**1. Business Intelligence (BI):**

* **Definição**: BI é um conjunto de processos e ferramentas que processam e transformam dados brutos em informações valiosas, apoiando decisões e fornecendo insights estratégicos.
* **Objetivo**: Facilitar a tomada de decisões informadas e baseadas em dados ao transformar dados em conhecimento útil.

**2. Etapas do Processo de BI:**

* **Semana 1**: Introdução ao BI e ao Power BI.
* **Semana 2**: Documentação de requisitos para BI.
* **Semana 3**: Desenvolvimento de ETLs (Extração, Transformação e Carga).
* **Semana 4**: Revisão de conceitos de BI e aplicação prática.

**3. Papéis na Área de Dados:**

* **Engenheiro de Dados**: Trabalha com ferramentas como Big Data, SQL, Spark, Airflow, entre outros, focando na infraestrutura de dados.
* **Analista de Dados**: Habilidade em análise de dados com Python, PowerBI, Tableau, Excel; transforma dados em insights práticos.
* **Cientista de Dados**: Foco em modelagem preditiva, estatística e machine learning, usando Python e ferramentas estatísticas.
* **Report Designer**: Especializa-se na criação de relatórios e dashboards utilizando Figma, Tableau, e Superset para visualização de dados.

**4. ETL (Extrair, Transformar, Carregar):**

* **Extração**: Coleta de dados de diversas fontes (bancos de dados, APIs, planilhas).
* **Transformação**: Limpeza de dados, tratamento de valores nulos, agregação (soma, média, mediana, etc.) e classificação.
* **Carga**: Armazenamento em Data Warehouses, que são bases para análises de dados complexas.

**5. OLAP (Processamento Analítico) vs OLTP (Processamento Transacional):**

* **OLAP**:
  + Projetado para análises complexas e leituras intensas.
  + Utilizado em dashboards de BI e Data Warehouses, evita consumir dados diretamente de bases OLTP.
* **OLTP**:
  + Projetado para operações em tempo real, como inserções e atualizações frequentes.
  + Essencial para sistemas que requerem consistência de dados e resposta rápida.

**6. Modelagem de Dados em Data Warehouses:**

* **Esquema Estrela (Star Schema)**:
  + Estrutura comum de modelagem com tabelas de fatos e dimensões.
  + **Tabela de Fatos**: Armazena eventos numéricos (ex.: vendas) e responde a “o quê, quanto e quando”.
  + **Tabela de Dimensões**: Fornece contexto (ex.: tempo, cliente, produto) e permite explorar dados por diferentes ângulos.
  + **Exemplos**: fVendas, fPagamentos para fatos; dProduto, dCliente para dimensões.

**7. Tipos de Processamento de Dados:**

* **Batch (Processamento em Lote)**:
  + Executa tarefas em grandes volumes de dados periodicamente.
  + Ferramentas: Spark e Airflow para ingestão e orquestração de dados.
* **Streaming (Processamento Contínuo)**:
  + Processa dados em tempo real conforme chegam ao sistema.
  + Ferramentas: Kafka, Spark e Google Dataflow (GCP), permitindo respostas rápidas a eventos.

**8. Tipos de Carga de Dados no ETL:**

* **Carga Incremental**: Extrai apenas dados alterados desde a última extração, economizando tempo e recursos.
* **Carga Completa**: Extrai todos os dados da fonte, útil para dados que mudam pouco, porém é mais lenta e consome mais recursos.

**9. Planejamento de BI e Documentação de Requisitos:**

* **Objetivos de Negócio**: Entender e definir os objetivos de negócio do cliente, focando em resultados (ex.: reduzir tempo de processamento de pedidos).
* **Indicadores de Desempenho (KPIs)**: Identificar indicadores que medem o sucesso dos objetivos (ex.: tempo médio de processamento de pedidos, Net Promoter Score).
* **Métricas**: Definir métricas derivadas dos indicadores (ex.: taxa de uso de recursos, percentual de estoque disponível).
* **Fontes de Dados**: Identificar fontes de dados para cada indicador (ex.: bancos de dados, logs de sistema, planilhas).
* **Data Driven Canvas**: Ferramenta visual para documentação rápida e colaborativa de requisitos.
* **BUS Matrix**: Matriz que relaciona indicadores com dimensões, facilitando a visualização e comunicação com o cliente.

**10. Transformações de Dados no ETL:**

* **Limpeza e Tratamento**: Remoção de dados nulos, correção de erros e tratamento ortográfico.
* **Agregação de Dados**: Realização de cálculos como soma, média, e desvio padrão.
* **Classificação**: Definir status (ex.: status = 1 -> pendente, status = 2 -> concluído).
* **Cálculos e Agregações**: Cálculo de valores com base em regras específicas, como o valor da venda multiplicado pela quantidade.

**11. Ferramentas e Tecnologias Utilizadas:**

* **Ferramentas de BI**: PowerBI, Tableau, QlikView.
* **Linguagens de Programação e Scripts**: Python e SQL.
* **Ferramentas para ETL e Orquestração**: Airbyte, Airflow, Jenkins.