In [21]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [22]:
```
df=pd.read_csv(r"D:\Python DA\Stock Market\52WeekHigh.csv")
```

In [23]:
```
df
```

Out[23]:

	Symbol	Series	LTP	%chng	New 52W/H price	Prev.High	Prev. High Date
0	ABBOTINDIA	EQ	31300.00	-1.25	32200.00	31900.00	05-Jun-2025
1	ABINFRA	BE	139.00	-1.64	144.00	144.00	03-Jun-2025
2	ANURAS	EQ	1045.00	0.08	1056.00	1048.50	05-Jun-2025
3	AONETOTAL	EQ	11.70	1.30	11.90	11.69	20-May-2025
4	APLAPOLLO	EQ	1922.00	0.37	1924.80	1920.00	05-Jun-2025
...	...	...	...	...	...	...	...
90	UNIONGOLD	EQ	96.25	-0.82	100.00	99.55	17-Apr-2025
91	UTIBANKETF	EQ	58.01	1.54	58.19	57.52	23-Apr-2025
92	VILAS	SM	580.00	7.40	583.95	566.50	07-Nov-2024
93	WELCORP	EQ	965.30	0.31	990.00	973.00	05-Jun-2025
94	WELINV	BE	1171.95	-1.22	1240.00	1199.00	05-Jun-2025

95 rows × 7 columns

In [27]:
```
df.describe()
```

Out[27]:

	LTP	%chng	New 52W/H price	Prev.High
count	95.000000	95.000000	95.000000	95.000000
mean	1795.725263	1.644632	1842.542421	1814.220000
std	4390.402242	3.609400	4513.541145	4460.669138
min	6.410000	-10.570000	6.410000	6.110000
25%	102.875000	-0.620000	104.385000	101.730000
50%	327.980000	1.340000	340.800000	311.450000
75%	1080.075000	3.860000	1131.150000	1119.200000
max	31300.000000	14.490000	32200.000000	31900.000000

In [9]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 95 entries, 0 to 94
Data columns (total 7 columns):
 #   Column            Non-Null Count  Dtype  
---  ------            --------------  -----  
 0   Symbol            95 non-null     object 
 1   Series            95 non-null     object 
 2   LTP               95 non-null     float64
 3   %chng             95 non-null     float64
 4   New 52W/H price   95 non-null     float64
 5   Prev.High         95 non-null     float64
 6   Prev. High Date   95 non-null     object 
dtypes: float64(4), object(3)
memory usage: 5.3+ KB
```

In [34]:
```
df.columns = df.columns.str.strip()
df['Prev. High Date'] = pd.to_datetime(df['Prev. High Date'], errors='coerce')
df['Prev. High Month'] = df['Prev. High Date'].dt.month
df['Prev. High Year'] = df['Prev. High Date'].dt.year
```

In [36]:
```
df['Distance_from_High'] = df['LTP'] - df['New 52W/H price']
df['Pct_change_from_prev_high'] = ((df['New 52W/H price'] - df['Prev.High']) / df['Prev.High']) * 100
```

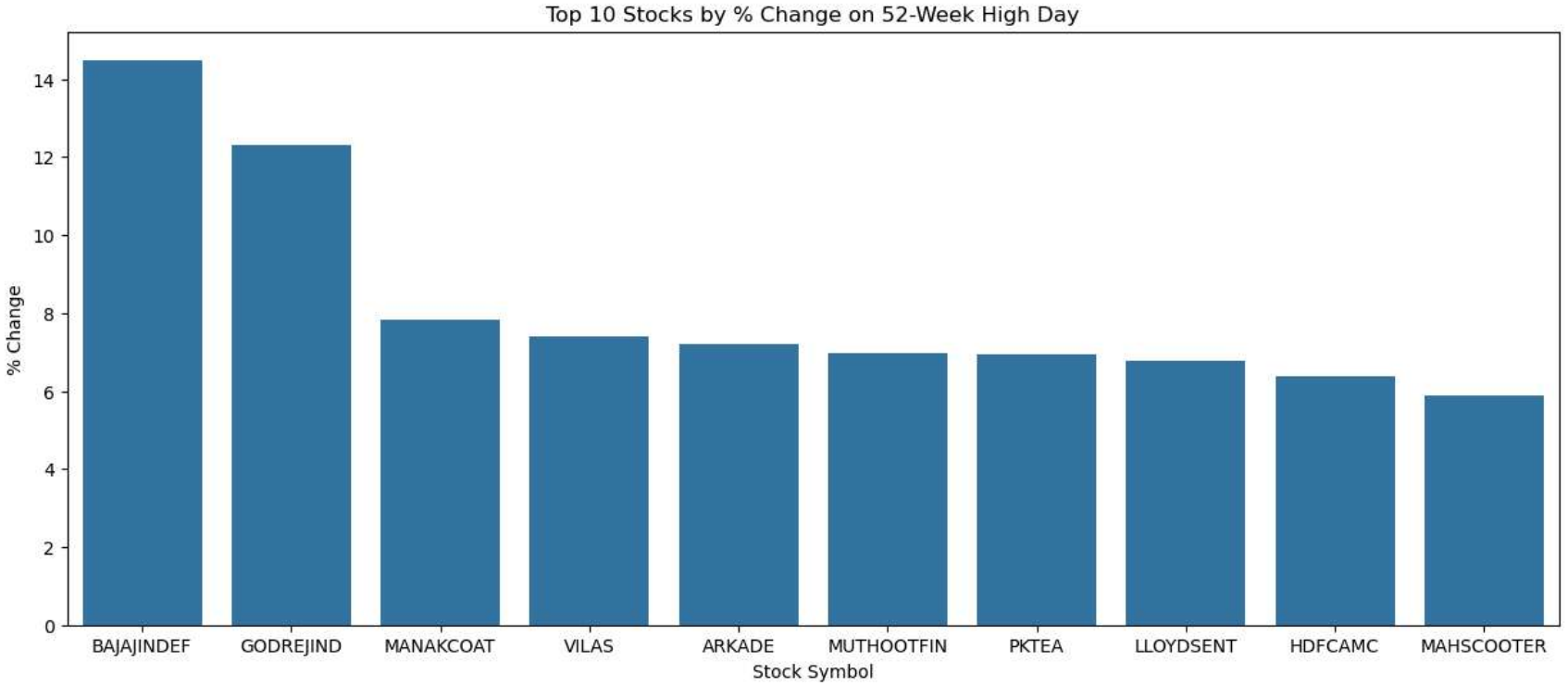
In [38]:
```
df
```

Out[38]:

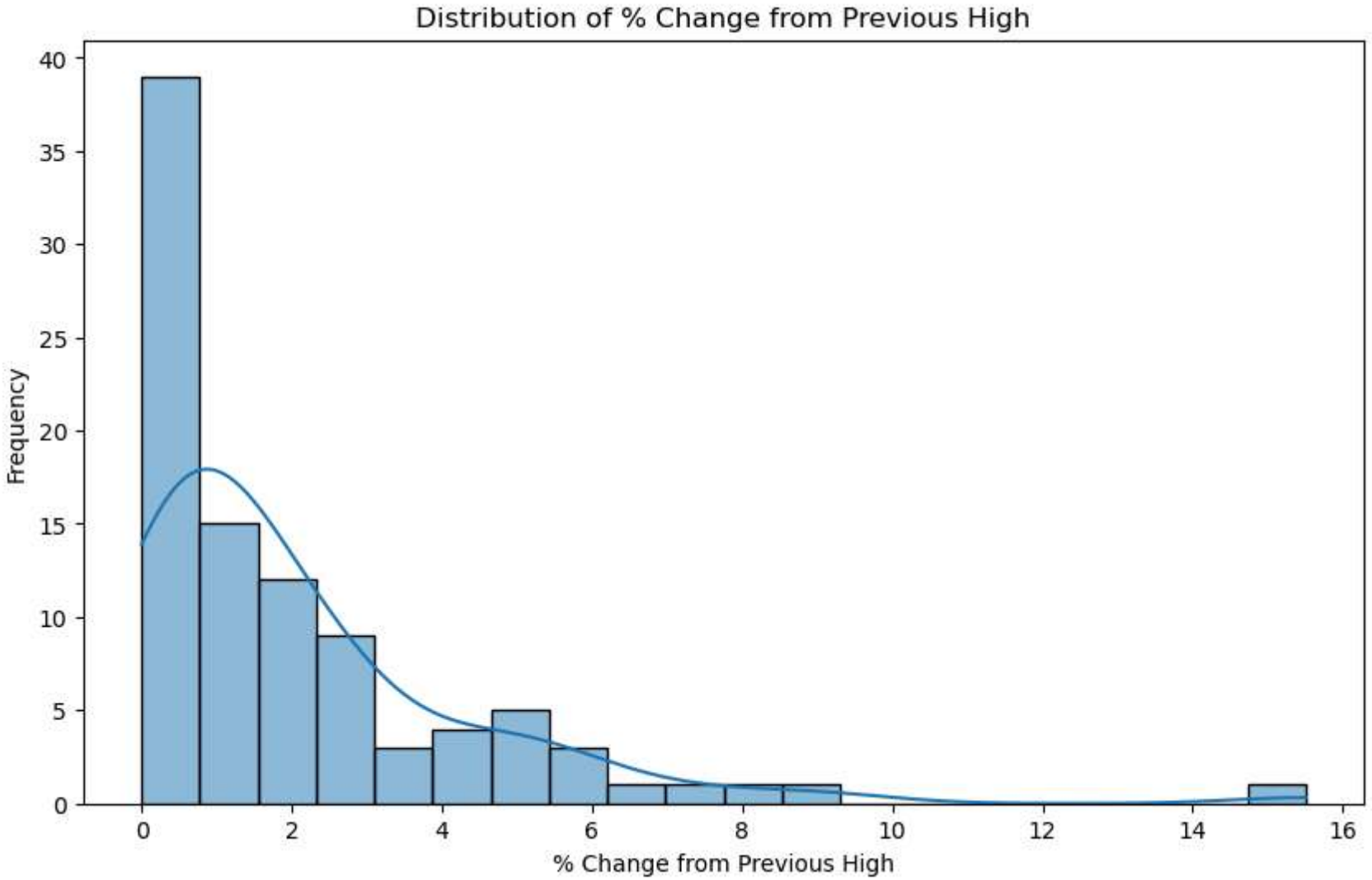
	Symbol	Series	LTP	%chng	New 52W/H price	Prev.High	Prev. High Date	Prev. High Month	Prev. High Year	Distance_from_High	Pct_change_from_pre
0	ABBOTINDIA	EQ	31300.00	-1.25	32200.00	31900.00	2025-06-05	6	2025	-900.00	0.9
1	ABINFRA	BE	139.00	-1.64	144.00	144.00	2025-06-03	6	2025	-5.00	0.0
2	ANURAS	EQ	1045.00	0.08	1056.00	1048.50	2025-06-05	6	2025	-11.00	0.7
3	AONETOTAL	EQ	11.70	1.30	11.90	11.69	2025-05-20	5	2025	-0.20	1.7
4	APLAPOLLO	EQ	1922.00	0.37	1924.80	1920.00	2025-06-05	6	2025	-2.80	0.2
...	...	...	...	...	...	...	...	...	...	...	
90	UNIONGOLD	EQ	96.25	-0.82	100.00	99.55	2025-04-17	4	2025	-3.75	0.4
91	UTIBANKETF	EQ	58.01	1.54	58.19	57.52	2025-04-23	4	2025	-0.18	1.1
92	VILAS	SM	580.00	7.40	583.95	566.50	2024-11-07	11	2024	-3.95	3.0
93	WELCORP	EQ	965.30	0.31	990.00	973.00	2025-06-05	6	2025	-24.70	1.7
94	WELINV	BE	1171.95	-1.22	1240.00	1199.00	2025-06-05	6	2025	-68.05	3.4

95 rows × 11 columns

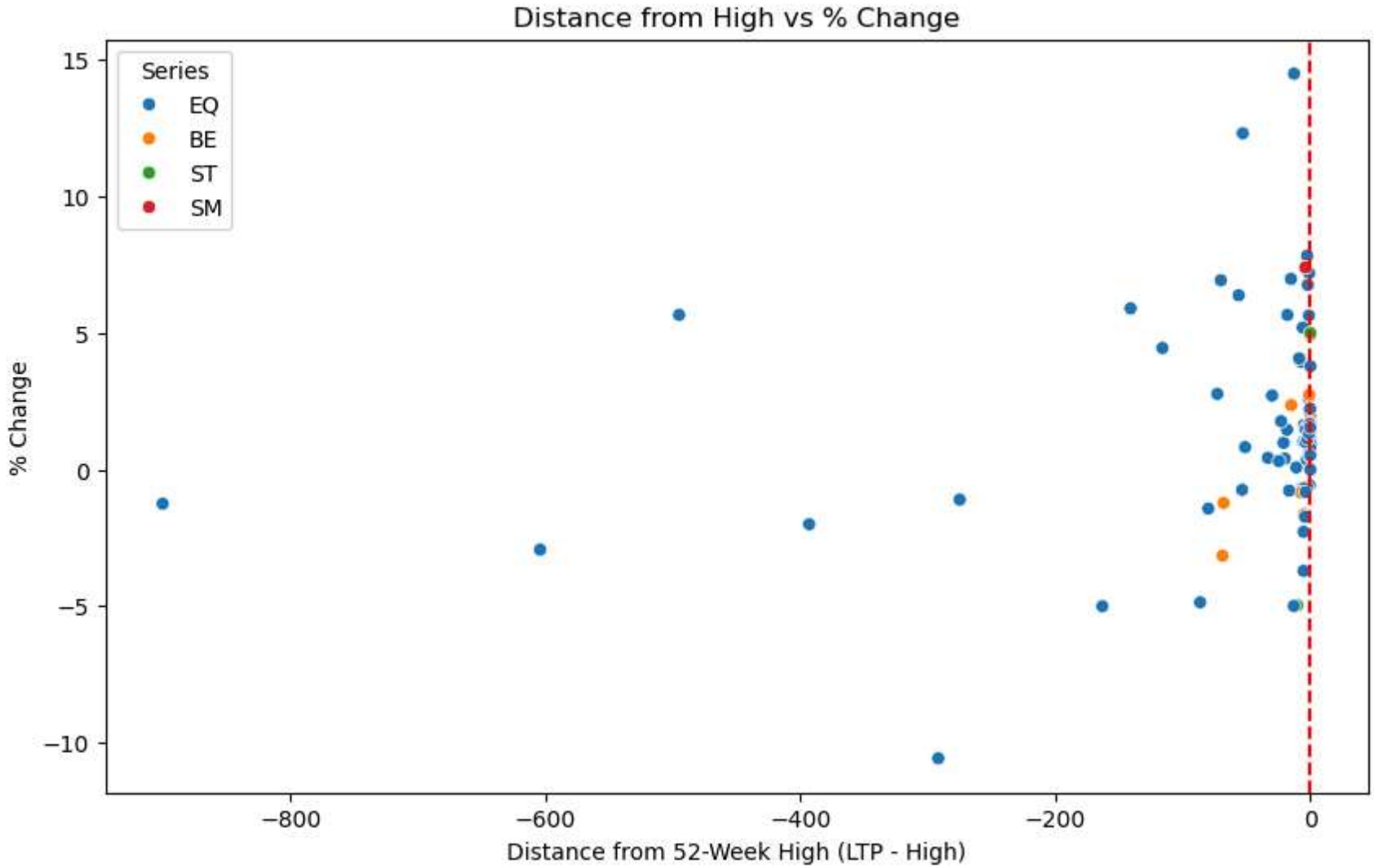
In [44]:
```
top10 = df.sort_values('%chng', ascending=False).head(10)
plt.figure(figsize=(15,6))
sns.barplot(data=top10, x='Symbol', y='%chng')
plt.title('Top 10 Stocks by % Change on 52-Week High Day')
plt.ylabel('% Change')
plt.xlabel('Stock Symbol')
plt.show()
```



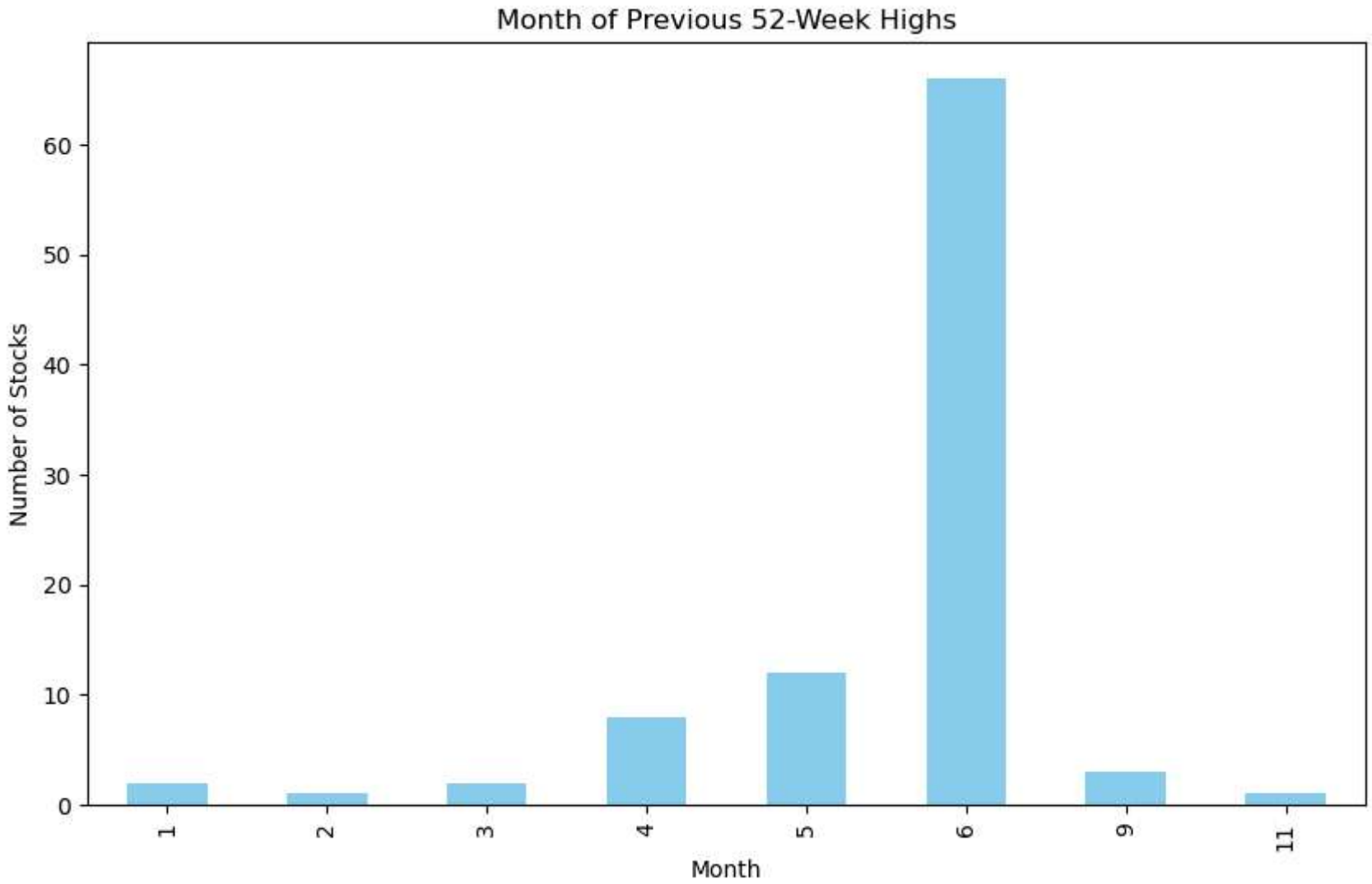
In [46]:
```
plt.figure(figsize=(10,6))
sns.histplot(df['Pct_change_from_prev_high'], bins=20, kde=True)
plt.title('Distribution of % Change from Previous High')
plt.xlabel('% Change from Previous High')
plt.ylabel('Frequency')
plt.show()
```



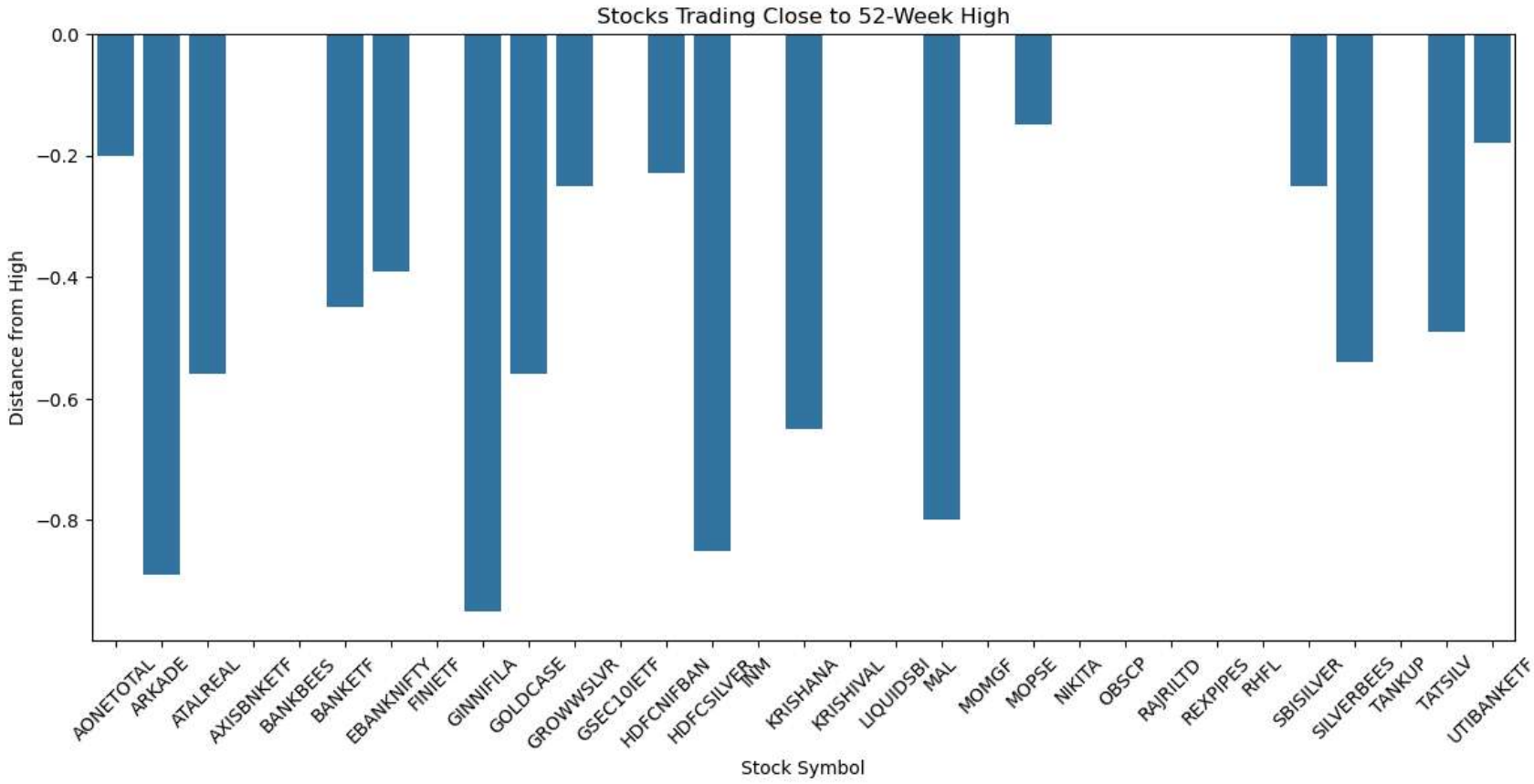
In [48]:
```
plt.figure(figsize=(10,6))
sns.scatterplot(data=df, x='Distance_from_High', y='%chng', hue='Series')
plt.axvline(0, color='red', linestyle='--')
