

# Data Engineer - data42

Job ID

REQ-10012035

Jul 18, 2024

Inde

## Résumé

The Data Engineer will work alongside data scientists and domain experts to enable teams to answer scientific questions using multi-modal data on the data42 platform. They will be involved in gathering use-case requirements, performing engineering activities for data services, building ETL processes/data pipelines in quick iterations to deliver data ready for analysis. The Data Engineer will integrate data engineering best practices and data quality checks and seek to continuously optimize efficiency.

## About the Role

Your responsibilities will include, but are not limited to:

- Collaborates with domain experts, data scientists and other stakeholders to fulfil use-case specific data needs.
- Designs, develops, tests, and maintains ETL processes/data pipelines to extract, prepare and iterate data for analysis in close alignment with TA / DA scientific leads and data scientists.
- Implements and maintains data checks to ensure accurate and high quality-data in close collaboration with domain experts.
- Identifies and rectifies data inconsistencies and irregularities.
- Promotes culture of transparency and communication regarding data modifications and lineage to all stakeholders.
- Implements and advocates for data engineering best practices, ensuring ETL processes/data pipelines are efficient, well-documented and well-tested.
- Plays a role in knowledge sharing across data42 and wider data engineering community at Novartis.
- Ensures compliance with Security and Governance Principles.

## Minimum Requirements:

- Bachelor's degree in computer science or other quantitative field (Mathematics, Statistics, Physics, Engineering, etc.) or equivalent practical experience.
- Proven experience as a data engineer, data wrangler or a similar role.
- Exceptional programming skills with expertise in Python, R and Spark.
- Experience and familiarity with a variety of data types, including but not limited to images, tabular, unstructured, and text.
- Experience in scalable data processing engines, data ingestion, extraction and modeling.
- Proficient knowledge in statistics, with an ability to assess data quality, errors, inconsistencies, etc.
- Good knowledge of data engineering best practices · Excellent communication and stakeholder management skills.
- Demonstrated ability to work independently and as part of global Agile teams.

Desirable additional skills in two or more of the following areas:

- Hands on experience on Palantir Foundry (Code Repository, Code Workbook, Contour, Data Lineage, etc.)
- Knowledge of CDISC data standards (SDTM, ADaM)
- Experience using AI (eg: GenAI/LLMs) for data wrangling.
- Experience with pooling of clinical trial data.
- High-level understanding of the drug discovery and development process.

**Why Novartis:** Helping people with disease and their families takes more than innovative science. It takes a community of smart, passionate people like you. Collaborating, supporting and inspiring each other. Combining to achieve breakthroughs that change patients' lives. Ready to create a brighter future together? <https://www.novartis.com/about/strategy/people-and-culture>

**Join our Novartis Network:** Not the right Novartis role for you? Sign up to our talent community to stay connected and learn about suitable career opportunities as soon as they come up: <https://talentnetwork.novartis.com/network>

Division

Biomedical Research

Business Unit

Pharma Research

Emplacement

Inde

Site

Hyderabad (Office)

Company / Legal Entity

IN10 (FCRS = IN010) Novartis Healthcare Private Limited

Functional Area

Recherche & Développement

Job Type

Full time

Employment Type

Regular

Shift Work

No

[Apply to Job](#)

Novartis is committed to building an outstanding, inclusive work environment and diverse teams' representative of the patients and communities we serve.

iframe{ width: 100%; margin-top: 3rem; } @media screen and (max-width: 767px){ iframe{ height: 30vh !important; } } @media screen and (min-width: 768px){ iframe{ height: 34vh !important; } }

Job ID

REQ-10012035

## Data Engineer - data42

[Apply to Job](#)

---

**Source URL:** <https://www.adacap.com/careers/careersearch/job/details/req-10012035-data-engineer-data42>

## List of links present in page

1. <https://www.novartis.com/about/strategy/people-and-culture>
2. <https://talentnetwork.novartis.com/network>
3. [https://novartis.wd3.myworkdayjobs.com/en-US/Novartis\\_Careers/job/Hyderabad-Office/Data-Engineer---data42\\_REQ-10012035](https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Hyderabad-Office/Data-Engineer---data42_REQ-10012035)
4. [https://novartis.wd3.myworkdayjobs.com/en-US/Novartis\\_Careers/job/Hyderabad-Office/Data-Engineer---data42\\_REQ-10012035](https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Hyderabad-Office/Data-Engineer---data42_REQ-10012035)