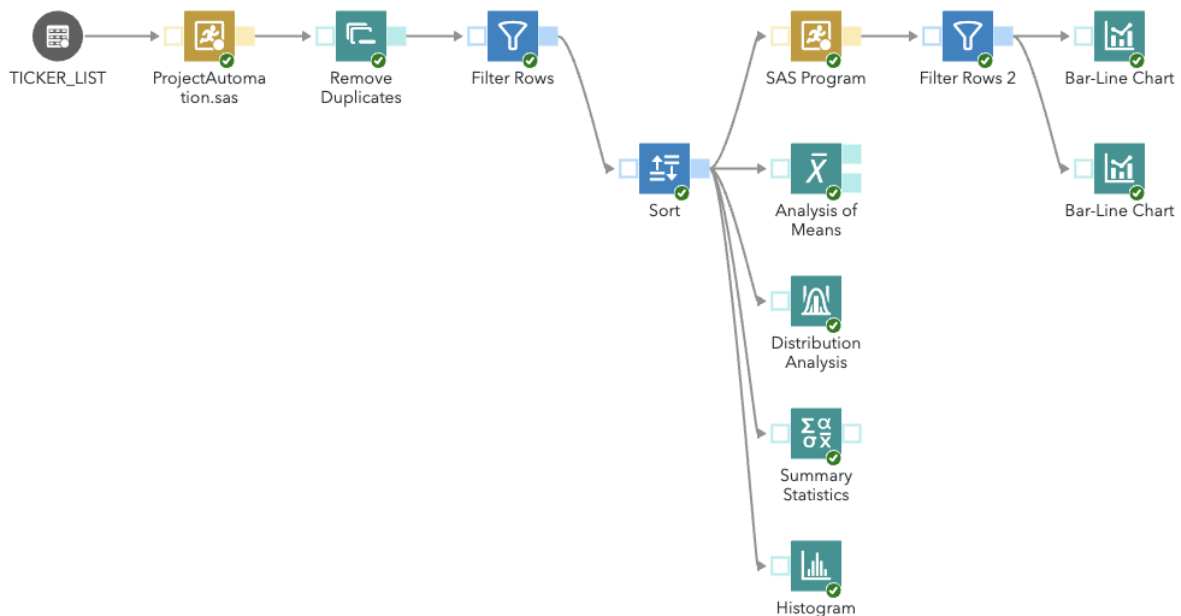


Automatyzacja procesów analitycznych w SAS

Projekt z użyciem danych akcyjnych Yahoo Finance



Szymon Czuszek

Uniwersytet Ekonomiczny w Katowicach

Wprowadzenie	4
I. Źródło Yahoo Finance	5
II. Języki w SAS Viya	7
III. Analiza w SAS Viya	9
1. Automatyzacja pobierania danych akcyjnych z Yahoo Finance	9
1.1. Wyjaśnienie celu szczegółowego	9
1.2. Opis procesu uzyskania danych wynikowych (kody)	9
1.3. Wykresy i tabele z danymi	12
1.4. Wnioski uzyskane z przeprowadzonej analizy	12
2. Analiza średniej objętości transakcji akcji	12
2.1. Wyjaśnienie celu szczegółowego	12
2.2. Opis procesu uzyskania danych wynikowych (kody)	13
2.3. Wykresy i tabele z danymi	14
2.4. Wnioski uzyskane z przeprowadzonej analizy	14
3. Obliczanie średniej ruchomej dla cen akcji	16
3.1. Wyjaśnienie celu szczegółowego	16
3.2. Opis procesu uzyskania danych wynikowych (kody)	16
3.3. Wykresy i tabele z danymi	17
3.4. Wnioski uzyskane z przeprowadzonej analizy	17
4. Obliczanie zmienności cen akcji	18
4.1. Wyjaśnienie celu szczegółowego	18
4.2. Opis procesu uzyskania danych wynikowych (kody)	19
4.3. Wykresy i tabele z danymi	20
4.4. Wnioski uzyskane z przeprowadzonej analizy	20
5. Wizualizacja i analiza zależności między wolumenem a ceną zamknięcia akcji	21
5.1. Wyjaśnienie celu szczegółowego	21
5.2. Opis procesu uzyskania danych wynikowych (kody)	21
5.3. Wykresy i tabele z danymi	22
5.4. Wnioski uzyskane z przeprowadzonej analizy	22
6. Automatyzacja analizy ilości transakcji akcyjnych	23
6.1. Wyjaśnienie celu szczegółowego	23
6.2. Opis procesu uzyskania danych wynikowych (kody)	23
6.3. Wykresy i tabele z danymi	24
6.4. Wnioski uzyskane z przeprowadzonej analizy	25
7. Analiza popularności spółki w porównaniu do innych	25
7.1. Wyjaśnienie celu szczegółowego	25

7.2. Opis procesu uzyskania danych wynikowych (kody)	26
7.3. Wykresy i tabele z danymi	26
7.4. Wnioski uzyskane z przeprowadzonej analizy	27
8. Wizualizacja i automatyzacja analizy rozkładów cen zamknięcia	27
8.1. Wyjaśnienie celu szczegółowego	27
8.2. Opis procesu uzyskania danych wynikowych (kody)	28
8.3. Wykresy i tabele z danymi	28
8.4. Wnioski uzyskane z przeprowadzonej analizy	29
9. Obliczanie statystyk opisowych cen akcji spółek giełdowych	31
9.1. Wyjaśnienie celu szczegółowego	31
9.2. Opis procesu uzyskania danych wynikowych (kody)	32
9.3. Wykresy i tabele z danymi	32
9.4. Wnioski uzyskane z przeprowadzonej analizy	33
10. Porównawcza analiza cen zamknięcia i wolumenu obrotu spółek	34
10.1. Wyjaśnienie celu szczegółowego	34
10.2. Opis procesu uzyskania danych wynikowych (kody)	35
10.3. Wykresy i tabele z danymi	35
10.4. Wnioski uzyskane z przeprowadzonej analizy	35
Podsumowanie	36
Bibliografia	38

Wprowadzenie

Współczesne podejście do analizy danych oraz automatyzacji procesów biznesowych wymaga zastosowania nowoczesnych narzędzi programistycznych i technologii, które umożliwiają efektywne przetwarzanie i analizowanie dużych zbiorów danych. W ramach niniejszego projektu skupiono się na zastosowaniu narzędzi takich jak Python, SQL, oraz technologii dostępnych w ramach środowiska SAS, w tym SAS Macro, SAS 4GL oraz SAS Flow. Celem projektu jest zautomatyzowanie procesów analizy danych oraz przetwarzania informacji w oparciu o dane finansowe pochodzące z Yahoo Finance.

Yahoo Finance jest jednym z najważniejszych źródeł publicznych danych finansowych, które umożliwia dostęp do notowań akcji, indeksów giełdowych, walut, surowców oraz innych instrumentów finansowych. Dane te, zarówno w czasie rzeczywistym, jak i historyczne, stanowią podstawę do przeprowadzania analiz rynku, prognozowania trendów oraz podejmowania decyzji inwestycyjnych. Platforma ta oferuje wygodny dostęp do szerokiego zakresu danych finansowych, co czyni ją szczególnie wartościowym źródłem do wykorzystania w projektach związanych z automatyzacją analiz oraz procesów biznesowych.

Projekt ten koncentruje się na integracji danych z Yahoo Finance z wykorzystaniem narzędzi programistycznych, które umożliwiają efektywną obróbkę danych, ich analizę, a także automatyzację całego procesu. Python, jako wszechstronny język programowania, umożliwia pobieranie danych z Yahoo Finance, ich przetwarzanie oraz implementację zaawansowanych analiz.

SAS Macro, SAS 4GL oraz SAS Flow stanowią kluczowe narzędzia w zakresie automatyzacji procesów analitycznych, umożliwiając tworzenie dynamicznych programów, uproszczoną manipulację danymi oraz budowanie wizualnych przepływów pracy, które automatyzują procesy analizy i raportowania.

Celem końcowym projektu jest stworzenie zautomatyzowanego rozwiązania, które pozwoli na szybkie i efektywne pobieranie, przetwarzanie i analizowanie danych finansowych z Yahoo Finance, a także generowanie wyników w formie raportów i prognoz. Dzięki wykorzystaniu nowoczesnych narzędzi programistycznych i automatyzacyjnych, projekt ten ma na celu usprawnienie procesów analitycznych i przyspieszenie podejmowania decyzji w oparciu o dane finansowe.

