Parag Badgujar

Data Analyst

+91-7840956506 badgujarparag1902@gmail.com linkedin.com/in/paragbadgujar

Summary

Data Analyst with a strong foundation in SQL, Python, Excel, and Power BI. Skilled in cleaning, analyzing, and visualizing data to generate insights and support data-driven decision-making. Interested in collaborating on meaningful analytical work in fast-paced environments.

Skills

SQL :	Joins, Subqueries, Window Functions
Python :	Pandas, NumPy, Matplotlib, Data Cleaning
Power BI :	DAX, Visualization
Excel :	XLOOKUP, VLOOKUP, HLOOKUP, Pivot Tables, Data Cleaning
Analytical Abilities	EDA, Insight Generation, Descriptive Statistics, Statistical Analysis

Experience

Data Analyst Intern **May 2024 – June 2024**

Tribal Development Department *Tools Used: Excel, Power BI*

- **Integrated over 15,000 beneficiary records, enhancing data consistency** and improving reporting **accuracy by 25%**
- Cleaned, transformed, and validated 15,000+ data rows using Excel and Python to ensure smooth and reliable analysis.
- Collaborated with 3 senior officers and non-technical team members to present key findings and visual insights, supporting policy-level decision-making.
  Conducted an extensive statistical analysis of the IPL using various data **visualization and statistical techniques**.

Projects

Phishing URL Detection | *Python, Machine Learning* **Jan 2025 – May 2025**

- **Analyzed over 11,000 website URLs** to identify patterns in phishing threats.
- Created and selected 20+ useful features to improve model performance.
- Used Recursive Feature Elimination (RFE) to reduce features by 35%, improving model efficiency and interpretability.
- Built a phishing detection model using Random Forest, that correctly identified harmful websites with 85.1% accuracy.

Statistical Analysis on IPL | *Python, Minitab, Microsoft Excel, Microsoft Word* **June 2024 – Dec 2024**

- Conducted an extensive statistical analysis of the IPL using various data **visualization and statistical techniques**.
- Utilized **Chi-Square** Test for analyzing categorical data and testing relationships between variables.
- Created **graphical representations** to illustrate findings and support data-driven conclusions.

Certifications

- **Data Analytics with Python – NPTEL :** Completed 3 practical case studies involving Python, data cleaning, dashboards, and presentations.
- **Machine Learning Engineer Certification – Symbiosis Institute of Technology, Pune**

Education

M.Sc. Statistics **Graduated: 2025**

KBC North Maharashtra University, Jalgaon

B.Sc. Statistics **Graduated: 2023**

KBC North Maharashtra University, Jalgaon