To become a **Data Engineer**, you need a mix of technical skills that allow you to design, build, and manage data pipelines, as well as maintain data systems that support analytics and machine learning. Here are the essential technical skills required:

**1. Programming Languages**

* **Python**: Widely used for data processing, scripting, and building data pipelines. Libraries like Pandas, NumPy, and PySpark are crucial.
* **SQL**: Essential for querying and managing relational databases. You should be comfortable writing complex queries, joining tables, and working with databases like MySQL, PostgreSQL, and Oracle.
* **Java/Scala**: Often used for working with big data frameworks like Apache Hadoop and Apache Spark.

**2. Database Management**

* **Relational Databases (RDBMS)**: Understand the fundamentals of relational databases, including database design, indexing, and query optimization. Examples: MySQL, PostgreSQL, Oracle.
* **NoSQL Databases**: Learn how to work with NoSQL databases, which are useful for handling unstructured data. Examples: MongoDB, Cassandra, HBase.
* **Data Warehousing**: Knowledge of data warehousing solutions like Amazon Redshift, Google BigQuery, or Snowflake. These platforms help store and analyze large datasets.

**3. Big Data Technologies**

* **Apache Hadoop**: Understand the Hadoop ecosystem, including HDFS (Hadoop Distributed File System), MapReduce, and YARN. Hadoop is used for storing and processing large datasets.
* **Apache Spark**: A powerful big data processing framework used for batch and real-time processing. Knowing Spark (especially PySpark) is crucial for a Data Engineer.
* **Apache Kafka**: Used for real-time data streaming and event-driven systems. It's vital for building data pipelines that need to process data in real-time.

**4. ETL (Extract, Transform, Load) Tools**

* **Airflow**: A popular open-source tool for scheduling and managing ETL workflows. It's used to automate data pipelines.
* **Talend**: A versatile ETL tool that can connect to various data sources and transform data as needed.
* **Apache Nifi**: Another tool for data flow automation and real-time data processing.

**5. Cloud Platforms**

* **AWS (Amazon Web Services)**: Familiarity with services like S3 (storage), EC2 (computing), RDS (database), and Redshift (data warehousing). AWS Glue is also used for ETL tasks.
* **Microsoft Azure**: Knowledge of Azure Data Lake, Azure Synapse Analytics, and Azure SQL Database.
* **Google Cloud Platform (GCP)**: Services like Google BigQuery, Cloud Storage, and Dataflow are useful for managing data pipelines.

**6. Data Modeling**

* **Schema Design**: Understanding how to design efficient database schemas, including normalization and denormalization.
* **Star and Snowflake Schemas**: Familiarity with data warehousing schemas is beneficial for efficient data storage and querying.
* **Dimensional Modeling**: Useful for creating data models that support business intelligence (BI) and reporting tools.

**7. Data Pipelines and Workflow Orchestration**

* **Building Data Pipelines**: Knowledge of how to design, implement, and maintain scalable data pipelines that can handle large volumes of data.
* **Workflow Orchestration**: Tools like Apache Airflow, Prefect, or Luigi are used to schedule, monitor, and manage complex data workflows.

**8. Version Control Systems**

* **Git**: Understanding how to use Git for version control is important for collaboration and maintaining codebase integrity.

**9. Data Processing Frameworks**

* **Apache Beam**: A unified programming model for both batch and streaming data processing. It can be used with backends like Apache Flink, Apache Spark, and Google Cloud Dataflow.
* **Dask**: A parallel computing library in Python, which can handle large datasets and distribute computations across clusters.

**10. Data Visualization Tools (Optional but Beneficial)**

* **Tableau/Power BI**: Understanding how to visualize data can help in creating dashboards for data monitoring and reporting.
* **Matplotlib/Seaborn**: Python libraries for basic data visualization.

**11. Operating Systems and Shell Scripting**

* **Linux/Unix**: Proficiency in Linux-based systems is necessary since many data processing tasks and big data tools run on Linux servers.
* **Shell Scripting**: Bash scripting is useful for automating tasks and managing servers.

**12. DevOps and Containerization (Bonus Skills)**

* **Docker**: Knowledge of containerization is important for deploying and scaling data applications.
* **Kubernetes**: Understanding how to manage and orchestrate containers can be beneficial for managing large-scale data solutions.
* **CI/CD**: Familiarity with Continuous Integration and Continuous Deployment pipelines to streamline code deployments.

**Learning Path:**

1. **Start with Python and SQL**, as they are foundational.
2. **Learn Big Data technologies** like Hadoop and Spark.
3. **Explore Cloud Platforms** and familiarize yourself with data services offered by AWS, Azure, or GCP.
4. **Understand ETL processes** and how to build efficient data pipelines.
5. **Practice building projects** that cover end-to-end data engineering workflows.

Mastering these skills will prepare you for a successful career as a Data Engineer, and there are many online courses and certifications (e.g., from Coursera, Udacity, or DataCamp) that can help you build expertise.