I. Źródło Yahoo Finance

Raport został przygotowany na podstawie danych akcyjnych z Yahoo finance. Yahoo Finance to jedna z najpopularniejszych platform internetowych oferujących dostęp do danych finansowych, w tym informacji o notowaniach akcji, indeksach giełdowych, walutach, surowcach, obligacjach i innych instrumentach finansowych. Platforma ta oferuje dane na żywo, historyczne notowania, analizy techniczne, oraz raporty finansowe firm. Posiada również możliwość pobierania danych za pomocą API, co czyni ją szczególnie interesującym źródłem dla projektów analitycznych i automatyzacji procesów. Yahoo Finance jest doskonałym źródłem danych do projektów, z kilku powodów:

1. **Dostępność szerokiego zakresu danych finansowych:** Yahoo Finance oferuje dane na temat akcji, indeksów giełdowych, walut, obligacji i surowców. Pozwala to na analizę szerokiego wachlarza rynków finansowych, co może być kluczowe w projektach związanych z analizą giełdy, inwestycji czy trendów rynkowych.
2. **Bezpośredni dostęp do danych historycznych:** Platforma udostępnia dane historyczne, co pozwala na przeprowadzanie analiz trendów w czasie, porównań między różnymi okresami oraz prognozowania na podstawie wcześniejszych wyników. Możliwość korzystania z danych historycznych jest niezwykle cenna przy budowie modeli predykcyjnych lub testowaniu teorii rynkowych.
3. **Łatwość integracji z SAS i 4GL:** Dzięki dostępności danych w formacie CSV oraz możliwości korzystania z API, dane z Yahoo Finance mogą być łatwo zintegrowane z SAS oraz innymi narzędziami analitycznymi. W SAS można używać makr i kodu 4GL do automatyzacji pobierania, przetwarzania i analizowania tych danych.
4. **Możliwości analizy technicznej:** Yahoo Finance dostarcza dane potrzebne do przeprowadzania analiz technicznych (np. wykresy świecowe, wskaźniki techniczne), które mogą być użyteczne przy analizie akcji i przewidywaniu ich ruchów na rynku. Możliwość automatyzacji obliczeń tych wskaźników za pomocą makr i kodów w SAS daje możliwość szybszego i bardziej efektywnego przetwarzania dużych zbiorów danych.
5. **Dane w czasie rzeczywistym i możliwość tworzenia powiadomień:** Yahoo Finance oferuje dane na żywo, co umożliwia śledzenie zmieniających się cen akcji oraz reagowanie na zmiany w czasie rzeczywistym. Możliwość automatycznego uruchamiania skryptów w SAS po wykryciu zmian w danych (np. alerty cenowe) zwiększa efektywność analizy i może wspierać decyzje inwestycyjne.

6. **Bezproblemowa automatyzacja procesów:** Zintegrowanie Yahoo Finance z SAS umożliwia automatyczne pobieranie danych, ich przetwarzanie, analizę oraz generowanie raportów lub prognoz bez potrzeby manualnej interwencji. Dzięki temu można skoncentrować się na analizach i wynikach, podczas gdy procesy techniczne są zautomatyzowane.

II. Języki w SAS Viya

Językami programowania użytymi w niniejszym raporcie są: Python, SQL oraz SAS Macro, SAS 4GL, oraz SAS flow. Każdy z tych języków odgrywa istotną rolę w procesie analizy i przetwarzania danych, a ich różnorodność umożliwia wykorzystanie szerokiego spektrum funkcji i technik.

Python: Python to interpretowany, obiektowy język programowania wysokiego poziomu z dynamiczną semantyką, opracowany przez Guido van Rossuma. Pierwotnie został wydany w 1991 roku i zyskał ogromną popularność ze względu na swoją prostotę oraz wszechstronność. Jego nazwa jest ukłonem w stronę brytyjskiej grupy komediowej Monty Python, co odzwierciedla filozofię języka – ma być on łatwy i zabawny w użyciu.

Python jest często wybierany jako pierwszy język programowania przez początkujących ze względu na jego czytelną i intuicyjną składnię, która pozwala skupić się na zrozumieniu podstawowych koncepcji programowania, zamiast na zawiłościach technicznych. Jest powszechnie używany w wielu dziedzinach, w tym w analizie danych, uczeniu maszynowym, web development oraz automatyzacji zadań. Dzięki bogatej bibliotece standardowej oraz licznym zewnętrznym pakietom, Python umożliwia szybkie i efektywne rozwiązywanie różnorodnych problemów programistycznych.

SQL (Structured Query Language): Strukturalny język zapytań (SQL) to język programowania przeznaczony do zarządzania i manipulowania danymi w relacyjnych bazach danych. Relacyjna baza danych organizuje informacje w formie tabelarycznej, gdzie wiersze i kolumny reprezentują różne atrybuty i relacje między danymi.

SQL pozwala użytkownikom na wykonywanie różnorodnych operacji na bazie danych, takich jak przechowywanie, aktualizowanie, usuwanie oraz wyszukiwanie i pobieranie informacji. Język SQL jest niezbędnym narzędziem w zarządzaniu bazami danych, ponieważ umożliwia tworzenie zapytań, które mogą być zarówno proste, jak i niezwykle złożone. Ponadto, SQL jest wykorzystywany do utrzymywania integralności danych oraz optymalizowania wydajności systemów bazodanowych, co jest kluczowe dla sprawnego działania aplikacji opartych na dużych zbiorach danych.

SAS Macro: SAS Macro to język makr w środowisku SAS, który umożliwia automatyzację i parametryzację kodu. Dzięki użyciu makr, możliwe jest generowanie dynamicznych programów, które mogą być dostosowywane do różnych danych i warunków. Makra w SAS pozwalają na tworzenie powtarzalnych bloków kodu, które mogą być łatwo używane w różnych częściach programu, co pozwala na zaoszczędzenie czasu i zmniejszenie

błędów w kodowaniu. Język ten jest szczególnie przydatny przy przetwarzaniu dużych zbiorów danych oraz w zadaniach wymagających wielokrotnego użycia tych samych operacji, co czyni go istotnym narzędziem w projektach automatyzacji procesów analizy danych.

SAS 4GL (Fourth-Generation Language): SAS 4GL to język programowania wysokiego poziomu opracowany specjalnie dla użytkowników pracujących w środowisku SAS. Jego celem jest umożliwienie łatwego przetwarzania danych i generowania raportów, a także umożliwienie użytkownikom pisania programów przy minimalnej konieczności zarządzania szczegółami technicznymi. 4GL różni się od tradycyjnych języków programowania, takich jak SAS Base, tym, że koncentruje się na zapewnieniu prostoty użytkowania i efektywności, eliminując konieczność pisania skomplikowanego kodu. Używa się go głównie do manipulacji danymi, w tym agregacji, transformacji i generowania podsumowań, co czyni go bardzo użytecznym narzędziem w analizie dużych zbiorów danych oraz automatyzacji powtarzających się zadań.

SAS Flow: SAS Flow to wizualne narzędzie wykorzystywane do projektowania i automatyzowania procesów analitycznych w środowisku SAS. Umożliwia użytkownikom tworzenie procesów analizy danych poprzez tworzenie przepływów pracy, w których definiowane są poszczególne kroki analizy, takie jak importowanie danych, ich przetwarzanie, analiza oraz prezentacja wyników. SAS Flow jest szczególnie użyteczny w kontekście projektów, które wymagają automatyzacji powtarzalnych zadań lub procesów przetwarzania danych. Dzięki jego zastosowaniu, procesy mogą być zaprojektowane w sposób graficzny, co sprawia, że nawet osoby bez zaawansowanej wiedzy programistycznej mogą łatwo zbudować i zarządzać skomplikowanymi analizami i procesami biznesowymi.

III. Analiza w SAS Viya

1. Automatyzacja pobierania danych akcyjnych z Yahoo Finance

1.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego kroku jest automatyczne pobieranie danych finansowych z Yahoo Finance, co pozwala na uzyskanie aktualnych informacji o notowaniach akcji w określonym przedziale czasowym.
- ❖ Dzięki temu, proces gromadzenia danych staje się bardziej efektywny i niezależny od ręcznych interwencji.
- ❖ Kod w języku Python, SAS 4GL, oraz SAS Macro:

1.2. Opis procesu uzyskania danych wynikowych (kody)

```
%macro open_csv_as_dataset(ticker, start_date, end_date, root_path);  
  /* Define the path for the CSV file */  
  %let csv_path = &root_path/&ticker..&start_date..&end_date..csv;  
  
  /* Check if the file exists */  
  %if %sysfunc(fileexist(&csv_path)) %then %do;  
    /* Import the CSV file as a SAS dataset */  
    proc import datafile="&csv_path"  
      out=work.&ticker._data  
      dbms=csv  
      replace;  
      getnames=yes; /* Use the first row as column names */  
      datarow=4; /* Skip the first two rows */  
    run;  
  
    /* Rename the first column to Date */  
    data work.&ticker._data;  
      set work.&ticker._data;  
      rename Price = Date; /* Rename the first column (Price) to Date */  
    run;
```

```

        /* Print the first few rows of the dataset to check */
        proc print data=work.&ticker._data(obs=10);
        run;
    %end;
    %else %do;
        %put ERROR: The file &csv_path does not exist.;
    %end;
%mend;

%let ticker = MSFT;
%let start_date = 2022-01-01;
%let end_date = 2023-01-01;
%let                                     root_path                                     =
/export/viya/homes/szymon.czuszek@edu.uekat.pl/casuser/Automatyzacja_procesow/dane_z
rodlowe/yfinance;

PROC PYTHON;
submit;
import yfinance as yf
import os

# Fetch the parameters from SAS macro variables
ticker = SAS.symget("ticker")
start_date = SAS.symget("start_date")
end_date = SAS.symget("end_date")
root_path = SAS.symget("root_path")

# Download the data for the specified ticker and date range
data = yf.download(ticker, start=start_date, end=end_date)

# Define the export path dynamically based on the root path, ticker, and date range
export_path = os.path.join(root_path, f'{ticker}.{start_date}.{end_date}.csv')

# Create the directory if it does not exist

```

```

os.makedirs(os.path.dirname(export_path), exist_ok=True)

# Save the data to CSV
data.to_csv(export_path)

# Print the first few rows to confirm data retrieval
print(data.head())
endsubmit;
RUN;

%open_csv_as_dataset(&ticker, &start_date, &end_date, &root_path);

```

❖ Wyjaśnienie technikaliów:

- Makro służy do zaimportowania pliku CSV jako zestawu danych SAS. Jego kluczowe elementy to:
 - Parametry makra:
 - ticker: Symbol giełdowy (np. MSFT dla Microsoftu).
 - start_date i end_date: Okres dat, dla którego dane giełdowe są pobierane.
 - root_path: Ścieżka bazowa, gdzie plik CSV będzie zapisany.
- Logika makra:
 - Definicja ścieżki pliku CSV.
 - Sprawdzenie, czy plik istnieje.
 - Import pliku CSV do SAS: Jeśli plik istnieje.
 - Dane są importowane do zestawu danych w bibliotece work.
 - Wyświetlany jest komunikat o błędzie, jeśli plik nie istnieje.
- Blok PROC PYTHON: Część w Pythonie jest odpowiedzialna za pobranie danych giełdowych z biblioteki yfinance i zapisanie ich jako pliku CSV.
- Logika PROC PYTHON:
 - Pobieranie zmiennych SAS w Pythonie z SAS.symget()
 - Pobranie danych z yfinance
 - Tworzenie ścieżki i zapisywanie danych do CSV
- Wywołanie makra

- Po wykonaniu bloku Pythona, makro `open_csv_as_dataset` jest uruchamiane, aby załadować data set do SAS.

