

## SUMMARY

Impact-driven Data analyst with **3+ years of experience** transforming complex datasets into actionable insights that drive smarter decisions. Skilled in uncovering patterns, optimizing workflows, and crafting solutions that enhance efficiency, mitigate risks, and foster innovation.

## PROFESSIONAL EXPERIENCE

### Illinois Department of Human Services

Illinois, United States

#### Database Analyst

Sept 2024 - Present

- Validated large-scale prescriber datasets (**2M+ records**), including licenses and DEA numbers, using advanced ETL processes with SQL and Python (Pandas) to maintain data integrity and regulatory compliance.
- Resolved data quality and user access issues by executing targeted SQL queries, **addressing 40+ access** and data discrepancy issues monthly, resulting in a **95% reduction in discrepancies** and improved overall data security.
- Delivered actionable insights Tableau dashboards (trend analyses, geographic heatmaps, prescriber performance metrics), enhancing **stakeholder engagement by 30%** and directly supporting informed public health decisions to address substance use disorder trends.

### University Of Illinois System

Illinois, United States

#### Graduate Research Assistant ( CSC 532 )

Nov 2023 - May 2024

- Developed a personalized show recommendation system using a hybrid model combining collaborative and content-based filtering, addressing limitations in recommendation accuracy, and improving prediction accuracy by 5%.
- Enhanced dataset quality by leveraging **Large Language Models (LLM) and Open AI APIs** to generate detailed show descriptions from minimal input data, significantly enriching dataset completeness and recommendation effectiveness.
- Created a Streamlit web site for real-time data uploads (up to 200MB), visualize audience insights (e.g., geographic density plots), and support ticket recommendation workflows, **reducing overall data processing time by 5%** and enhancing stakeholder decision-making.

### Comviva

Delhi, India

#### Product Data Analyst

Jan 2022 - Apr 2023

- Collaborated with the development team to design key web pages on the media streaming platform using Angular, JavaScript, and HTML/CSS, enhancing data collection for user behavior tracking and streaming analytics during the FIFA 2022 World Cup.
- Extracted and analyzed large streaming datasets (**over 10 million records**) using Google Analytics, SQL, and Python (Pandas, Matplotlib) to identify key trends, such as **15% peak traffic surges** and audience drop-off points.
- Delivered detailed reports with actionable insights, enabling clients to optimize content delivery strategies leading to a **10% increase in viewer engagement**, improved streaming stability, and a **20% reduction in latency during high-traffic events**.

### TAM Enterprises

Remote, India

#### Data Analyst

Aug 2020 - Aug 2021

- Collected and **managed data for 100+ students**, **improving data accuracy by 10%** through advanced cleaning, feature engineering, and resolution of inconsistencies, enhancing data quality for reliable insights.
- Conducted detailed analysis of student performance trends delivering reports and dashboards that drove improvements in career mentoring, free learning programs, and placement support.

## TECHNICAL SKILLS

**Programming / Databases:** Python (pandas, NumPy, Streamlit, scikit-learn), R, SQL

**Data Visualization:** Tableau, Power BI, Google Analytics, MS Excel

**ETL & Automation:** Automated ETL pipelines, data integration & reporting

**Data Analysis/ Modelling:** Predictive modeling, trend analysis, feature engineering

**Machine Learning:** Regression, Clustering, LLM, RAG, Open AI API, gradient boosting

**Tools/ Frameworks:** AWS, Microsoft Fabric, Cypress, Agile Workflow, JIRA

## PROJECTS

### Real vs Fake Job Listings Fraud Detection Analysis: [ProjectLink](#)

- Built a predictive classification model using Python (Pandas, Scikit-learn, NLP) to analyze **17,000+ job postings**, identifying fraudulent listings with **98% accuracy**, effectively mitigating hiring risks and improving decision-making for recruitment processes.

### NYC Taxi Data Pipeline and Analytics Solution with Microsoft Fabric: [ProjectLink](#)

- Developed a scalable data pipeline for the NYC Taxi & Limousine Commission using Microsoft Fabric, optimizing ETL processes and real-time data flow to enhance **Power BI** reporting, reducing data processing time by **12%** and improving analytics efficiency.

### House Sales Analysis Tableau Dashboard: [ProjectLink](#)

- Analyzed over 20,000 records from the King County House Sales dataset, performing data cleaning, outlier detection, and standardization, and created a Tableau dashboard visualizing **5+ key metrics**, accelerating real estate decision-making by **30%**.

## EDUCATION

### University of Illinois System

Illinois, United States

#### Master of Science in Computer Science

GPA: 4/4

Coursework: Data Visualization(562), Advance Database Concepts(572), Machine Learning (532)