Data (raw features)

In [2]:
```
combined_data.head()
```

Out[2]:

	NVDA_Close_NVDA	NVDA_High_NVDA	NVDA_Low_NVDA	NVDA_Open_NVDA	NVDA_Volume_NVDA	^GSPC_Close_^GSPC	^G
Date							
2022-02-14	24.226711	24.833702	23.715560	23.894264	440424000	4401.669922	44
2022-02-15	26.451010	26.500926	24.742850	24.907575	699869000	4471.069824	44
2022-02-16	26.466984	26.537867	25.509576	26.215403	732676000	4475.009766	44
2022-02-17	24.466309	25.742186	24.124877	25.587443	810595000	4380.259766	44
2022-02-18	23.602745	24.944513	23.061645	24.627041	761255000	4348.870117	43

Plots

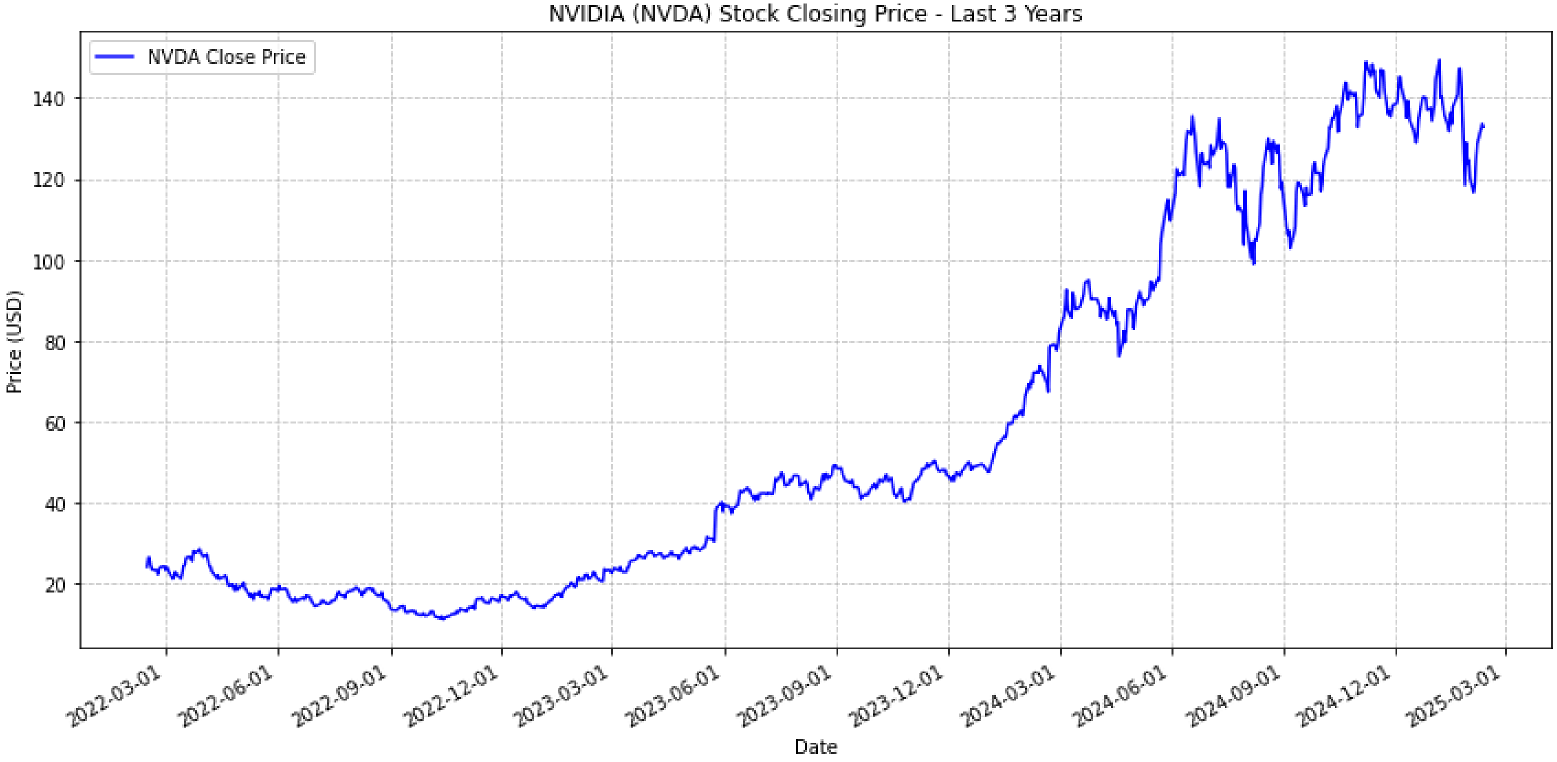
Price

Varies quite a bit over the time range

In [4]:

```
fig
```

Out[4]:

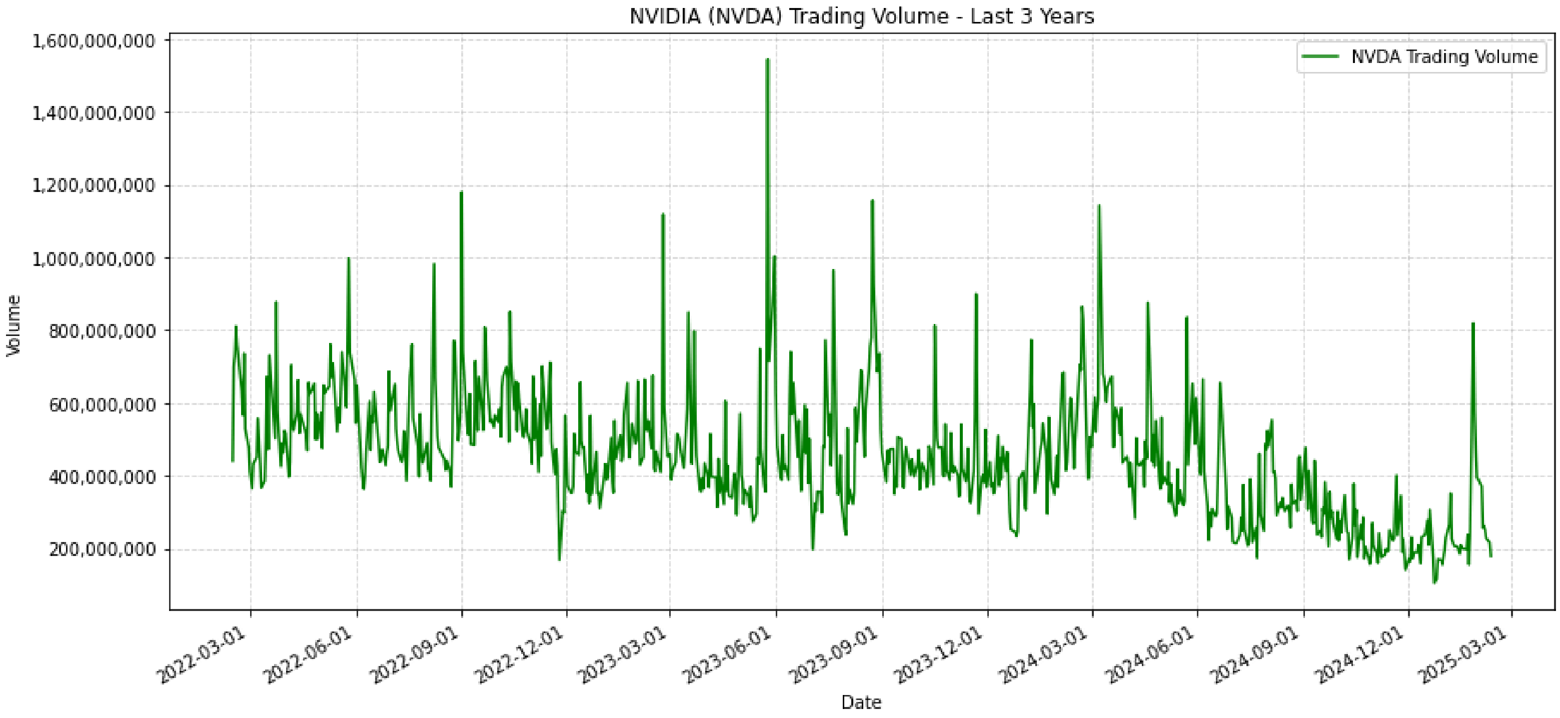


NVIDIA (NVDA) Stock Closing Price - Last 3 Years
NVDA Close Price
Price (USD)
140
120
100
80
60
40
20
2022-03-01
2022-06-01
2022-09-01
2022-12-01
2023-03-01
2023-06-01
2023-09-01
2023-12-01
2024-03-01
2024-06-01
2024-09-01
2024-12-01
2025-03-01
Date

Volume, in shares

In [6]:
```
fig
```

Out[6]:

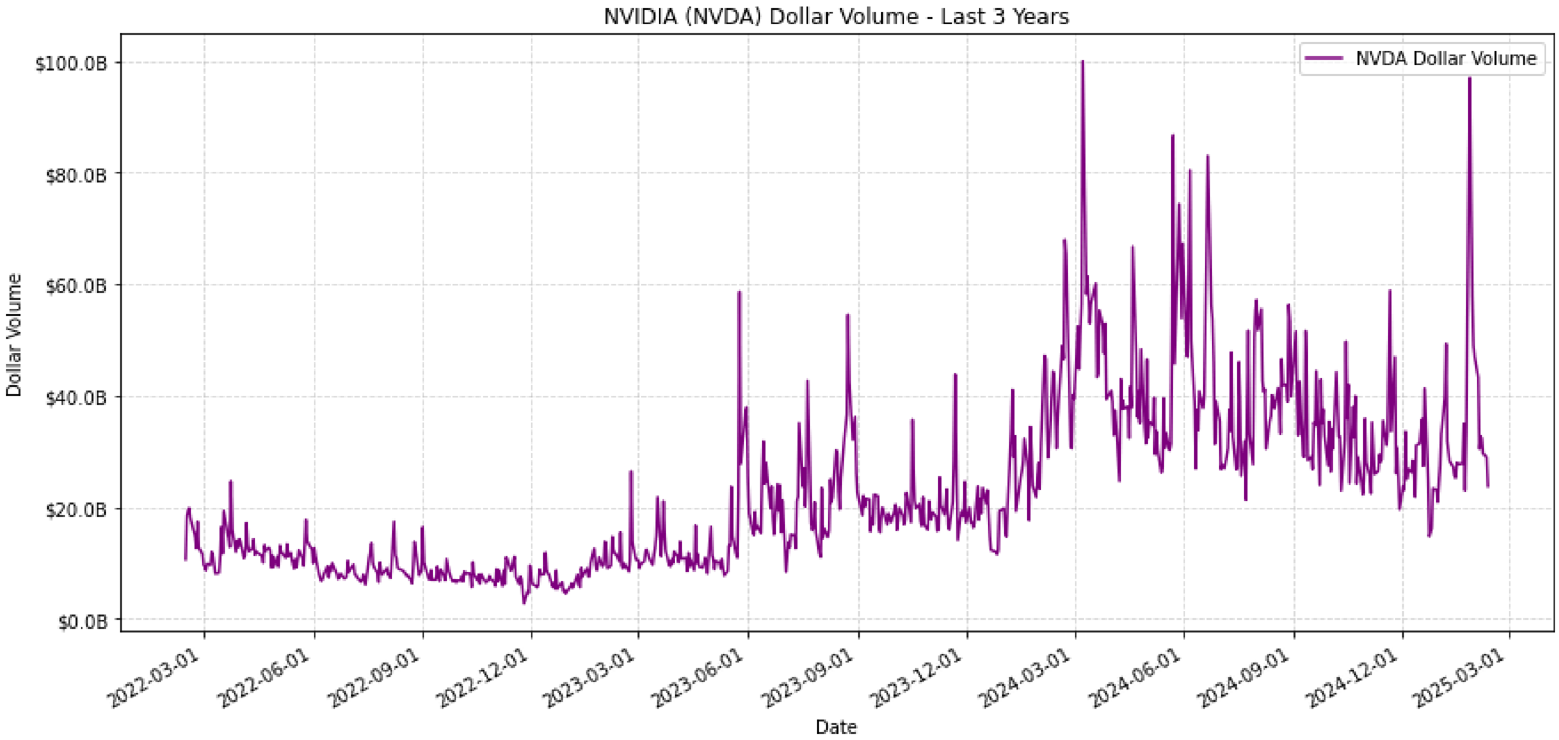


NVIDIA (NVDA) Trading Volume - Last 3 Years
NVDA Trading Volume
1,600,000,000
1,400,000,000
1,200,000,000
1,000,000,000
800,000,000
600,000,000
400,000,000
200,000,000
Volume
2022-03-01
2022-06-01
2022-09-01
2022-12-01
2023-03-01
2023-06-01
2023-09-01
2023-12-01
2024-03-01
2024-06-01
2024-09-01
2024-12-01
2025-03-01
Date