Latest New Technologies:-

The tech industry in India is rapidly evolving, and certain skills are in high demand, offering lucrative salaries. Here are some of the latest technologies and fields that can lead to well-paying jobs:

**1. Artificial Intelligence (AI) & Machine Learning (ML)**

* **Skills**: Deep Learning, Neural Networks, Natural Language Processing (NLP), Computer Vision.
* **Tools**: TensorFlow, PyTorch, Scikit-Learn.
* **Career Roles**: Data Scientist, AI Engineer, Machine Learning Engineer.
* **Why**: AI/ML professionals are in high demand across various sectors, including finance, healthcare, e-commerce, and tech.

**2. Data Science & Analytics**

* **Skills**: Data Analysis, Big Data, Data Visualization, Statistics.
* **Tools**: Python, R, SQL, Hadoop, Apache Spark, Tableau, Power BI.
* **Career Roles**: Data Scientist, Data Analyst, Data Engineer, Business Intelligence Analyst.
* **Why**: Organizations rely on data-driven decisions, making data science a critical area with high-paying job opportunities.

**3. Cloud Computing**

* **Skills**: Cloud Architecture, Cloud Security, DevOps, Networking.
* **Platforms**: Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP).
* **Career Roles**: Cloud Engineer, Cloud Architect, DevOps Engineer.
* **Why**: Companies are moving their infrastructure to the cloud, increasing the demand for cloud professionals with skills in scalable and secure cloud architecture.

**4. Cybersecurity**

* **Skills**: Ethical Hacking, Network Security, Risk Management, Encryption.
* **Tools**: Wireshark, Metasploit, Nessus, Splunk.
* **Career Roles**: Cybersecurity Analyst, Ethical Hacker, Security Consultant.
* **Why**: As data breaches and cyber threats increase, skilled cybersecurity professionals are essential, and companies are willing to pay a premium for security expertise.

**5. Full Stack Development**

* **Skills**: Front-end (HTML, CSS, JavaScript, React, Angular), Back-end (Node.js, Django, Ruby on Rails).
* **Tools**: Git, Docker, Jenkins.
* **Career Roles**: Full Stack Developer, Software Engineer, Web Developer.
* **Why**: Full stack developers can build complete web applications, and startups and tech companies highly value their versatility.

**6. DevOps & Automation**

* **Skills**: Continuous Integration/Continuous Deployment (CI/CD), Infrastructure as Code (IaC), Scripting.
* **Tools**: Docker, Kubernetes, Jenkins, Ansible, Terraform.
* **Career Roles**: DevOps Engineer, Site Reliability Engineer (SRE), Automation Engineer.
* **Why**: DevOps enables faster software development and deployment, leading to higher efficiency. The demand for automation and seamless integration is growing.

**7. Blockchain**

* **Skills**: Blockchain Development, Smart Contracts, Cryptography.
* **Platforms**: Ethereum, Hyperledger, Solana, Binance Smart Chain.
* **Career Roles**: Blockchain Developer, Blockchain Architect, Crypto Analyst.
* **Why**: Blockchain is not just about cryptocurrencies; it's also used in finance, supply chain management, and healthcare, with high demand for skilled developers.

**8. Internet of Things (IoT)**

* **Skills**: Embedded Systems, Sensors, Wireless Communication, Data Analytics.
* **Tools**: Arduino, Raspberry Pi, MQTT, ThingSpeak.
* **Career Roles**: IoT Developer, Embedded Software Engineer, IoT Architect.
* **Why**: IoT connects physical devices to the internet, and there is increasing demand for smart, connected devices in various industries.

**9. Robotic Process Automation (RPA)**

* **Skills**: Process Automation, Scripting, Workflow Automation.
* **Tools**: UiPath, Automation Anywhere, Blue Prism.
* **Career Roles**: RPA Developer, Process Automation Engineer, RPA Consultant.
* **Why**: Automation helps businesses streamline operations, reduce costs, and improve efficiency, creating opportunities for RPA experts.

**10. Augmented Reality (AR) & Virtual Reality (VR)**

* **Skills**: 3D Modeling, Game Development, Interactive Design.
* **Tools**: Unity, Unreal Engine, Blender.
* **Career Roles**: AR/VR Developer, Game Developer, Immersive Experience Designer.
* **Why**: AR and VR are increasingly used in gaming, education, real estate, and marketing, leading to high demand for developers.

**11. Quantum Computing (Emerging Field)**

* **Skills**: Quantum Mechanics, Quantum Algorithms, Cryptography.
* **Tools**: IBM Qiskit, Microsoft Quantum Development Kit, Google Cirq.
* **Career Roles**: Quantum Research Scientist, Quantum Software Developer, Quantum Data Scientist.
* **Why**: Though still emerging, quantum computing is poised to revolutionize industries, and professionals in this field are rare and well-compensated.

**Conclusion**

These technologies represent some of the most promising fields for career growth and high salaries in India. Building skills in any of these areas can lead to lucrative job opportunities. Consider choosing a field that aligns with your interests and investing time in hands-on projects, certifications, and real-world applications to improve your employability.