

Senior Data Engineer

A Senior Data Engineer leads the inventorying, gathering, cleaning, processing, and programming of automated datasets for use by analysts and others. They may work with other data engineers and act as mentor, expert, and project manager to implement data engineering projects. They are concerned with establishing standard practices that enable continuous improvement and data operations, including but not limited to documentation, tooling, and project intake, prioritization, and management.

A Senior Data Engineer must be able to work with analysts and subject matter experts to define and build requirements for automated datasets, including but not limited to defining business transformations, validation rules, and frequency of updates. The ability to map the requirements on to an efficient and sustainable solution is necessary. Expertise in ETL/ELT, data profiling, and other data quality tools and approaches is required.

A Senior Data Engineer may also be involved in the identification of datasets with lines of business, particularly around a project or program need. They may be involved in helping to make a catalog of raw datasets accessible to analysts, data scientists and others for more exploratory analysis to inform the development of data products and services. They will need to work closely with analysts, data modelers, warehouse architects, and IT staff to design and implement proposed solutions and architectures that meet business needs.

A Senior Data Engineer should be fluent in one or more data manipulation languages with an emphasis on frameworks appropriate for production code and depending on need (e.g., Java, Scala, Python, Advanced SQL, etc.). They should also be able to implement data automations within existing frameworks as opposed to writing one off scripts. They may spend less time on writing code or constructing pipelines as they will lead on ensuring work products meet quality standards through a mix of methods including but not limited to peer review, pair programming, and automated testing. This is dependent on program needs and complexity of projects.

While they should have proficiency in code and best practices in engineering management, they should also be comfortable implementing in low-code ETL/ELT platforms if that is the more sustainable approach for the organization.

Experience

This Classification must have a minimum of three (3) years of relevant experience in authoring and monitoring data pipelines.

And Education

This classification requires the possession of a bachelor's or advanced degree in computer science, engineering, information systems, math or technology-related field. Additional qualifying experience may be substituted for the required education on a

year-for-year basis. Successful completion of a data engineering bootcamp plus an additional year of experience may be substituted for the degree requirement.