

Data Engineer

Analitik, Programista czy Administrator?



01. **Data Engineer**

SQL Developer v2.0?

Data
Analyst

Data
Designer

Data
Engineer

DataOps

Data
Scientist



Data Engineer

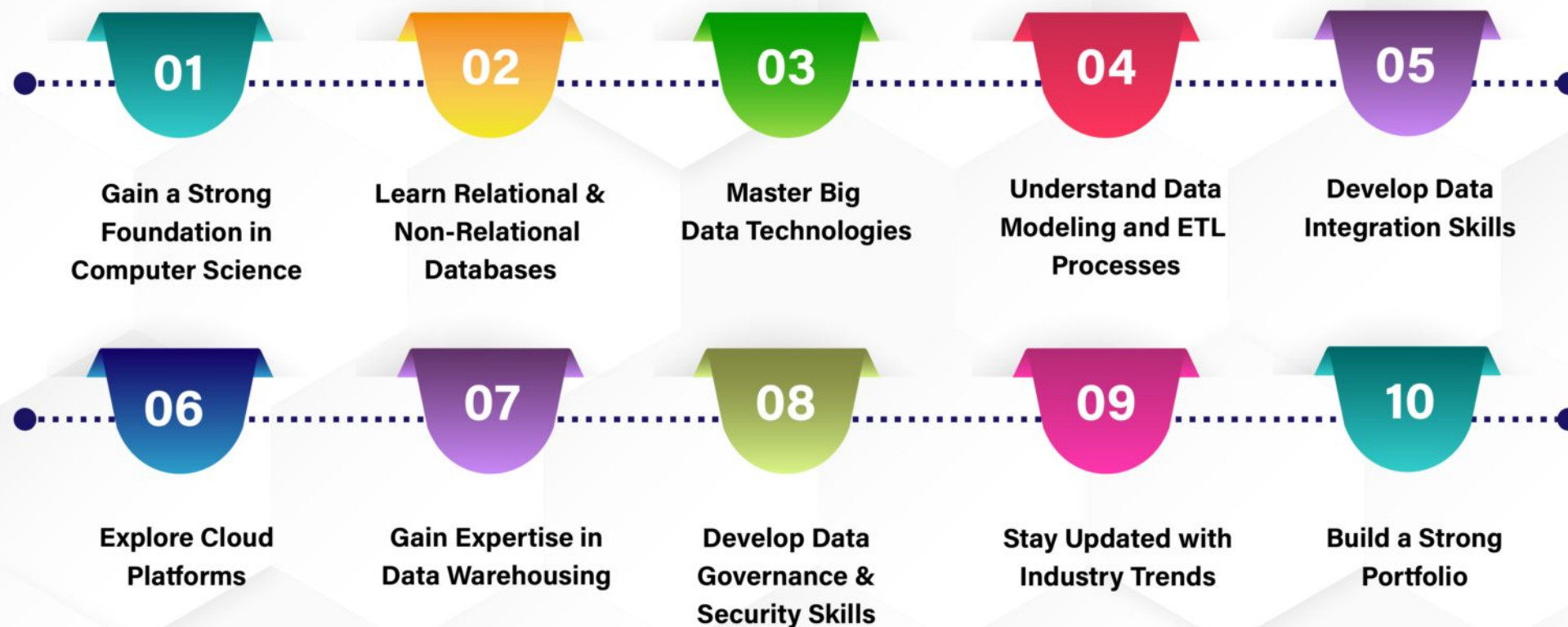
- **Data Engineer:** przetwarzanie/procesowanie danych
- **Data Analyst:** analiza oraz wyciąganie wniosków z danych
- **Data Scientist:** predykcja danych, wyciąganie "extra" informacji
- **DataOps:** tworzenie i zarządzanie infrastrukturą pod dane
- **Data Designer:** projektowanie struktury danych (sposoby mapowania, itd..)



01. Roadmap

Co należy poznać?

● ROADMAP FOR A THRIVING ● **DATA ENGINEERING CAREER**



DATA ENGINEERING ROADMAP

8. Data Governance & Security

- Data Masking / Encryption
- Data Policies / Compliance
- Authentication & Authorization

6. Data Processing

- Batch Processing: Spark
- Delta Lake & Delta Live Tables with Spark SS
- Stream Processing: Kafka

4. Storage

- Data warehousing
- Lakehouse
- OLAP & OLTP
- Storage Tools: S3/Azure Blob
- Warehousing Tools: Snowflake/BigQuery/Synapse

1. Programming

- Python, R, Java or Scala
- SQL
- Linux Commands

9. Infrastructure as Code

Kubernetes,
Terraforms & CI/CD

7. Big Data Technologies

- MapReduce
- Massive Parallel processing
- Distributed Systems
- Tools: Apache Hadoop, Apache Spark

5. Data Pipelines

- ETL Pipelines / Tools
- Scheduling & Orchestration
- Tools: Apache Airflow, Azure Data Factory, Dagster, Apache NiFi

3. Cloud Fundamentals

- Microsoft Azure
- Google Cloud
- AWS

2. Databases

- Relational & NoSQL Databases (MySQL, MongoDB, etc)
- Data modelling (Fact & Dimension Tables)

Must Have:

- Język programowania: Python lub Scala
- Język SQL + relacyjna baza danych (np. PostgreSQL)
- Znajomość baz nierelacyjnych
- Apache Spark
- Formaty danych: JSON, Parquet, AVRO, Delta, CSV

Oczekiwana znajomość:

- Umiejętność budowania procesów ETL/ELT: Apache Airflow, Prefect, Apache NiFi
- Przechowywanie danych: ADSL, S3, HDFS, itd..
- Narzędzia do streamingu: Apache Kafka
- Narzędzia chmurowe
- ! Databricks !

A jak to wygląda w praktyce?

infoShareAcademy.com