

# Risshikaa Yantrapati

## Data Engineer

Location: WA | Phone: (405)673-6464 | Email: [yantrapatirisshikaa@gmail.com](mailto:yantrapatirisshikaa@gmail.com) | LinkedIn: [linkedin.com/in/risshikaa-yantrapati](https://www.linkedin.com/in/risshikaa-yantrapati)

### SUMMARY

- With over 4+ years of professional experience in Data Engineering, adept at leveraging a comprehensive suite of technical skills to build and optimize robust data solutions.
- Proficient in programming languages including R, Python, SQL, and PEARL, excelling in designing algorithms and models that refine data into actionable insights.
- Expertise extends to sophisticated data manipulation and analysis packages like NumPy, Pandas, and TensorFlow, enhancing the ability to handle complex datasets efficiently.
- Skilled in deploying and managing databases such as MongoDB, MySQL, and PostgreSQL, ensuring data integrity and accessibility.
- Proficiency with big data technologies like Hadoop, Apache Spark, and Kafka allows for the implementation of scalable data processing systems that support enterprise-level data workflows.
- Solid background in cloud technologies, with practical experience in AWS, Azure, and IBM Cloud, empowers the architecture and maintenance of secure, cost-effective cloud-based data solutions. The approach is rooted in a thorough understanding of SDLC methodologies, particularly Agile and Waterfall, guiding project management and collaborative efforts.

### PROFESSIONAL EXPERIENCE

#### Charter Communications

Denver, CO

#### System Engineer Intern

May 2023 – Aug 2023

- Revamped Python codebase by implementing performance optimizations and enhancing code readability; reduced script execution time by 45% and increased developer efficiency by simplifying code structure and logic flow.
- Orchestrated the implementation of Grafana and Heavi AI to craft dynamic graphical representations of data insights; optimized data visualization processes, resulting in a 40% increase in data-driven decision-making efficiency.
- Implemented DNS service discovery to optimize network efficiency, reducing downtime. Utilized Ansible pull for streamlined configuration management, leading to improved operational efficiency.
- Coordinated the integration of AWS services to fortify data security by securely managing and distributing sensitive passkeys.

#### Cognizant Technology Solutions

Hyderabad, India

#### Big Data Developer

Jan 2022 – Aug 2022

- Enhanced business decision-making processes by developing and optimizing data processing pipelines with Python and SQL, integrating advanced analytical tools like NumPy, Pandas, and Scikit-learn, which facilitated sophisticated predictive modeling.
- Ensured data accuracy and availability across organizational platforms by implementing and managing robust ETL processes using SSIS and MySQL, optimizing data extraction, transformation, and loading activities. Reduced data processing time by 30% through SSIS package optimization, improving data availability and timeliness.
- Improved operational efficiency and strategic planning by designing and creating dynamic data visualizations and interactive dashboards with Power BI and MS Excel, enabling stakeholders to extract critical insights through real-time data analytics.
- Boosted data storage scalability and processing capabilities by configuring and maintaining environments using Hadoop and Hive, streamlining data flow management with Apache Nifi and orchestrating containerized applications in AWS with Kubernetes.
- Achieved high availability, fault tolerance, and scalable performance in data applications by leveraging AWS cloud services, aligning deployment, monitoring, and management processes with the organization's growth and strategic data-driven objectives.

#### Virtusa Consulting Services

Chennai, India

#### Big Data Developer

Jun 2019 – Jan 2022

- Enhanced predictive analytics capabilities by developing and optimizing complex data algorithms with R and Python, utilizing libraries like NumPy, Pandas, and Seaborn, which led to a 30% increase in data analysis efficiency and supported strategic business decision-making.
- Facilitated real-time business insight access by designing and implementing versatile data visualizations with Tableau and MS Excel, and developing comprehensive reports using SSRS, which improved operational decision-making by 25% and enhanced tactical planning.
- Improved data accuracy and accessibility for analytics by managing and optimizing ETL processes using SSIS, integrating data from diverse sources into MongoDB and PostgreSQL databases, resulting in a 40% enhancement in data reliability.
- Supported big data initiatives and enhanced data-driven decision-making by engineering scalable data processing solutions with MapReduce and Apache Spark, which increased data handling efficiency by 35% for large-scale datasets.
- Ensured high availability and optimal performance in the Azure cloud environment, aligning with organizational goals and cloud-first strategies, which reduced operational costs by 20% while improving resource scalability.
- Continuously improved data solution processes within an Agile framework, actively contributing to sprint planning and retrospectives, leading to a 15% increase in project delivery efficiency and alignment with business objectives.
- Maintained codebase integrity and enhanced team collaboration by utilizing Git for version control, which secured data projects and facilitated a 10% improvement in development team productivity.

