



Análise de Dados com Linguagem Python

Projeto 9

Definição do Problema

Análise de Dados com Linguagem Python

Projeto 9

Pré-Processamento de Dados de Texto Extraídos de Data Lake



Copyright © Data Science Academy



Toda empresa que tem presença na internet pode estar interessada em extrair insights de comentários de usuários sobre seus produtos ou serviços, sobre concorrentes, sobre o mercado ou sobre suas próprias preferências.

Neste projeto teremos em mãos dados de texto simulando comentários de usuários. Dados de texto são dados em formato não estruturado e seria muito difícil usar um Data Warehouse para armazenar os dados e realizar análise. Por essa razão, usaremos um Data Lake, onde os dados serão carregados em seu formato bruto e então faremos o pré-processamento dos dados em tempo de análise para entregar insights aos tomadores de decisão.