consultant

- Engineered data pipelines using PySpark to process large datasets efficiently for ESG Team.
- Automated workflows with Apache Airflow by creating DAGs, improving data extraction speed and MIS Reporting.
- Conducted ESG data analysis, providing insights that informed clients sustainability strategies as per UN norms.
- Co-developed and deployed a machine learning model to predict Entity's loan defaults with AI Team.
- Responsible for choosing the closest Regression Model to be used for the prediction.
- Handling a team of 5+, responsible for collecting Evidences through Research for different Generic codes given by client.

Mar 2022 - Feb 2023
Aimleap | Bengaluru, India
Data Analyst

- Developed various web scrapers and bots to extract data from multiple sources using Selenium and Scrapy.
- Automated the process of data collection be used every fortnight using Python script.
- Created visualizations in Power BI tailored to client specifications, enhancing data comprehension and decision-making for different products.
- Deployed code on AWS, ensuring scalability and reliability for data solutions.
- Collaborated closely with Data Science team to define data requirements and deliver actionable insights, aligning technical solutions according to the business requirements.

EDUCATION

Bachelor of Engineering
GHRCE, NAGPUR - 2021

- Electronics & Telecommunication
- CGPA : 8.03/10

Data Science Bootcamp
JOVIAN - 2020

- Python
- SQL
- Machine Learning

TECHNICAL SKILLS

- Data Analysis
- Big Data (PySpark)
- Airflow
- Machine Learning Algorithm
- ETL Tools
- Docker
- Power Bi
- Data Lake
- Databricks
- Selenium Automation
- SSMS

SOFT SKILLS

- Observation
- Decision-making
- Communication
- Multitasking

LANGUAGES

- Python
- SQL
- JavaScript

Aug 2021- Feb 2022
Electro-Tech Services Pvt.ltd l Nagpur, India
Data Analyst Intern

- Gathering data using SQL, Selenium, present data analysis reports to the lead.
- Loading the data and using ML methods to obtain predictions on employees holidays.
- Generating insightful dashboards using tableau for the company' employee data to understand the patterns.

PROJECTS (VIEW ALL)

Exploratory Data Analysis on Brazilian E-Commerce Olist (View)

- Analysis of 9 different CSV files containing an average of 50k rows and 5 columns using Pandas.
- Visualizations (scatter plots, folium charts, geo heatmaps, etc.) using Seaborn & Plotly.
- Uncovered insights e.g. e-commerce customers tend to shop most frequently at noon

Unlocking Titanic's Secrets: Predicting Survival with Logistic Regression using Pyspark (View)

- Utilized PySpark for distributed computing, handling largescale data processing efficiently.
- Implemented Logistic Regression for survival prediction based on Titanic dataset features.
- Constructed Pipelines in PySpark for streamlined workflow from data processing for evaluation.

Automation of Scraping Using Selenium with Apache Airflow (View)

- Managing dependencies between tasks ensures that tasks are executed in the correct order and handle data flow smoothly.
- Using Python code within the DAG definition, the project dynamically generates tasks to handle different scraping scenarios, such as varying URLs, data fields, and scraping intervals.
- Stores the scraped data for further analysis or processing in Postgres SQL local databse and visualizing it in Power Bi for dashboard creation.

CERTIFICATES

- Career Essentials in Data Analysis by Microsoft and LinkedIn (Link)
- Level Up : Advanced SQL (Link)
- SQL Essential Training (2019) (Link)
- Data Analysis with Python : Zero to Pandas (Link)
- Mastering Intermediate DAX - Power BI, Power Pivot & SSAS (Link)

AWARDS

- Winner of Regression Hackathon (2022)
- Winner of High Five Award (Client - 2023)
- Runner-up of Above & Beyond (Client - 2024)