Trend analysis on IN Petrol Price (Incomplete)

A collaborative report by Rishikesh Sheernam and Yeshvanth Kurapati

In [91]:

```
import numpy as np
import pandas as pd
pd.plotting.register_matplotlib_converters()
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

```
/kaggle/input/d/rhilok/new-data-hyd-ptrl/diesel.csv
/kaggle/input/d/rhilok/new-data-hyd-ptrl/new_data_hyd.csv
/kaggle/input/d/rhilok/new-data-hyd-ptrl/final_petrol_data.csv
/kaggle/input/d/rhilok/new-data-hyd-ptrl/petrol.csv
/kaggle/input/d/rhilok/new-data-hyd-ptrl/new_data_all.csv
```

Petrol rates in 2021

Petrol rates of different cities from Jan 2021 to present *(27/2/21)*

In [132]:

```
new_data=pd.read_csv('/kaggle/input/d/rhilok/new-data-hyd-ptrl/new_data_all.csv',index_
col='date',parse_dates=True)

new_data=new_data.drop(["Unnamed: 0"],axis=1)
new_data
```

Out[132]:

date	city	rate
2021-01-01	Delhi	83.71
2021-01-02	Delhi	83.71
2021-01-03	Delhi	83.71
2021-01-04	Delhi	83.71
2021-01-05	Delhi	83.71
...	...	...
2021-02-23	Bengaluru	93.95
2021-02-24	Bengaluru	93.95
2021-02-25	Bengaluru	93.95
2021-02-26	Bengaluru	93.95
2021-02-27	Bengaluru	94.20

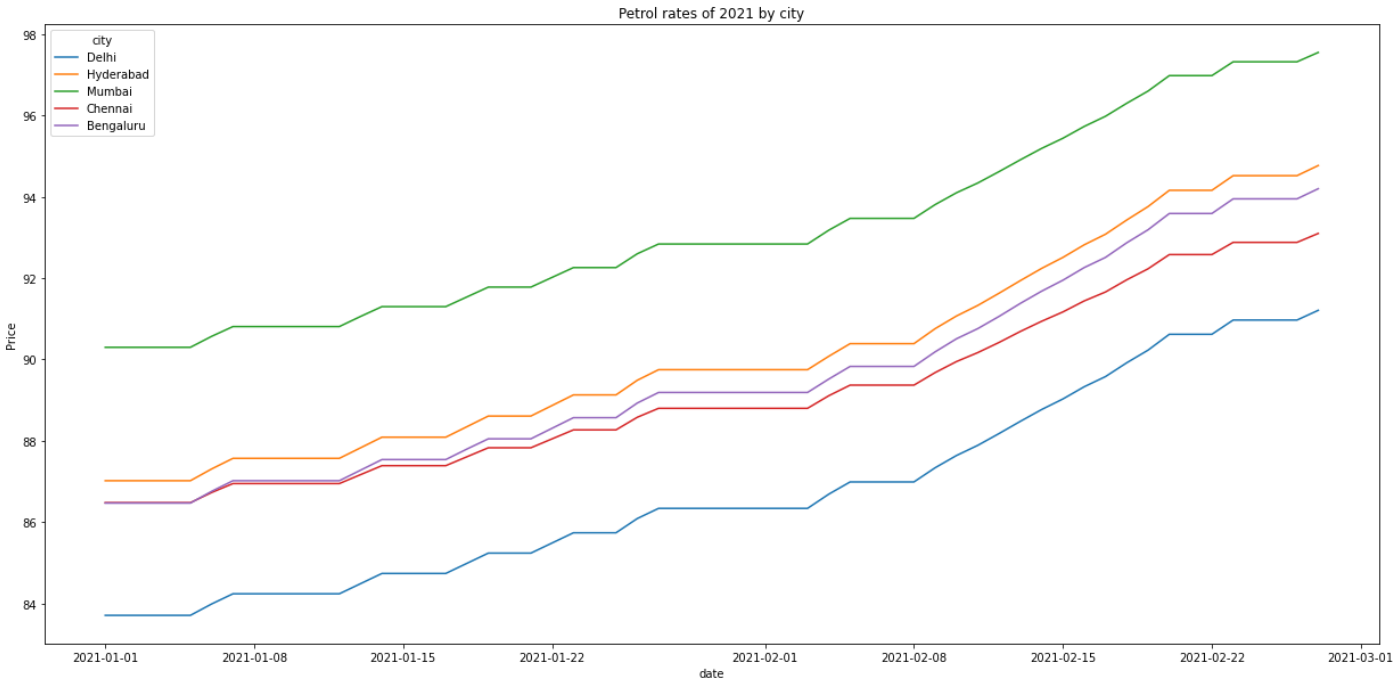
290 rows × 2 columns

In [140]:

```
plt.figure(figsize=(21,10))
sns.lineplot(data=new_data , x=new_data.index,y='rate',hue='city')
plt.title("Petrol rates of 2021 by city")
plt.ylabel("Price")
```

Out[140]:

```
Text(0, 0.5, 'Price')
```



Petrol rates till 2021

(Shaded region is one standard deviation from the mean)

Loading petrol data till 2021

In [110]:

```
old_data=pd.read_csv('/kaggle/input/d/rhilok/new-data-hyd-ptrl/petrol.csv',index_col='date',parse_dates=True)
old_data=old_data.drop(columns='state')
old_data.head()
```

Out[110]:

date	city	rate
2011-12-01	**Hyderabad**	**73.07**
2012-04-01	**Hyderabad**	**73.08**
2012-05-24	**Hyderabad**	**81.44**
2012-06-03	**Hyderabad**	**78.11**
2012-06-29	**Hyderabad**	**74.89**

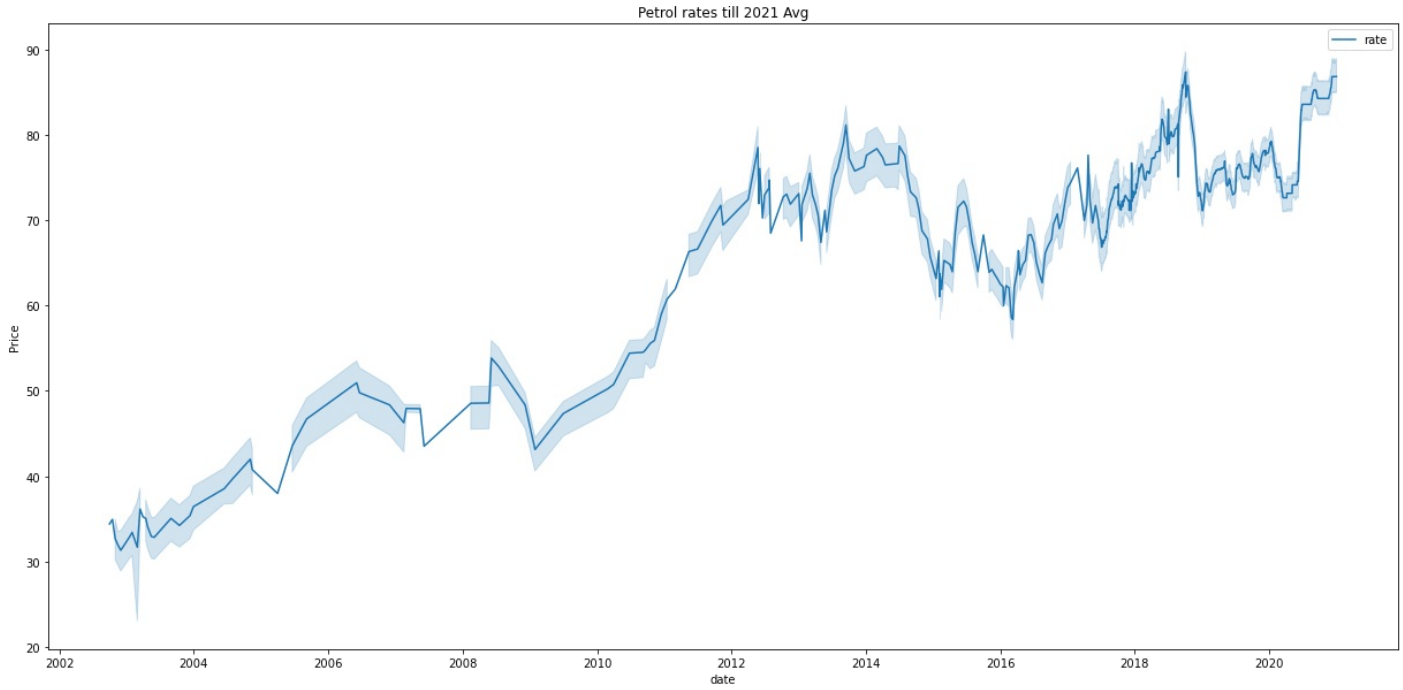
In [114]:

```
plt.figure(figsize=(21,10))

sns.lineplot(data=old_data)
plt.title("Petrol rates till 2021 Avg")
plt.ylabel("Price")
```

Out[114]:

```
Text(0, 0.5, 'Price')
```



Now let us work with data from 2019 to present

In [123]:

```
full_data=pd.read_csv('/kaggle/input/d/rhilok/new-data-hyd-ptrl/final_petrol_data.csv',
index_col='date',parse_dates=True)
full_data=full_data.drop(["Unnamed: 0",'state'],axis=1)
full_data
```

Out[123]:

date	city	rate
2011-12-01	Hyderabad	73.07
2012-04-01	Hyderabad	73.08
2012-05-24	Hyderabad	81.44
2012-06-03	Hyderabad	78.11
2012-06-29	Hyderabad	74.89
...	...	...
2021-02-23	Bengaluru	93.95
2021-02-24	Bengaluru	93.95
2021-02-25	Bengaluru	93.95
2021-02-26	Bengaluru	93.95
2021-02-27	Bengaluru	94.20

5338 rows × 2 columns

In [139]:

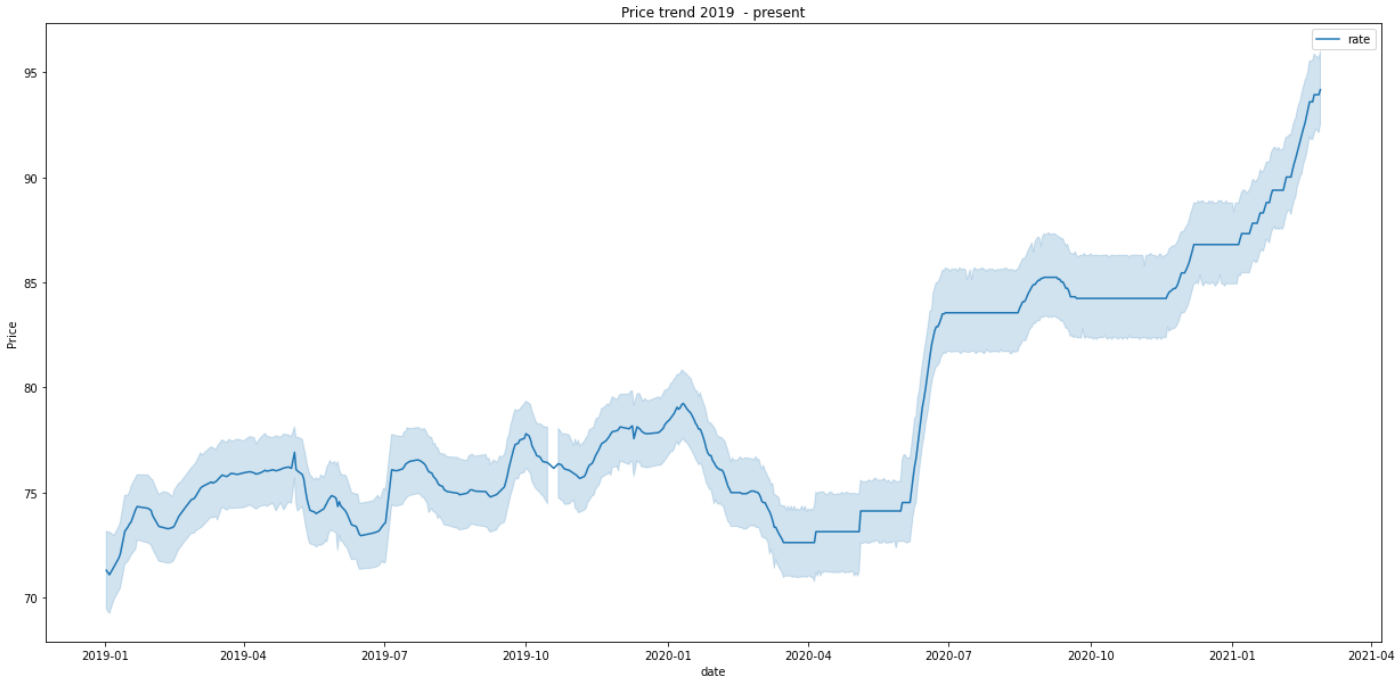
```
plt.figure(figsize=(21,10))
sns.lineplot(data=full_data.loc["2019-01-01":"2021-02-27"])
plt.title("Price trend 2019  - present")
plt.ylabel("Price")
```

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:2: FutureWarning: Value ba
sed partial slicing on non-monotonic DatetimeIndexes with non-existing keys is deprecate
```

```
d and will raise a KeyError in a future Version.
```

Out[139]:

```
Text(0, 0.5, 'Price')
```



In [ ]: