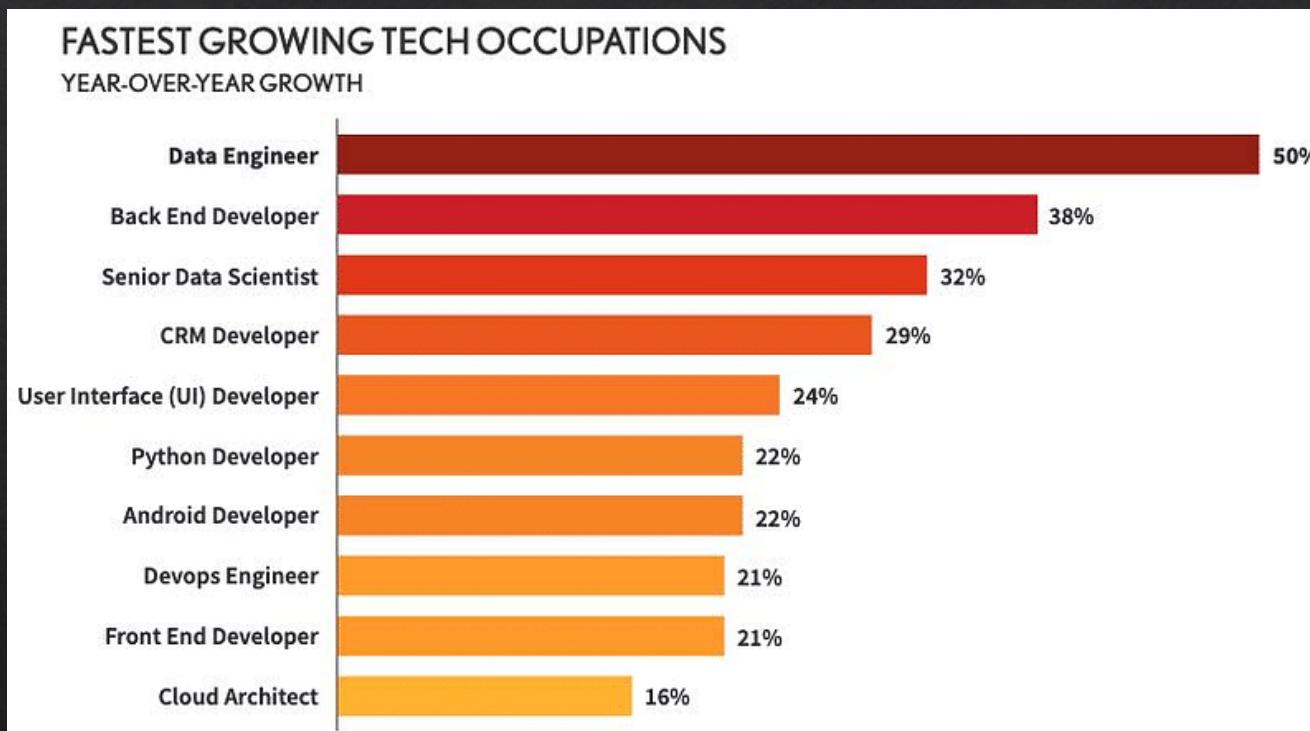


Engenharia de Dados



Fonte: Dice



Tutor

Fernando Amaral

Quantos Bancos de Dados Relacionais Existem?

- ❖ Mais de 100!
- ❖ O curso é focado em fundamentos!
 - Estruturas de dados
 - Conceitos
 - Linguagens: SQL e Python
- ❖ Porém precisamos utilizar ferramentas!



Nuvem!



VOCÊ NÃO PRECISA
INSTALAR NADA!

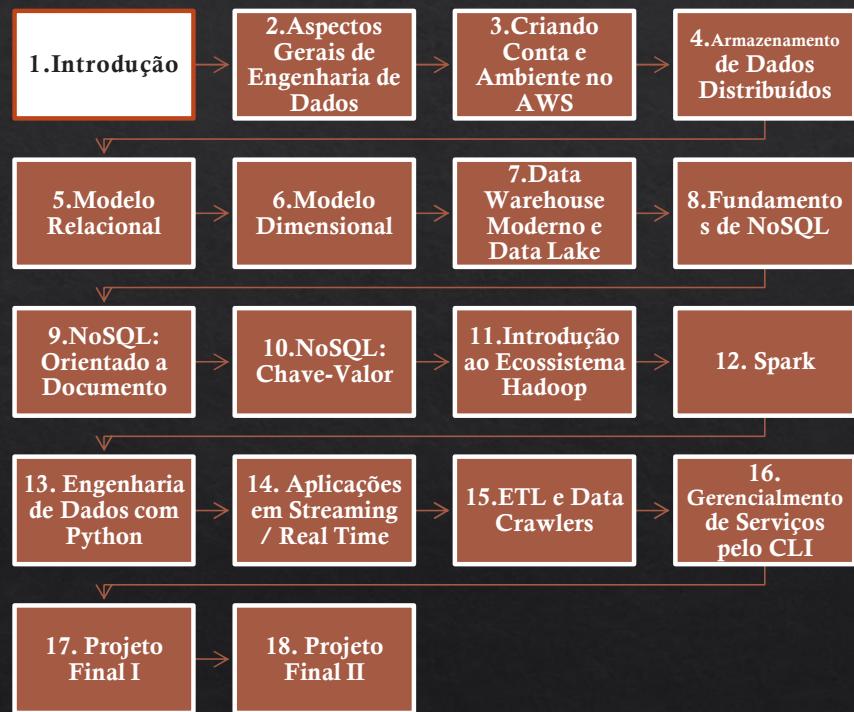


QUALQUER
COMPUTADOR COM
ACESSO A INTERNET!



A MAIORIA DOS
SERVIÇOS POSSUEM FREE
TIER!

O Que Vamos Estudar!



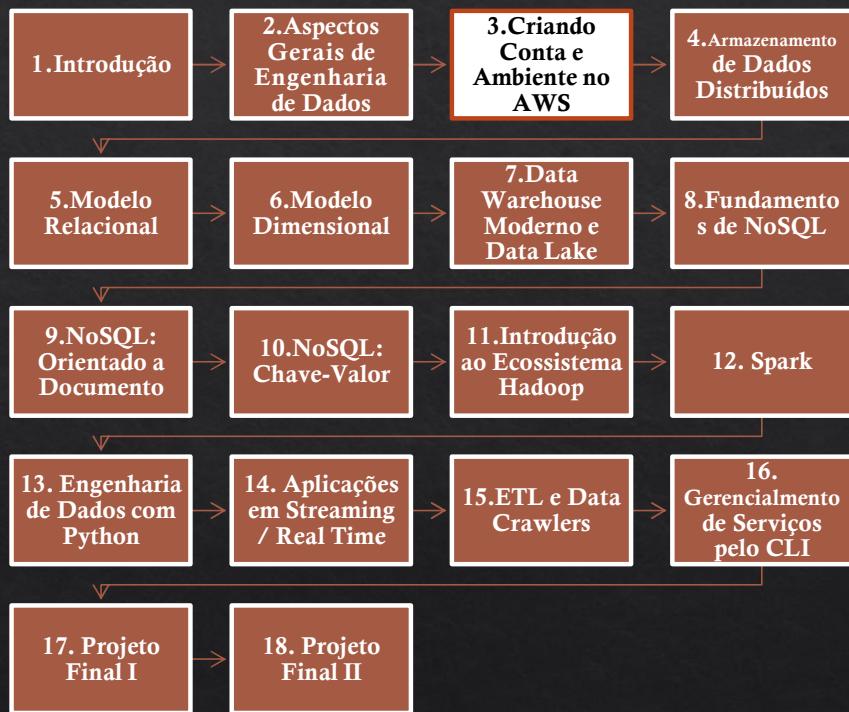
- Orientações
- O que Big Data
- Conceitos
- Teorema CAP
- Operacional vs Analítico
- Etc.

2. Aspectos Gerais de Engenharia de Dados

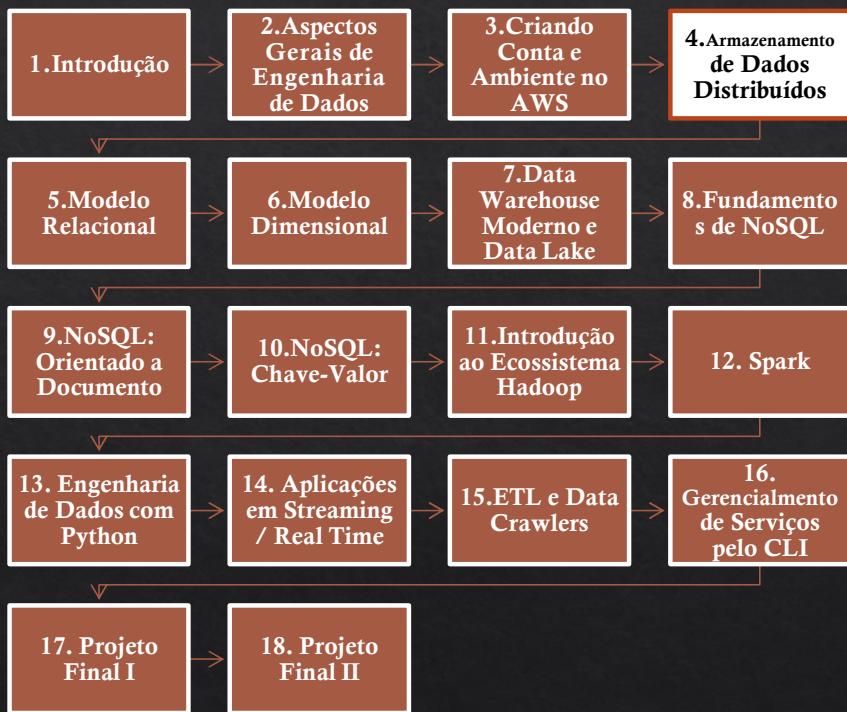


- Data Lakes
- Segurança
- Riscos
- Nuvem
- Machine Learning
- Etc.

