O que é ciência de dados.

São técnicas para produção de informação e conhecimento a partir de dados produzidos.

Atualmente temos uma produção de dados em grande volume (Big Data). Temos sistemas produzindo dados, sensores, internet, Inteligência das Coisas (IOT), máquinas em linhas de produção, celulares e etc.

Dados são estratégicos para as empresas, pois propiciam mudanças sociais, tais como: curas de doenças, viver mais, qualidade de vida, aquecimento global, etc.

Devemos distinguir o que é produção e análise de dados. Produção são todos os meios que produzem dados (câmeras, dispositivos eletrônicos, teclados, satélites, celulares, etc.).

Apenas uma pequena parte destes dados são analisados, permitindo realizar análises que geram dados que propiciam um diferencial competitivo (lucratividade, reduzir a fome, melhorar a qualidade de vida, etc.).

O nosso estudo será dividido em duas partes principais: Explorar e Prever.

Na etapa de Exploração, nós buscaremos conhecer os dados. Na Exploração dos dados nós já iremos conhecer bastante sobre os dados.

Na etapa de Previsão iremos estimas o que deverá acontecer no futuro. Ex.: Quanto uma empresa irá vender no futuro de acordo um determinado investimento, qual será o seu crescimento estimado.