

Data Engineer/Big Data Team Internship Job Description

Overview

Join our innovative Data Engineer/Big Data Team Internship, a 5-month program designed for students and recent graduates passionate about transforming raw data into powerful insights to support community initiatives. This part-time, hybrid role offers the flexibility to work remotely or on-site in Pune and Mumbai, providing hands-on experience in data pipeline development, big data processing, and analytics. No prior data engineering or big data experience is required, as comprehensive training in tools like Python, SQL, Hadoop, and Spark will be provided under the mentorship of experienced professionals.

Program Highlights

- Duration: 5 months, with the first 2 months unpaid and a stipend (₹10,000–₹15,000/month) for the remaining 3 months; extension possible based on performance and project needs.
- Type: Part-time (15–20 hours/week, flexible schedule to accommodate academic commitments).
- Location: Hybrid (remote with occasional on-site work in Pune or Mumbai, requiring a reliable internet connection for remote tasks).
- Mentorship: Personalized guidance from seasoned data engineers and big data professionals with regular feedback sessions.
- Training Provided: No prior skills required; comprehensive training in Python, SQL, Hadoop, Spark, and cloud-based data tools.
- Hands-On Projects: Contribute to real-world projects, including building data pipelines and analytics solutions for community-focused platforms.

Key Responsibilities

As a Data Engineer/Big Data Team Intern, you will:

- Assist in designing and building data pipelines to collect, process, and store large datasets for community initiatives in Pune and Mumbai.
- Support data extraction, transformation, and loading (ETL) processes using tools like Python, SQL, or Apache Spark.
- Collaborate with the team to integrate data from various sources into centralized systems for analysis.
- Assist in optimizing data workflows for scalability, performance, and reliability.
- Create and update documentation for data pipelines, schemas, and workflows to ensure clarity and maintainability.
- Support data quality checks and validation to ensure accurate and reliable insights.
- Contribute to building dashboards or reports to visualize insights for community stakeholders.
- Assist in exploring big data technologies (e.g., Hadoop, Spark) to handle large-scale data processing.

Learning Opportunities

Under expert mentorship, you will gain:

- Data Engineering:
 - Master designing and building scalable data pipelines with training in Python, SQL, and ETL processes.
 - Learn to manage and process large datasets using big data frameworks like Hadoop and Spark.
- Big Data Processing:
 - Gain experience in handling large-scale data with distributed computing tools.
 - Develop skills in optimizing data workflows for performance and scalability.
- Data Analysis and Visualization:
 - Enhance proficiency in data cleaning, transformation, and visualization using tools like Pandas and Tableau.
 - Learn to create actionable insights for community-focused initiatives.
- Digital Tools:
 - Build proficiency in data engineering tools (e.g., Apache Spark, Airflow) and cloud platforms (e.g., AWS, Google Cloud) with training.
 - Learn to use version control systems (e.g., Git) for collaborative development.
- Career Development:
 - Receive personalized mentorship to guide your growth in data engineering and big data.
 - Build a professional portfolio showcasing contributions to data pipelines and analytics projects.

Required Skills and Qualifications

We welcome candidates passionate about data engineering and big data. No prior skills are required, and training will be provided.

- Technical and Communication Skills:
 - Basic understanding of programming or data processing through coursework, projects, or self-study.
 - Willingness to learn data engineering tools and big data technologies through provided training.
- General Skills:
 - Strong attention to detail for ensuring data quality and pipeline reliability.
 - Excellent communication skills for collaborating with data teams and stakeholders.
 - Proactive problem-solving abilities and a self-motivated attitude toward learning.
 - Ability to work independently and collaboratively in a hybrid team under mentorship.

Requirements

- Current enrollment in computer science, data science, engineering, or related fields at an accredited institution.
- Availability for 15–20 hours per week with a flexible schedule.
- Access to a reliable internet connection and a computer capable of running data engineering tools (e.g., Python, SQL, Hadoop).
- Willingness to travel occasionally to Pune or Mumbai for on-site project meetings or data reviews.

Preferred Qualifications

While not mandatory, these qualifications strengthen your application:

- Personal or academic projects in data processing, ETL, or analytics (e.g., building a data pipeline, analyzing datasets).
- Basic familiarity with Python, SQL, or any data processing tool.
- Exposure to big data concepts or tools (e.g., Hadoop, Spark) through coursework or self-study.
- Familiarity with remote collaboration tools (e.g., Slack, Trello, Zoom) or version control systems (e.g., Git).

Who Should Apply

- Students: Undergraduates or postgraduates in computer science, data science, engineering, or related fields eager to gain experience in data engineering and big data, with no technical background required.
- Freshers: Recent graduates passionate about launching careers in data engineering or big data analytics, with training provided to bridge skill gaps.

Work Environment

- Collaborative Culture: Join a supportive, inclusive hybrid team fostering creativity and continuous learning in data engineering and big data.
- Mentorship: Receive one-on-one guidance from experienced data engineers to accelerate your skills, with tailored training.
- Real-World Impact: Work on data-driven solutions supporting community initiatives in Pune and Mumbai.
- Flexible Arrangements: Hybrid work (remote with occasional on-site tasks in Pune or Mumbai) with a flexible 15–20 hours/week schedule.
- Duration and Compensation: 5 months, with the first 2 months unpaid and a stipend (₹10,000–₹15,000/month) for the remaining 3 months; extension possible.

Benefits

- Hands-On Experience: Gain practical skills in building data pipelines, processing big data, and generating insights, with comprehensive training.
- Mentorship: Learn from experts in data engineering and big data, with personalized feedback.
- Portfolio Development: Build a professional portfolio showcasing contributions to data engineering and analytics projects.
- Academic and Career Support: Earn academic credit (subject to approval) or reference letters for academic and professional advancement.
- Networking: Connect with data engineering and big data professionals, opening future career opportunities.

How to Apply

Take the first step toward a rewarding career in data engineering and big data! Our three-stage selection process ensures we find passionate candidates:

1. Application Submission: Submit your resume, along with any relevant project portfolios or links to work (e.g., data pipelines, analytics projects, GitHub repositories) to support@rbcurexia.com.
2. Assessment Task: Complete a short task on data processing or pipeline design to demonstrate your abilities (no prior skills required; guidance provided).
3. Interview: Participate in a virtual or in-person interview to discuss your skills, passion, and fit, with opportunities to learn more about the role.

Apply today to gain invaluable experience in data engineering and big data, with all necessary training provided, and contribute to impactful community solutions in Pune and Mumbai!

Contact Us

For more information, reach out to us:

- Website: www.rbcurexia.com
- Email: support@rbcurexia.com