1.3. Wykresy i tabele z danymi

7 Jan 2025, 18:26:51

Log Results Output Data (1)

MSFT_DATA Table rows: 251 Columns: 6 of 6 Rows 1 to 200

Enter expression

	Date	Close	High	Low	Open	Volume
1	2022-01-03	326.94082642	330.1150092	322.08676725	327.52683535	28865100
2	2022-01-04	321.33474731	327.38034718	318.51215195	327.01895363	32674300
3	2022-01-05	308.99935913	318.46330912	308.60869646	318.25818663	40054300
4	2022-01-06	306.55764771	311.26521137	304.22338857	305.8446668	39646100
5	2022-01-07	306.71398926	309.1165933	302.85612412	306.82140884	32720000
6	2022-01-10	306.93856812	307.37808225	297.58206745	302.27007913	44289500

1.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Automatyzacja pobierania danych pozwala na szybkie pozyskiwanie informacji o akcjach w wybranym okresie, co jest kluczowe w dynamicznych rynkach finansowych. Proces ten eliminuje konieczność ręcznego wprowadzania danych, co zmniejsza ryzyko błędów i oszczędza czas.
- Wystarczy zmienić tylko zmienne `ticker`, `start_date`, `end_date`, aby otrzymać obrobione dane dla innej spółki i określonego zakresu dat.

2. Analiza średniej objętości transakcji akcji

2.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego kroku jest obliczenie średniej objętości transakcji dla danej akcji, co pozwala na ocenę popularności i aktywności rynkowej w danym okresie. Jest to istotne w analizach związanych z płynnością rynku.
- ❖ Informacje te są przydatne z kilku powodów:
 - Płynność rynku: Wysoka średnia objętość obrotu wskazuje na dobrą płynność akcji, co zmniejsza ryzyko zatorów podczas zawierania transakcji i pozwala inwestorom szybko nabywać lub sprzedawać akcje.

- Sentyment rynkowy: Skoki w wolumenie transakcji mogą wskazywać na ważne wydarzenia rynkowe, takie jak publikacja wyników finansowych, zmiany w strategii firmy, decyzje regulacyjne czy reakcje na wydarzenia makroekonomiczne.
- Identyfikacja trendów: Analiza objętości w czasie pozwala na zidentyfikowanie trendów lub okresów o podwyższonej aktywności, co jest kluczowe w strategiach inwestycyjnych.

❖ Kod w języku SAS 4GL i SAS Macro:

2.2. Opis procesu uzyskania danych wynikowych (kody)

```
%macro generate_volume_chart(dataset);
    /* Sprawdzenie, czy przekazano parametr zbioru danych */
    %if &dataset=%then
        %do;
            %put ERROR: Dataset parameter is missing.;
            %return;
        %end;

    /* Ustawienie dynamicznego tytułu dla wykresu */
    title "Average Volume for Stock: &dataset";

    PROC CHART DATA=&dataset;
        /* Generowanie poziomego wykresu słupkowego dla średniej objętości
transakcji */
        HBAR DATE / SUMVAR=Volume TYPE=MEAN;
    RUN;

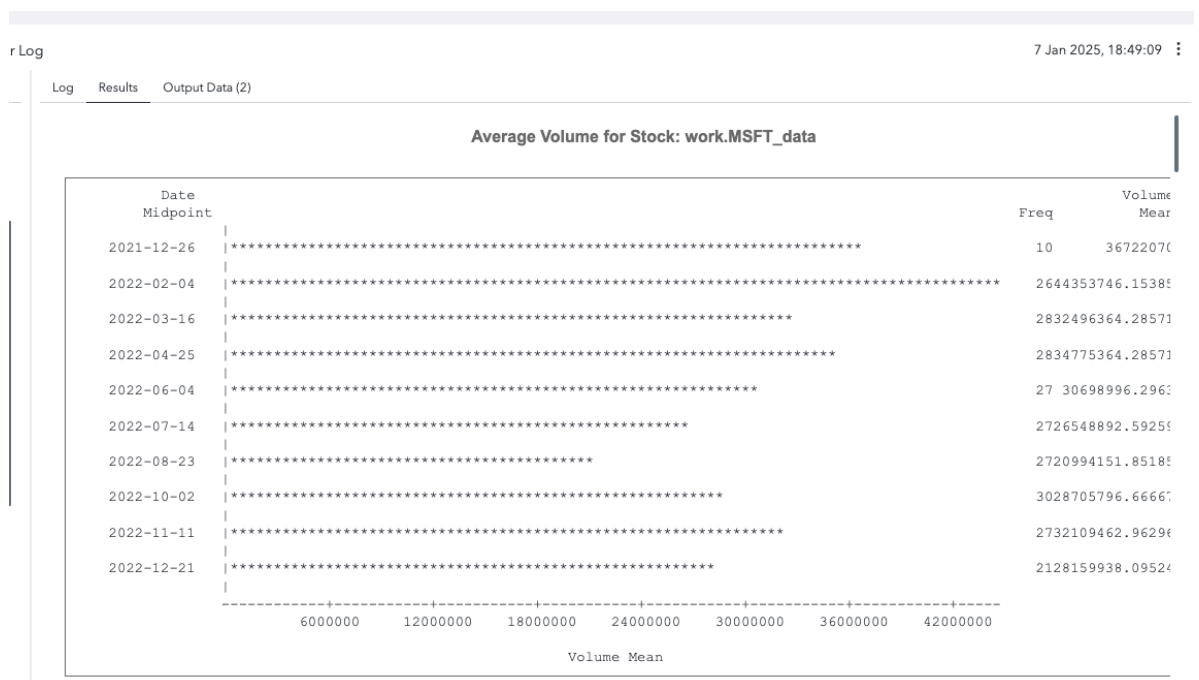
    /* Wyczyszczenie tytułu po zakończeniu procedury */
    title;
%mend generate_volume_chart;

%generate_volume_chart(work.&ticker._data);
```

❖ Wyjaśnienie technikaliów:

- Makro w SAS to rodzaj automatyzacji pozwalający wykonywać powtarzalne operacje na danych.
- Jeśli użytkownik nie poda nazwy zbioru danych, makro wypisze komunikat błędu (ERROR: Dataset parameter is missing.).
- Dynamicznie ustawia tytuł wykresu na podstawie nazwy przekazanego zbioru danych (&dataset).
- PROC CHART: Procedura SAS, która generuje wykresy. W tym przypadku wykres słupkowy (HBAR) przedstawia średnią wartość objętości transakcji.
- Po zakończeniu działania procedury PROC CHART, usuwa tytuł ustawiony wcześniej. Jeśli nie zostanie usunięty, tytuł pozostanie ustawiony globalnie i może pojawić się w innych wykresach.
- Jeśli użytkownik chce wygenerować wykres dla danych akcji Microsoftu znajdujących się w zbiorze work.MSFT_data, wystarczy wywołać makro w następujący sposób: %generate_volume_chart(work.MSFT_data);

2.3. Wykresy i tabele z danymi



2.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Zastosowanie w automatyzacji:

- Uniwersalność: Makro można stosować do dowolnego zbioru danych, wystarczy podać jego nazwę jako parametr.
 - Łatwość analizy: Automatycznie generowane wykresy eliminują konieczność ręcznego tworzenia wizualizacji.
 - Skalowalność: Można zautomatyzować analizę wielu tickerów akcji, wywołując makro dla każdego z nich w pętli.
- Na podstawie wykresu, który pokazuje średni wolumen obrotu dla akcji Microsoftu (work.MSFT_data) w różnych przedziałach dat, można wyciągnąć następujące wnioski:
- Wolumen obrotu: Średni wolumen obrotu różni się w poszczególnych okresach czasu. Maksymalne wartości osiągnęte są w niektórych przedziałach, szczególnie w 2022 roku, a inne przedziały wykazują niższe średnie wartości.
 - Sezonowość lub zmienność: Wartości wolumenu nie są równomierne, co może wskazywać na sezonowe trendy lub zdarzenia rynkowe wpływające na aktywność handlową.
 - Najwyższa średnia: Najwyższą średnią wolumenową widać w okresie około 2022-03-16 (ok. 28,3 mln), co może być związane z określonymi wydarzeniami finansowymi lub gospodarczymi.
 - Ogólny trend: Średni wolumen obrotu spada w drugiej połowie 2022 roku. Wartość z końca grudnia (2022-12-21) jest znacznie niższa niż na początku roku.
 - Zdarzenia zewnętrzne: Aby dokładniej zrozumieć przyczyny tych zmian, można przeanalizować dane historyczne i wydarzenia, które mogły wpływać na wolumen obrotu, np. publikacje wyników finansowych, decyzje związane z polityką monetarną czy makroekonomiczne zmiany.
- Automatyzacja generowania wykresów pozwala na szybkie wizualizowanie danych o objętości transakcji, co umożliwia efektywną ocenę aktywności na rynku bez konieczności ręcznego tworzenia wykresów. Dzięki temu proces analizy staje się bardziej skalowalny i czasowo efektywny.