plt.title('Distance from High vs % Change')
plt.xlabel('Distance from 52-Week High (LTP - High)')
plt.ylabel('% Change')
plt.show()
```



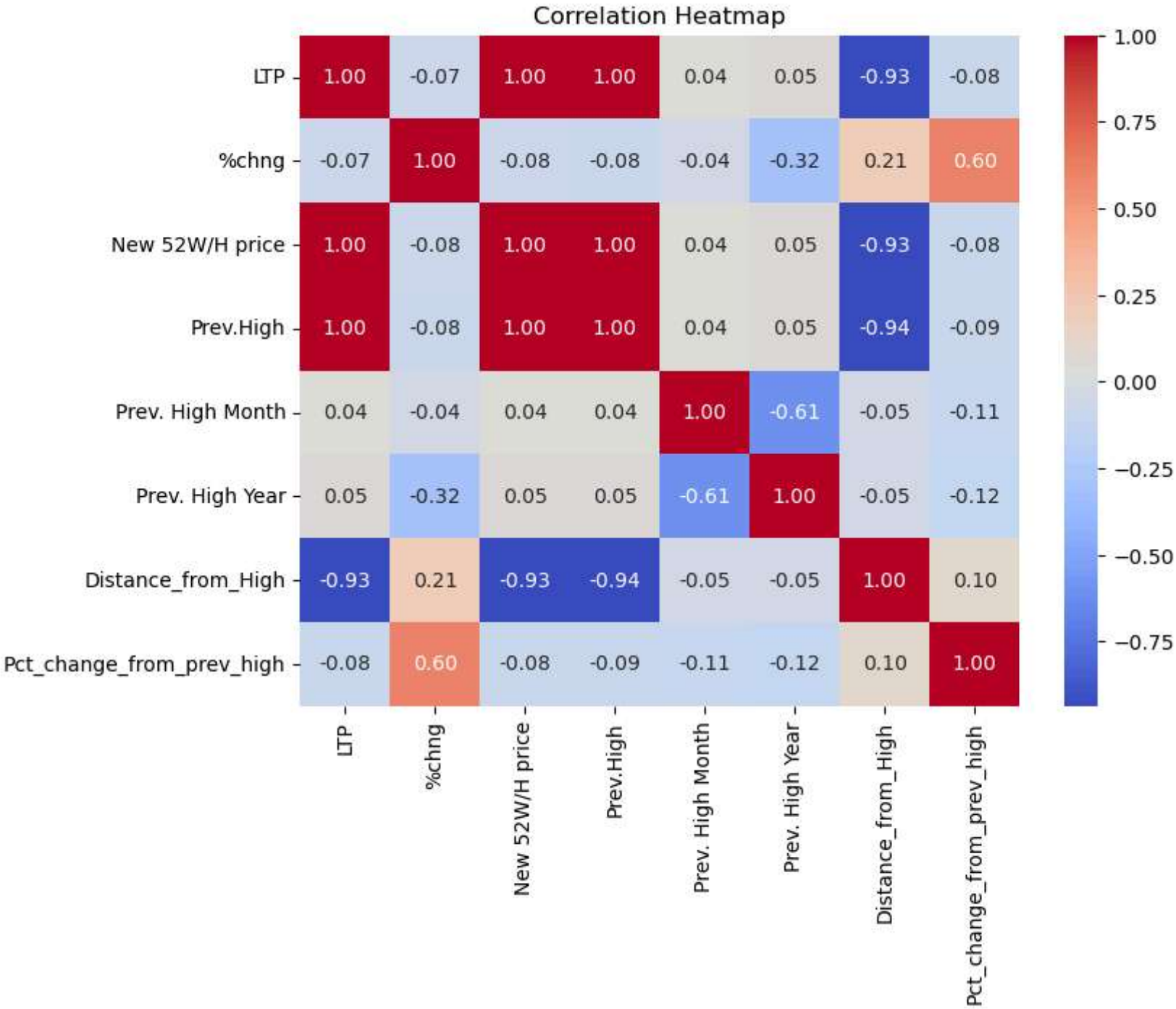
In [50]:
```
plt.figure(figsize=(10,6))
df['Prev. High Month'].value_counts().sort_index().plot(kind='bar', color='skyblue')
plt.title('Month of Previous 52-Week Highs')
plt.xlabel('Month')
plt.ylabel('Number of Stocks')
plt.show()
```



In [68]:
```
close_to_high = df[np.abs(df['Distance_from_High']) < 1]
plt.figure(figsize=(14,6))
sns.barplot(data=close_to_high, x='Symbol', y='Distance_from_High')
plt.xticks(rotation=45)
plt.title('Stocks Trading Close to 52-Week High')
plt.ylabel('Distance from High')
plt.xlabel('Stock Symbol')
plt.show()
```

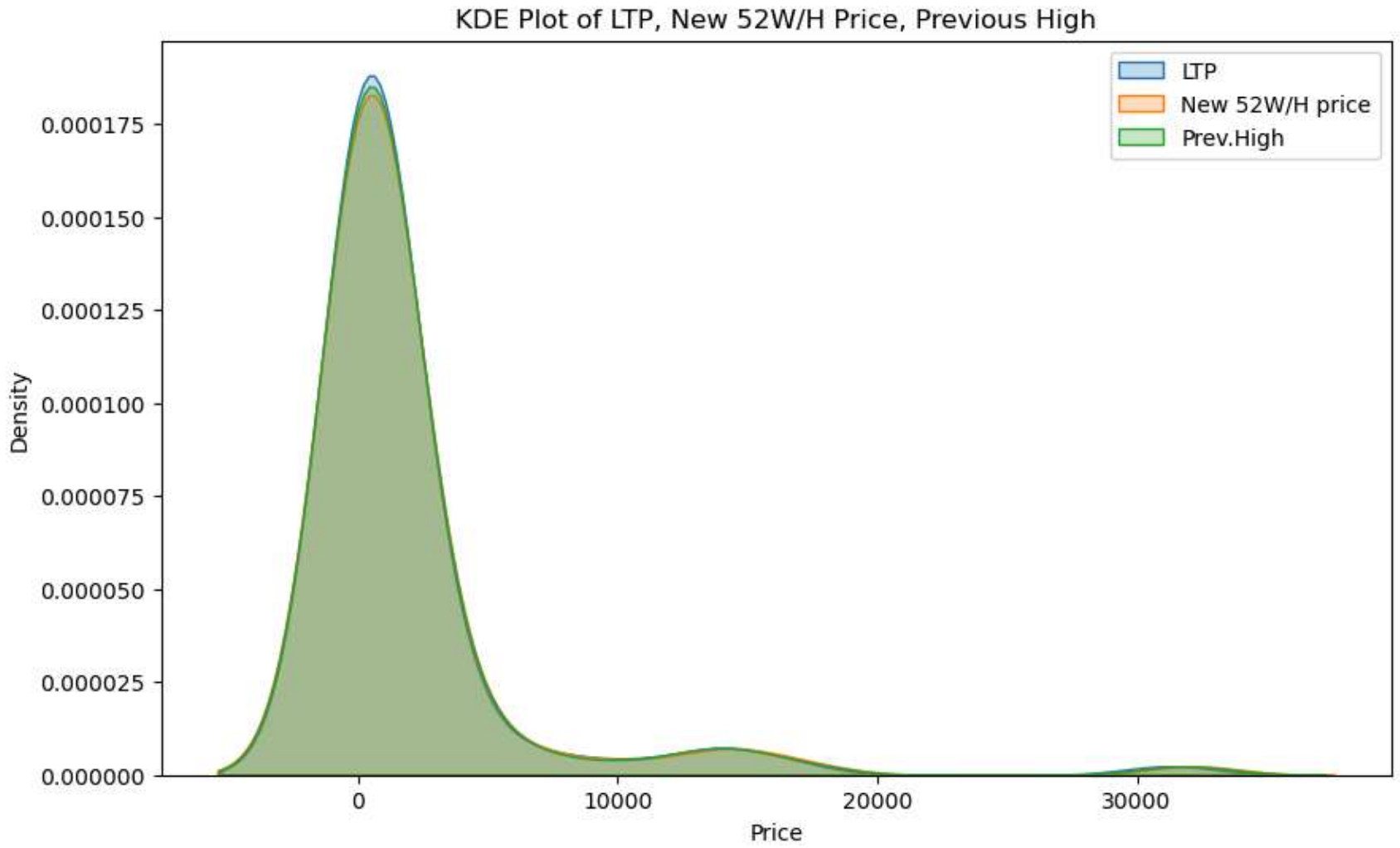


In [77]:
```
plt.figure(figsize=(8,6))
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Heatmap')
plt.show()
```

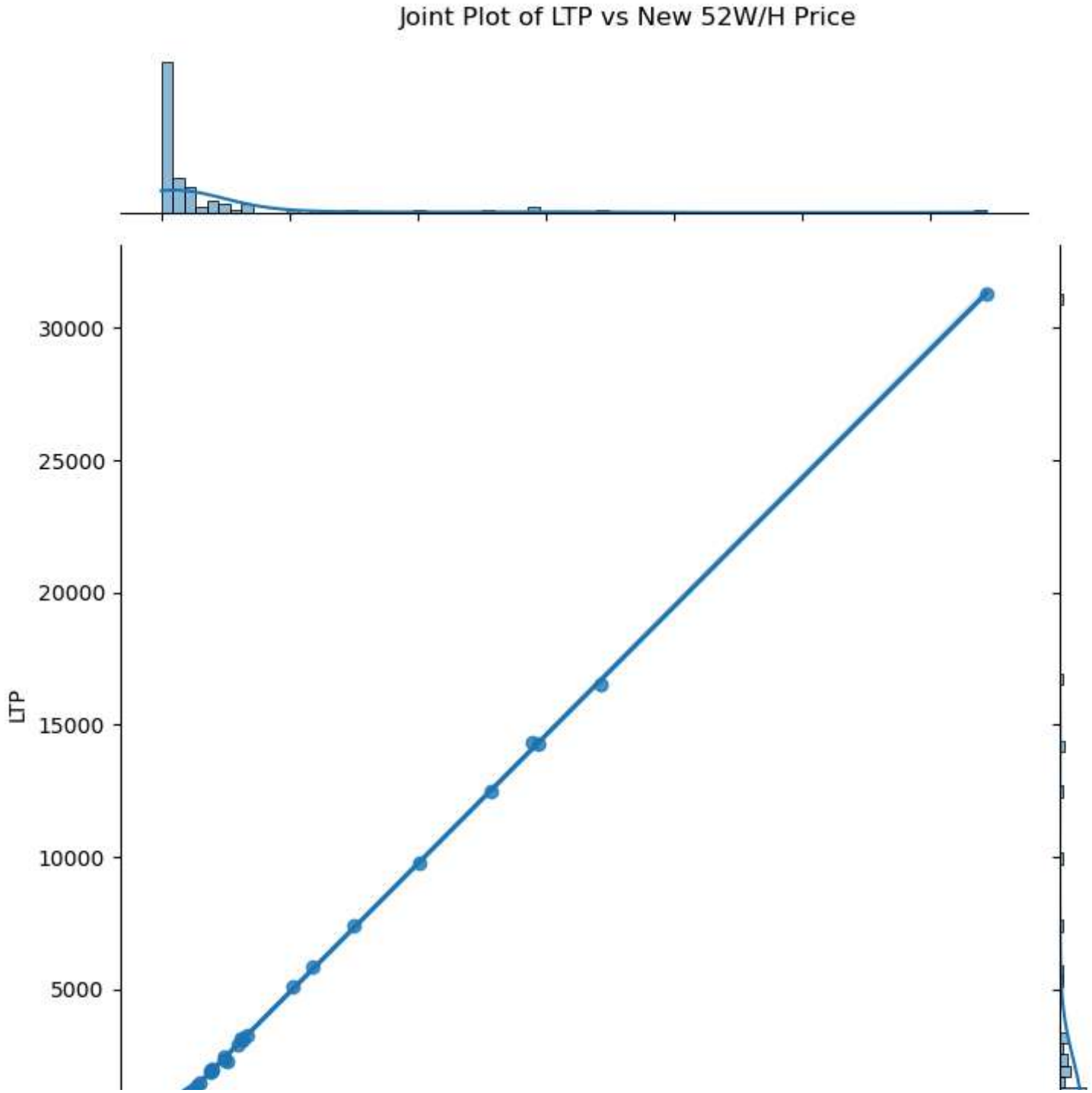