- Boosted team productivity and enhanced project documentation by proficiently using MS Office tools, creating detailed documentation and presentations that improved communication clarity by 20% and documented project scopes and outcomes effectively.

**Cipla**

**India**

**Data Analyst**

**Oct 2018 – May 2019**

- Developed Python and SQL scripts to automate ETL tasks, streamlining data extraction, transformation, and loading workflows, enhancing analysis efficiency by 30% and maintaining high accuracy.
- Wrote optimized SQL queries with indexing and query tuning techniques to speed up database performance by 25%, enabling faster retrieval of key datasets and improving overall analysis timelines.
- Built comprehensive Tableau and Excel dashboards to monitor KPIs, providing stakeholders with interactive and visually appealing views of business metrics for quick and informed decision-making.
- Utilized advanced SQL queries for data blending, integrating disparate datasets to provide comprehensive views of business performance and uncovering critical trends that guided strategic planning.
- Designed dynamic charts and visualizations in Excel and Tableau to represent KPI trends, providing clear reports that helped stakeholders understand revenue performance, customer behavior, and market trends.
- Implemented AWS services like EC2, S3, and Lambda to securely store and automate data processing tasks, reducing manual work and improving the scalability of data analysis workflows for large datasets.

## TECHNICAL SKILLS

<b>Programming Language:</b>	R, Python, SQL, Java
<b>Packages:</b>	NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow,
<b>Reporting Tools:</b>	Tableau, Power BI, MS Excel, SSRS
<b>ETL Tools:</b>	SSIS
<b>Database:</b>	MongoDB, MySQL, PostgreSQL
<b>Big Data Ecosystem:</b>	Hadoop, MapReduce, Hive, Apache Spark, Pig, Kafka, Apache Nifi, Kubernetes
<b>Cloud Technologies:</b>	AWS, Azure, IBM Cloud
<b>Methodology</b>	SDLC, Agile, Waterfall
<b>Other Tools:</b>	Git, MS Office
<b>IDE's:</b>	PyCharm, Jupyter Notebook
<b>Operating Systems:</b>	Windows, Linux

## EDUCATION

**University of Colorado Boulder, Boulder**

**Master's in Science, Data Science**

**Aug 2022 – May 2024**

**Gudlavalleru Engineering College, JNTUK, Gudivada**

**Bachelor of Technology, Computer Science & Engineering**

**Aug 2015 - May 2019**

## PROJECTS

**Capstone Project, Q2 Impact Company**

- Spearheaded a comprehensive analysis of hurricane crisis response planning in New York State, leveraging diverse datasets and collaborative stakeholder engagement.
- Developed actionable recommendations to address socioeconomic vulnerabilities and enhance resilience, utilizing tools such as Pandas, NumPy, and Tableau for data analysis and visualization.

**Car Rental Analysis**

- Utilizing a scalable pipeline integrating databases, Kafka messaging system, and visualization tools for insightful data exploration and problem resolution.

**Fake Job prediction**

- Developed a machine learning classifier leveraging natural language processing to accurately distinguish between genuine fraudulent job postings, employing data-driven approaches and rigorous model evaluation techniques.

**Trends in US Major Pollutants**

- Using R Language, Based on R plots I visualized how the four major pollutants changed over a 16-year period (2000-2016).

## CERTIFICATIONS

- **Data Science and Big Data Analytics**
- **Associate Cloud Architect from Amazon**
- **Cloud Infrastructure and Services**
- **Web Services (AWS) conducted by APSSDC**
- **ML Building Blocks by AWS**