3. Obliczanie średniej ruchomej dla cen akcji

3.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego celu jest obliczenie średniej ruchomej (Moving Average) dla cen akcji w określonym oknie czasowym.
- ❖ Średnia ruchoma jest wykorzystywana do wygładzania zmienności cen akcji i identyfikacji trendów rynkowych.
- ❖ Kod SAS 4GL i SAS Macro:

3.2. Opis procesu uzyskania danych wynikowych (kody)

```
%macro moving_average(dataset, window);  
    DATA &dataset._moving_avg;  
        SET &dataset;  
        MA_Close=MEAN(LAG1(Close),    LAG2(Close),    LAG3(Close),  
LAG4(Close),  
        LAG5(Close));  
        FORMAT MA_Close 8.2;  
    RUN;  
  
    PROC PRINT DATA=&dataset._moving_avg;  
        TITLE "Stock Prices with &window-Day Moving Average for  
&dataset";  
        VAR Date Close MA_Close;  
    RUN;  
  
%mend moving_average;  
  
%moving_average(work.&ticker._data, 5);
```

- ❖ Wyjaśnienie technikaliów:
 - Makro SAS Służy do obliczania średniej kroczącej dla zestawu danych oraz generowania tabeli wynikowej z wyliczonymi wartościami.
 - Tworzy nowy zbiór danych o nazwie bazującej na wejściowym zbiorze (dodaje _moving_avg do nazwy).

- W przypadku, gdy dataset=work.MSFT_data, nowy zbiór danych będzie miał nazwę work.MSFT_data_moving_avg.
- Tworzy nową zmienną MA_Close, która oblicza średnią kroczącą z 5 poprzednich dni.
- MEAN(...): Oblicza średnią arytmetyczną podanych wartości. Jeśli dane są brakujące (np. w pierwszych obserwacjach, gdzie nie ma pełnych 5 poprzednich dni), wynik również będzie brakujący.
- FORMAT MA_Close 8.2;: Określa format zmiennej MA_Close. Wartość ma 8 znaków szerokości i 2 miejsca po przecinku.
- DATA=&dataset._moving_avg określa, że wydruk dotyczy zbioru danych utworzonego w poprzednim kroku.

3.3. Wykresy i tabele z danymi

Log 7 Jan 2025, 18:49:09

Log Results Output Data (2)

Stock Prices with 5-Day Moving Average for work.MSFT_data

Obs	Date	Close	MA_Close
1	2022-01-03	326.94082642	.
2	2022-01-04	321.33474731	326.94
3	2022-01-05	308.99935913	324.14
4	2022-01-06	306.55764771	319.09
5	2022-01-07	306.71398926	315.96
6	2022-01-10	306.93856812	314.11
7	2022-01-11	307.63198853	310.11
8	2022-01-12	310.84527588	307.37
9	2022-01-13	297.68945313	307.74
10	2022-01-14	302.96353149	305.96
11	2022-01-18	295.58966064	305.21
12	2022-01-19	296.2538147	302.94
13	2022-01-20	294.56414795	300.67
14	2022-01-21	289.12408447	297.41
15	2022-01-24	289.45614624	295.70
16	2022-01-25	281.75997925	293.00
17	2022-01-26	289.78820801	290.23
18	2022-01-27	292.84521484	288.94
19	2022-01-28	301.06887817	288.59
20	2022-01-31	303.7253418	290.98
21	2022-02-01	301.55712891	293.84
22	2022-02-02	306.14746094	297.80
23	2022-02-03	294.22232056	301.07
24	2022-02-04	298.80288696	301.34
25	2022-02-07	293.92929077	300.89

3.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Średnia krocząca jako wskaźnik trendu:

- Średnia krocząca (MA_Close) pomaga wygładzić zmienność dziennych cen akcji, co umożliwia lepsze zrozumienie krótkoterminowych trendów.
- Wartości MA_Close są dostępne dopiero od 5. obserwacji, ponieważ do jej obliczenia potrzebne są dane z poprzednich 5 dni.
- Ogólny trend w analizowanym okresie:
 - Cena akcji spadała w analizowanym okresie, co widać zarówno po wartości "Close", jak i "MA_Close". Pod koniec okresu wartości te obniżyły się z około 326 USD do mniej więcej 290 USD.
- Korelacja między ceną akcji a średnią kroczącą:
 - MA_Close zawsze podąża za ceną "Close", ale zmienia się wolniej. To wynika z faktu, że średnia krocząca jest opóźnionym wskaźnikiem bazującym na poprzednich danych.
- Przykłady istotnych momentów:
 - Wartość "Close" z dnia 2022-01-07 wynosiła 306.71, natomiast MA_Close osiągnęła wartość 315.96. Średnia krocząca była wyższa od aktualnej ceny, co może wskazywać na spadkowy trend.
 - Pod koniec okresu (np. 2022-01-27), cena akcji spadła do 288.94, a MA_Close wynosiła 288.94, co świadczy o kontynuacji trendu spadkowego.
- Obliczanie średniej ruchomej za pomocą makr w SAS umożliwia łatwe dostosowanie parametrów okna czasowego, co pozwala na przeprowadzanie analizy w różnych okresach. Automatyzacja tego procesu sprawia, że jest on szybki i prosty, niezależnie od liczby analizowanych akcji.

4. Obliczanie zmienności cen akcji

4.1. Wyjaśnienie celu szczegółowego

- ❖ Zmienność cen akcji jest miarą fluktuacji cen w określonym czasie.
- ❖ Obliczenie zmienności jest kluczowe w analizach ryzyka i ocenie potencjalnych zmian cen na rynku.
- ❖ Kod w języku SAS 4GL i SAS Macro:

4.2. Opis procesu uzyskania danych wynikowych (kody)

```
%macro calculate_volatility(dataset, window);  
    PROC          EXPAND          DATA=&dataset          OUT=&dataset._volatility  
METHOD=none;  
    CONVERT      Close=Volatility    /    TRANSFORMOUT=(MOVSTD  
&window);  
    RUN;  
  
    PROC PRINT DATA=&dataset._volatility;  
        TITLE "Volatility of Stock Price for &dataset (Window: &window  
Days)";  
        VAR Date Close Volatility;  
    RUN;  
  
%mend calculate_volatility;  
  
%calculate_volatility(work.&ticker._data, 5);
```

❖ Wyjaśnienie technikaliów:

- Makro `calculate_volatility` oblicza zmienność (ang. volatility) cen akcji w oparciu o ruchome odchylenie standardowe (ang. moving standard deviation) z określonym oknem czasowym. Wynik jest zapisywany do nowego zestawu danych i wyświetlany w raporcie.
- `PROC EXPAND`: Procedura SAS używana do transformacji i interpolacji danych czasowych.
- `TRANSFORMOUT=(MOVSTD &window)`: Transformacja oblicza ruchome odchylenie standardowe dla zmiennej `Close`, korzystając z okna czasowego określonego przez parametr `window`.
- Przykładowo, jeśli mamy zestaw danych `work.MSFT_data` i chcemy obliczyć zmienność z oknem 5-dniowym, możemy wywołać makro w następujący sposób: `%calculate_volatility(work.MSFT_data, 5);`

4.3. Wykresy i tabele z danymi

Log 7 Jan 2025, 18:49:09

Log Results Output Data (2)

Volatility of Stock Price for work.MSFT_data (Window: 5 Days)

Obs	Date	Close	Volatility
1	2022-01-03	326.94082642	0
2	2022-01-04	321.33474731	3.9640965486
3	2022-01-05	308.99935913	9.1786543393
4	2022-01-06	306.55764771	9.7693601581
5	2022-01-07	306.71398926	9.4165376878
6	2022-01-10	306.93856812	6.3529234398
7	2022-01-11	307.63198853	0.9999793385
8	2022-01-12	310.84527588	1.7851695591
9	2022-01-13	297.68945313	4.9139961515
10	2022-01-14	302.96353149	5.0550837186
11	2022-01-18	295.58966064	6.4440738012
12	2022-01-19	296.2538147	6.3834299662
13	2022-01-20	294.56414795	3.3041283767
14	2022-01-21	289.12408447	4.9399037346
15	2022-01-24	289.45614624	3.4395422038
16	2022-01-25	281.75997925	5.6708535276
17	2022-01-26	289.78820801	4.5881226542
18	2022-01-27	292.84521484	4.0996061732
19	2022-01-28	301.06887817	6.9644759748
20	2022-01-31	303.7253418	8.8502845812
21	2022-02-01	301.55712891	6.0961036333
22	2022-02-02	306.14746094	5.017863818
23	2022-02-03	294.22232056	4.4604780999
24	2022-02-04	298.80288696	4.6081206663
25	2022-02-07	293.92929077	5.1521604306

4.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Początkowy brak zmienności: Na początku okresu (3 stycznia 2022), zmienność wynosi 0, co wynika z faktu, że nie ma wystarczającej liczby danych do obliczenia 5-dniowej zmienności.
- Wzrost zmienności w pierwszym tygodniu: Wartość zmienności stopniowo rośnie, osiągając najwyższą wartość (około 9.77) 6 stycznia 2022 r. Świadczy to o większych wahaniami cen w tym okresie.
- Spadek zmienności po dużym wzroście: Po pierwszym tygodniu zmienność zaczyna spadać, osiągając niskie wartości w okolicach 11 stycznia 2022 r. (około 0.99). Oznacza to, że ceny akcji ustabilizowały się w tym czasie.
- Okresowe wzrosty i spadki zmienności: Dane pokazują, że zmienność cen akcji fluktuuje – raz wzrasta, raz spada. Najwyższe wartości (np. 8.85 w dniu 31 stycznia) wskazują na momenty większej niepewności na rynku, podczas gdy niższe wartości oznaczają większą stabilność.

- Długoterminowe trendy: Ogólnie można zauważyć, że zmienność nie jest stała i zależy od różnych czynników rynkowych (np. wydarzeń makroekonomicznych, publikacji raportów finansowych).
- Automatyczne obliczanie zmienności umożliwia łatwe monitorowanie ryzyka w danym okresie, co jest kluczowe przy podejmowaniu decyzji inwestycyjnych. Dzięki automatyzacji procesu, analiza zmienności staje się szybka i może być regularnie wykonywana na bieżąco.

5. Wizualizacja i analiza zależności między wolumenem a ceną zamknięcia akcji

5.1. Wyjaśnienie celu szczegółowego

- ❖ Celem jest analiza korelacji między wolumenem transakcji a ceną zamknięcia akcji, co pozwala na identyfikację relacji pomiędzy tymi dwiema zmiennymi i ich wpływ na rynek.
- ❖ Kod w języku SAS 4GL i SAS Macro:

5.2. Opis procesu uzyskania danych wynikowych (kody)

```
%macro correlation_analysis(dataset);
    PROC CORR DATA=&dataset;
        TITLE "Correlation Between Volume and Close Price for &dataset";
        VAR Volume Close;
    RUN;
```

```
%mend correlation_analysis;
```

```
%correlation_analysis(work.&ticker._data);
```

- ❖ Wyjaśnienie technikaliów:
 - Makro w języku SAS, które automatyzuje analizę korelacji między zmiennymi Volume i Close w podanym zbiorze danych.
 - PROC CORR to wbudowana procedura w SAS służąca do obliczania korelacji między zmiennymi.
 - DATA=&dataset wskazuje, że procedura ma pracować na zbiorze danych przekazanym do makra jako parametr (&dataset).

- Symbol & oznacza, że odwołujemy się do parametru makra.
- Wynikiem działania będzie raport z procedury PROC CORR, który zawiera:
 - Macierz korelacji między zmiennymi Volume i Close.
 - Statystyki, takie jak współczynniki korelacji, wartość p (dla testów istotności), liczba obserwacji itp.

5.3. Wykresy i tabele z danymi

7 Jan 2025, 19:32:20

Correlation Between Volume and Close Price for work.MSFT_data

The CORR Procedure

2 Variables: Volume Close

Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
Volume	251	31219322	11483873	7836049700	9200800	90428900
Close	251	263.74952	24.65857	66201	210.61905	326.94083

Pearson Correlation Coefficients, N = 251 Prob > r under H0: Rho=0		
	Volume	Close
Volume	1.00000	0.14782 0.0191
Close	0.14782 0.0191	1.00000

5.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Wartość korelacji jest stosunkowo niska, co sugeruje, że istnieje słaba dodatnia korelacja między wolumenem obrotu (Volume) a ceną zamknięcia (Close) akcji.
- Korelacja wynosząca 0.14782 wskazuje, że wzrost wolumenu obrotu jest nieznacznie związany z wzrostem ceny akcji, ale zależność ta jest na tyle słaba, że nie ma silnego wpływu na jedną zmienną w odniesieniu do drugiej.
- Wartość $p < 0.05$ oznacza, że współczynnik korelacji jest statystycznie istotny. Oznacza to, że istnieje wystarczająca dowody na to, że obserwowany współczynnik korelacji nie wynika z przypadku.

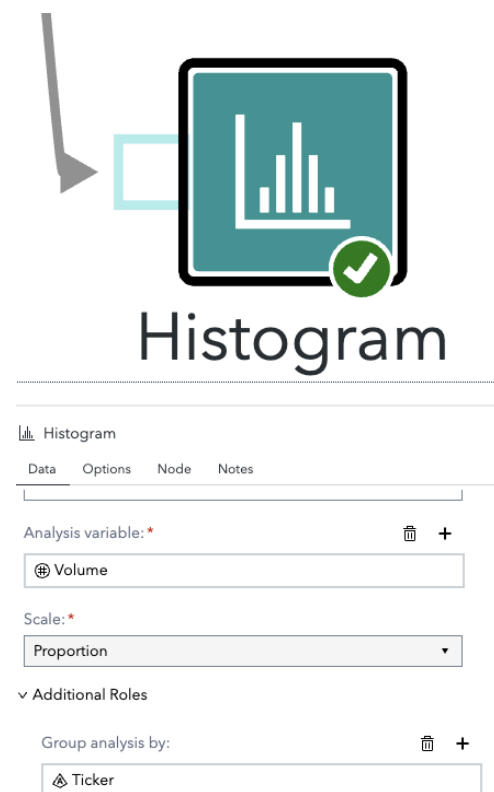
- Mimo iż korelacja jest dodatnia, jej siła jest niewielka, co sugeruje, że zmiany w wolumenie obrotu mają tylko minimalny wpływ na zmiany ceny akcji Mikrosoft w badanym okresie.
- Automatyzacja analizy korelacji pozwala na szybkie identyfikowanie potencjalnych zależności między zmiennymi bez konieczności ręcznego przetwarzania danych. To narzędzie wspiera podejmowanie decyzji inwestycyjnych w oparciu o zależności rynkowe, co jest szczególnie istotne w dynamicznych warunkach rynkowych.

6. Automatyzacja analizy ilości transakcji akcyjnych

6.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego kroku jest automatyczne tworzenie histogramów dla spółek w danych wejściowych, oraz wyciąganie wniosków z ich wizualizacji.
- ❖ Rozkład wolumenu jest typowy dla danych finansowych, gdzie wahania wolumenu mogą być związane z wydarzeniami rynkowymi.
- ❖ Kod w języku SAS Flow:

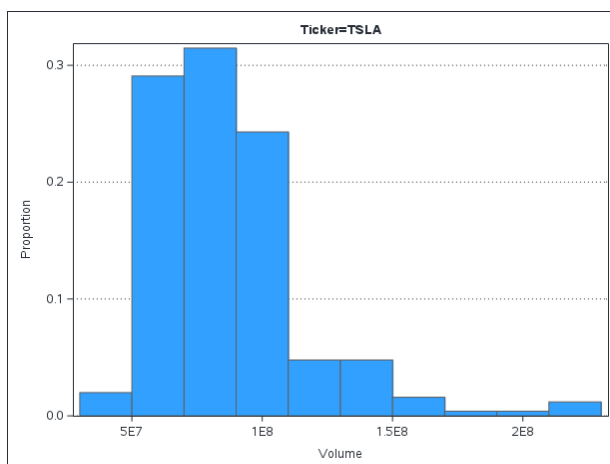
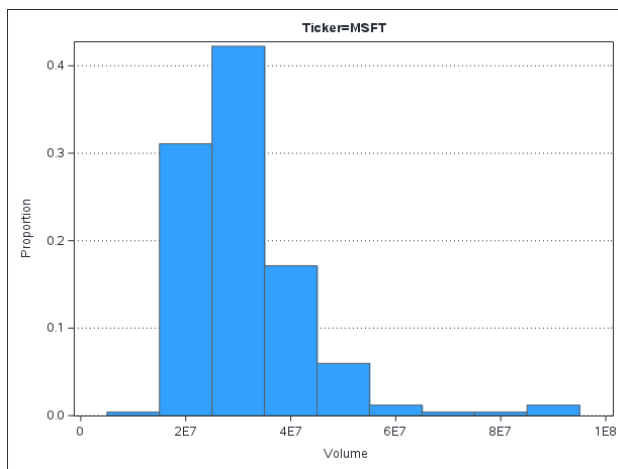
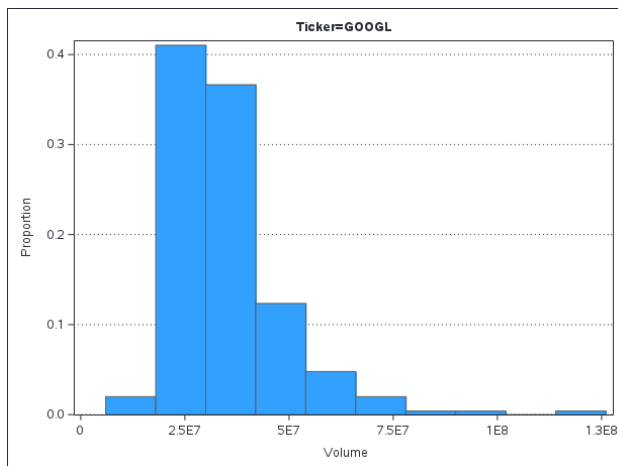
6.2. Opis procesu uzyskania danych wynikowych (kody)



❖ Wyjaśnienie technikaliów:

- Użyto narzędzia no code, o nazwie Histogram, aby wygenerować histogramy dla każdej spółki w danych wejściowych, na podstawie zmiennej Volume.
- Rozkład wolumenu jest typowy dla danych finansowych, gdzie wahania wolumenu mogą być związane z wydarzeniami rynkowymi.