Correlation Heatmap
LTP
%chng
New 52W/H price
Prev.High
Prev. High Month
Prev. High Year
Distance_from_High
Pct_change_from_prev_high
1.00 -0.07 1.00 1.00 0.04 0.05 -0.93 -0.08
-0.07 1.00 -0.08 -0.08 -0.04 -0.32 0.21 0.60
1.00 -0.08 1.00 1.00 0.04 0.05 -0.93 -0.08
1.00 -0.08 1.00 1.00 0.04 0.05 -0.94 -0.09
0.04 -0.04 0.04 0.04 1.00 -0.61 -0.05 -0.11
0.05 -0.32 0.05 0.05 -0.61 1.00 -0.05 -0.12
-0.93 0.21 -0.93 -0.94 -0.05 -0.05 1.00 0.10
-0.08 0.60 -0.08 -0.09 -0.11 -0.12 0.10 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

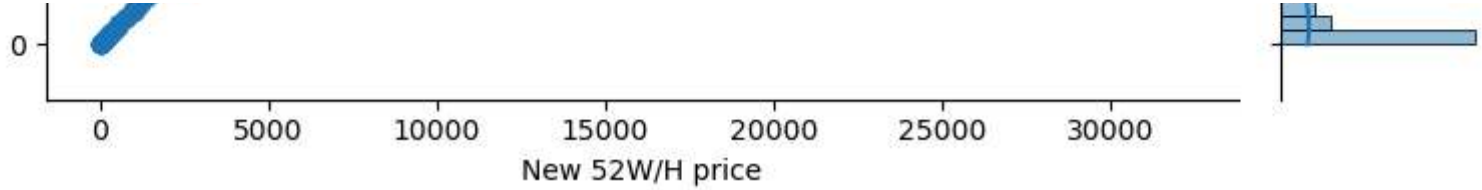
In [79]:
```
plt.figure(figsize=(10,6))
sns.kdeplot(df['LTP'], label='LTP', shade=True)
sns.kdeplot(df['New 52W/H price'], label='New 52W/H price', shade=True)
sns.kdeplot(df['Prev.High'], label='Prev.High', shade=True)
plt.title('KDE Plot of LTP, New 52W/H Price, Previous High')
plt.xlabel('Price')
plt.ylabel('Density')
plt.legend()
plt.show()
```