4. Armazenamento de Dados Distribuídos

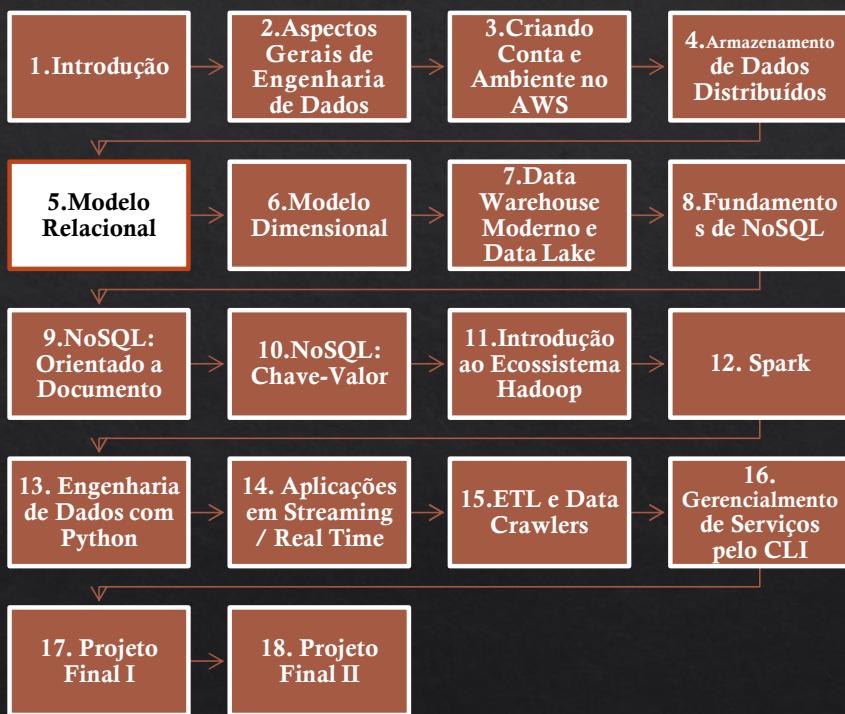


5. Modelo Relacional



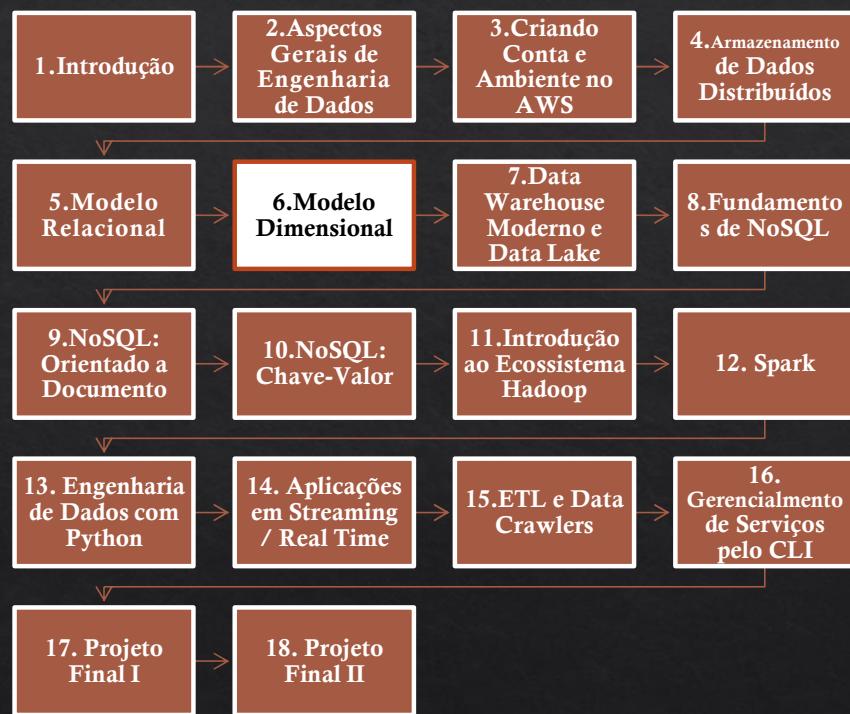
- Criando Buckets e Pastas
- Gerenciamento de Ciclo de Vida
- Segurança

6. Modelo Dimensional



- Arquitetura
- ACID
- Níveis de Isolamento
- Formas Normais
- SQL
- Projeto Prático

6. Modelo Dimensional



- Granulidade
- Cubos
- Fatos
- Dimensões
- Hierarquia
- Chave Substituta
- Caso Prático

7. Data Warehouse Moderno e Data Lake

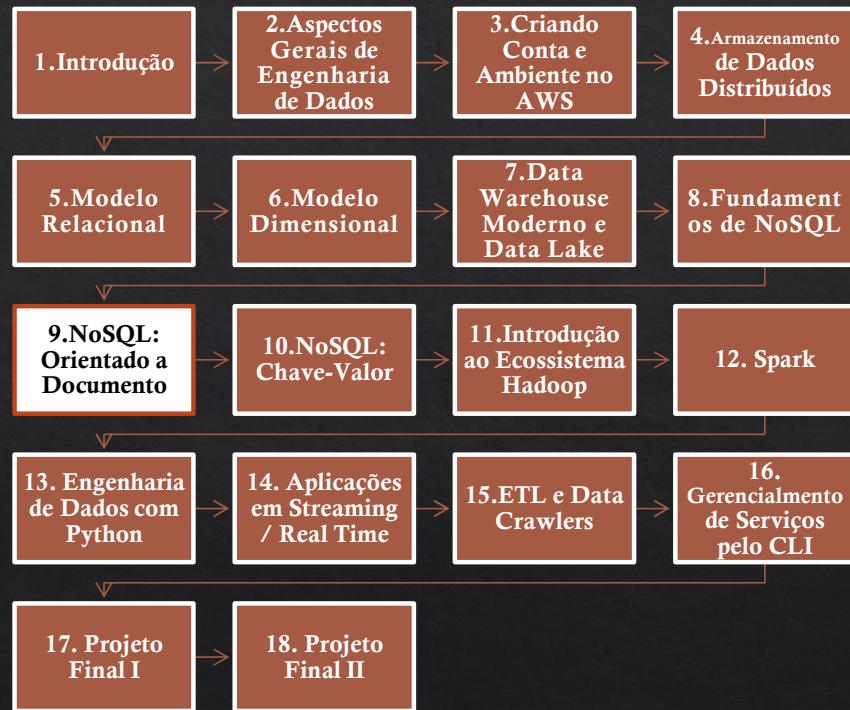


8. Fundamentos de NoSQL



- Introdução
- KVP
- Orientado a Colunas
- Orientado a Documentos
- Grafos

9.NoSQL: Orientado a Documento



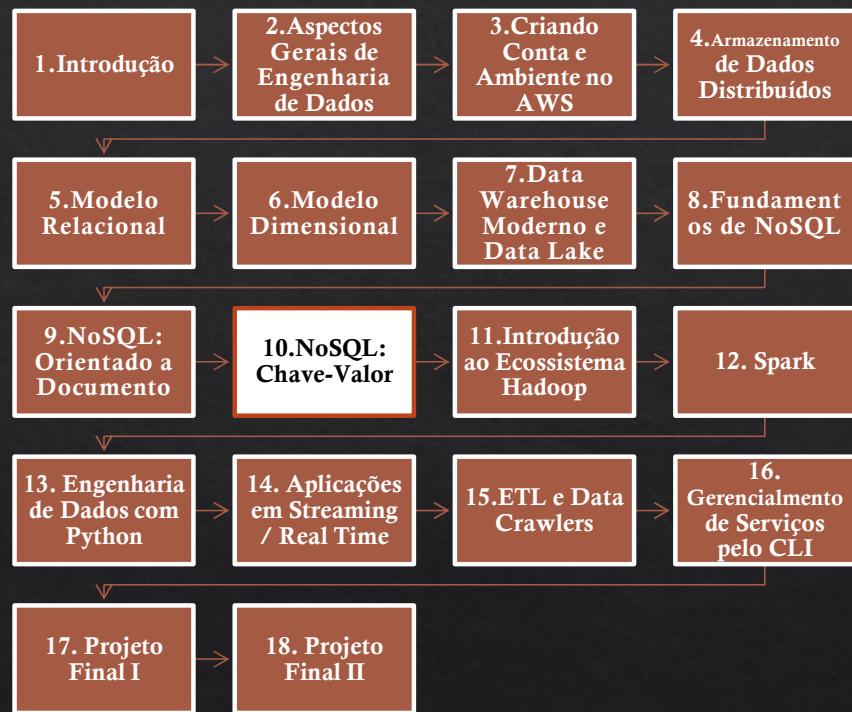
mongoDB®



Amazon
EC2

- Instalação
- Criando Coleções
- Inserindo Documentos
- Buscando Documentos
- Importando Dados

10.NoSQL: Chave-Valor



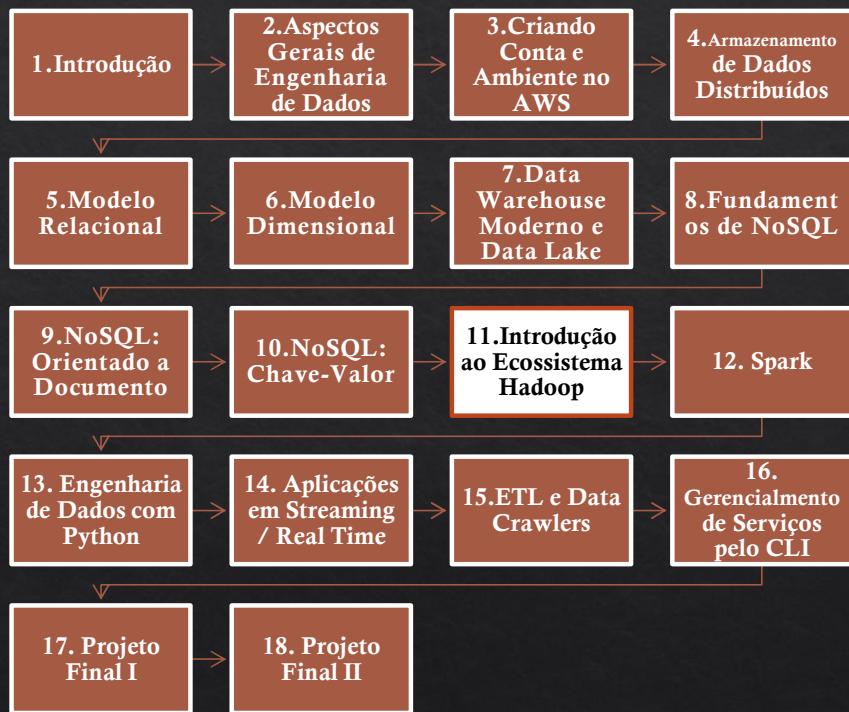
redis



Amazon
EC2

- Instalação
- Strings
- Hash
- List
- Set
- Transações

11.Introdução ao Ecossistema Hadoop



- Map Reduce
- Hdfs
- Yarn

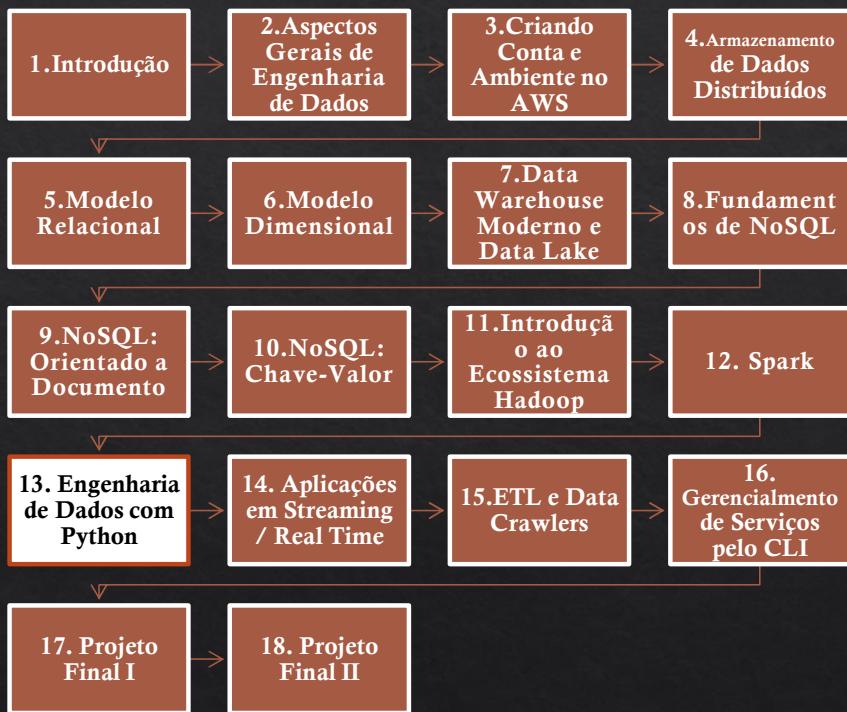
12. Spark



databricks

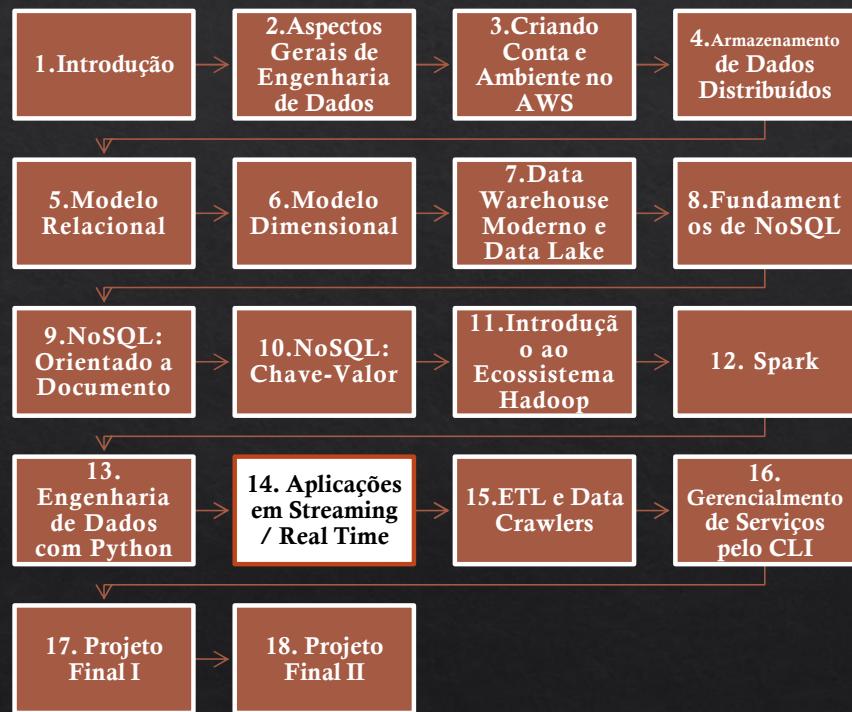
- Criando Cluster
- Importando Dados
- Explorando Dados
- Machine Learning

13. Engenharia de Dados com Python



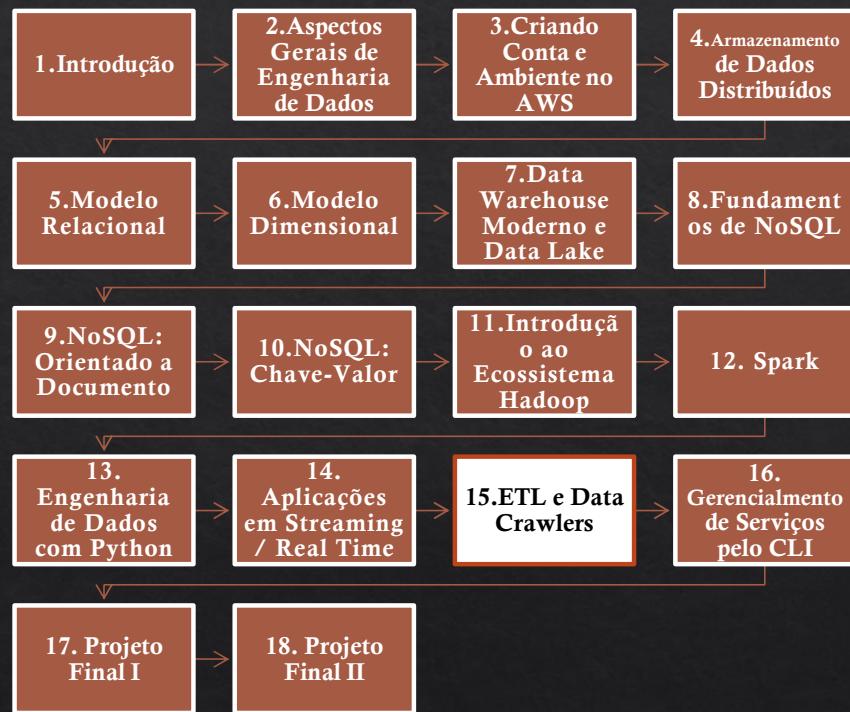
- Criando Instância do RDS
- Criando Banco de Dados
- Criando Objetos
- Inserindo e Consultando Dados

14. Aplicações em Streaming / Real Time

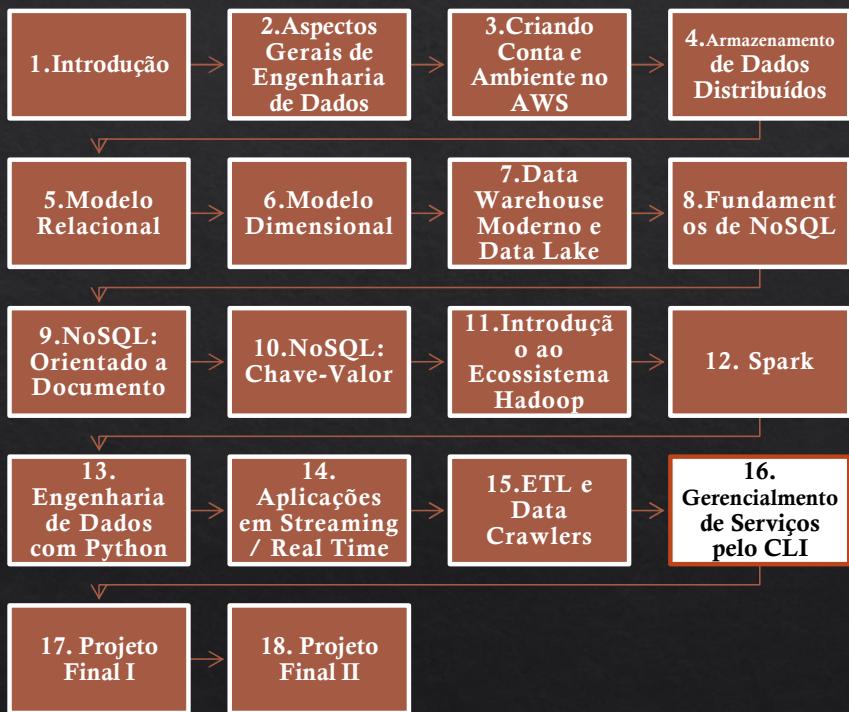


- Criar Streaming de Dados
- Criando Consumidor e Produtor
- Criando Serviço de Entrega

15.ETL e Data Crawlers

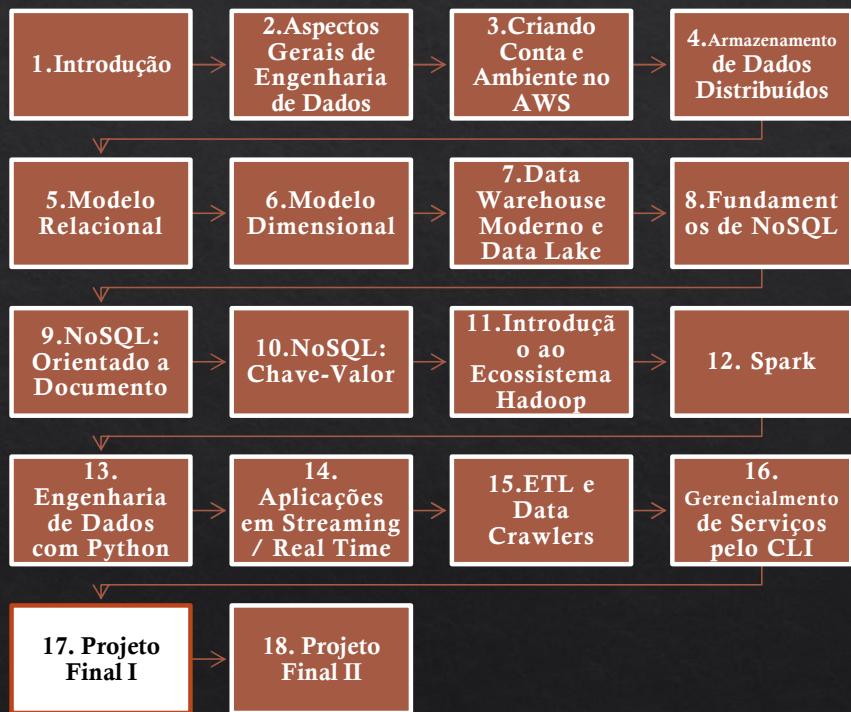


16. Gerencialmento de Serviços pelo CLI

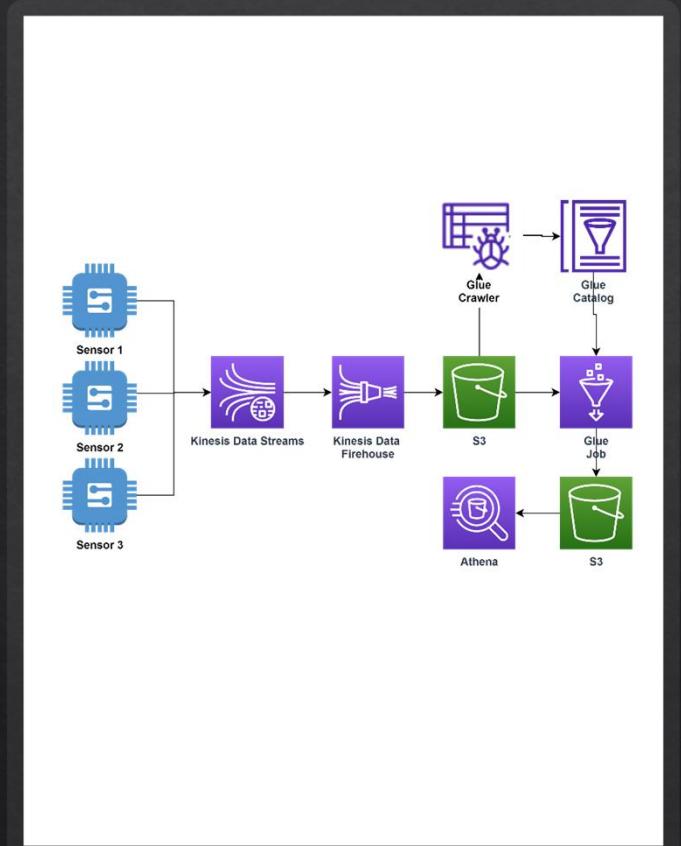


- O que é
- Configurando um Bucket
- Copiando e Movendo Objetos
- Sincronização

17. Projeto Final I



18. Projeto Final II



Material para Baixar

- ◊ Slides
- ◊ Dados
- ◊ Scripts
- ◊ Diagramas
- ◊ Aula de Orientação (texto)
- ◊ Alguns materiais específicos poderão ser baixados
por http diretamente no ambiente!