6.3. Wykresy i tabele z danymi



6.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

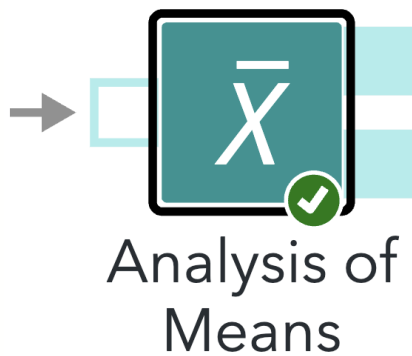
- W każdym przypadku rozkład wolumenu jest asymetryczny prawostronnie (skośny), co sugeruje, że większość dni charakteryzuje się niższymi wartościami wolumenu, a tylko w niewielkiej liczbie dni odnotowano znacząco wyższe wolumeny.
- Rozkład wolumenu jest typowy dla danych finansowych, gdzie wahania wolumenu mogą być związane z wydarzeniami rynkowymi.
- GOOGL: Większość dni wolumen oscyluje wokół wartości $2,5E7$ - $5E7$. Niewiele dni przekracza poziom $1E8$.
- MSFT: Wolumen jest podobny do GOOGL w zakresie $2E7$ - $4E7$, z wyraźnymi dniami o niższym wolumenie w stosunku do GOOGL.
- TSLA: Wolumen transakcji dla TSLA jest wyraźnie większy. Najczęstsze wartości oscylują wokół $5E7$ - $1E8$, a w niektórych przypadkach osiąga wartości powyżej $2E8$.
- Spółka TSLA cieszy się największym zainteresowaniem inwestorów pod względem obrotów, co może wynikać z jej wyższej zmienności, medialności, lub liczby aktywnych inwestorów indywidualnych.
- GOOGL i MSFT mają bardziej stabilny wolumen, co może świadczyć o większej aktywności ze strony dużych inwestorów instytucjonalnych.
- W przypadku TSLA, większy wolumen może wskazywać na wyższe ryzyko, ale i większe możliwości zysku dla inwestorów krótkoterminowych.
- MSFT i GOOGL mogą być bardziej odpowiednie dla strategii długoterminowych, z uwagi na bardziej przewidywalne ruchy wolumenu.

7. Analiza popularności spółki w porównaniu do innych

7.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego kroku jest stworzenie wykresu pudełkowego, który poda nam wiele informacji, o medianie wolumenu, rozstępie międzykwartylowym, wartości odstających, zakresie wartości, oraz średniej.
- ❖ Wszystkie powyższe informacje pomagają w podjęciu odpowiedniej i świadomej decyzji inwestycyjnej.
- ❖ Kod w języku SAS Flow:

7.2. Opis procesu uzyskania danych wynikowych (kody)



\bar{X} Analysis of Means

Data Options Output Node Notes

☐ Sort data before chart creation

Response variable: *

@ Volume

Group variable: *

📌 Ticker

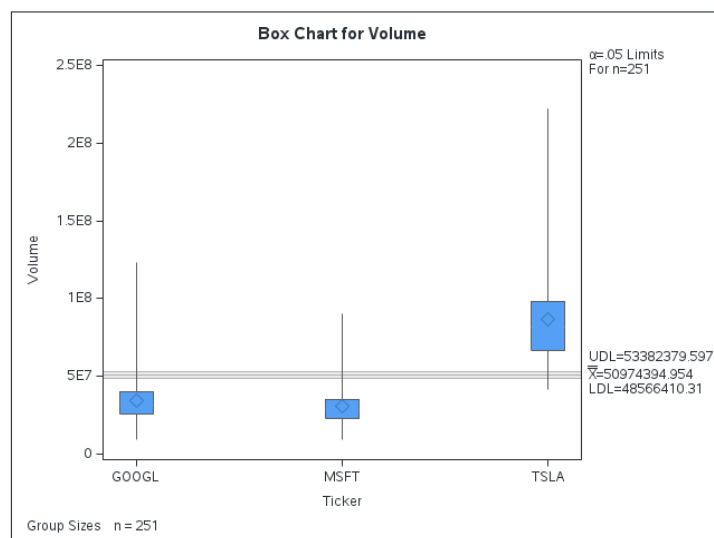
Chart type: *

Mean chart with boxes

❖ Wyjaśnienie technikaliów:

- Użyto narzędzia no code, o nazwie Analysis of Means, aby wygenerować wykres rodzaju pudełkowego ze średnią dla każdej spółki w danych wejściowych, na podstawie zmiennej Volume i grupowania Ticker.
- Wykres ten pozwala na wizualne określenie jak popularna jest spółka względem innych.
- Wykres ukazuje również obserwacje nietypowe.

7.3. Wykresy i tabele z danymi



7.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- GOOGL i MSFT: Mają zbliżone mediany wolumenu, które znajdują się na poziomie ok. $2E7$ - $3E7$.
- TSLA: Charakteryzuje się znacznie wyższą medianą wolumenu ($\sim 1E8$), co wskazuje na większe zainteresowanie akcjami tej spółki w porównaniu do pozostałych.
- TSLA: Ma największy rozstęp międzykwartylowy (obszar między dolnym a górnym kwartyłem), co świadczy o dużej zmienności wolumenu.
- GOOGL i MSFT: Mają węższy IQR, co sugeruje stabilniejsze wolumeny w porównaniu do TSLA.
- Wszystkie spółki mają widoczne wartości odstające (outliers) w górnych rejestrach danych:
 - Dla TSLA wartości odstające są bardziej znaczące i częstsze, co podkreśla jej podatność na duże skoki wolumenu.
 - Dla GOOGL i MSFT wartości odstające występują rzadziej i mają mniejszą skalę.
- TSLA: Maksymalny wolumen ($\sim 2.5E8$) jest wyraźnie większy niż w przypadku GOOGL i MSFT.
- GOOGL i MSFT: Maksymalne wartości wolumenu są znacznie niższe i bardziej zbliżone.
- Średnia wolumenu dla każdej spółki jest wyższa niż mediana, co potwierdza prawostronną asymetrię danych (istnienie wysokich wartości odstających).
- TSLA: Wysoka zmienność wolumenu i większa liczba wartości odstających mogą sugerować większe ryzyko inwestycyjne, ale też większe możliwości dla aktywnych traderów.
- GOOGL i MSFT: Niższa zmienność wolumenu może być bardziej atrakcyjna dla inwestorów szukających stabilności.

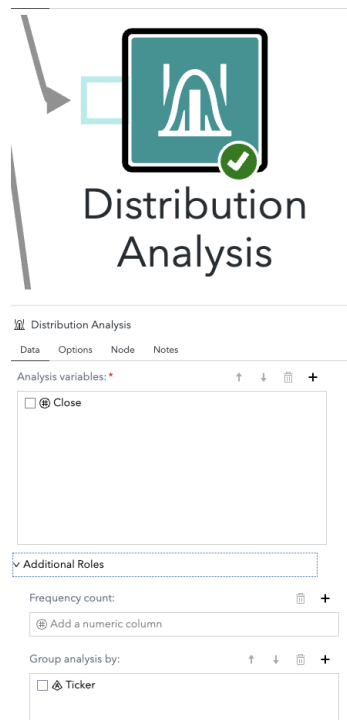
8. Wizualizacja i automatyzacja analizy rozkładów cen zamknięcia

8.1. Wyjaśnienie celu szczegółowego

- ❖ Celem tego celu jest wizualne określenie w jakich zakresach w danym okresie czasowym wahają się ceny spółki.

- ❖ Wykres ukazuje również jakie ceny są częste, a jakie mają niskie prawdopodobieństwo i znajdują się w ogonach.
- ❖ Jest to przydatne o tyle, że ukazuje jaka cena jest relatywnie wysoka a jaka relatywnie niska, jeśli zastanawiamy się nad wejściem na rynek.
- ❖ Kod SAS Flow:

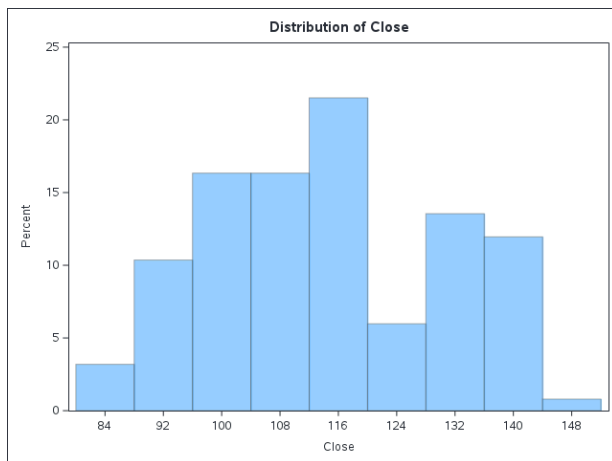
8.2. Opis procesu uzyskania danych wynikowych (kody)



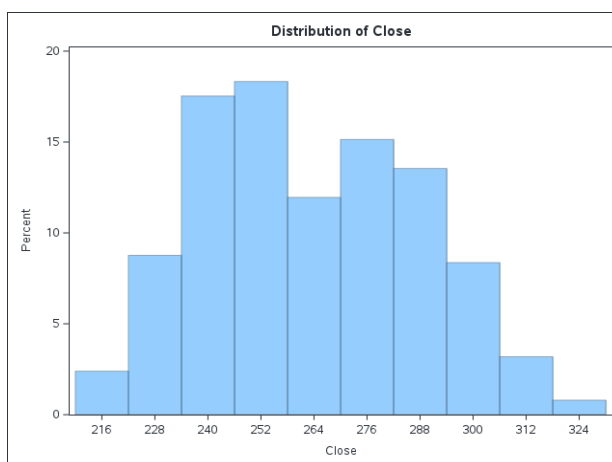
- ❖ Wyjaśnienie technikaliów:
 - Użyto narzędzia no code, o nazwie Distribution Analysis, aby wygenerować wykres przedstawiający rozkład danych wejściowych, na podstawie zmiennej Close (Price) i grupowania Ticker.
 - Wykres ten pozwala na wizualne określenie w jakich zakresach w danym okresie czasowym wahają się ceny spółki.
 - Wykres ukazuje również jakie ceny są częste, a jakie mają niskie prawdopodobieństwo i znajdują się w ogonach.

8.3. Wykresy i tabele z danymi

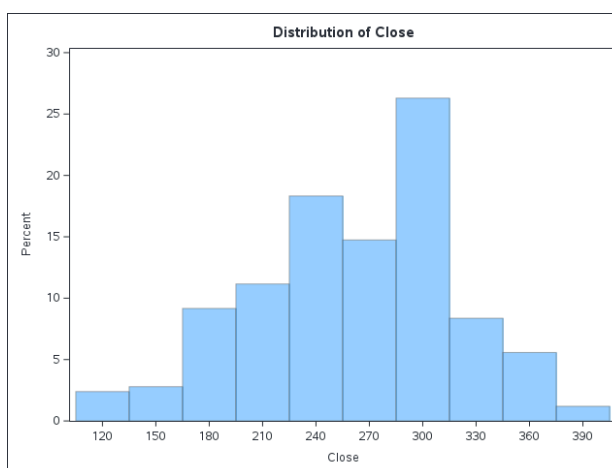
Ticker=GOOGL



Ticker=MSFT



Ticker=TSLA



8.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

➤ GOOGL (Alphabet Inc.)

- Rozkład: Rozkład cen zamknięcia jest zbliżony do normalnego, z lekkim przesunięciem w kierunku wyższych wartości.
- Zakres cen: Ceny zamknięcia wahają się w przedziale od ok. 84 do 148, przy czym największa liczba dni przypada na przedział 116–124.
- Wnioski:
 - Większość dni ceny zamknięcia oscylują wokół wartości w środku zakresu.
 - Niskie ceny (<100) oraz bardzo wysokie ceny (>140) są rzadkością.
 - GOOGL zajmuje miejsce pośrednie pod względem rozkładu i zakresu cen w porównaniu z MSFT i TSLA.

➤ MSFT (Microsoft Corporation)

- Rozkład: Rozkład jest symetryczny i przypomina rozkład normalny.
- Zakres cen: Ceny zamknięcia mieszczą się w przedziale od ok. 216 do 324, z najwyższą częstotliwością w przedziale 240–264.
- Wnioski:
 - Ceny zamknięcia dla MSFT są bardziej stabilne niż dla TSLA i koncentrują się w średnim zakresie.
 - Bardzo wysokie ceny (>300) występują sporadycznie.
 - MSFT ma bardziej stabilny rozkład cen zamknięcia, co czyni go bardziej przewidywalnym dla inwestorów.

➤ TSLA (Tesla Inc.)

- Rozkład: Rozkład jest asymetryczny prawostronnie, z większym skupieniem w dolnym zakresie cen.
- Zakres cen: Ceny zamknięcia wahają się od ok. 120 do 390, z największą liczbą dni przypadającą na przedział 240–280.
- Wnioski:
 - Ceny TSLA są bardziej zmienne, co wskazuje na wyższą nieprzewidywalność i ryzyko.
 - Wyższe ceny (>300) są mniej częste, ale ich obecność wpływa na asymetrię rozkładu.

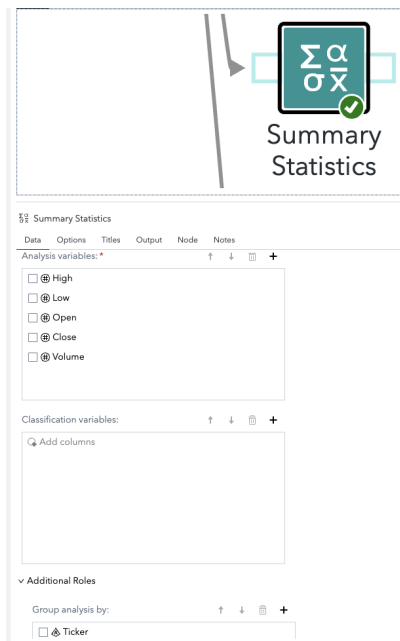
- TSLA charakteryzuje się najszerszym zakresem cen i większą zmiennością, co może wynikać z jej dynamicznego charakteru i wysokiej spekulacji na rynku.
- Dla traderów krótkoterminowych: TSLA może oferować większe możliwości zysku (ale także większe ryzyko) ze względu na szerszy zakres i większe odchylenia cen.
- Dla inwestorów długoterminowych: MSFT wydaje się być bardziej stabilnym wyborem, a GOOGL oferuje balans między zmiennością a stabilnością.

9. Obliczanie statystyk opisowych cen akcji spółek giełdowych

9.1. Wyjaśnienie celu szczegółowego

- ❖ Celem analizy jest przeprowadzenie szczegółowego porównania kluczowych statystyk opisowych dla cen akcji wybranych spółek giełdowych (GOOGL, MSFT, TSLA). Analiza ta pozwala:
 - Zidentyfikować zmienność cen akcji: Ocena odchyłeń standardowych oraz zakresów cen dostarcza informacji o stabilności i ryzyku inwestycyjnym dla każdej spółki.
 - Porównać aktywność rynkową: Analiza średniego wolumenu oraz jego zmienności wskazuje na zainteresowanie inwestorów danym aktywem w wybranym okresie.
 - Wskazać potencjalne trendy inwestycyjne: Dzięki analizie minimalnych, maksymalnych i średnich wartości cen możliwe jest oszacowanie atrakcyjności akcji w długim i krótkim okresie.
 - Opracować rekomendacje inwestycyjne: Wnioski z analizy pomagają podjąć decyzję, które akcje mogą być odpowiednie dla inwestorów długoterminowych, a które dla krótkoterminowych spekulantów.
- ❖ Kod w języku SAS Flow:

9.2. Opis procesu uzyskania danych wynikowych (kody)



❖ Wyjaśnienie technikaliów:

- Użyto narzędzia no code, o nazwie Summary Statistics, aby wygenerować tabelę przedstawiającą statystyki opisowe danych wejściowych, na podstawie zmiennych Close, Low, Open, High (Price), Volume, oraz grupowania Ticker.
- Analiza została zbudowana poprzez dodawanie kolejnych kroków w przepływie.

9.3. Wykresy i tabele z danymi

Ticker=GOOGL

Variable	Mean	Std Dev	Minimum	Maximum	N
High	116.0630755	16.2386043	86.2074899	150.9991072	251
Low	112.7921836	15.8952235	83.0389756	144.9968790	251
Open	114.4647380	16.2101323	85.0915348	150.7036846	251
Close	114.3458581	16.0509550	83.1286545	147.4654236	251
Volume	34767530.28	13619267.27	9701400.00	123200000	251

Ticker=MSFT

Variable	Mean	Std Dev	Minimum	Maximum	N
High	266.7248450	24.7935253	216.2418303	329.4555925	251
Low	259.6725267	24.3284022	209.3938178	321.4433872	251
Open	263.4055508	24.8946209	213.4359147	326.8725886	251
Close	263.2226425	24.6093090	210.1983185	326.2877502	251
Volume	31219321.51	11483873.11	9200800.00	90428900.00	251

Ticker=TESLA

Variable	Mean	Std Dev	Minimum	Maximum	N
High	270.5677160	58.9185248	116.2699966	402.6666565	251
Low	256.5073573	56.5672387	108.2399979	378.6799927	251
Open	264.1352322	57.8490291	110.3499985	396.5166626	251
Close	263.0930810	57.8143564	109.0999985	399.9266663	251
Volume	86936333.07	29496489.44	41864700.00	221923300	251

9.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

➤ GOOGL (Alphabet Inc.)

■ Średnie wartości (Mean):

- Cena zamknięcia wynosi średnio 114.35, a średni wolumen to 34,767,530.28.

■ Wahania cen (Standard Deviation):

- Wartości cen mają niewielkie wahania (ok. 16), co wskazuje na stabilność cen GOOGL.
- Wolumen transakcji jest bardziej zmienny (odchylenie standardowe: 13,619,267.27), co może wynikać z okresowych wzrostów aktywności na rynku.

■ Zakres cen:

- Minimalna cena zamknięcia wynosi 83.13, a maksymalna 147.47, co daje zakres cen wynoszący 64.34.

■ Wnioski:

- GOOGL wydaje się stabilną spółką pod względem cen akcji, ale zmienność wolumenu wskazuje na okresowe wzrosty zainteresowania inwestorów.

➤ MSFT (Microsoft Corporation)

■ Średnie wartości (Mean):

- Cena zamknięcia wynosi średnio 263.22, a średni wolumen to 31,219,321.51.

■ Wahania cen (Standard Deviation):

- Odchylenie standardowe cen wynosi ok. 24.6, co świadczy o umiarkowanej zmienności.
- Wolumen transakcji jest mniej zmienny niż w przypadku GOOGL (odchylenie standardowe: 11,483,873.11).

■ Zakres cen:

- Minimalna cena zamknięcia wynosi 210.20, a maksymalna 326.29, co daje zakres cen wynoszący 116.09.

■ Wnioski:

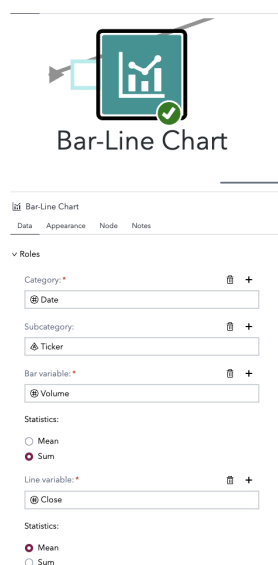
- MSFT ma umiarkowaną zmienność zarówno w zakresie cen, jak i wolumenu, co czyni ją stabilnym wyborem inwestycyjnym.
- TSLA (Tesla Inc.)
 - Średnie wartości (Mean):
 - Cena zamknięcia wynosi średnio 263.09, a średni wolumen to 86,936,333.07.
 - Wahania cen (Standard Deviation):
 - TSLA wykazuje najwyższe odchylenia standardowe cen (57.81) oraz wolumenu (29,496,489.44), co wskazuje na dużą zmienność.
 - Zakres cen:
 - Minimalna cena zamknięcia wynosi 109.10, a maksymalna 399.93, co daje zakres cen wynoszący 290.83.
 - Wnioski:
 - TSLA jest znacznie bardziej zmienna niż GOOGL i MSFT, co może przyciągać spekulantów i inwestorów krótkoterminowych.

