# **Capítulo 3: Prática de termos e conceitos**

* Vêm de uma variedade de fontes como pessoas imagens text sensores e sites: **dados**
* Tipos diferentes de dados: **variedade**
* Dados que entram rapidamente: **velocidade**
* Grande volume de dados: **volume**
* Criada para manter os dados mais próximas da origem para pré-processamento: **computação em neblina**
* Uma coleção de data centers ou grupos de servidores conectados: **a nuvem**
* Software de fonte aberta criado para distribuir e processar grandes conjuntos de dados em quantidades menores: **Hadoop**
* Grande volume de dados dividido em partes menores e armazenados em diferentes servidores: **banco de dados distribuído**
* Dados criados por aplicativos que usam entrada no formato "fixo", como planilhas ou formulários médicos: **dados estruturados**
* Um tipo de arquivo de texto simples que usa vírgulas para separar colunas em uma tabela de dados: **CSV - Valores Separados por Vírgula**
* Dados gerados em um estilo "livre", como áudio vídeo páginas na Web e tweets: **dados não estruturados**
* O processo de transformar dados brutos em informações significativas, detectando padrões e relacionamentos em grandes conjuntos de dados: **mineração de dados**
* O processo de inspeção limpeza transformação e modelagem de dados para descobrir informações úteis: **análise de dados**