Introduction

A Foray into Financial Data Analysis

Contents

Note: This document is knitted to pdf for easier viewing.

Introduction

This group project focuses on exploring a chosen application domain through the lens of data science. The primary objective is to gain hands-on experience with data science techniques while developing collaborative working skills and expanding programming knowledge. We have chosen to focus on finance data, specifically exploring areas such as credit card fraud detection, financial time series analysis, and other relevant financial datasets. The primary objective is to gain practical experience with data science techniques within the financial domain, while fostering collaborative teamwork and enhancing programming skills. The exploration of this application domain will guide our choice of dataset for analysis, which will be carried out in Assessment 1.

Our challenge involves conducting a literature review of data science resources that aid in analyzing finance data. This will include:

- examining the types of financial data that exist: We will search for sources of data and assess their suitability for use to explore a classification or regression problem.
- exploring the key resources available for analysis: We will perform exploratory data analysis to understand the data structure. We will use resources to explore and understand methods to analyse such data.
- judge suitability of such datasets: We will consider the statistical methods that are used to analyse the data and check whether they fall within the scope of Assessment 1. We will consider the ways of evaluating such methods, taking care to adapt the methods as needed to the particular dataset.

Additionally, we will explore the use of general data science resources, assessing their applicability to financial datasets. This will help us build a repertory of resources that we can then use during the next assessment, and even further beyond. Our team will identify relevant resources, run and modify code examples, and create original visualizations, culminating in a structured report that synthesizes our findings and insights.

References

1. Project brief, accessible on Github. This link was accessed 18:00 on 26/09/2024.