ADS506 Assignment 1.2 sp500_stocks

Gabriel E. Mancillas Gallardo

2024-10-28

Assignment: Propose a Time Series Dataset for Your Final Project ### Data Source

The data source is the S&P 500 stock data from Kaggle. The dataset contains the stock prices of the S&P 500 companies from 2013 to 2018. The dataset contains the following columns:

```r
#library(<PACKAGE_DEPENDENCIES>)
library(readr)
library(dplyr)
library(tsibble)

# Load data
sp500_data <- read_csv("/Users/gabrielmancillas/Documents/GitHub/S-P-500/data/sp500_stocks_filtered.csv

# Show data
head(sp500_data)
```

```
## # A tibble: 6 x 8
##   Date       Symbol `Adj Close` Close  High   Low  Open  Volume
##   <date>     <chr>        <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
## 1 2022-01-03 MMM           131.  149.  150.  147.  149. 2309117
## 2 2022-01-04 MMM           132.  151.  152.  149.  149. 3016551
## 3 2022-01-05 MMM           132.  150.  152.  148.  148. 3531070
## 4 2022-01-06 MMM           131.  149.  152.  148.  151. 2996458
## 5 2022-01-07 MMM           132.  150.  151.  148.  149. 3349039
## 6 2022-01-10 MMM           130.  148.  150.  148.  150. 2772328
```

```r
#total number of rows and columns
dim(sp500_data)
```

```
## [1] 356124      8
```

Data Source

Include public links to data if it is too large to upload (do not upload datasets larger than 50MB).

Importing the Data

Time Series Plot

```r
#time series plot...
```

Discussion

...