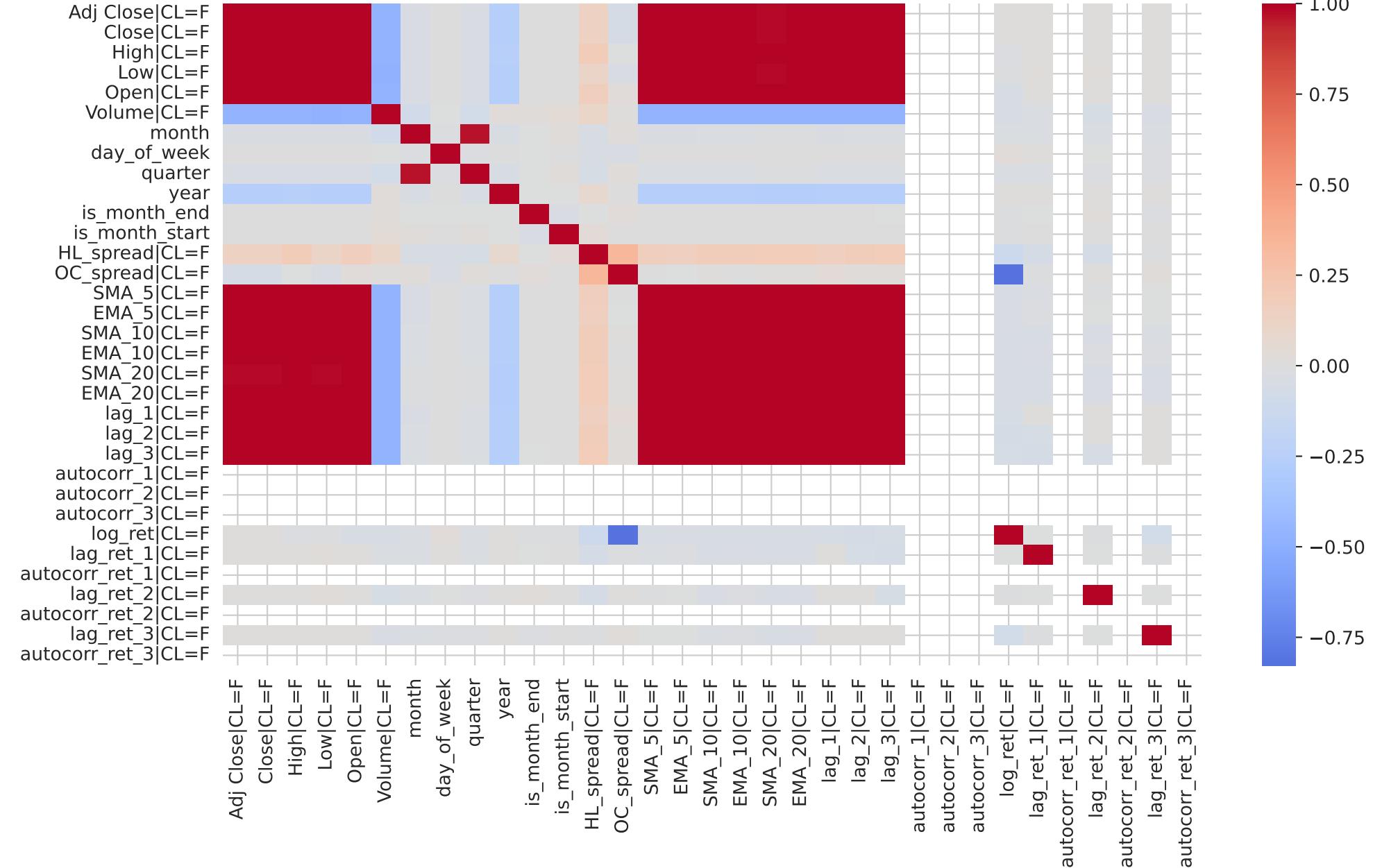


Engineered Feature EDA

Tickers: PSX, VLO, MPC, ^GSPC, CL=F, GC=F, ^VIX

Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close CL=F	float64	3770	0	0.0	2992	96.97
Close CL=F	float64	3770	0	0.0	2992	96.97
High CL=F	float64	3770	0	0.0	2924	99.96
Low CL=F	float64	3770	0	0.0	2945	96.71
Open CL=F	float64	3770	0	0.0	2939	98.5
Volume CL=F	int64	3770	0	0.0	3751	262001.0
month	int32	3770	0	0.0	12	2.0
day_of_week	int32	3770	0	0.0	5	0.0
quarter	int32	3770	0	0.0	4	1.0
year	int32	3770	0	0.0	16	2011.0
is_month_end	int64	3770	0	0.0	2	1.0
is_month_start	int64	3770	0	0.0	2	0.0
HL_spread CL=F	float64	3770	0	0.0	1718	3.25
OC_spread CL=F	float64	3770	0	0.0	2004	1.53
SMA_5 CL=F	float64	3770	0	0.0	3693	96.97
EMA_5 CL=F	float64	3770	0	0.0	3770	96.97
SMA_10 CL=F	float64	3770	0	0.0	3736	96.97
EMA_10 CL=F	float64	3770	0	0.0	3770	96.97
SMA_20 CL=F	float64	3770	0	0.0	3759	96.97
EMA_20 CL=F	float64	3770	0	0.0	3770	96.97
lag_1 CL=F	float64	3769	1	0.03	2991	96.97
lag_2 CL=F	float64	3768	2	0.05	2991	96.97
lag_3 CL=F	float64	3767	3	0.08	2990	96.97
autocorr_1 CL=F	float64	3770	0	0.0	1	0.9958
autocorr_2 CL=F	float64	3770	0	0.0	1	0.993
autocorr_3 CL=F	float64	3770	0	0.0	1	0.9905
log_ret CL=F	float64	3767	3	0.08	3754	0.0271
lag_ret_1 CL=F	float64	3766	4	0.11	3753	0.0271
autocorr_ret_1 CL=F	float64	3770	0	0.0	1	0.0006
lag_ret_2 CL=F	float64	3765	5	0.13	3752	0.0271
autocorr_ret_2 CL=F	float64	3770	0	0.0	1	-0.008
lag_ret_3 CL=F	float64	3764	6	0.16	3751	0.0271
autocorr_ret_3 CL=F	float64	3770	0	0.0	1	-0.0806

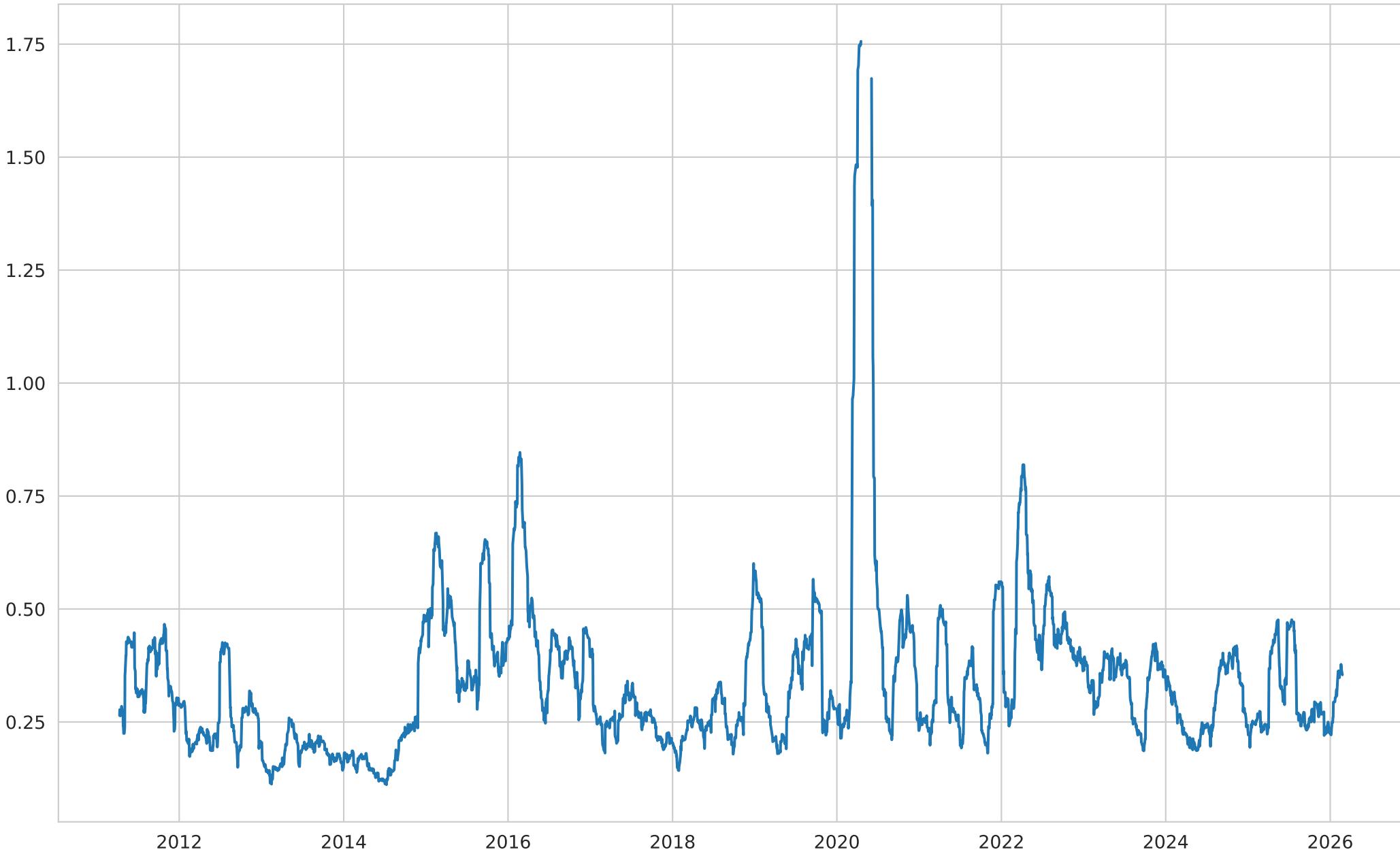
CL=F • Correlation



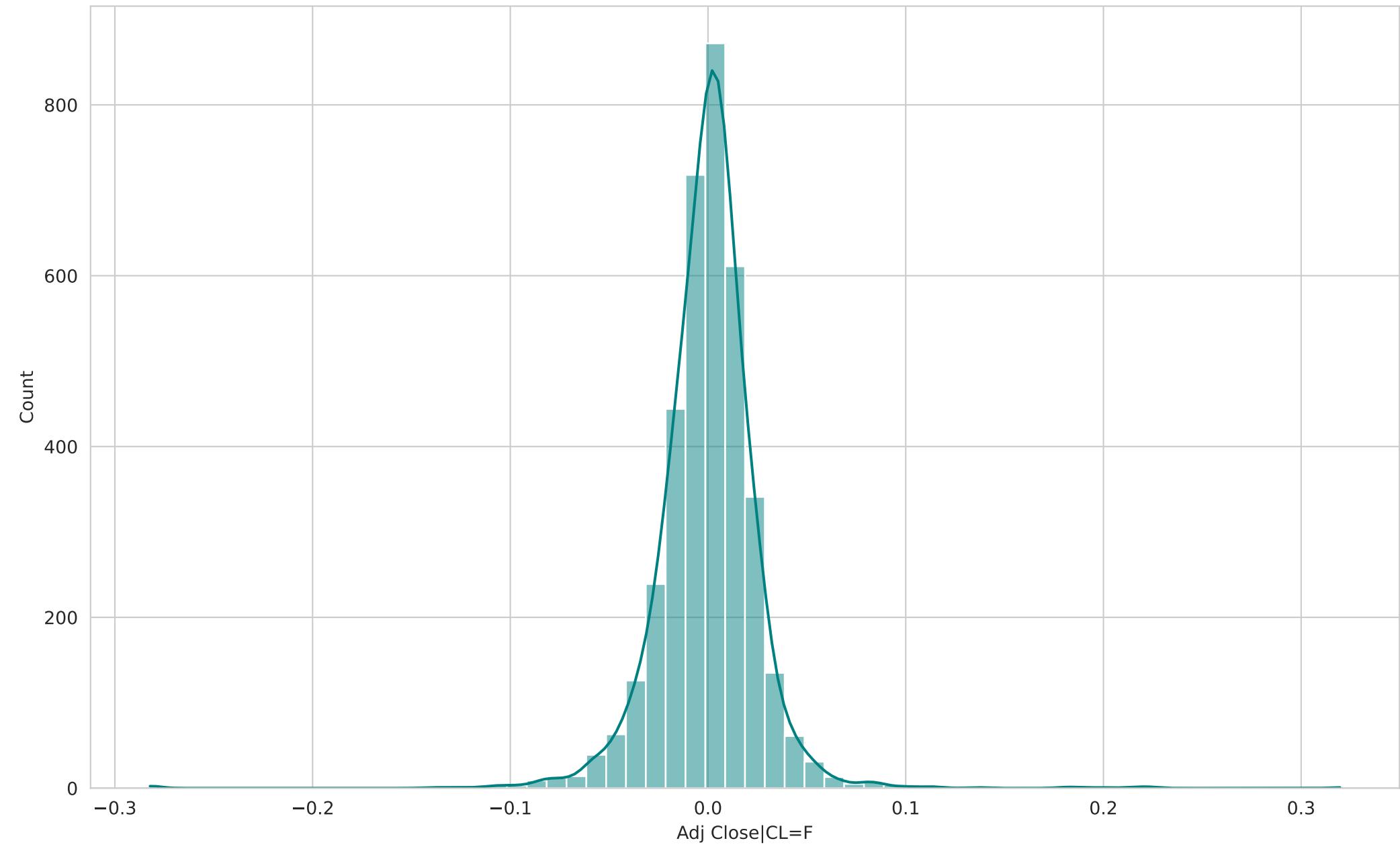
CL=F • Price (Adj Close|CL=F)



CL=F • 30-Day Rolling Vol

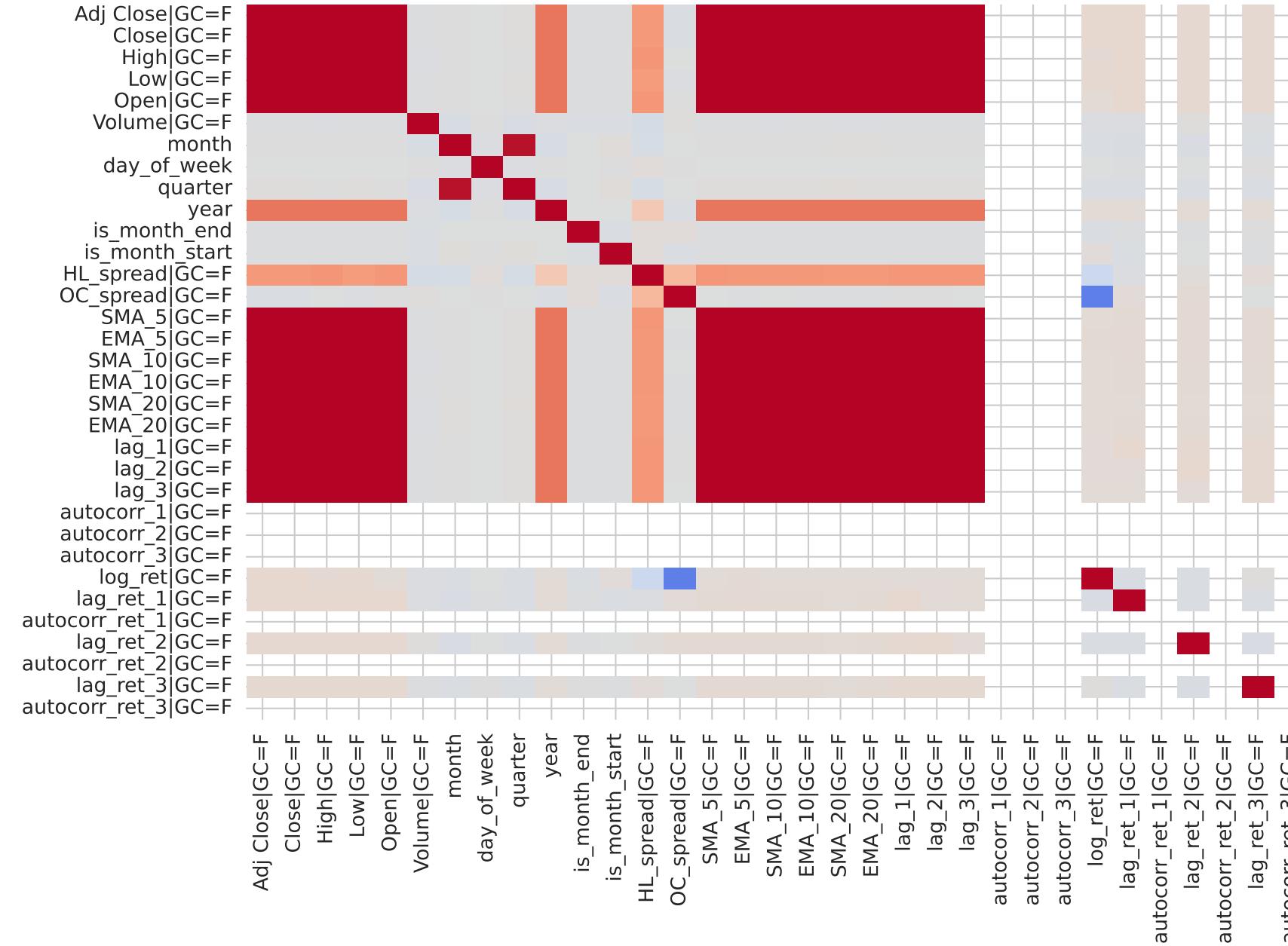


CL=F • Return Distribution



Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close GC=F	float64	3769	0	0.0	3106	1409.3
Close GC=F	float64	3769	0	0.0	3106	1409.3
High GC=F	float64	3769	0	0.0	3116	1415.7
Low GC=F	float64	3769	0	0.0	3118	1407.7
Open GC=F	float64	3769	0	0.0	3101	1410.7
Volume GC=F	int64	3769	0	0.0	1319	205.0
month	int32	3769	0	0.0	12	2.0
day_of_week	int32	3769	0	0.0	5	0.0
quarter	int32	3769	0	0.0	4	1.0
year	int32	3769	0	0.0	16	2011.0
is_month_end	int64	3769	0	0.0	2	1.0
is_month_start	int64	3769	0	0.0	2	0.0
HL_spread GC=F	float64	3769	0	0.0	1165	8.0
OC_spread GC=F	float64	3769	0	0.0	1350	1.3999
SMA_5 GC=F	float64	3769	0	0.0	3714	1409.3
EMA_5 GC=F	float64	3769	0	0.0	3769	1409.3
SMA_10 GC=F	float64	3769	0	0.0	3741	1409.3
EMA_10 GC=F	float64	3769	0	0.0	3769	1409.3
SMA_20 GC=F	float64	3769	0	0.0	3758	1409.3
EMA_20 GC=F	float64	3769	0	0.0	3769	1409.3
lag_1 GC=F	float64	3768	1	0.03	3105	1409.3
lag_2 GC=F	float64	3767	2	0.05	3104	1409.3
lag_3 GC=F	float64	3766	3	0.08	3103	1409.3
autocorr_1 GC=F	float64	3769	0	0.0	1	0.9993
autocorr_2 GC=F	float64	3769	0	0.0	1	0.9987
autocorr_3 GC=F	float64	3769	0	0.0	1	0.9983
log_ret GC=F	float64	3768	1	0.03	3752	0.0151
lag_ret_1 GC=F	float64	3767	2	0.05	3751	0.0151
autocorr_ret_1 GC=F	float64	3769	0	0.0	1	-0.0335
lag_ret_2 GC=F	float64	3766	3	0.08	3750	0.0151
autocorr_ret_2 GC=F	float64	3769	0	0.0	1	-0.0218
lag_ret_3 GC=F	float64	3765	4	0.11	3749	0.0151
autocorr_ret_3 GC=F	float64	3769	0	0.0	1	0.0143

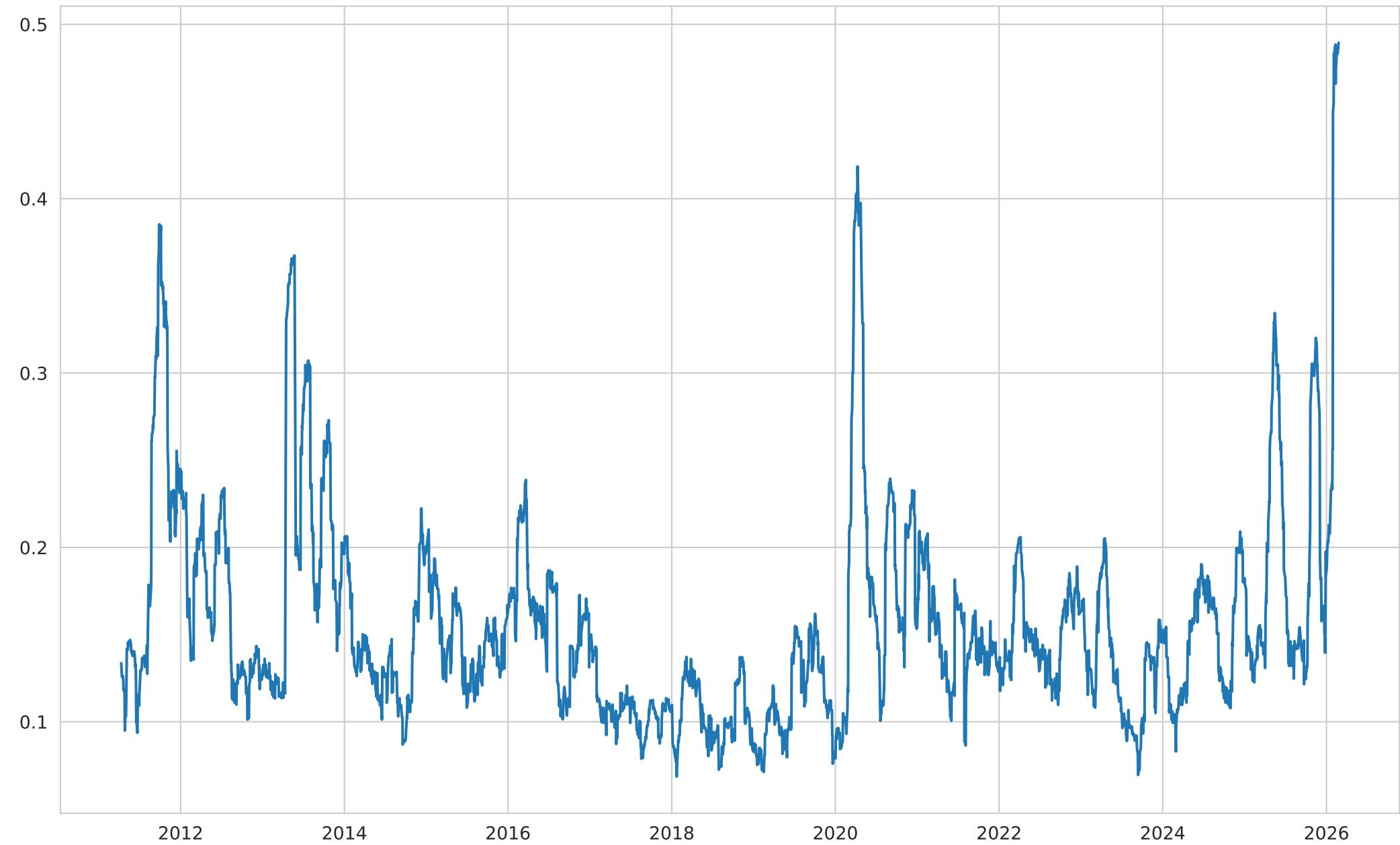
GC=F • Correlation



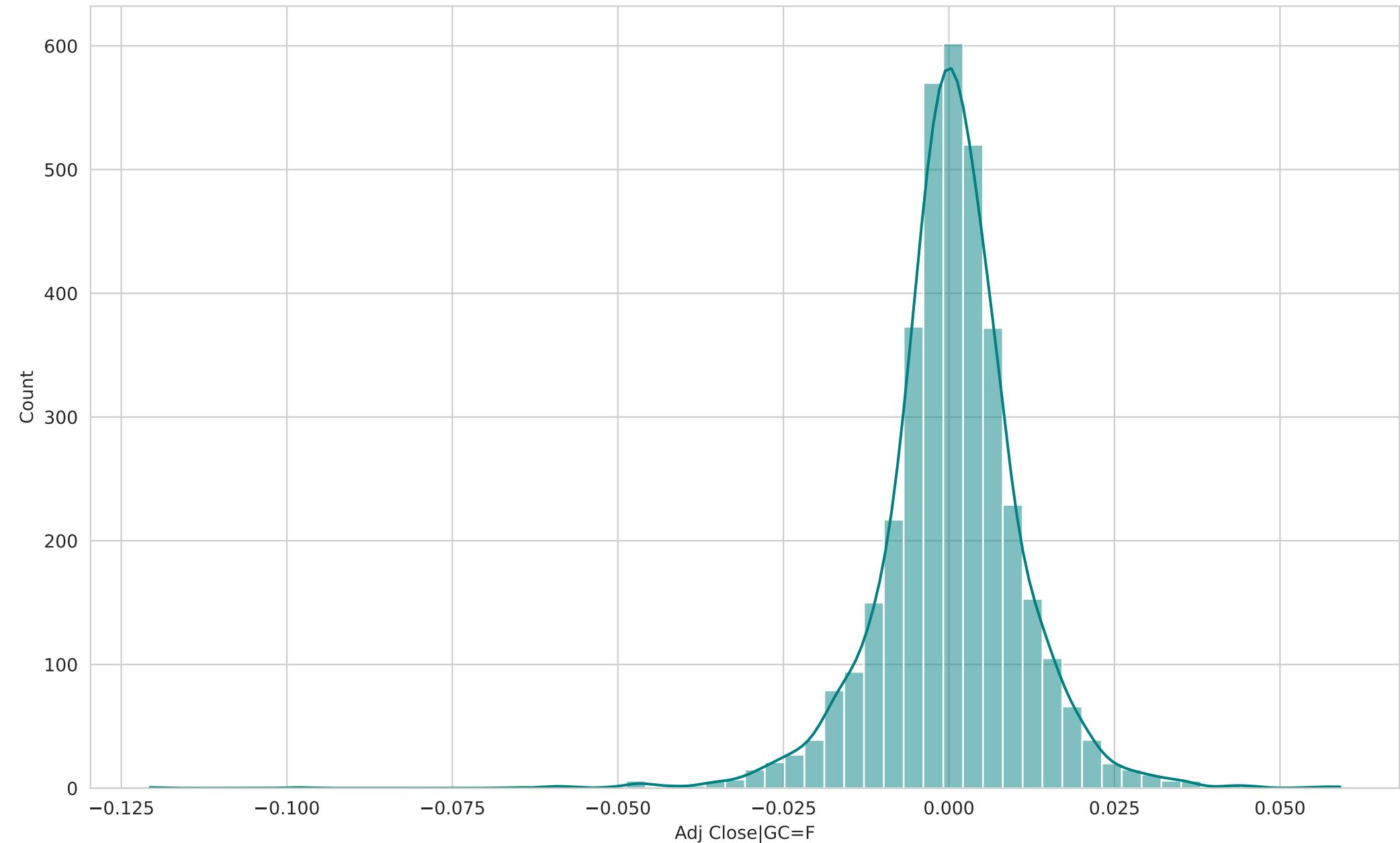
GC=F • Price (Adj Close|GC=F)



GC=F • 30-Day Rolling Vol

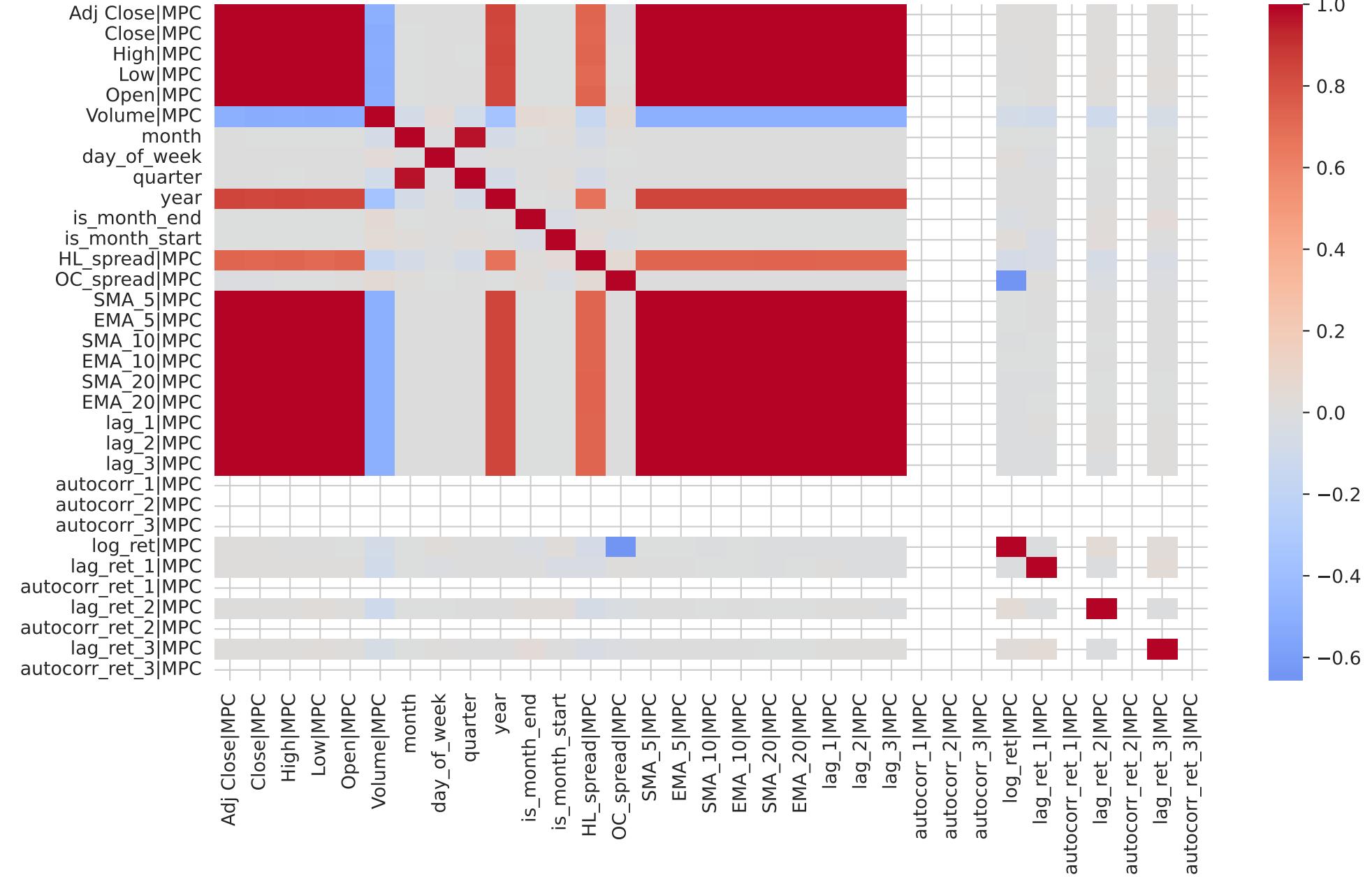


GC=F • Return Distribution



Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close MPC	float64	3688	0	0.0	3580	12.2291
High MPC	float64	3688	0	0.0	3160	18.625
Low MPC	float64	3688	0	0.0	3196	18.42
Open MPC	float64	3688	0	0.0	3162	19.375
Volume MPC	int64	3688	0	0.0	3612	1675400.0
month	int32	3688	0	0.0	12	6.0
day_of_week	int32	3688	0	0.0	5	4.0
quarter	int32	3688	0	0.0	4	2.0
year	int32	3688	0	0.0	16	2011.0
is_month_end	int64	3688	0	0.0	2	0.0
is_month_start	int64	3688	0	0.0	2	0.0
HL_spread MPC	float64	3688	0	0.0	2227	1.155
OC_spread MPC	float64	3688	0	0.0	2498	0.75
SMA_5 MPC	float64	3688	0	0.0	3684	12.2291
EMA_5 MPC	float64	3688	0	0.0	3688	12.2291
SMA_10 MPC	float64	3688	0	0.0	3685	12.2291
EMA_10 MPC	float64	3688	0	0.0	3688	12.2291
SMA_20 MPC	float64	3688	0	0.0	3686	12.2291
EMA_20 MPC	float64	3688	0	0.0	3688	12.2291
lag_1 MPC	float64	3687	1	0.03	3579	12.2291
lag_2 MPC	float64	3686	2	0.05	3578	12.2291
lag_3 MPC	float64	3685	3	0.08	3577	12.2291
autocorr_1 MPC	float64	3688	0	0.0	1	0.9995
autocorr_2 MPC	float64	3688	0	0.0	1	0.9989
autocorr_3 MPC	float64	3688	0	0.0	1	0.9984
log_ret MPC	float64	3687	1	0.03	3679	0.0107
lag_ret_1 MPC	float64	3686	2	0.05	3678	0.0107
autocorr_ret_1 MPC	float64	3688	0	0.0	1	-0.01
lag_ret_2 MPC	float64	3685	3	0.08	3677	0.0107
autocorr_ret_2 MPC	float64	3688	0	0.0	1	0.048
lag_ret_3 MPC	float64	3684	4	0.11	3676	0.0107
autocorr_ret_3 MPC	float64	3688	0	0.0	1	0.0284

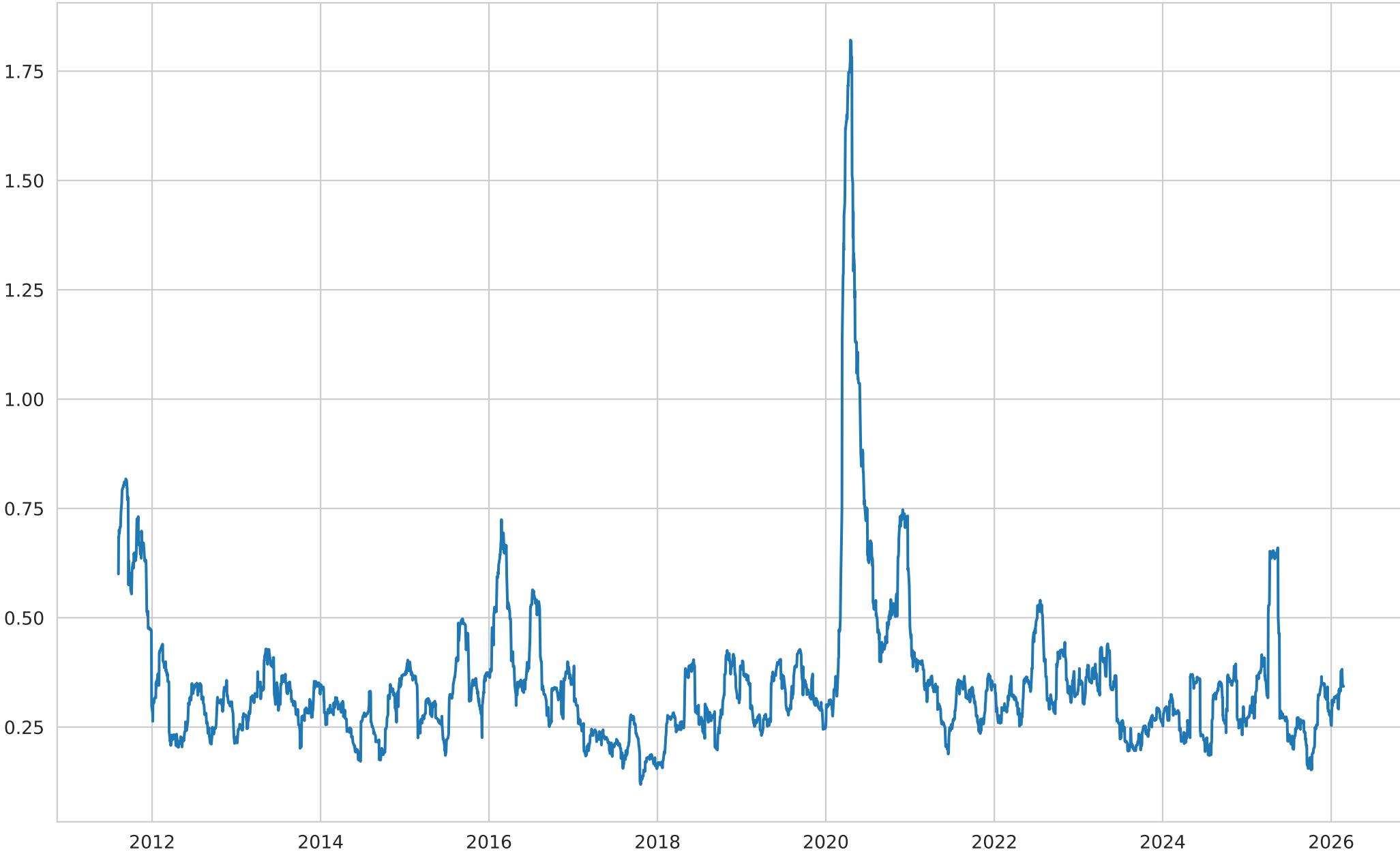
MPC • Correlation



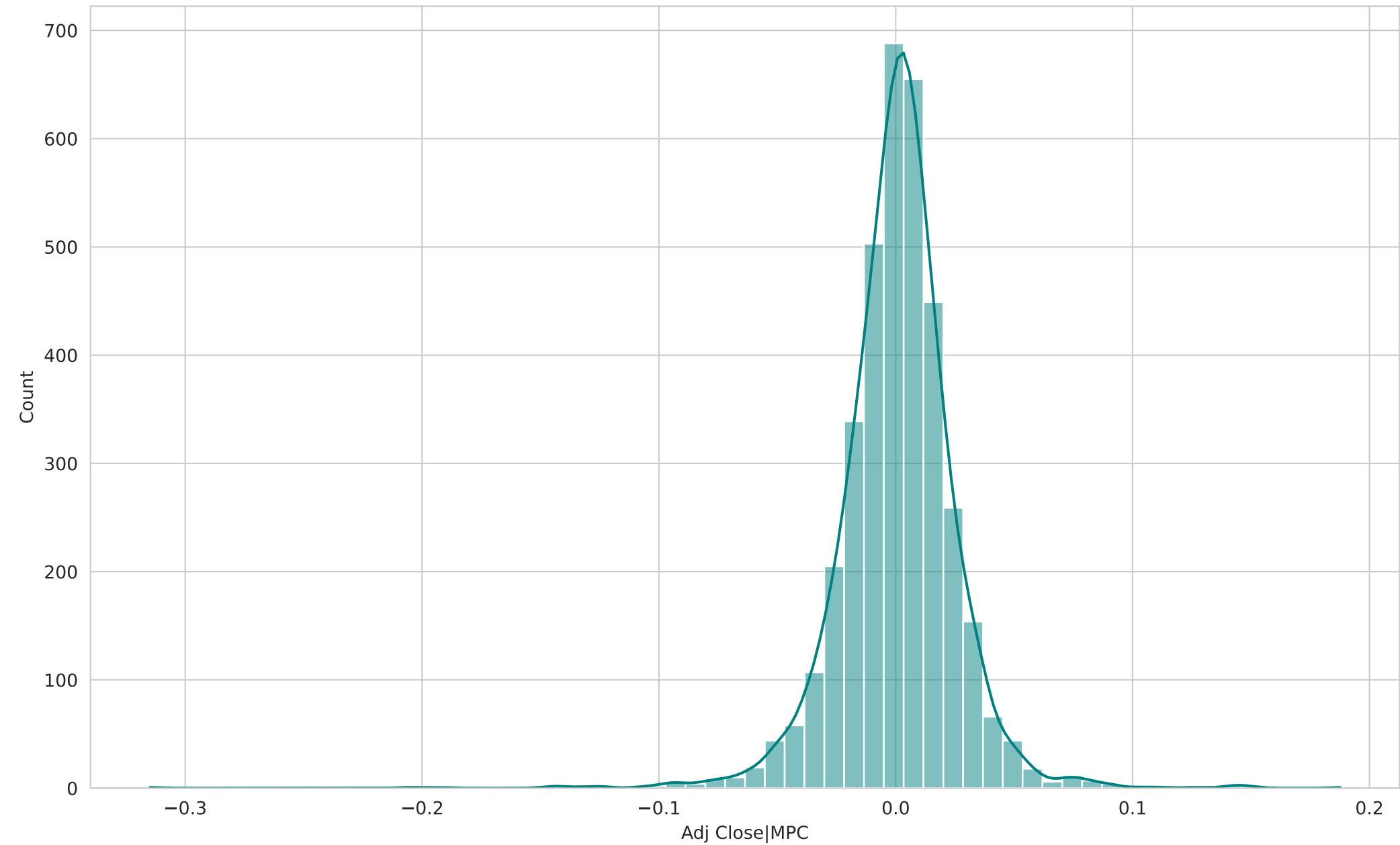
MPC • Price (Adj Close|MPC)



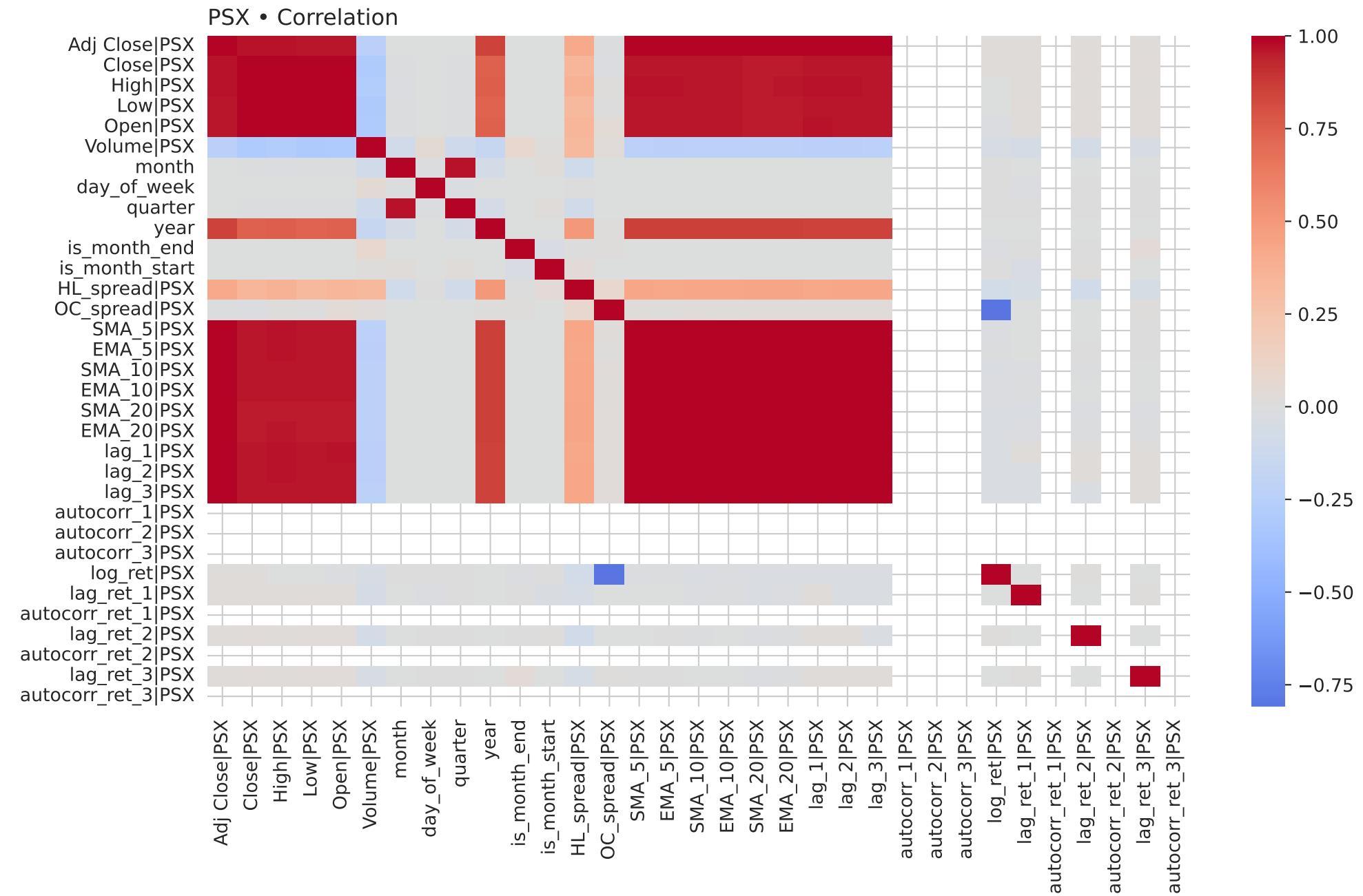
MPC • 30-Day Rolling Vol



MPC • Return Distribution



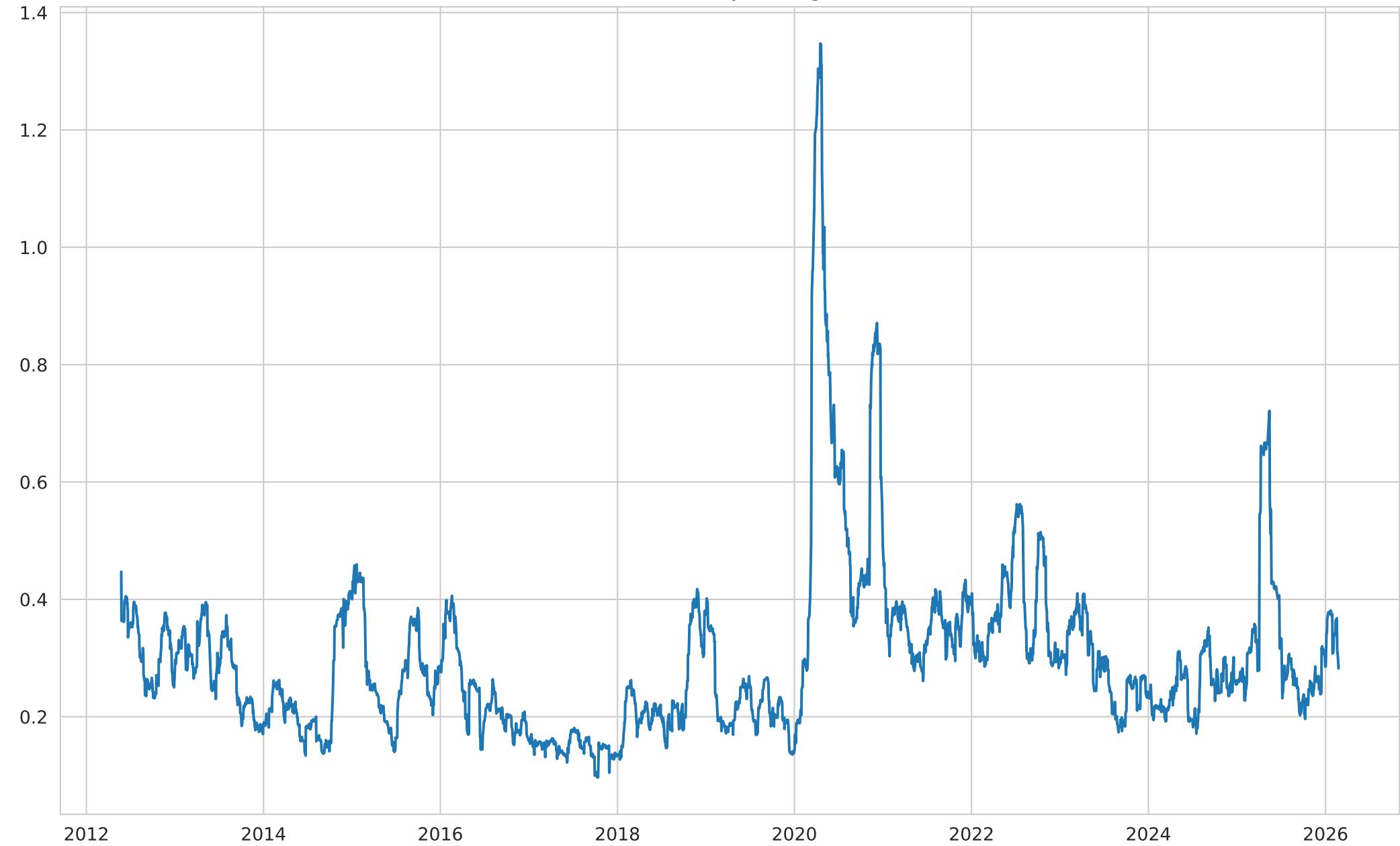
Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close PSX	float64	3487	0	0.0	3385	21.4859
High PSX	float64	3487	0	0.0	2774	34.0
Low PSX	float64	3487	0	0.0	2768	32.0
Open PSX	float64	3487	0	0.0	2727	33.25
Volume PSX	int64	3487	0	0.0	3333	654700.0
month	int32	3487	0	0.0	12	4.0
day_of_week	int32	3487	0	0.0	5	3.0
quarter	int32	3487	0	0.0	4	2.0
year	int32	3487	0	0.0	15	2012.0
is_month_end	int64	3487	0	0.0	2	0.0
is_month_start	int64	3487	0	0.0	2	0.0
HL_spread PSX	float64	3487	0	0.0	1691	2.21
OC_spread PSX	float64	3487	0	0.0	1853	-0.75
SMA_5 PSX	float64	3487	0	0.0	3482	21.4859
EMA_5 PSX	float64	3487	0	0.0	3487	21.4859
SMA_10 PSX	float64	3487	0	0.0	3487	21.4859
EMA_10 PSX	float64	3487	0	0.0	3487	21.4859
SMA_20 PSX	float64	3487	0	0.0	3486	21.4859
EMA_20 PSX	float64	3487	0	0.0	3487	21.4859
lag_1 PSX	float64	3486	1	0.03	3384	21.4859
lag_2 PSX	float64	3485	2	0.06	3383	21.4859
lag_3 PSX	float64	3484	3	0.09	3382	21.4859
autocorr_1 PSX	float64	3487	0	0.0	1	0.9987
autocorr_2 PSX	float64	3487	0	0.0	1	0.9974
autocorr_3 PSX	float64	3487	0	0.0	1	0.996
log_ret PSX	float64	3486	1	0.03	3471	0.0846
lag_ret_1 PSX	float64	3485	2	0.06	3470	0.0846
autocorr_ret_1 PSX	float64	3487	0	0.0	1	-0.0074
lag_ret_2 PSX	float64	3484	3	0.09	3469	0.0846
autocorr_ret_2 PSX	float64	3487	0	0.0	1	0.0142
lag_ret_3 PSX	float64	3483	4	0.11	3468	0.0846
autocorr_ret_3 PSX	float64	3487	0	0.0	1	-0.0073



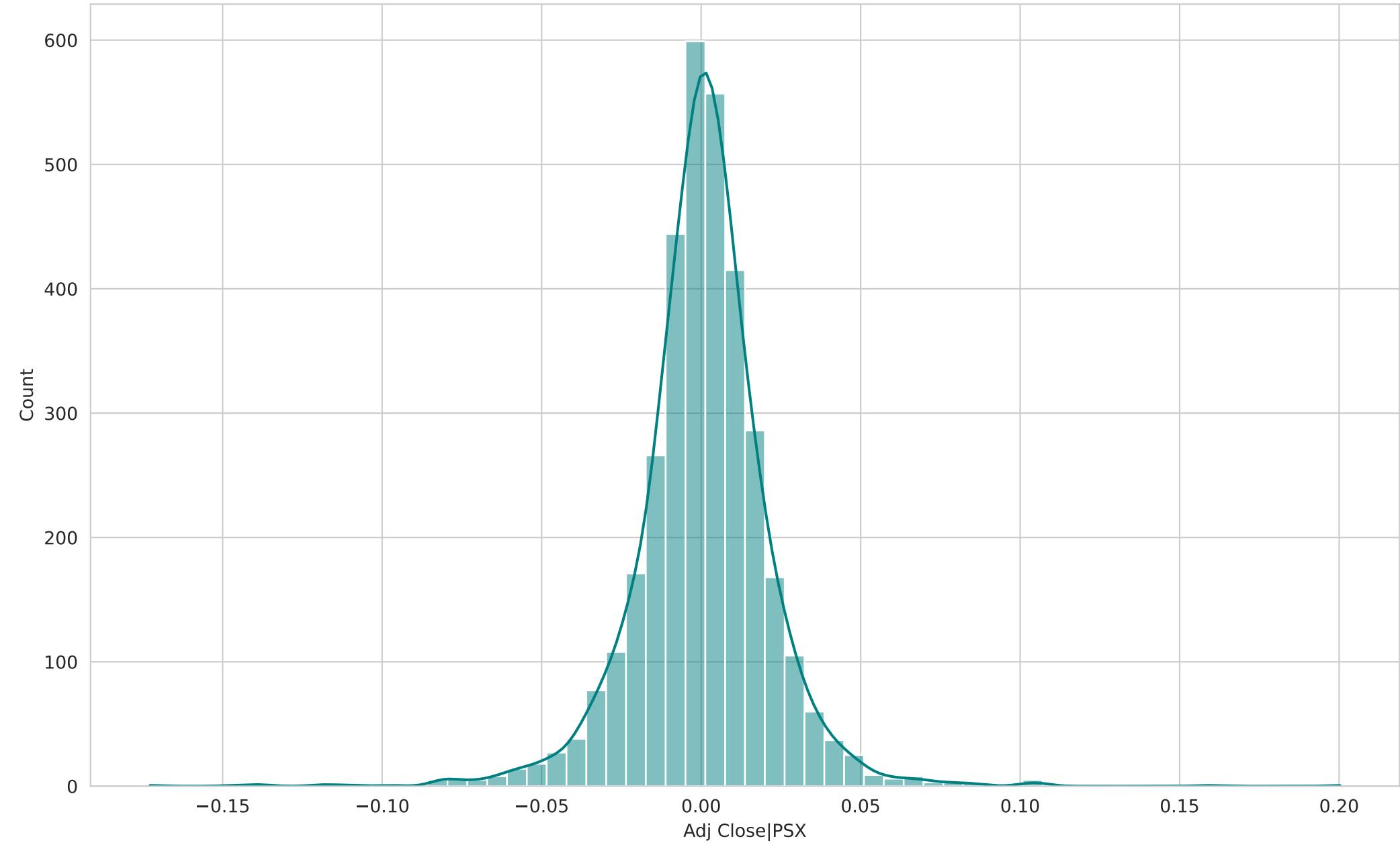
PSX • Price (Adj Close|PSX)



PSX • 30-Day Rolling Vol



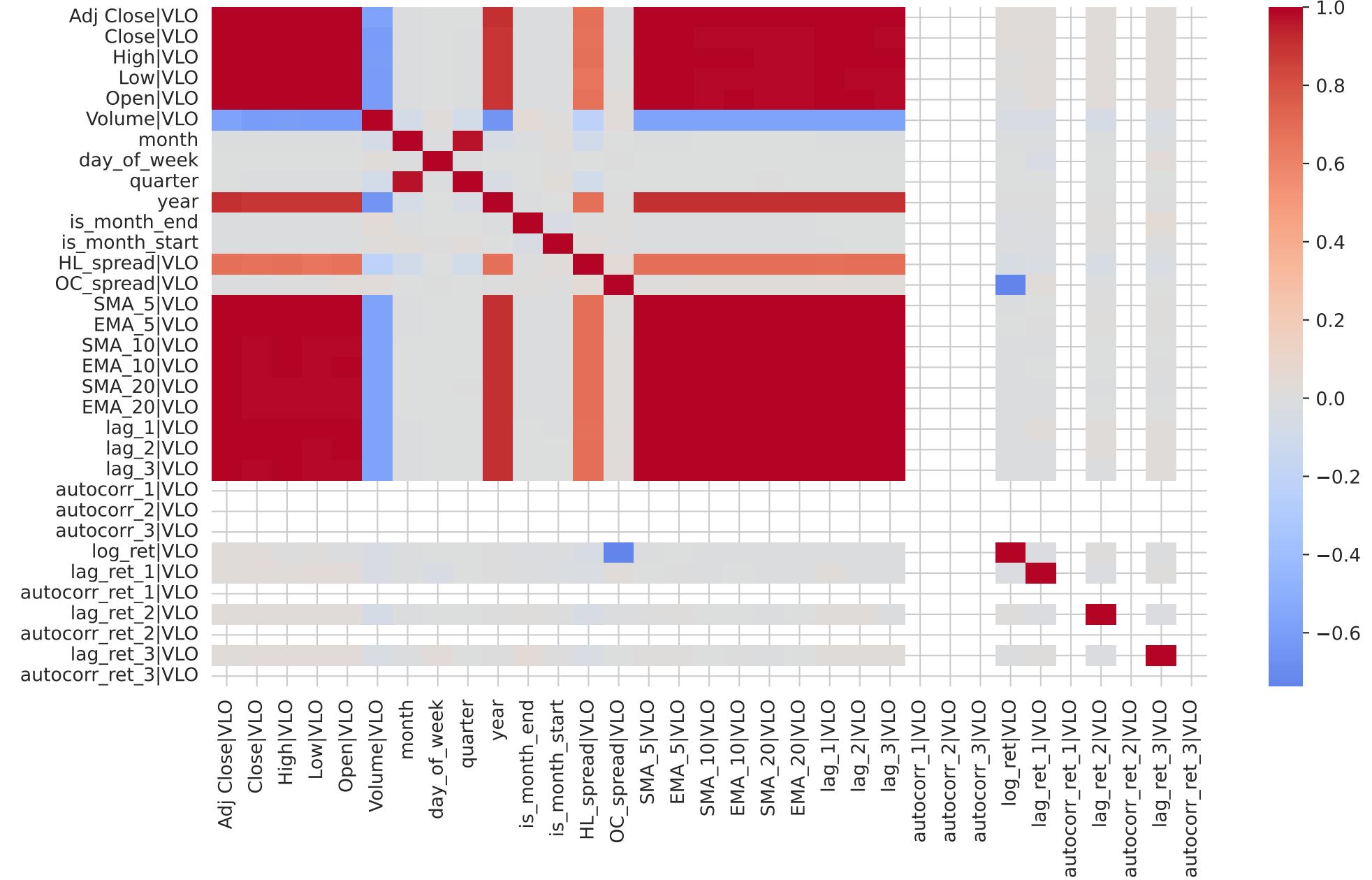
PSX • Return Distribution



Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close VLO	float64	3770	0	0.0	3684	15.468
High VLO	float64	3770	0	0.0	3258	25.7587
Low VLO	float64	3770	0	0.0	3209	25.457
Open VLO	float64	3770	0	0.0	3212	26.2614
Volume VLO	int64	3770	0	0.0	3700	12430356.0
month	int32	3770	0	0.0	12	2.0
day_of_week	int32	3770	0	0.0	5	0.0
quarter	int32	3770	0	0.0	4	1.0
year	int32	3770	0	0.0	16	2011.0
is_month_end	int64	3770	0	0.0	2	1.0
is_month_start	int64	3770	0	0.0	2	0.0
HL_spread VLO	float64	3770	0	0.0	2296	0.9324
OC_spread VLO	float64	3770	0	0.0	2544	0.5027
SMA_5 VLO	float64	3770	0	0.0	3759	15.468
EMA_5 VLO	float64	3770	0	0.0	3770	15.468
SMA_10 VLO	float64	3770	0	0.0	3767	15.468
EMA_10 VLO	float64	3770	0	0.0	3770	15.468
SMA_20 VLO	float64	3770	0	0.0	3770	15.468
EMA_20 VLO	float64	3770	0	0.0	3770	15.468
lag_1 VLO	float64	3769	1	0.03	3683	15.468
lag_2 VLO	float64	3768	2	0.05	3682	15.468
lag_3 VLO	float64	3767	3	0.08	3681	15.468
autocorr_1 VLO	float64	3770	0	0.0	1	0.9992
autocorr_2 VLO	float64	3770	0	0.0	1	0.9984
autocorr_3 VLO	float64	3770	0	0.0	1	0.9976
log_ret VLO	float64	3769	1	0.03	3758	-0.0435
lag_ret_1 VLO	float64	3768	2	0.05	3757	-0.0435
autocorr_ret_1 VLO	float64	3770	0	0.0	1	-0.021
lag_ret_2 VLO	float64	3767	3	0.08	3756	-0.0435
autocorr_ret_2 VLO	float64	3770	0	0.0	1	0.0146
lag_ret_3 VLO	float64	3766	4	0.11	3755	-0.0435
autocorr_ret_3 VLO	float64	3770	0	0.0	1	-0.0109

VLO • Schema & Missingness

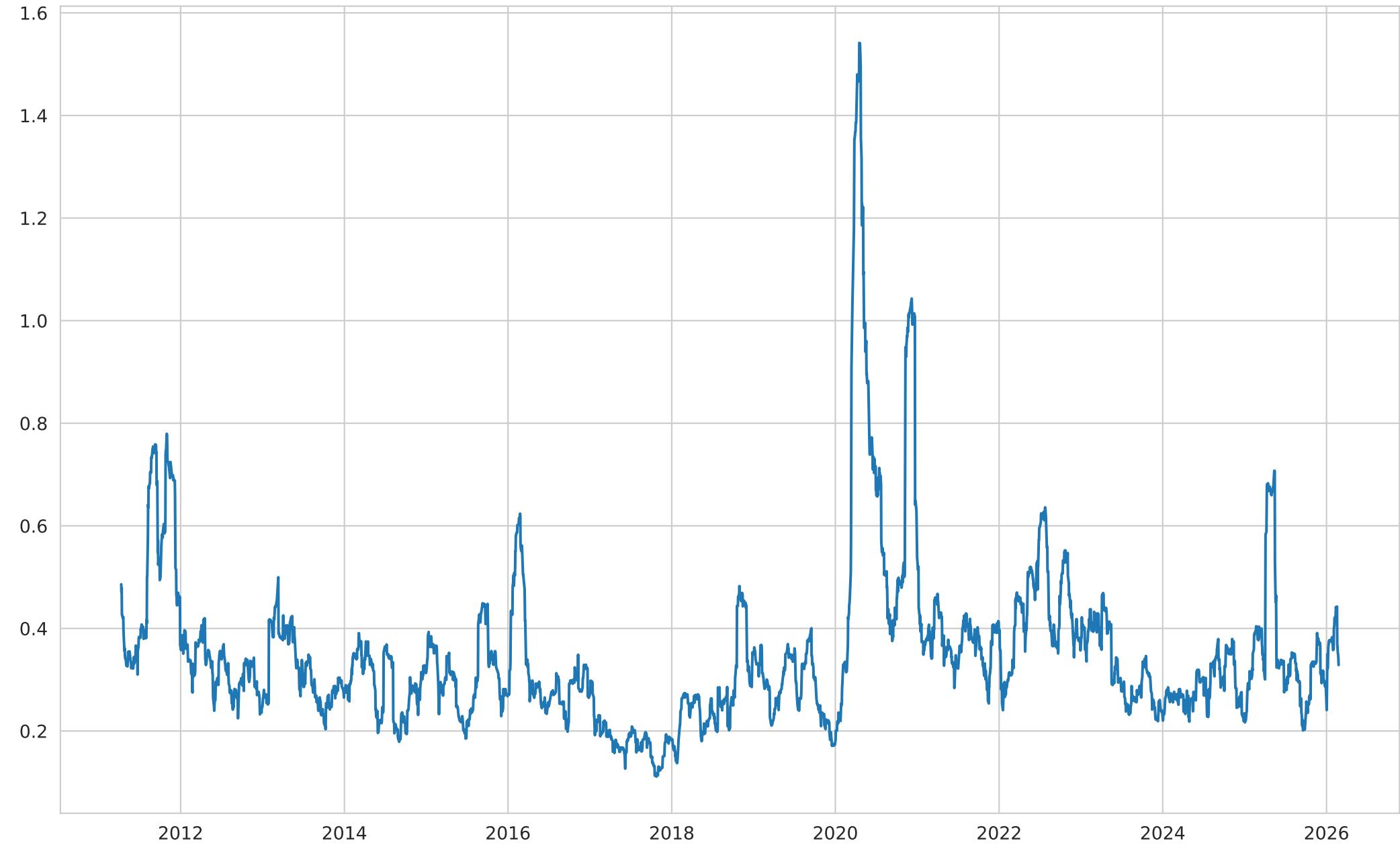
VLO • Correlation



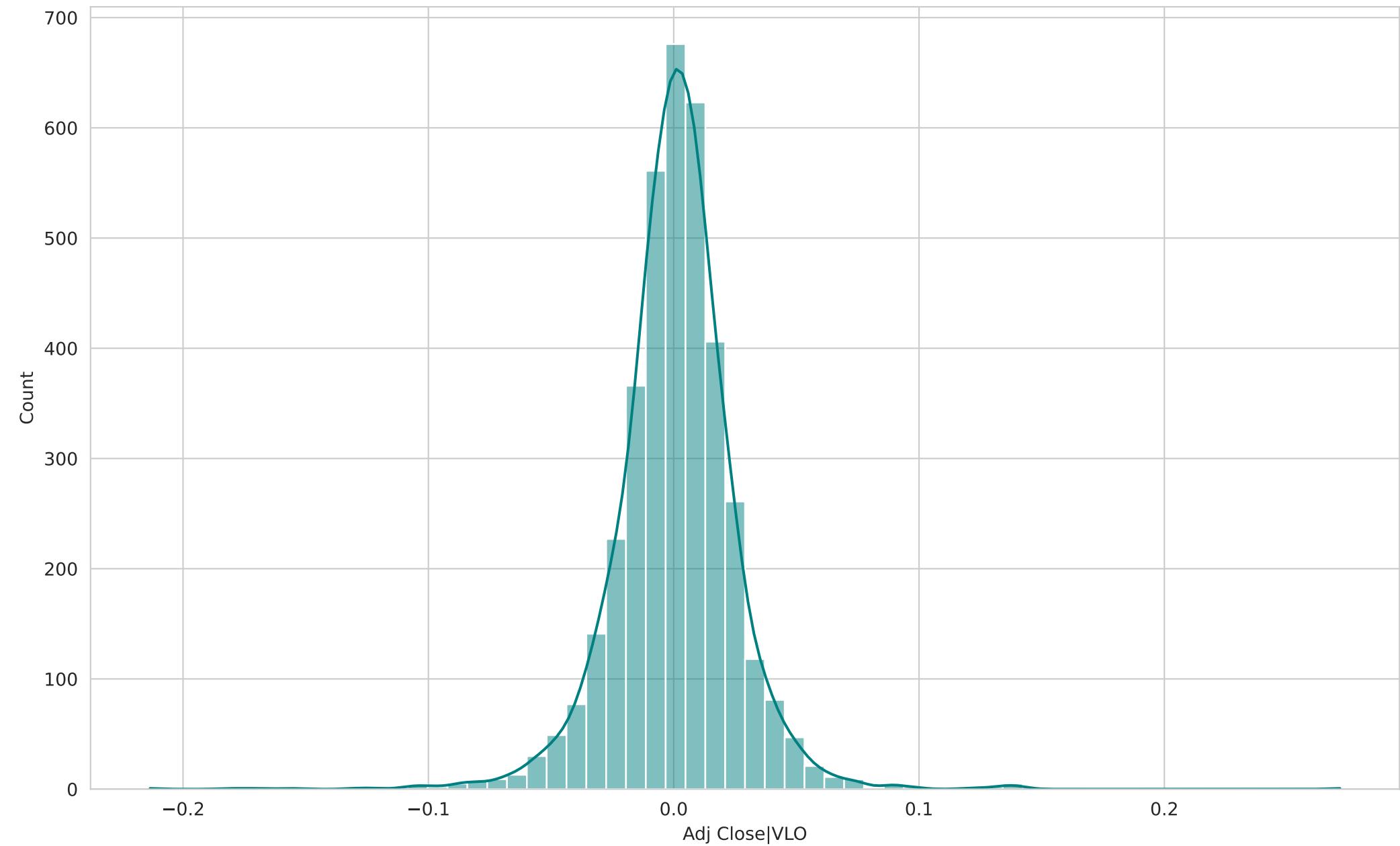
VLO • Price (Adj Close|VLO)



VLO • 30-Day Rolling Vol

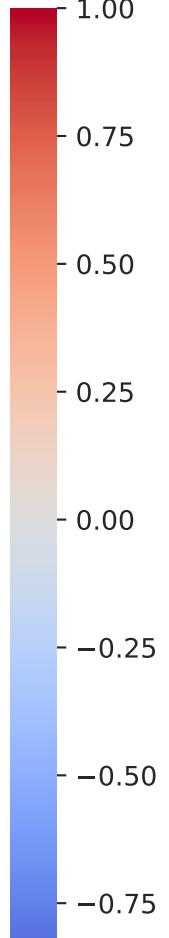
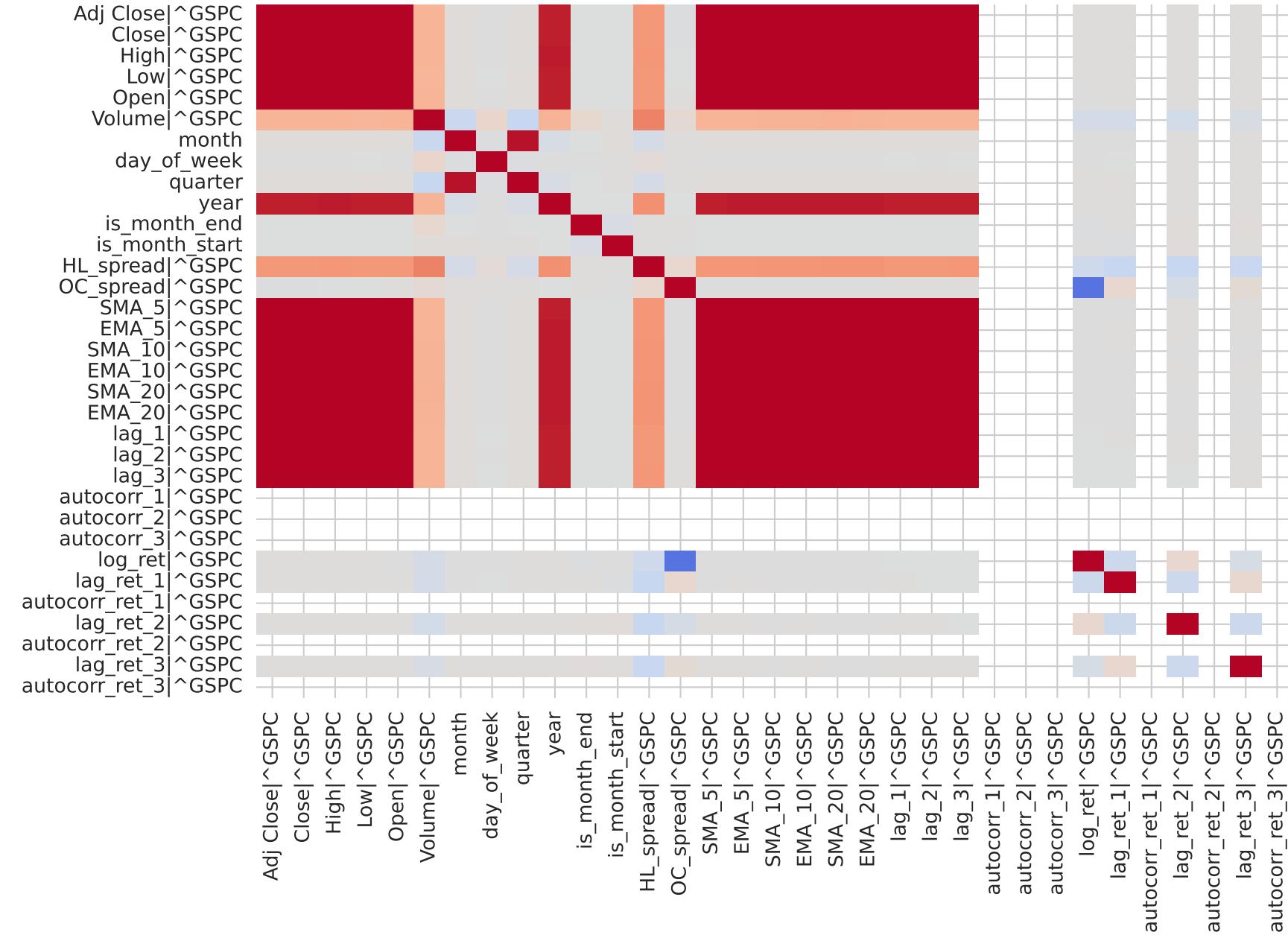


VLO • Return Distribution



Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close ^GSPC	float64	3770	0	0.0	3756	1327.22
Close ^GSPC	float64	3770	0	0.0	3756	1327.22
High ^GSPC	float64	3770	0	0.0	3743	1329.38
Low ^GSPC	float64	3770	0	0.0	3751	1320.55
Open ^GSPC	float64	3770	0	0.0	3746	1321.61
Volume ^GSPC	int64	3770	0	0.0	3722	1252850000.0
month	int32	3770	0	0.0	12	2.0
day_of_week	int32	3770	0	0.0	5	0.0
quarter	int32	3770	0	0.0	4	1.0
year	int32	3770	0	0.0	16	2011.0
is_month_end	int64	3770	0	0.0	2	1.0
is_month_start	int64	3770	0	0.0	2	0.0
HL_spread ^GSPC	float64	3770	0	0.0	3312	8.83
OC_spread ^GSPC	float64	3770	0	0.0	3414	-5.61
SMA_5 ^GSPC	float64	3770	0	0.0	3765	1327.22
EMA_5 ^GSPC	float64	3770	0	0.0	3770	1327.22
SMA_10 ^GSPC	float64	3770	0	0.0	3770	1327.22
EMA_10 ^GSPC	float64	3770	0	0.0	3770	1327.22
SMA_20 ^GSPC	float64	3770	0	0.0	3769	1327.22
EMA_20 ^GSPC	float64	3770	0	0.0	3770	1327.22
lag_1 ^GSPC	float64	3769	1	0.03	3755	1327.22
lag_2 ^GSPC	float64	3768	2	0.05	3754	1327.22
lag_3 ^GSPC	float64	3767	3	0.08	3753	1327.22
autocorr_1 ^GSPC	float64	3770	0	0.0	1	0.9997
autocorr_2 ^GSPC	float64	3770	0	0.0	1	0.9995
autocorr_3 ^GSPC	float64	3770	0	0.0	1	0.9992
log_ret ^GSPC	float64	3769	1	0.03	3769	-0.0159
lag_ret_1 ^GSPC	float64	3768	2	0.05	3768	-0.0159
autocorr_ret_1 ^GSPC	float64	3770	0	0.0	1	-0.1193
lag_ret_2 ^GSPC	float64	3767	3	0.08	3767	-0.0159
autocorr_ret_2 ^GSPC	float64	3770	0	0.0	1	0.0823
lag_ret_3 ^GSPC	float64	3766	4	0.11	3766	-0.0159
autocorr_ret_3 ^GSPC	float64	3770	0	0.0	1	-0.0476

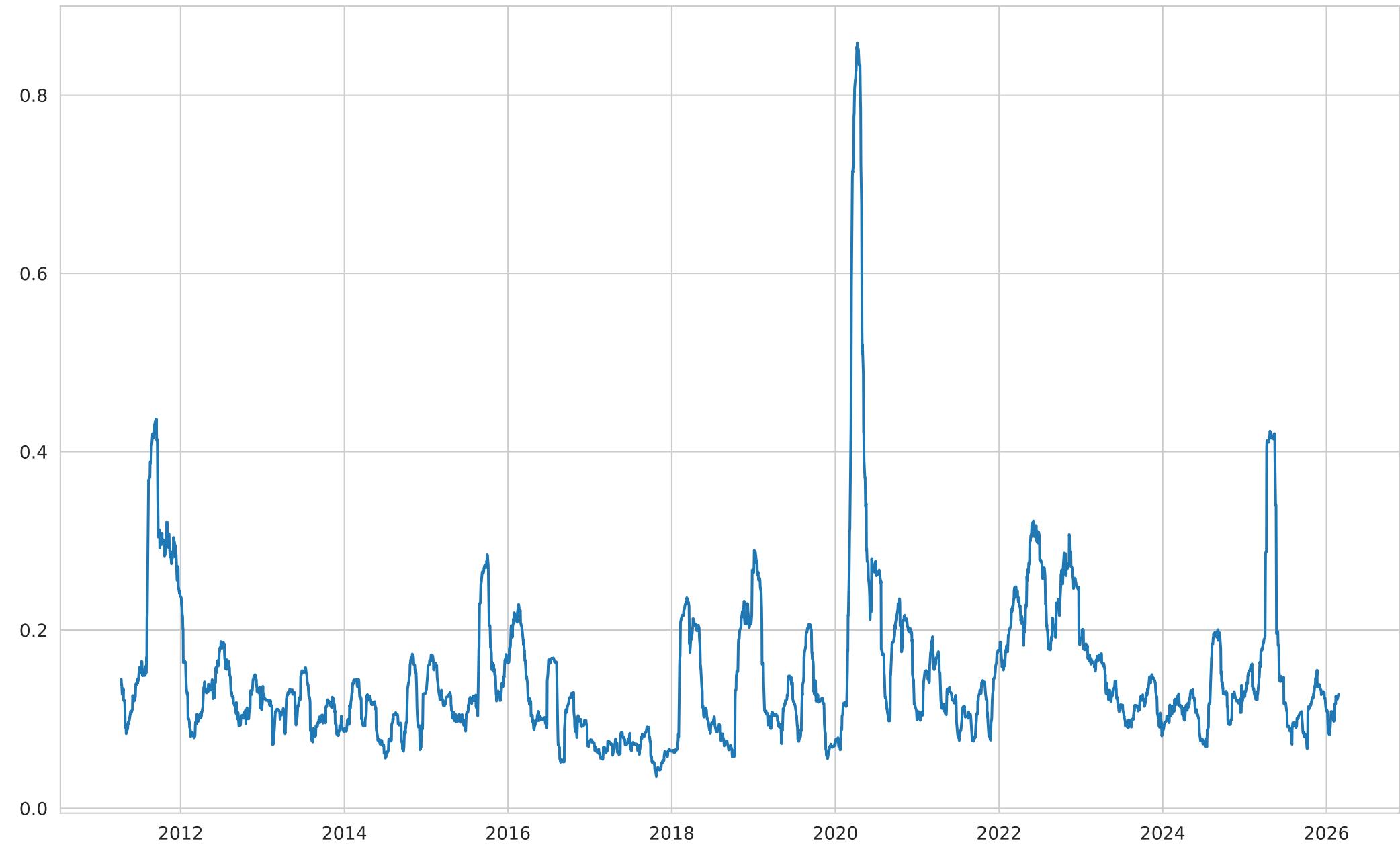
^GSPC • Correlation



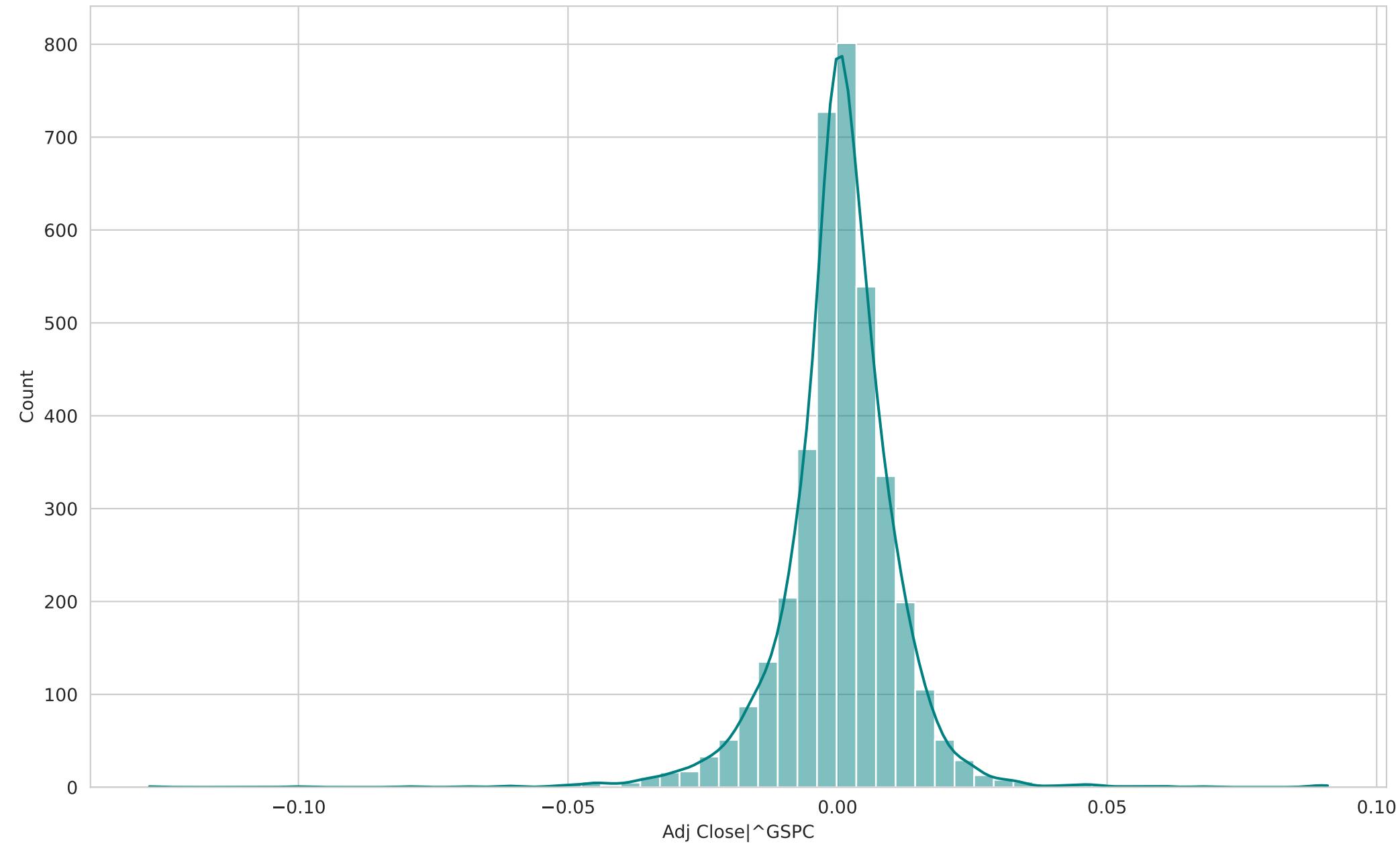
\wedge GSPC • Price (Adj Close| \wedge GSPC)



\wedge GSPC • 30-Day Rolling Vol

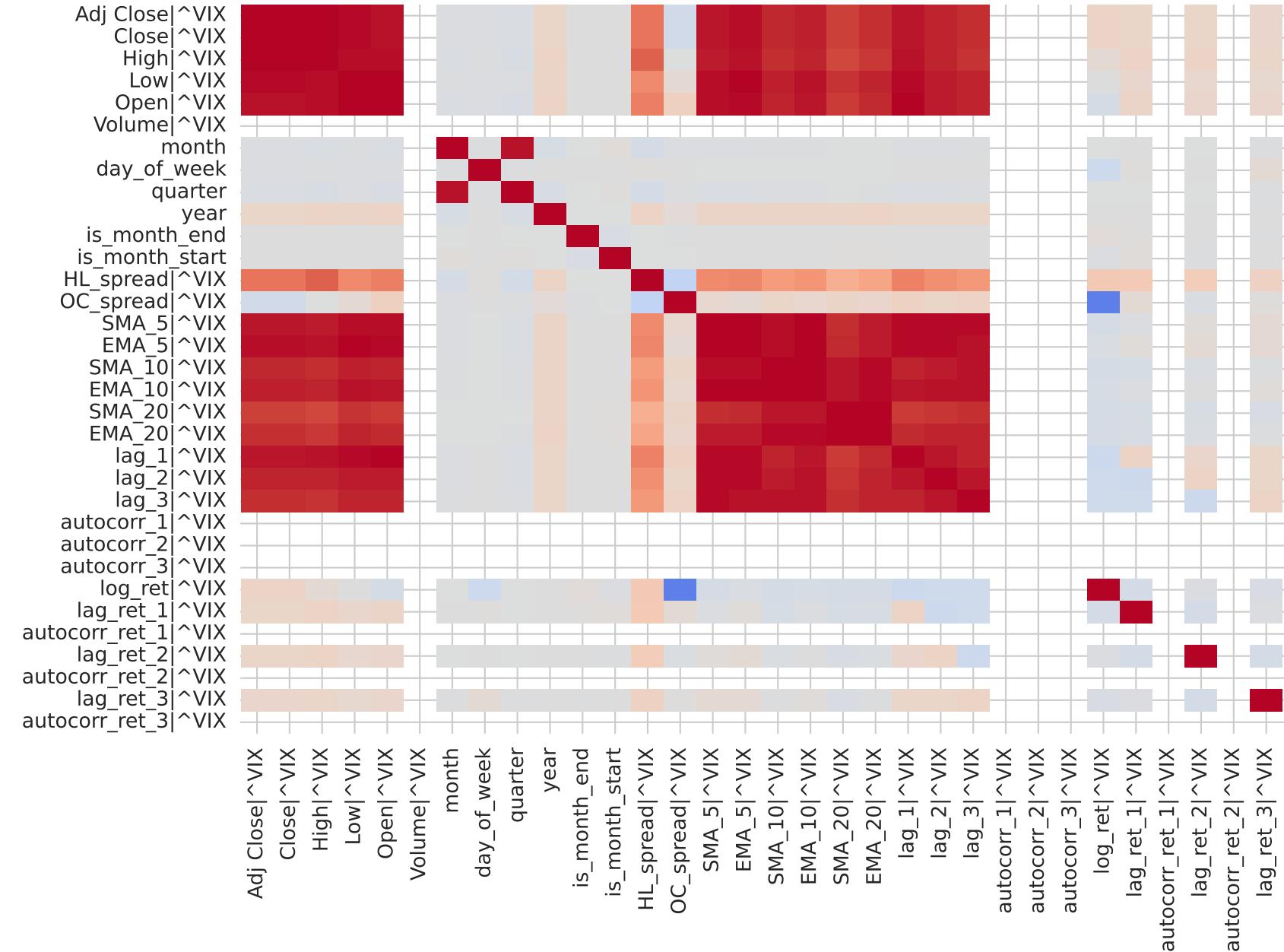


\wedge GSPC • Return Distribution

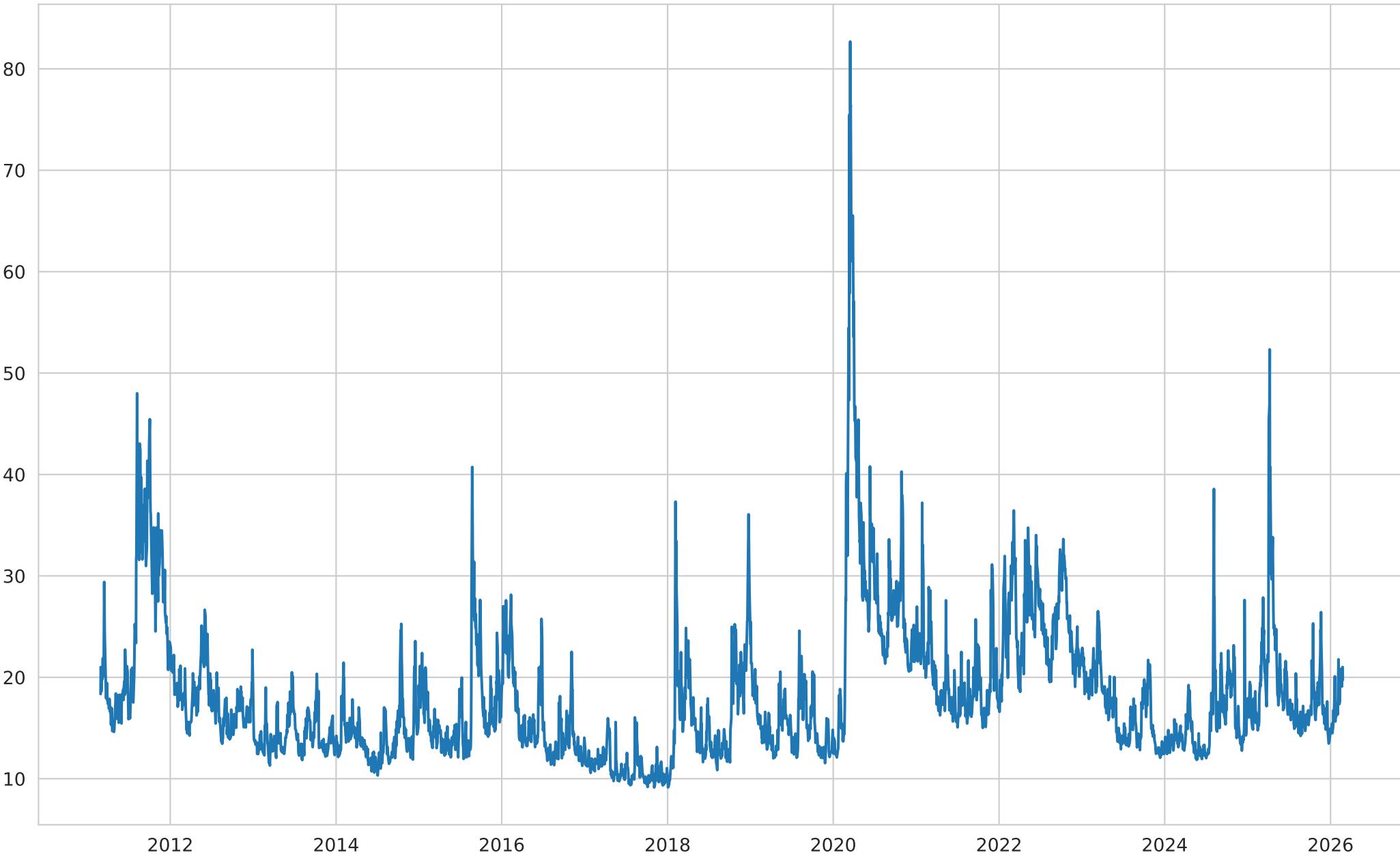


Column	dtype	non_null	missing	missing_%	unique	sample
Adj Close ^{^VIX}	float64	3770	0	0.0	1663	18.35
Close ^{^VIX}	float64	3770	0	0.0	1663	18.35
High ^{^VIX}	float64	3770	0	0.0	1736	19.27
Low ^{^VIX}	float64	3770	0	0.0	1596	18.14
Open ^{^VIX}	float64	3770	0	0.0	1653	19.12
Volume ^{^VIX}	int64	3770	0	0.0	1	0.0
month	int32	3770	0	0.0	12	2.0
day_of_week	int32	3770	0	0.0	5	0.0
quarter	int32	3770	0	0.0	4	1.0
year	int32	3770	0	0.0	16	2011.0
is_month_end	int64	3770	0	0.0	2	1.0
is_month_start	int64	3770	0	0.0	2	0.0
HL_spread ^{^VIX}	float64	3770	0	0.0	1971	1.13
OC_spread ^{^VIX}	float64	3770	0	0.0	2021	0.77
SMA_5 ^{^VIX}	float64	3770	0	0.0	3530	18.35
EMA_5 ^{^VIX}	float64	3770	0	0.0	3770	18.35
SMA_10 ^{^VIX}	float64	3770	0	0.0	3664	18.35
EMA_10 ^{^VIX}	float64	3770	0	0.0	3770	18.35
SMA_20 ^{^VIX}	float64	3770	0	0.0	3726	18.35
EMA_20 ^{^VIX}	float64	3770	0	0.0	3770	18.35
lag_1 ^{^VIX}	float64	3769	1	0.03	1663	18.35
lag_2 ^{^VIX}	float64	3768	2	0.05	1663	18.35
lag_3 ^{^VIX}	float64	3767	3	0.08	1663	18.35
autocorr_1 ^{^VIX}	float64	3770	0	0.0	1	0.9633
autocorr_2 ^{^VIX}	float64	3770	0	0.0	1	0.9388
autocorr_3 ^{^VIX}	float64	3770	0	0.0	1	0.9111
log_ret ^{^VIX}	float64	3769	1	0.03	3738	0.1354
lag_ret_1 ^{^VIX}	float64	3768	2	0.05	3737	0.1354
autocorr_ret_1 ^{^VIX}	float64	3770	0	0.0	1	-0.0777
lag_ret_2 ^{^VIX}	float64	3767	3	0.08	3736	0.1354
autocorr_ret_2 ^{^VIX}	float64	3770	0	0.0	1	-0.0222
lag_ret_3 ^{^VIX}	float64	3766	4	0.11	3735	0.1354
autocorr_ret_3 ^{^VIX}	float64	3770	0	0.0	1	-0.0337

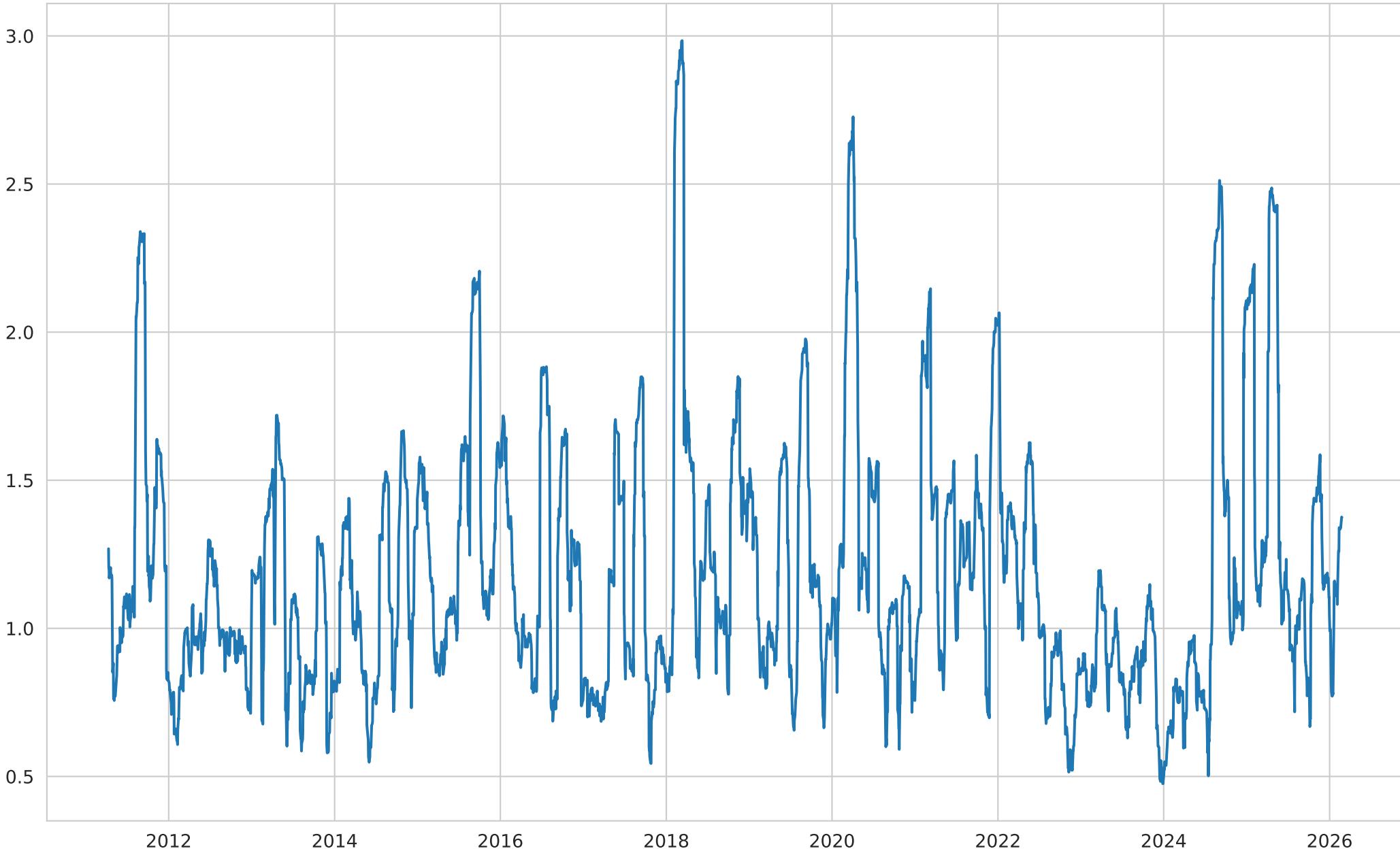
^VIX • Correlation



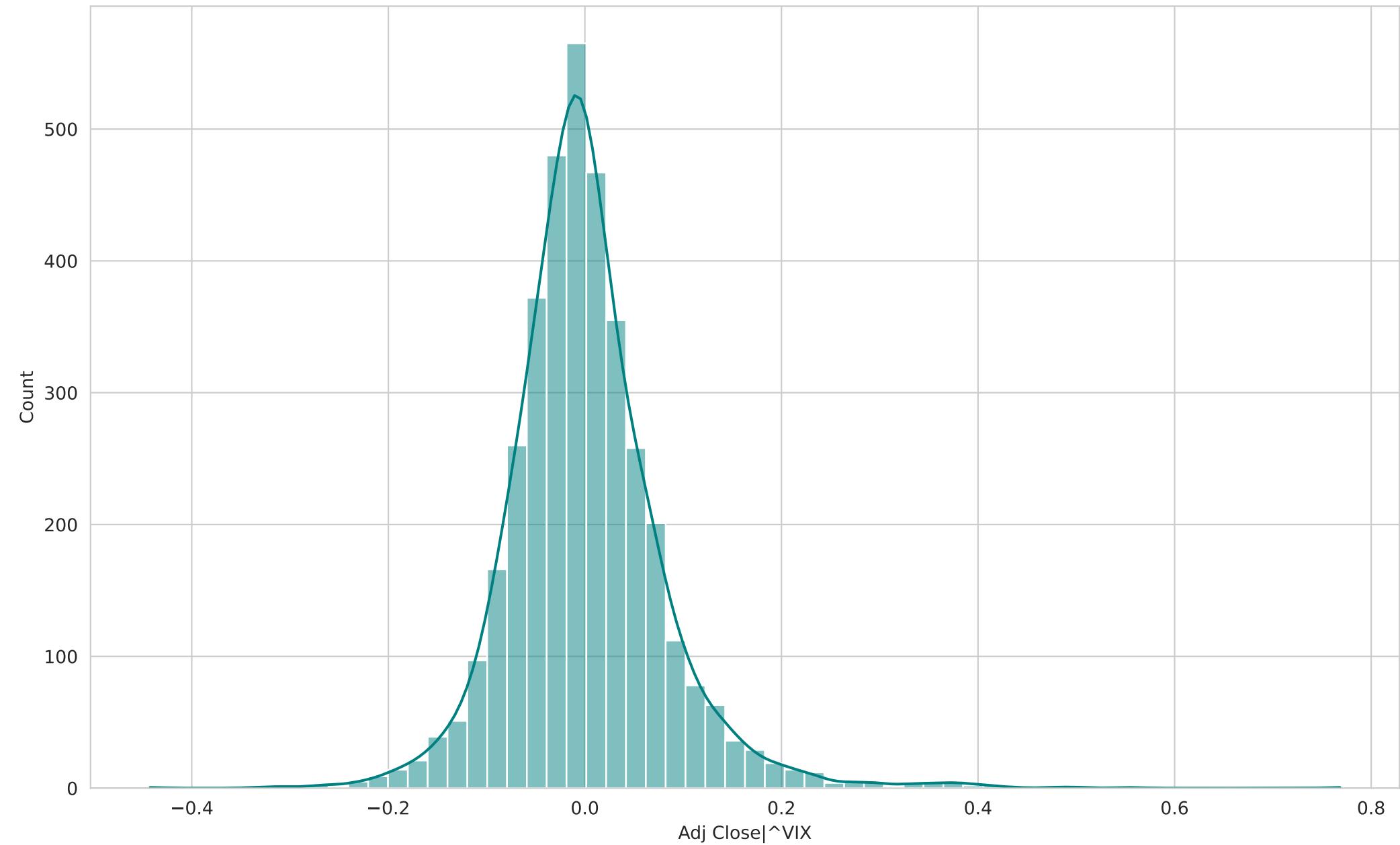
\wedge VIX • Price (Adj Close| \wedge VIX)



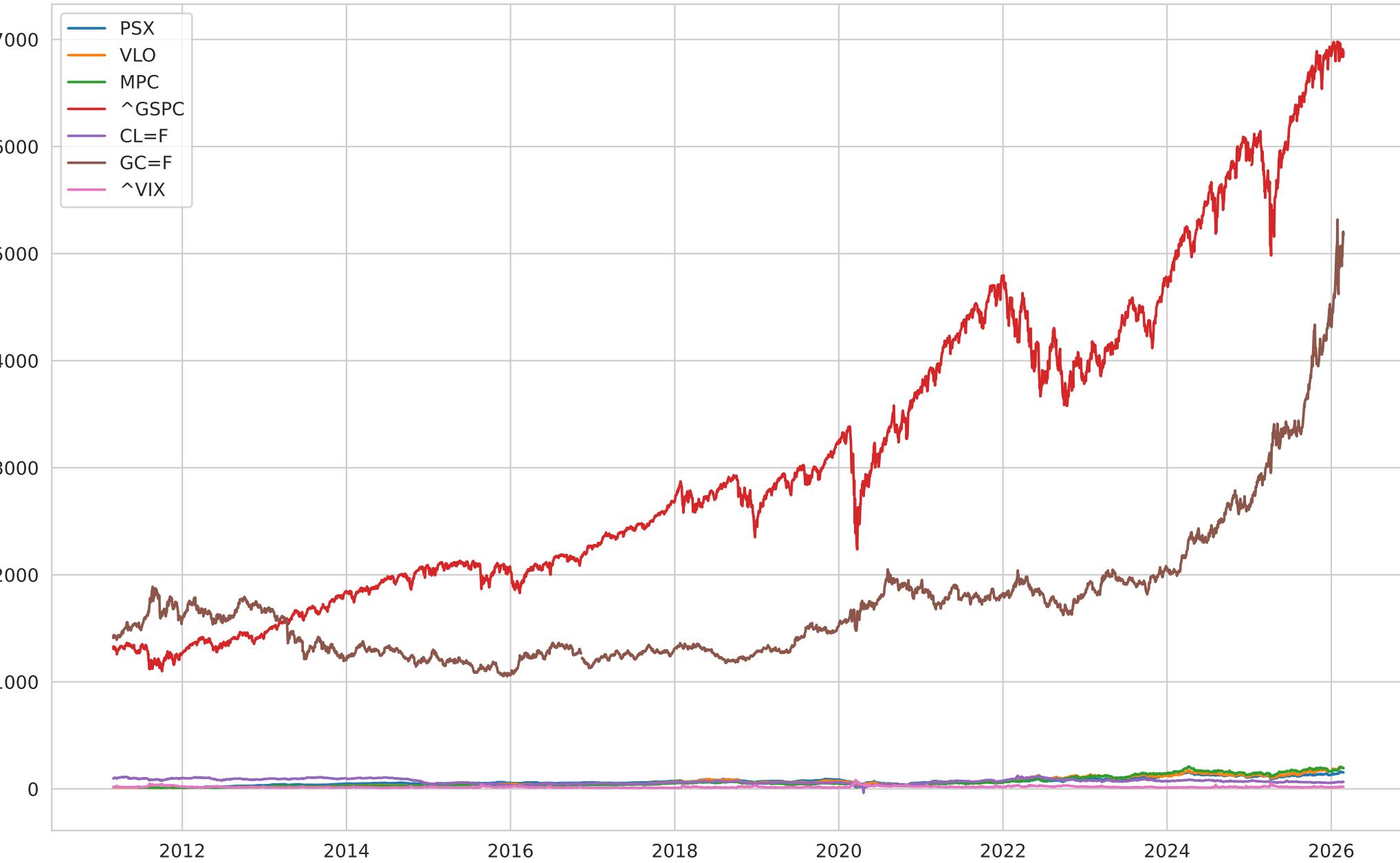
\wedge VIX • 30-Day Rolling Vol



\wedge VIX • Return Distribution



Cross-Ticker • Price Overlay



Cross-Ticker • Log Return Corr

