Akash Singh

Data Analyst

Experienced IT professional with 8.8 years of proven expertise as a Data analyst. Skilled in leveraging technical understanding and analytical prowess to drive business growth and enhance operational efficiencies.

✉ aks.1313769@gmail.com   📱 +61-432622588   📍 Australia

SKILLS

Data Cleansing & Preprocessing   R Scripting (dplyr, tidyr, stringr, data.table)

Data Profiling & Anomaly Detection Automation using R   Data Quality Management

SQL & Relational Databases Data Wrangling & Transformation   Data Mining / Data Modeling   Data Visualization

Business Intelligence   Time Series Forecasting   ETL (Extract, Transform, Load)   DataStage and Banking

Data Warehousing   Machine Learning   SDLC/STLC   Agile Methodology   EDA   Data Mapping / Data Cleaning

TOOLS

R Studio   Python   PySpark   MySQL/ MongoDB/ PostgreSQL   DBT   Hadoop   Scikit-learn

Jupyter Notebook   PowerBI   Google Cloud Platform (GCP)   Oracle SQL Developer   SAS

JIRA/BitBucket/ Confluence   ArcsGIS   ServiceNow   Excel   Microsoft Visual Studio   Splunk / AppDynamics

BigQuery

WORK EXPERIENCE

Sr. Consultant
EXL Analytics

01/2024 - Present
Clients : IAG Australia, Beyond Group, HSBC
Tasks

- Developed **automated Python scripts** to streamline ETL processes and improve data pipeline efficiency. Designed and developed **data warehousing** solutions to meet client requirements and support business objectives.
- Designed and implemented **Unix shell scripts** for data extraction, transformation, and job scheduling.
- Analyzed data to identify **trends, patterns, and insights** , providing valuable information to support decision-making processes.
- Identified and resolved **data gaps** that impacted the fulfillment of the business's functional requirements, while also providing innovative analytical insights for key projects.
- Created and maintained **data models** to support analytical projects. Performed **data transformations** and integrated data from Google Cloud Platform and DBT sources for comprehensive analysis. Integrated data from multiple sources into **BigQuery,** ensuring data accuracy and consistency.
- Cleansed and preprocessed data to ensure accuracy and reliability. Performed **exploratory data analysis (EDA)** to identify trends, patterns, and **outliers**. Utilized statistical techniques and **data visualization tools** to effectively present findings.
- Led the migration of legacy SQL-based systems to **PySpark-based distributed frameworks** , enabling scalability for a high-volume transactional dataset.
- Designed and automated data cleansing workflows using **R scripts** (dplyr, tidyr, stringr) to ensure high-quality data pipelines. Performed **data profiling, anomaly detection, and validation** on complex GCP/DBT sources datasets.
- Designed **data validation pipelines** using **PySpark** to ensure data quality and integrity during migration processes.

WORK EXPERIENCE

Sr. Analyst
Mobile Programming

11/2023 - 01/2024
Client : Poonawalla Fincorp Bank (Worked as a Data Analyst for the client)
Tasks

- Developed **Python-based data wrangling scripts** to preprocess and clean financial datasets for loan evaluations.
- Developed and implemented **Extract Transform Load (ETL) processes** and Business Intelligence solutions to enable data-driven decision-making. Designed an **automated log analysis tool** using Python and Unix shell scripting to improve debugging and error tracking.
- Developed an in-house application service for analytics on **Bank Statement transactions**, helping to streamline and accelerate the **loan evaluation process**.
- Conducted **database analysis** and design services to **optimize data structures** and improve data quality. Created m**odular and reusable R scripts** for ETL pipelines, boosting automation and reusability across teams.
- Worked with the business to define and create actionable **insights-to-action** (**data storytelling**) reports, dashboards, and general data analysis across Finance.
- Performed **data analysis** to understand loan performance metrics, customer behavior, and trends, enabling **data-driven decision-making** for loan lending strategies.
- Built **automated data transformation and cleansing logic** in **R** for bank transaction evaluations.
- Worked within an **Agile** environment, building products iteratively, and further supported the business by executing change management activities regarding data analysis and data products across key stakeholders.
- Automated **data aggregation pipelines** using PySpark for business intelligence reports, delivering critical insights in near-real-time.
- Conducted **performance tuning** for Spark jobs by optimizing partitioning, caching, and shuffling strategies to achieve significant efficiency gains.

System Engineer
Tata Consultancy Services

09/2016 - 03/2023
Clients : Fiat Chrysler Automobile, Barclays, Mitsubishi UFG Financial Group, Morgan Stanley, Westpac
Tasks

- Worked in the **BFSI domain,** spearheading requirement-gathering sessions with stakeholders to capture business needs and objectives, ensuring a clear understanding of client expectations.
- Played a pivotal role in **data integration projects**, ensuring seamless synchronization and integrity of data across various systems. Conducted thorough data analysis to identify gaps, trends, and opportunities, facilitating data-driven decision-making processes.
- Managed and executed the migration of business processes and data to **cloud platforms**, enhancing operational efficiency and scalability.
- Led **data mapping** exercises, identifying source and target data structures for various integration projects, optimizing data flows, and minimizing data discrepancies. Defined data transformation rules and processes, facilitating smooth **data migrations** and integrations across systems.
- Developed **Python-based automation scripts** to handle data migrations and transformations. Managed **Unix-based servers** for executing batch processes and job automation. Worked on **BFSI domain applications**, implementing Python scripts for financial risk modeling and fraud detection.
- Performed **data validation** and **reconciliation assessments** to ensure data precision and uniformity across various systems and interfaces. Utilized testing tools and frameworks to optimize testing procedures and expand test scope.
- Conducted detailed **data analysis** on data used across the business to evaluate business processes and improve or create new features, solutions, tools, and capabilities to enable self-service frameworks for data consumers in the analytics engine. Provided training and support to team members and stakeholders on **BigQuery** best practices.
- Ensured **data quality** and integrity during data extraction, transformation, and loading (**ETL**) processes. Demonstrated a strong understanding of **Agile principles** and methodologies, and maintained a strong team-player mindset with excellent communication skills.
- Provided technical expertise within **PowerBI** and **SQL** applications, including **Application Administration**, **Technical Design**, **Configuration**, **Data Migration**, **Database Analysis**, and **Database Modelling** interfaces.

WORK EXPERIENCE

Data Scientist
Coles Group (Internship)

Tasks

- Developed **Python-based customer segmentation models**, enhancing marketing campaigns and customer engagement.
- Gathered relevant data from various sources, such as **customer transaction history** , **demographic information** , and **interaction logs** , and prepared the data for analysis by cleaning, transforming, and organizing it.
- Conducted **exploratory data analysis** to gain insights into **customer behavior patterns**, identifying key factors that contribute to **customer churn** and discovering potential correlations and trends within the data.
- Built predictive models, such as **logistic regression**, **decision trees**, **random forests**, or more advanced **machine learning techniques** like gradient boosting or neural networks, to forecast and identify customers who are likely to churn. This involved training and validating the models using historical data.
- Assessed the performance of the developed models using appropriate evaluation metrics (e.g., **accuracy, precision-recall, F1-score**) and iterated on the models to improve their predictive power and reduce false positives/negatives.
- Utilized clustering algorithms or other segmentation techniques to group customers based on their behavior and other relevant characteristics. This provided insights into different customer segments and helped tailor retention strategies accordingly.

EDUCATION

Masters. of Science: Data Analytics
Royal Melbourne Institute of Technology (RMIT)

03/2021 - 12/2022   *Melbourne, Australia*

Bachelors. of Science: Information Technology
Pillai College (PCACS)

03/2013 - 04/2016   *Mumbai, India*

CERTIFICATES

Oracle Cloud Infrastructure 2020 Certified Architect Associate (Oracle)

Certified Associate: Generative AI LLMs (NVIDIA)

Microsoft Power BI Desktop for Business Intelligence (Udemy)

Azure AI Fundamentals (Microsoft)

HONOR AWARDS

Mumbai Univeristy (03/2013 - 04/2016)
Pillai's Institute of Technology (Maharashtra)

- I was recognized and honoured by **Mumbai University** for my outstanding performance in my Graduation stream.

Federation Games of India (06/2008 - 01/2019)
YMCA

- Represented for **INDIA: Sports Federation of India Basketball**.