```
C:\Users\HP\AppData\Local\Temp\ipykernel_18476\2944212408.py:2: FutureWarning:

`shade` is now deprecated in favor of `fill`; setting `fill=True`.
This will become an error in seaborn v0.14.0; please update your code.

  sns.kdeplot(df['LTP'], label='LTP', shade=True)
C:\Users\HP\AppData\Local\Temp\ipykernel_18476\2944212408.py:3: FutureWarning:

`shade` is now deprecated in favor of `fill`; setting `fill=True`.
This will become an error in seaborn v0.14.0; please update your code.

  sns.kdeplot(df['New 52W/H price'], label='New 52W/H price', shade=True)
C:\Users\HP\AppData\Local\Temp\ipykernel_18476\2944212408.py:4: FutureWarning:

`shade` is now deprecated in favor of `fill`; setting `fill=True`.
This will become an error in seaborn v0.14.0; please update your code.

  sns.kdeplot(df['Prev.High'], label='Prev.High', shade=True)
```

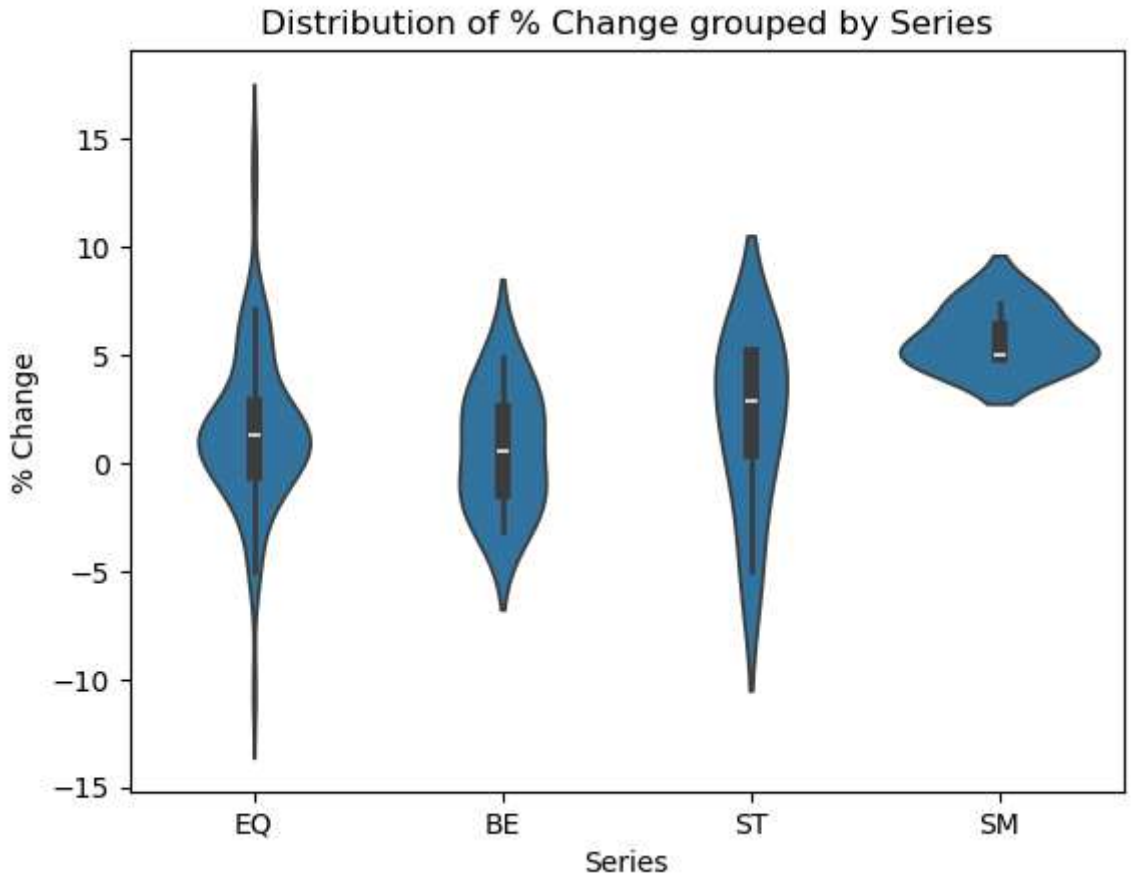


In [81]:
```
sns.jointplot(data=df, x='New 52W/H price', y='LTP', kind='reg', height=8)