Volume (in dollars)

```
In [8]: fig
```

Out[8]:

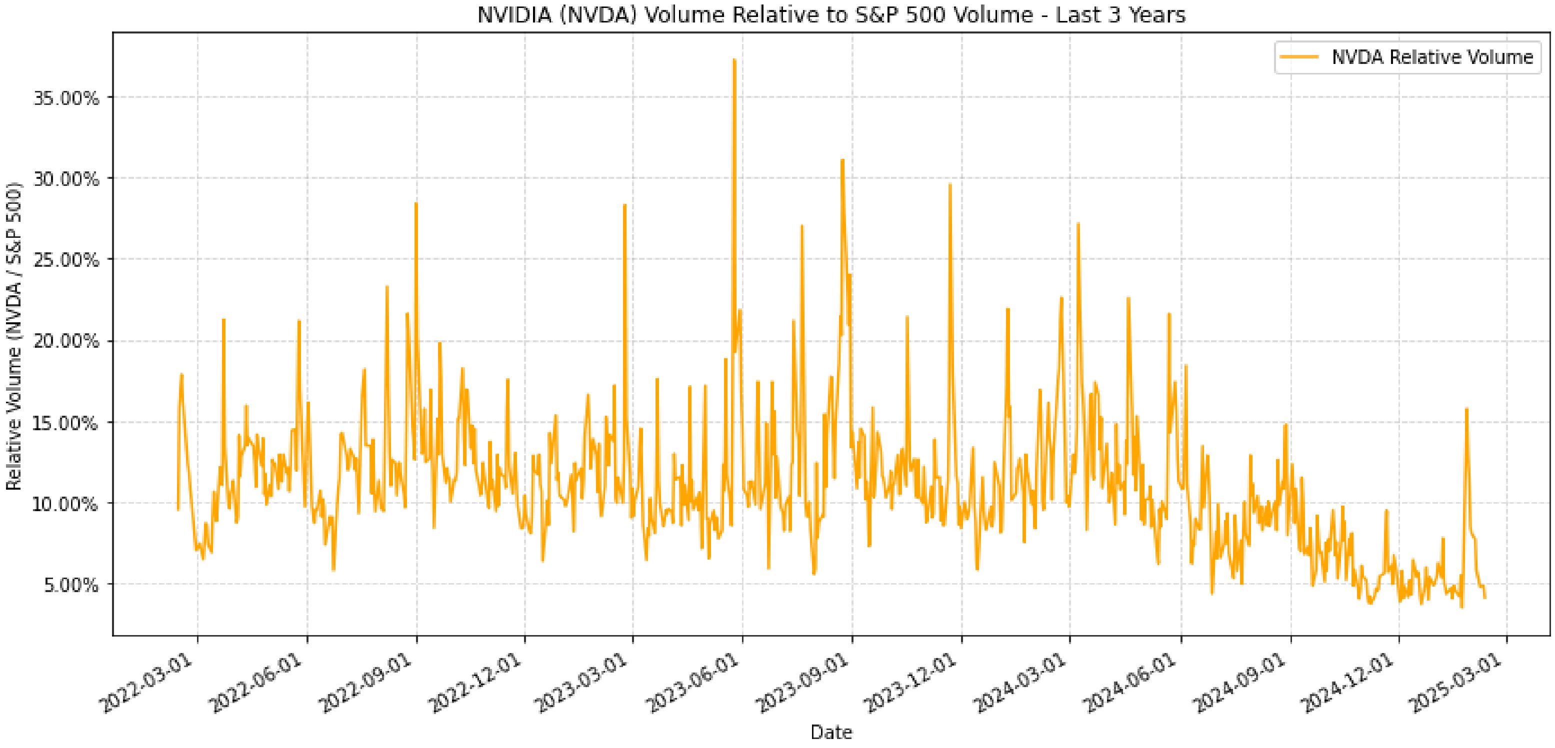


NVIDIA (NVDA) Dollar Volume - Last 3 Years
NVDA Dollar Volume
Dollar Volume
$100.0B
$80.0B
$60.0B
$40.0B
$20.0B
$0.0B
2022-03-01
2022-06-01
2022-09-01
2022-12-01
2023-03-01
2023-06-01
2023-09-01
2023-12-01
2024-03-01
2024-06-01
2024-09-01
2024-12-01
2025-03-01
Date

Volume (relative to total market volume)

In [10]:
```
fig
```

Out[10]:



Let's assume that the training data and test data are from the first/second half.

Are the distributions of training/test data similar ?

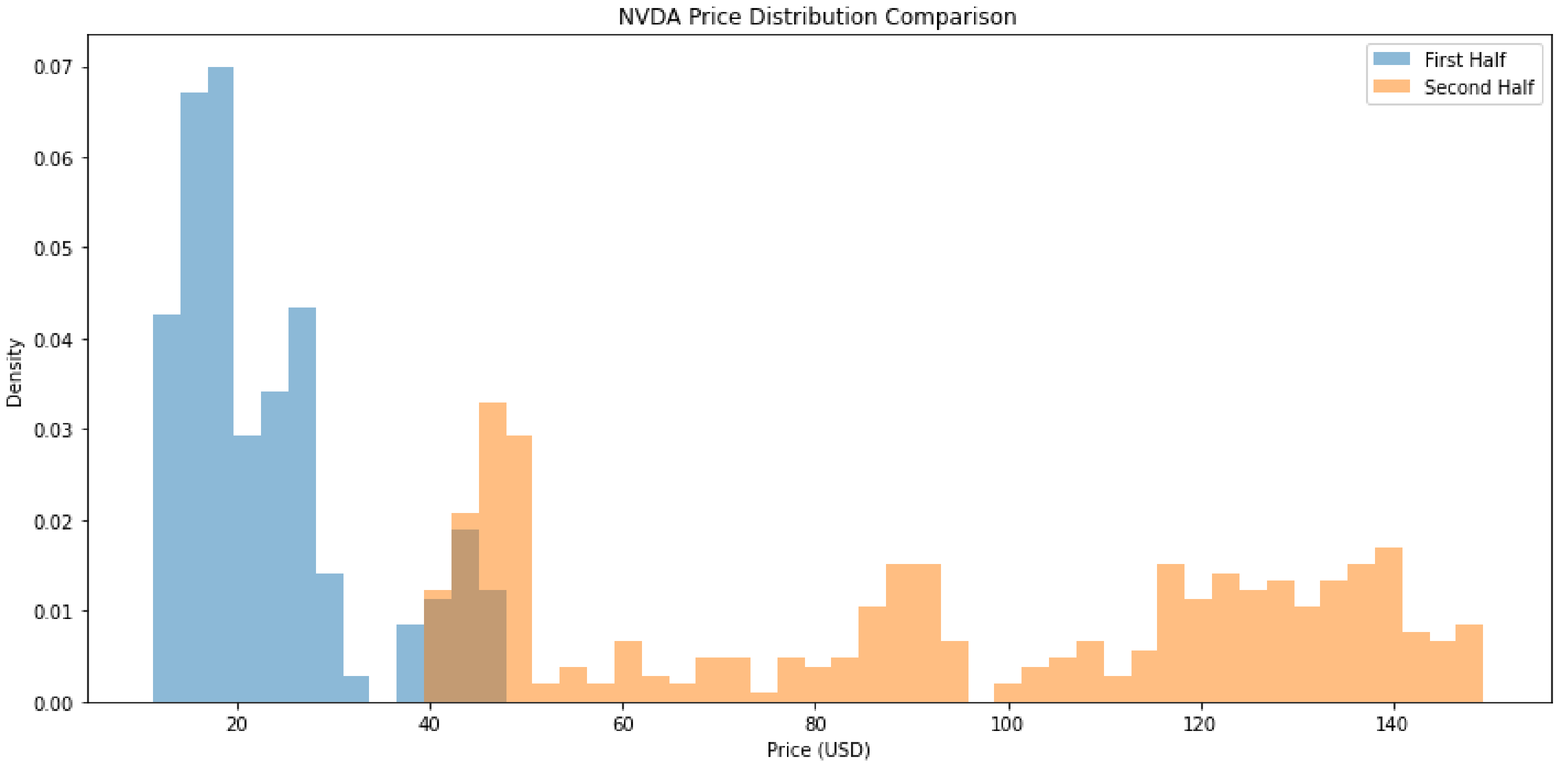
Note: we are assuming the second half if closer to the present, and therefore a more accurate proxy for unseen data that will be encountered in the near future.

Price distributions

In [12]:

```
fig
```

Out[12]:

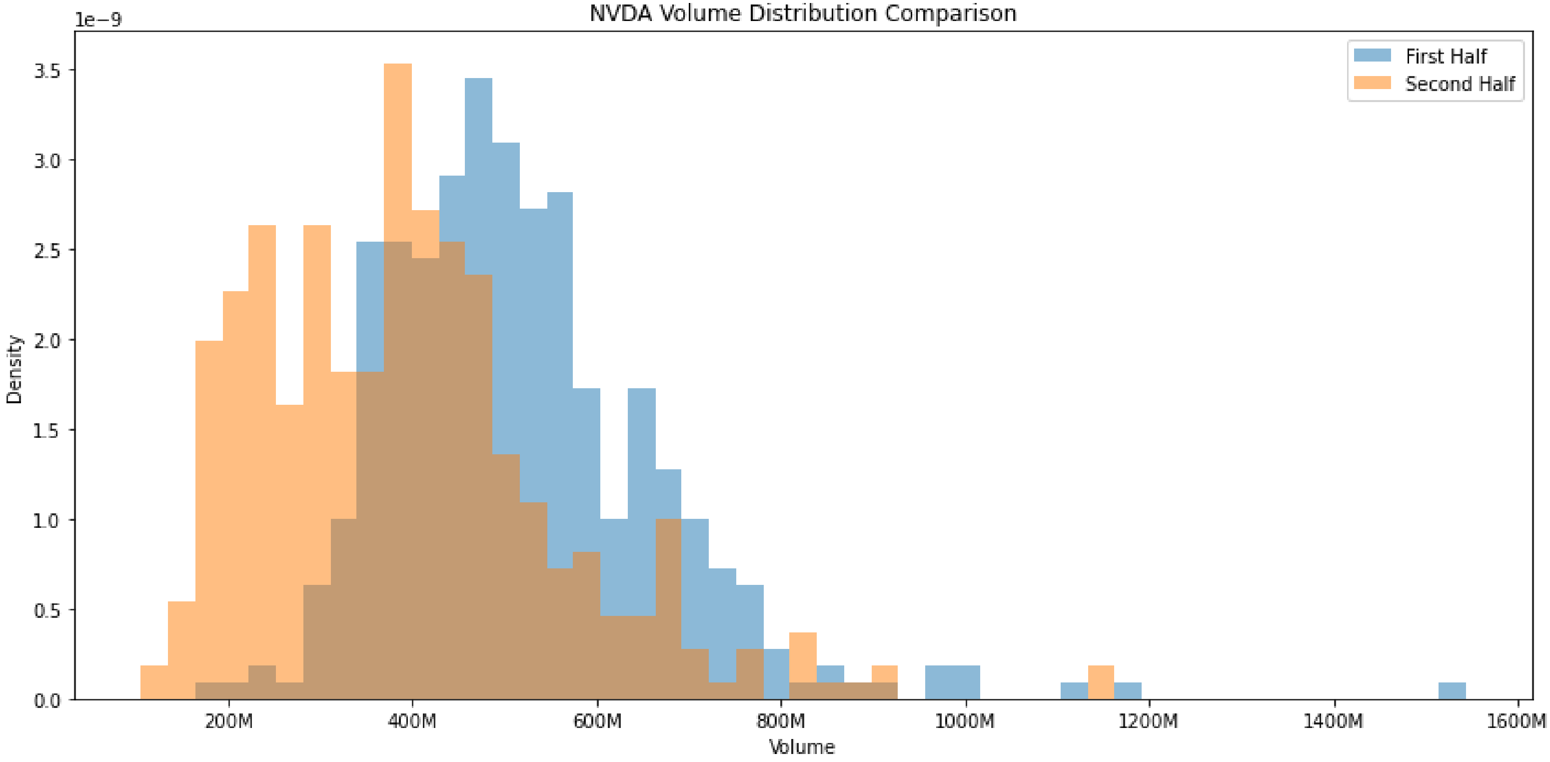


Volume (shares) distributions

In [14]:

```
fig
```

Out[14]:

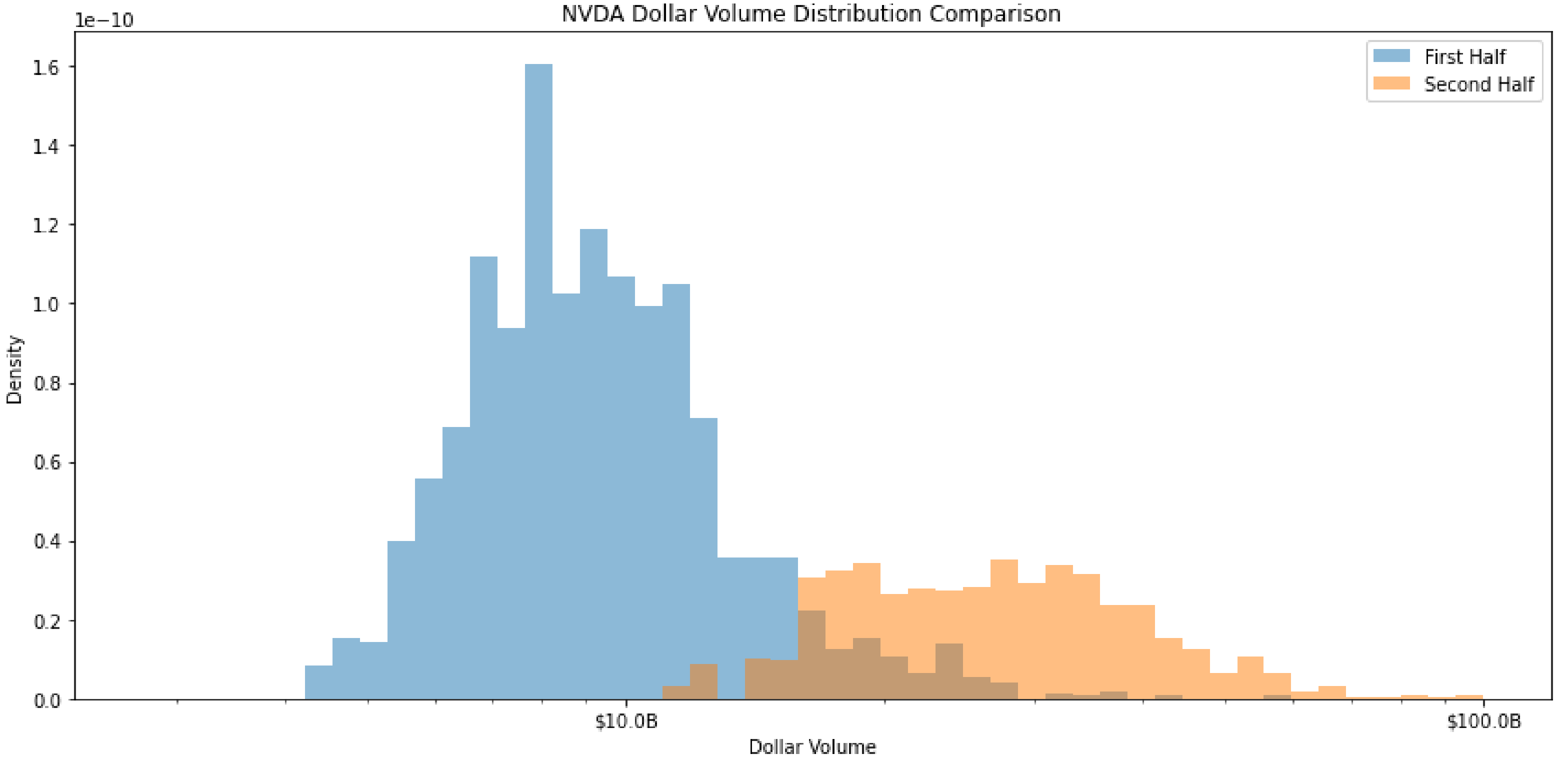


Volume (dollars) distributions

In [16]:

```
fig
```

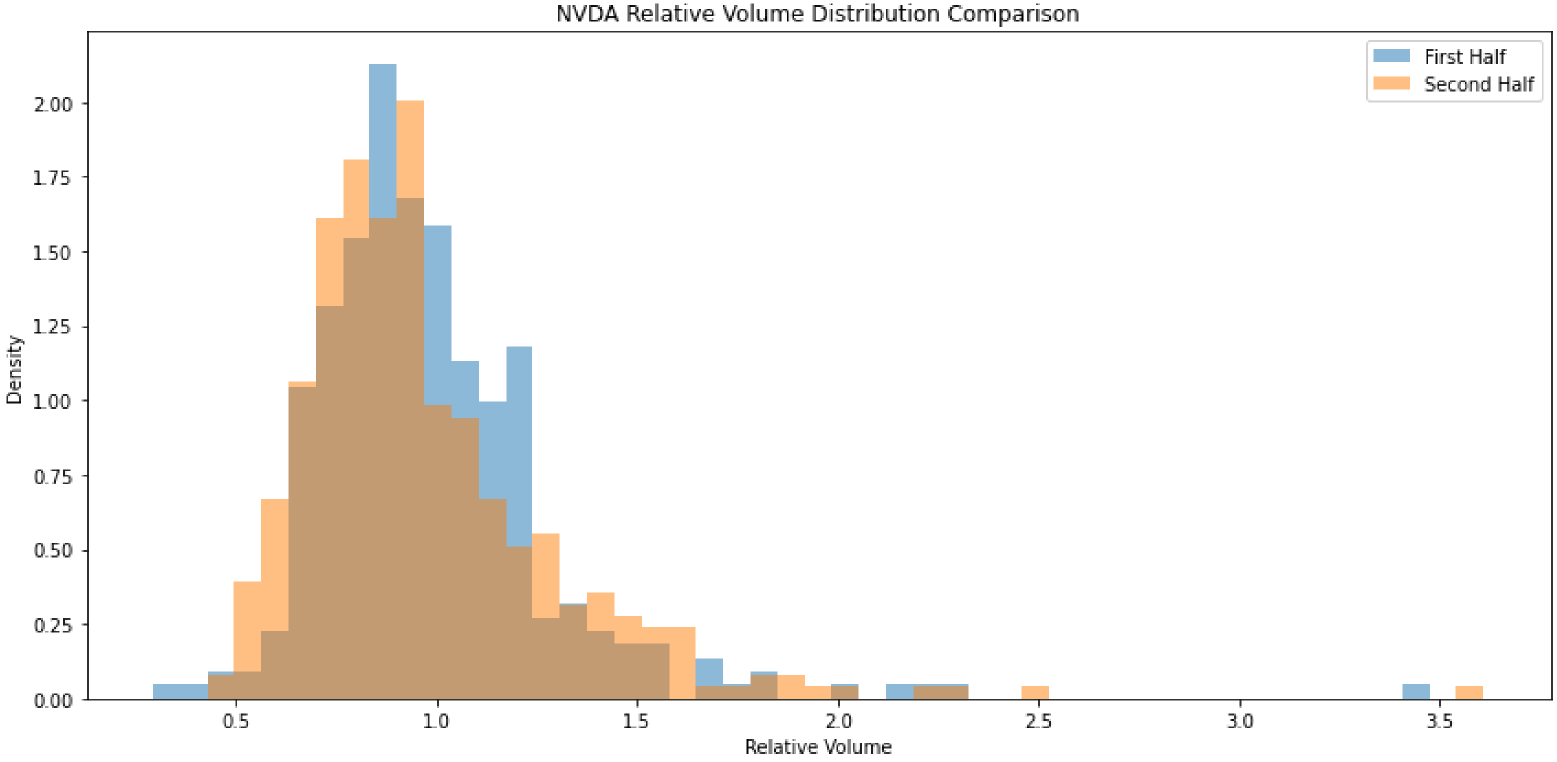
Out[16]:



Volume (relative to market) distributions

In [18]:
```
fig
```

Out[18]:



```
print("Done")
```