10. Porównawcza analiza cen zamknięcia i wolumenu obrotu spółek

10.1. Wyjaśnienie celu szczegółowego

- ❖ Celem analizy jest porównanie trzech spółek giełdowych (GOOGL, MSFT, TSLA) pod względem:
 - Średnich cen zamknięcia (Close).
 - Wolumenu obrotu (Volume).
- ❖ Identyfikacja różnic i trendów między spółkami, co może pomóc w ocenie ryzyka i atrakcyjności inwestycyjnej poszczególnych akcji.
- ❖ Kod w języku SAS Flow:

10.2. Opis procesu uzyskania danych wynikowych (kody)



❖ Wyjaśnienie technikaliów:

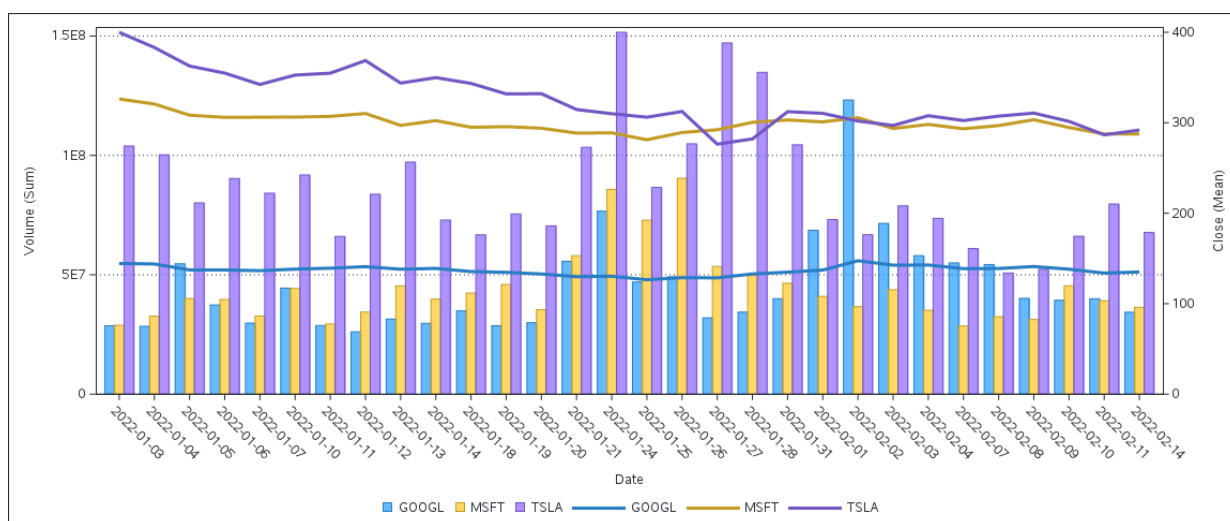
➤ Wybrano kombinowany wykres bar + plot:

■ Wykres słupkowy (bar): Dla wolumenu obrotu, reprezentujący ilościowe dane.

■ Wykres liniowy (plot): Dla cen zamknięcia, ukazujący trend cen.

➤ Po skonfigurowaniu wykresu użytkownik uruchamia proces wizualizacji, a narzędzie generuje gotowy wykres.

10.3. Wykresy i tabele z danymi



10.4. Wnioski uzyskane z przeprowadzonej analizy

❖ Wnioski:

- Ceny zamknięcia:
 - Spółka MSFT wykazuje najbardziej stabilne ceny zamknięcia, z najmniejszą zmiennością.
 - TSLA ma najwyższą średnią cenę zamknięcia, ale ceny są bardziej zmienne w porównaniu z MSFT i GOOGL.
 - GOOGL ma najniższą średnią cenę zamknięcia w analizowanym okresie.
- Wolumen obrotu:
 - TSLA ma wyraźnie największy wolumen obrotu, co wskazuje na wysoką aktywność handlową.
 - Wolumen obrotu dla MSFT i GOOGL jest znacznie mniejszy, przy czym GOOGL ma najniższy wolumen.
- Zależności między wolumenem a ceną:
 - Wysoki wolumen w przypadku TSLA może być związany z dużą zmiennością ceny, co sugeruje większe zainteresowanie inwestorów tą spółką.
 - Stabilność wolumenu dla MSFT może wskazywać na bardziej ustabilizowaną bazę inwestorów.

Podsumowanie

W ramach przeprowadzonej analizy skupiono się na automatyzacji procesów analizy danych giełdowych oraz ocenie zależności pomiędzy cenami akcji a wolumenem obrotu na przykładzie trzech spółek: Microsoft (MSFT), Tesla (TSLA) oraz Alphabet (GOOGL). Celem badania było wykorzystanie dostępnych narzędzi no-code oraz technik statystycznych do stworzenia procesu analitycznego, który pozwoliłby na szybkie i efektywne pozyskiwanie oraz analizowanie danych rynkowych, z jednoczesnym zminimalizowaniem manualnych działań.

Pierwszym etapem było pobranie danych giełdowych z publicznie dostępnych źródeł za pomocą narzędzi automatyzujących procesy zbierania danych. Wybrane zmienne obejmowały ceny akcji (Open, High, Low, Close) oraz wolumen obrotu (Volume) w określonym przedziale czasowym. W kolejnym kroku przeprowadzono wstępne statystyki opisowe, które umożliwiły wgląd w podstawowe właściwości analizowanych danych, takie jak średnia, odchylenie standardowe czy maksymalna i minimalna wartość zmiennych.

Następnie skupiono się na ocenie zależności pomiędzy wolumenem obrotu a ceną zamknięcia akcji. Wyniki analizy korelacji wykazały, że istnieje słaba, lecz statystycznie istotna dodatnia zależność między tymi zmiennymi ($r = 0.14782$). Pomimo istotności statystycznej, siła tej korelacji była zbyt niska, aby uznać wolumen obrotu za istotny czynnik wpływający na wartość akcji w analizowanym okresie.

Dalsza część analizy dotyczyła zmienności cen akcji, którą zbadano przy użyciu wskaźnika 5-dniowej zmienności. Wyniki pokazały, że zmienność była najwyższa w pierwszych dniach analizowanego okresu, co mogło wynikać z początkowej niepewności na rynku. W kolejnych dniach zauważono okresową stabilizację cen, przerywaną momentami zwiększonej zmienności, które mogły być efektem wydarzeń makroekonomicznych lub publikacji wyników finansowych.

Ostatnim elementem analizy było wykorzystanie średniej kroczącej jako wskaźnika trendu cen akcji. Analiza ta potwierdziła ogólny spadkowy trend wartości akcji we wszystkich trzech spółkach w analizowanym okresie. Wartości średniej kroczącej (MA_Close) pozwoliły na wygładzenie krótkoterminowych wahań cen, co umożliwiło lepsze zrozumienie kierunku zmian na rynku.

Na zakończenie stworzono wizualizację porównawczą, która prezentowała ceny akcji oraz wolumen obrotu wszystkich trzech spółek na jednym wykresie. Dzięki zastosowaniu narzędzi no-code możliwe było stworzenie przejrzystego wykresu łączącego dane dotyczące zarówno cen, jak i aktywności rynkowej, co podkreśla różnice w charakterystyce poszczególnych spółek.

Słaba korelacja między ceną a wolumenem obrotu: Analiza wykazała, że wolumen obrotu akcji ma jedynie minimalny wpływ na ceny zamknięcia, co sugeruje, że inne czynniki, takie jak wydarzenia makroekonomiczne, mają większe znaczenie w kształtowaniu wartości akcji.

Charakterystyka zmienności: Największą zmienność cen zaobserwowano w akcjach Tesla, co może wynikać z ich spekulacyjnego charakteru. Microsoft i Alphabet cechowały się większą stabilnością.

Średnia krocząca jako narzędzie analityczne: Zastosowanie średniej kroczącej pozwoliło na skuteczne identyfikowanie trendów rynkowych, co ma istotne znaczenie przy podejmowaniu decyzji inwestycyjnych.

Zalety automatyzacji: Wykorzystanie narzędzi no-code umożliwiło szybkie przetwarzanie danych oraz ich analizę w sposób powtarzalny i bezbłędny. Dzięki temu

możliwe jest regularne monitorowanie rynków finansowych przy minimalnym nakładzie pracy.

Realizacja projektu pokazała, że automatyzacja procesów analitycznych może znacząco usprawnić analizę danych giełdowych, zarówno pod względem czasu, jak i jakości uzyskiwanych wyników. Wykorzystanie narzędzi no-code oraz technik statystycznych daje możliwość szybkiego reagowania na zmieniające się warunki rynkowe oraz wsparcia w podejmowaniu decyzji inwestycyjnych. Wyniki analizy podkreślają istotność systematycznego monitorowania zmienności i trendów na rynku, co jest kluczowe dla efektywnego zarządzania portfelem inwestycyjnym.

Bibliografia

- Komentarze do kodu i określanie problemów biznesowych zostały stworzone przy pomocy narzędzi LLM, ale pomysły na ich utworzenie są oryginalnymi utworami autorów.
- Baza danych Yahoo Finance: <https://pypi.org/project/yfinance/>;
<https://www.geeksforgeeks.org/how-to-use-yfinance-api-with-python/>
- Repozytorium kodu: <https://github.com/Szymon-Czusek/SAS-Automation>