plt.suptitle('Joint Plot of LTP vs New 52W/H Price', y=1.02)
plt.show()
```

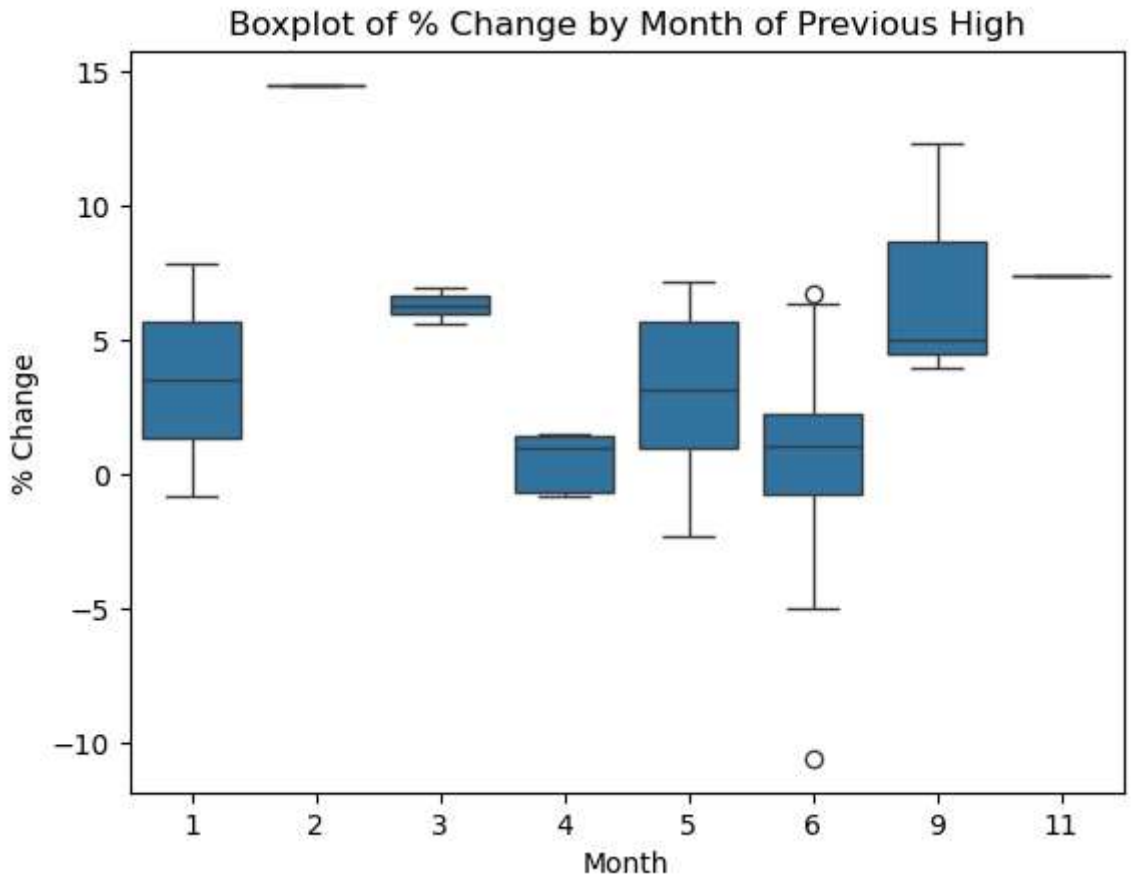

Joint Plot of LTP vs New 52W/H Price
LTP
30000
25000
20000
15000
10000
5000



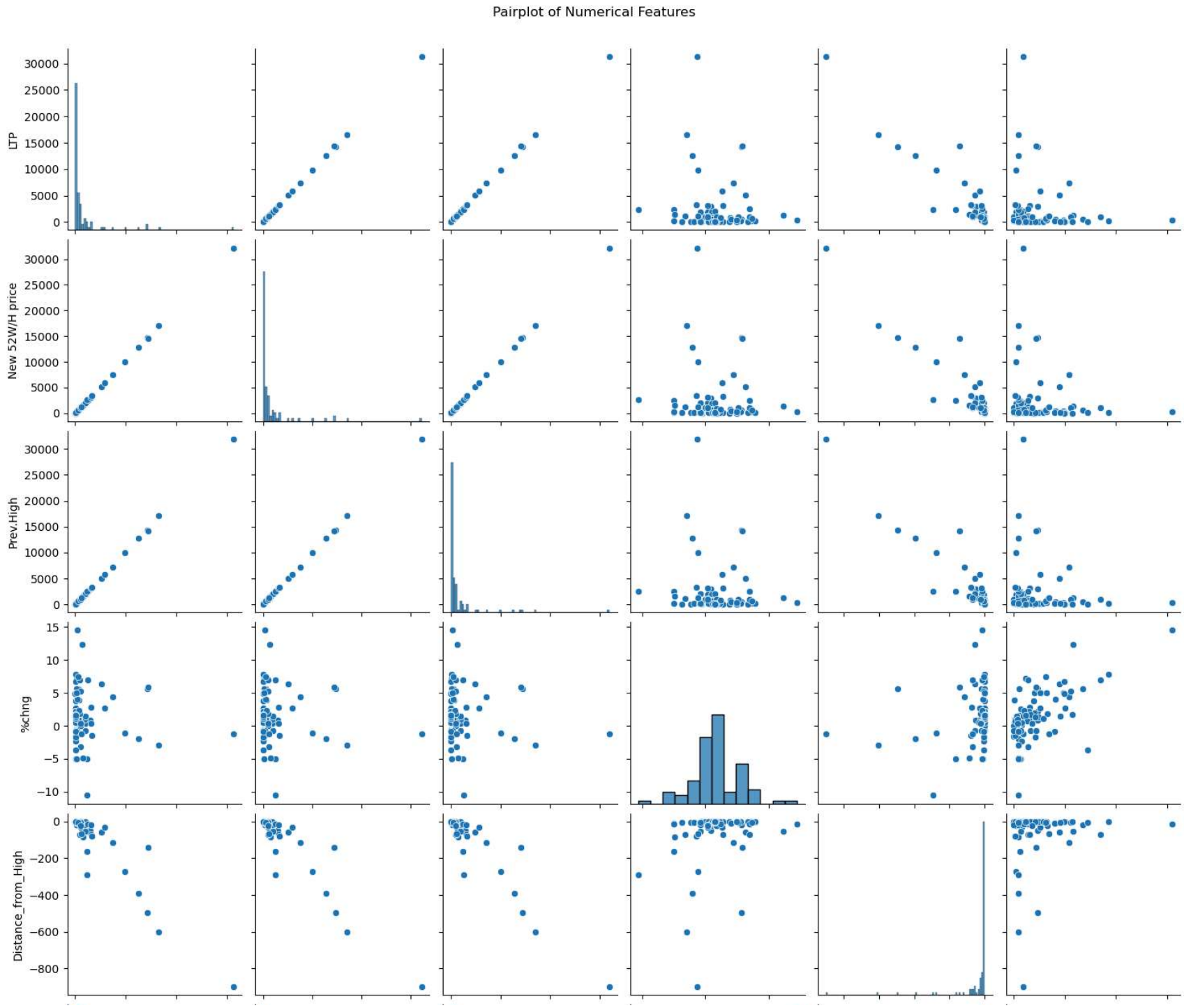
In [83]:
```
sns.violinplot(data=df, x='Series', y='%chng')
plt.title('Distribution of % Change grouped by Series')
plt.xlabel('Series')
plt.ylabel('% Change')
plt.show()
```

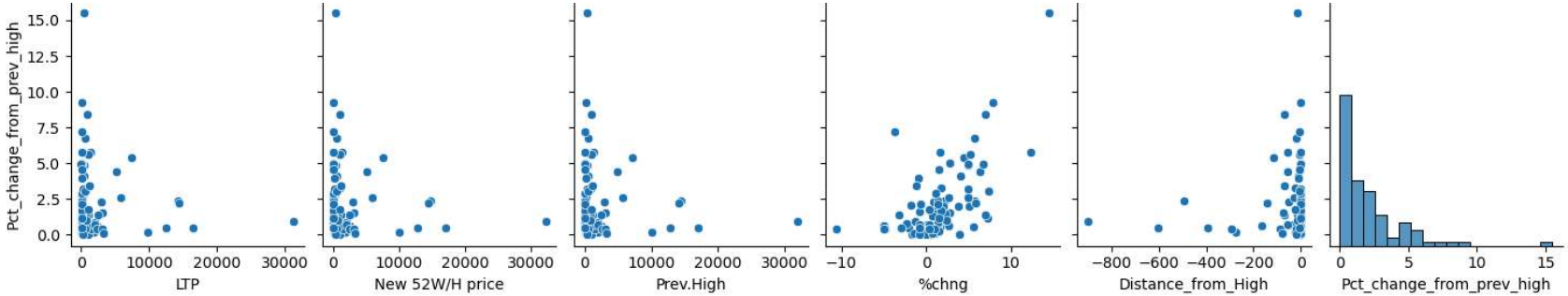


In [93]:
```
sns.boxplot(data=df, x='Prev. High Month', y='%chng')
plt.title('Boxplot of % Change by Month of Previous High')
plt.xlabel('Month')
plt.ylabel('% Change')
plt.show()
```



In [94]:
```
sns.pairplot(df[['LTP', 'New 52W/H price', 'Prev.High', '%chng', 'Distance_from_High', 'Pct_change_from_prev_high']])
plt.suptitle('Pairplot of Numerical Features', y=1.02)
plt.show()
```


Pairplot of Numerical Features
LTP
New 52W/H price
Prev.High
%chng
Distance_from_High



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: