Chapter 6
Datalyse:
Integrating Statistics and Computation for Streamlined Financial Data Analysis, Linear Model Building, and Time Series Analysis

Tejal Alwyn Menezes
https://orcid.org/0009-0008-8326-7822
Christ University, India

Tina Dokaniya
https://orcid.org/0009-0001-2588-5875
Christ University, India

Elainne William Desouza
Christ University, India

Ishaan Rai
Christ University, India

A. Vijayalakshmi
Christ University, India

ABSTRACT

Having robust financial markets is important for economic development. Using statistical analysis and data visualization on financial data helps identify trends and understand the underlying factors affecting various aspects of the economy. This allows policymakers to make better decisions. In this chapter, we aim to in-

DOI: 10.4018/979-8-3693-6215-0.ch006

troduce 'Datalyse,' an R programming developed website capable of performing various statistical tests and analyses. 'Datalyse' intends to solve this complexity by integrating statistics and web development. It accepts user-input data, computes the summary statistics, validates assumptions, and performs t-tests, F-tests, and z-tests. It checks the basic assumptions and generates a regression model, both simple and multiple linear regression. Additionally, residual analysis is performed to check model adequacy. The chapter demonstrates the application of 'Datalyse' in analyzing data and deriving insights through statistical methods, mainly regression analysis.

INTRODUCTION

Data analysis involves inspecting, cleansing, transforming, and modeling data. The goal is to discover helpful information that will support decision-making. To make proper inferences, the statistical theory for various concepts states certain assumptions, terminologies, and aspects of the data that define the kind of statistical inference one can use. Most statistical concepts require certain assumptions for the analysis results to be accurate, which forms the basis of hypothesis testing.

Hypothesis testing is a fundamental aspect of statistical analysis. It is a structured method to make informed decisions based on the data and provides scientific evidence to prove or disprove the theories about the data. Some methods to do this include performing tests such as T-test, F-test, and Z-test. Each test helps test different assumptions and requirements to obtain adequate and accurate results. For example, a one-sample t-test can be used to test if the value of the sample mean is significantly different from the known value of the population mean. Two sample t-tests can also be used to compare if the two groups' means are statistically different.

Meanwhile, the F-test can compare various models and determine if adding more predictors significantly improves the model fit. Moreover, the financial sector is the one with a high-risk factor. A structured method to make data-based decisions can be beneficial in various situations, such as checking if price changes significantly impact sales.

Once the assumptions are satisfied, various modeling techniques can be used to model the data. One such technique is linear regression, widely used in multiple fields such as economics, social sciences, and finance to predict and model relationships between variables. After modeling the data, the other assumptions, such as the assumption for normality and heteroscedasticity of errors, need to be tested to ensure that the model generated is suitable for the given dataset. Another such technique is time series analysis, which enables studying past data and helps forecast future possible values using these past values.