

## 1. Stationaritätstests (ADF)

--- ADF Test für: Price\_Log ---

ADF Statistic: -1.6235

p-value: 0.4709

Critical Values:

1%: -3.4353

5%: -2.8637

10%: -2.5679

-> Ergebnis: Nicht stationär ( $p > 0.05$ )

--- ADF Test für: Log\_Return ---

ADF Statistic: -15.4836

p-value: 0.0000

Critical Values:

1%: -3.4353

5%: -2.8637

10%: -2.5679

-> Ergebnis: Stationär ( $p \leq 0.05$ )

--- ADF Test für: Volume ---

ADF Statistic: -7.6369

p-value: 0.0000

Critical Values:

1%: -3.4353

5%: -2.8637

10%: -2.5679

-> Ergebnis: Stationär ( $p \leq 0.05$ )

--- ADF Test für: Avg\_Sentiment\_Final ---

ADF Statistic: -6.7831

p-value: 0.0000

## Statistische Relevanzanalyse GME & Reddit

```
Critical Values:
  1%: -3.4353
  5%: -2.8637
 10%: -2.5679
-> Ergebnis: Stationär (p <= 0.05)

--- ADF Test für: Post_Count ---
ADF Statistic: -4.2604
p-value: 0.0005
Critical Values:
  1%: -3.4353
  5%: -2.8637
 10%: -2.5679
-> Ergebnis: Stationär (p <= 0.05)

--- ADF Test für: Total_Posts ---
ADF Statistic: -2.4659
p-value: 0.1240
Critical Values:
  1%: -3.4353
  5%: -2.8637
 10%: -2.5679
-> Ergebnis: Nicht stationär (p > 0.05)
```

### Erläuterung zur Stationarität:

- Stationär: Eine Zeitreihe ist stationär, wenn ihr statistischen Eigenschaften (Mittelwert, Varianz, Autokorrelation) über die Zeit konstant bleiben. Sie schwankt um einen Mittelwert, ohne langfristigen Trend oder systematisch veränderliche Streuung. Stationarität ist oft eine Voraussetzung für Zeitreihenmodelle wie Granger-Kausalität.
- Nicht stationär: Eine Zeitreihe ist nicht stationär, wenn sich Mittelwert oder Varianz über die Zeit ändern. Dies ist häufig bei Aktienkursen (Trends) oder bei Metriken mit starkem Wachstum (wie Nutzerzahlen) der Fall. Granger-Tests mit nicht-stationären Reihen können zu irreführenden Ergebnissen (Scheinkorrelationen) führen.

## 2. Deskriptive Statistiken

## Statistische Relevanzanalyse GME & Reddit

Metrik	Price_Log	Volume	Price_Change	Log_Return	Avg_Sentiment_Final	Post_Count	Total_Posts	Median_Upvotes
count	1334.0	1334.0	1334.0	1334.0	1334.0	1334.0	1334.0	1334.0
mean	3.2617	23081386.8066	0.0152	0.0013	0.1352	1.4723	880.7196	175.9824
std	0.4283	66666961.6423	3.1248	0.0816	0.2949	1.9484	1057.1686	2143.3825
min	1.4861	1122700.0	-38.48	-0.9163	-0.9542	0.0	47.0	0.0
25%	3.0174	3627350.0	-0.5375	-0.0226	0.0	0.0	55.0	0.0
50%	3.2233	6331600.0	0.0	0.0	0.0	1.0	538.0	0.0
75%	3.5895	14550000.0	0.35	0.0149	0.3312	2.0	1290.0	0.0
max	4.4645	788631600.0	49.89	0.8539	0.9976	13.0	9104.0	63424.0

## 3. Korrelationsanalyse

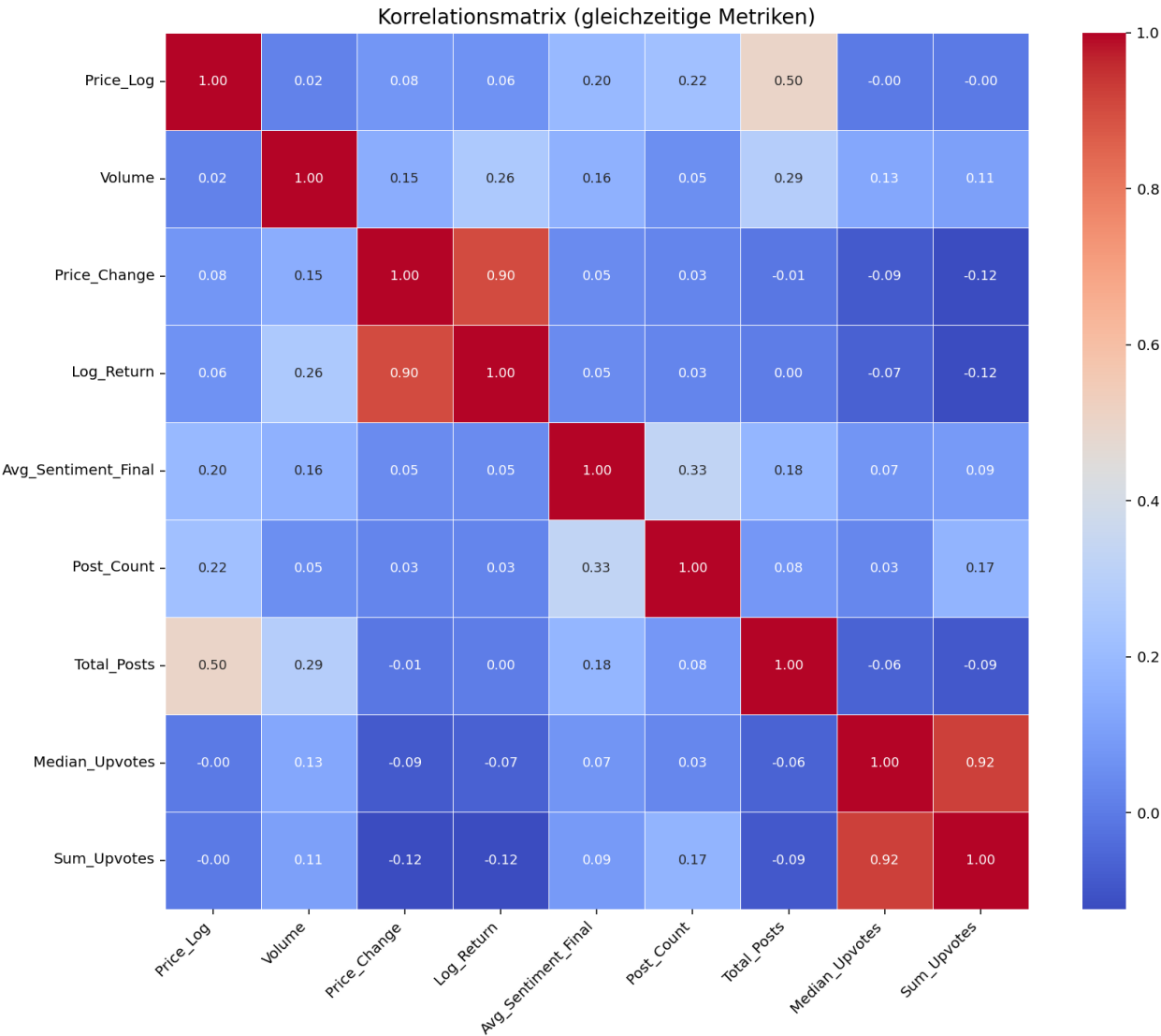
### 3.1 Gleichzeitige Korrelationen

#### Korrelationsmatrix (gleichzeitig)

Metrik	Price_Log	Volume	Price_Change	Log_Return	Avg_Sentiment_Final	Post_Count	Total_Posts	Median_Upvotes	Sum_Upvotes
Price_Log	1.0	0.02	0.077	0.062	0.196	0.221	0.503	-0.003	-0.001
Volume	0.02	1.0	0.152	0.259	0.163	0.053	0.285	0.127	0.106
Price_Change	0.077	0.152	1.0	0.9	0.046	0.03	-0.012	-0.086	-0.123
Log_Return	0.062	0.259	0.9	1.0	0.047	0.027	0.005	-0.067	-0.119
Avg_Sentiment_Final	0.196	0.163	0.046	0.047	1.0	0.332	0.181	0.07	0.089
Post_Count	0.221	0.053	0.03	0.027	0.332	1.0	0.079	0.031	0.173
Total_Posts	0.503	0.285	-0.012	0.005	0.181	0.079	1.0	-0.062	-0.09
Median_Upvotes	-0.003	0.127	-0.086	-0.067	0.07	0.031	-0.062	1.0	0.921
Sum_Upvotes	-0.001	0.106	-0.123	-0.119	0.089	0.173	-0.09	0.921	1.0

Heatmap (gleichzeitig):

# Statistische Relevanzanalyse GME & Reddit



Signifikanztests (Pearson r):  
Avg\_Sentiment\_Final vs Log\_Return: r=0.047, p=0.0834

## Statistische Relevanzanalyse GME & Reddit

Post\_Count vs Log\_Return:  $r=0.027$ ,  $p=0.3239$

Post\_Count vs Volume:  $r=0.053$ ,  $p=0.0513$

Total\_Posts vs Volume:  $r=0.285$ ,  $p=0.0000$  (SIGNIFIKANT)

## Statistische Relevanzanalyse GME & Reddit

### 3.2 Gelaggte Korrelationen (Lag 1 vs. Aktuell)

#### Korrelationsmatrix (gelaggt)

Metrik	Avg_Sentiment_Final_Lag1	Avg_Sentiment_VADER_Lag1	Post_Count_Lag1	Total_Posts_Lag1	Total_Comments_Lag1	Total_Subscribers_Lag1	Median_Upvotes_Lag1	Sum_Upvotes_Lag1	Volume_Lag1	Log_Return_Lag1
Price_Change	-0.001	-0.0	-0.039	-0.003	0.001	-0.008	0.016	-0.006	0.135	0.004
Log_Return	0.014	0.015	-0.03	0.015	0.013	-0.021	0.004	-0.019	0.151	0.04
Volume	0.15	0.151	0.05	0.286	0.253	-0.268	0.057	0.04	0.864	0.218

Signifikanztests (Pearson r, gelaggt):

Avg\_Sentiment\_Final\_Lag1 vs Log\_Return:  $r=0.014$ ,  $p=0.6042$

Post\_Count\_Lag1 vs Log\_Return:  $r=-0.030$ ,  $p=0.2767$

Post\_Count\_Lag1 vs Volume:  $r=0.050$ ,  $p=0.0676$

Total\_Posts\_Lag1 vs Volume:  $r=0.286$ ,  $p=0.0000$  (SIGNIFIKANT)

## 4. Granger-Kausalitätstests

```
--- Test: Beeinflusst 'Avg_Sentiment_Final' die Variable 'Log_Return'? (maxlag=5) ---
-> Keine signifikante Granger-Kausalität gefunden (alle p >= 0.05) für getestete Lags.
--- Test: Beeinflusst 'Post_Count' die Variable 'Log_Return'? (maxlag=5) ---
-> Keine signifikante Granger-Kausalität gefunden (alle p >= 0.05) für getestete Lags.
--- Test: Beeinflusst 'Total_Posts' die Variable 'Log_Return'? (maxlag=5) ---
-> WARNUNG: Mind. eine Zeitreihe (Log_Return: True, Total_Posts: False) nicht stationär. Ergebnisse evtl. unzuverlässig!
-> Keine signifikante Granger-Kausalität gefunden (alle p >= 0.05) für getestete Lags.
--- Test: Beeinflusst 'Avg_Sentiment_Final' die Variable 'Volume'? (maxlag=5) ---
-> Signifikante Granger-Kausalität gefunden (p < 0.05) bei folgenden Lags:
    Lag 2: p=0.0425
    Lag 3: p=0.0371
    Lag 5: p=0.0220
--- Test: Beeinflusst 'Post_Count' die Variable 'Volume'? (maxlag=5) ---
-> Keine signifikante Granger-Kausalität gefunden (alle p >= 0.05) für getestete Lags.
--- Test: Beeinflusst 'Total_Posts' die Variable 'Volume'? (maxlag=5) ---
-> WARNUNG: Mind. eine Zeitreihe (Volume: True, Total_Posts: False) nicht stationär. Ergebnisse evtl. unzuverlässig!
-> Signifikante Granger-Kausalität gefunden (p < 0.05) bei folgenden Lags:
    Lag 1: p=0.0024
    Lag 2: p=0.0125
    Lag 3: p=0.0104
    Lag 4: p=0.0100
    Lag 5: p=0.0351
```