

Arpit Gupta

Data Enthusiast

Gurugram, Haryana
+91 7983047963
arpittguptta@outlook.com



ABOUT

Analyst with a strong foundation in data analysis, statistical techniques, and data visualization. Adept at extracting meaningful insights from data to support informed decision-making. Eager to apply academic knowledge and contribute to data-driven solutions in a professional setting. Proficient in tools such as **Python, R, SQL**, and data visualization -**Tableau, Power BI**. Familiarity with cloud-based data storage and processing - **AWS, Azure, Google Cloud**.

KEY SKILLS: Python, SQL, Excel, Tableau, Alteryx, Statistics :: Data Analysis, Relational databases, Dashboards, Data Visualization

PROFESSIONAL EXPERIENCE

Amdocs Development Center, India — Software Engineer Associate

September 2022 - PRESENT

- Conducted **Statistical /Defect /Root Cause - Analysis** to extract valuable insights from complex datasets, driving data-informed decision-making within the organization using Python, SQL, Linux & other analysis tools and strategies.
- Lead the team in billing operations to deliver high-quality products. Delivered 13+ automation to accelerate business processes.
- Assessed large volumes of structured and unstructured data to identify trends, patterns, and correlations. Interpreted findings to provide actionable insights, facilitating data-driven decision-making within the organization.
- Collaborated with cross-functional teams to develop and implement data-driven strategies, resulting in a measurable improvement in operational efficiency and cost reduction.

Highradius Corporation — Full Stack Intern

June 2021 – May 2022

- Delivered 7+ applications, developing multiple modules as per Customer requirements.
- Worked on machine learning models leveraging various modelling techniques such as linear regression, random forest etc.
- Unit testing and integration testing of owned use cases to improve quality and deliver quality work product.
- Estimation, design, and development of modules including Regression ML models.

PROJECTS

Telefonica Argentina — Amdocs

- Demonstrated time reduction of 18% of billing operations by leveraging data-driven insights to find and fix defected data while upholding data quality.
- Demonstrated Expertise, by delivering 13+ in automations of manual tasks through development and implementation of efficient Shell Scripts.
- Conducted rigorous data quality assessments, data cleansing, and data validation to ensure the accuracy and reliability of information used for critical business operations.

Highradius Invoicing —Highradius Corporation

- Accomplished successful delivery of 7+ applications, exhibiting expertise in designing, estimating, and developing multiple modules.
- Developed a Prediction model with an accuracy of 96% - Payment Date while using machine learning algorithms such as Random Forest, enabling better insights in client’s payment behavior.
- Engaged in continuous maintenance activities, ensuring smooth functionality, and enhancing the overall user experience.

Covid Exploratory Analysis — Portfolio

- Executed **complete data** workflows involving **collection, cleaning, analysis, and visualization** with utilization of Kaggle for datasets, SQL for management, and Tableau for presentation and visualization.
- Leveraged SQL queries to analyze and extract COVID-19 data, generating crucial metrics on infection rates, fatalities, and vaccination trends for insightful regional insights.
- Computed detailed statistics on total cases, fatality ratios, and population-affected percentages, providing comprehensive analytics on the pandemic's geographical impact.

Nashville House Pricing — Portfolio

- Performed comprehensive data workflows, covering collection, cleaning, analysis, and visualization using Kaggle for datasets, SQL for management, and Tableau for visual representation.
- Analyzed housing data trends concerning factors like bedrooms, age of structures, and occupancy status, revealing patterns in average selling prices.
- Explored price variations over different years, offering insights into the market fluctuations and housing sales dynamics.

- Analyzed a movie industry dataset from Kaggle, using Python libraries like Pandas, Matplotlib, NumPy, and Seaborn for data cleaning and exploration.
- Discovered a strong correlation between budget and gross earnings, focusing on the relationship's impact on a movie's success.
- Identified top-performing movies based on their profit percentages concerning the budget, visualizing these relationships through various plots.
- Unpacked the correlation matrix to reveal highly correlated attributes and highlighted movies with the highest profit percentages through bar plots.

EDUCATION

Bachelor of Engineering (Hons. Cloud Computing) | CGPA: 7.1
Chandigarh University, Mohali
July 2018 - June 2022

Higher Secondary Certificate
Khalsa Public School, Saharanpur
April 2018 - June 2018

CERTIFICATIONS

- Google Data Analytics
- AWS Cloud Technical Essential
- The Joy of Computing Using Python
- Mastery Cloud Application developer
- Alteryx 2023: The complete Alteryx course for all levels

SKILLS

Python, SQL, Excel, Tableau, Pandas, R, Alteryx,
PowerBI, Data visualization, Statistical analysis,
Data warehousing, Database management, Dashboard,
Business intelligence, PL/SQL, ETL, Data Warehouse, Data Modelling,
Google Cloud Platform (GCP), Jupyter Notebook, Google Big Query, Data cleaning,
A/B testing, Git/GitHub, C/C++, Ext Js, Linux, HTML/CSS,
JavaScript

AWARDS

- Medal, District Level Current Affairs
- Bronze, Inter-department Football Tournament.
- Perform Foundational Data, ML, and AI Tasks: Google Cloud.

INTERESTS

- Photography
- Problem Solving
- Fitness
- Travel

LANGUAGES

- English
- Hindi