

Abstract:

In this data-driven analysis project, we explore a large insurance dataset to uncover the key drivers of insurance costs. Through descriptive analyses, correlation analysis, and regression modelling, we identify the factors that strongly influence insurance charges, including variables such as age, BMI, smoking status, and number of children. Furthermore, we develop predictive models to estimate charges for new patients, providing valuable insights for insurance companies to understand the factors contributing to insurance costs and make informed pricing decisions.