

OIL AND GAS PRICES ANALYSIS FROM 1932-2014

Oil & Gas Price Analysis:

1. Import the required libraries
2. Import the dataset and review the data frame
3. Plotting a general graph from the dataset
4. Histograms and price distribution (pct.change)
5. percentage change distribution (+-%1, +-%5 change)
6. Extreme and normal condition probabilities
7. Investigating Extremes — Minimum and Maximum values and their position

```
In [107...  ## importing required libraries
import pandas as pd
import numpy as np
##use pylab inline to numpy,matplotlib
%pylab inline
import matplotlib.pyplot as plt
from matplotlib import style
from matplotlib import rcParams
plt.rcParams['figure.figsize'] = (30, 10)
import matplotlib.pylab as pylab
params = {'legend.fontsize': 'xx-large',
          'figure.figsize': (30, 10),
          'axes.titlesize': 'xx-large',
          'xtick.labelsize': 'xx-large',
          'ytick.labelsize': 'xx-large'}
pylab.rcParams.update(params)
```

Populating the interactive namespace from numpy and matplotlib

```
In [108...  ## Importing the data

df=pd.read_csv("Oil and Gas 1932-2014.csv")

## SElecting the features for data
df1= pd.DataFrame(df["year"])
df2=df["oil_price_2000"]
df3=df1.join(df2)
df3=pd.DataFrame(df3)
print(df3.shape)

df1= pd.DataFrame(df["year"])

df4=df["gas_price_nom"]

df5=df1.join(df4)
df5=pd.DataFrame(df5)
```

(15521, 2)

```
In [109...  ## checking missing values in Oil dataset and replace it by mean values
df3.isnull().sum()
x = df3.mean()
```

```
y = round(x,2)
df3= df3.fillna(y)

df3.isnull().sum()
```

```
Out[109... year          0
oil_price_2000      0
dtype: int64
```

```
In [110... ## checking missing values in Gas dataset and replace it by mean values
df5.isnull().sum()
x = df5.mean()
y = round(x,2)
df5= df5.fillna(y)

df5.isnull().sum()
```

```
Out[110... year          0
gas_price_nom       0
dtype: int64
```

```
In [111... ## View of the Oildata
df3.head()
```

```
Out[111...   year  oil_price_2000
0  1932      10.860860
1  1933       8.815239
2  1934      12.715790
3  1935      12.039690
4  1936      13.383210
```

```
In [112... ## View of GasData
df3.head()
```

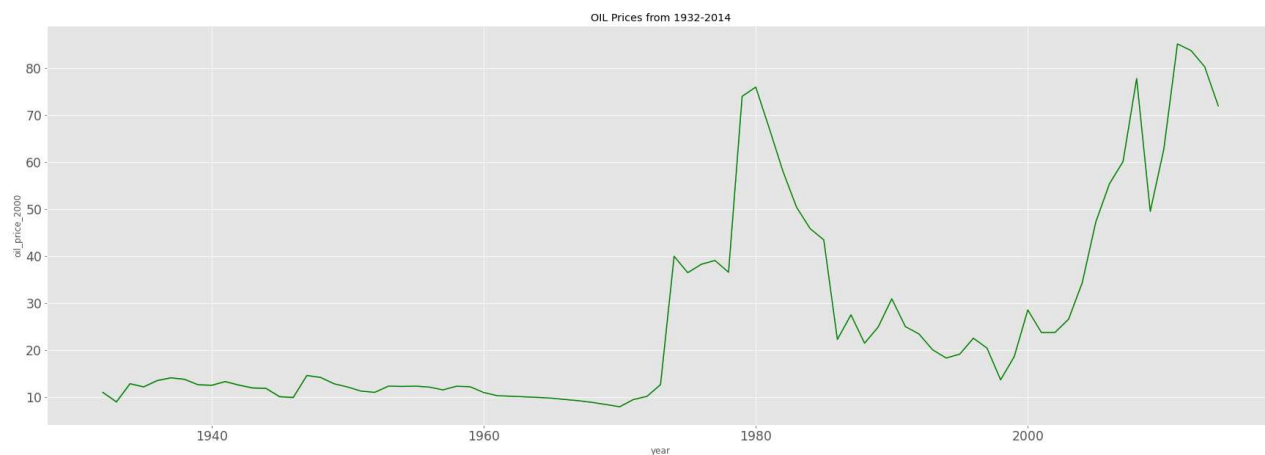
```
Out[112...   year  oil_price_2000
0  1932      10.860860
1  1933       8.815239
2  1934      12.715790
3  1935      12.039690
4  1936      13.383210
```

Visualization of OilData set

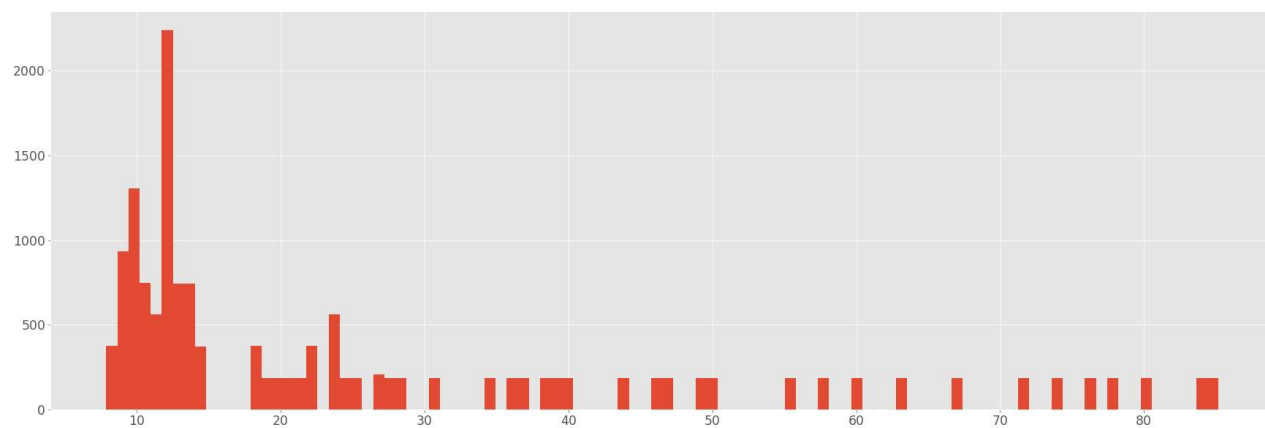
```
In [145... ## Visualising the OILDataset

import seaborn as sns
style.use("ggplot")
sns.lineplot(x='year',y='oil_price_2000',data=df3,color='green')
```

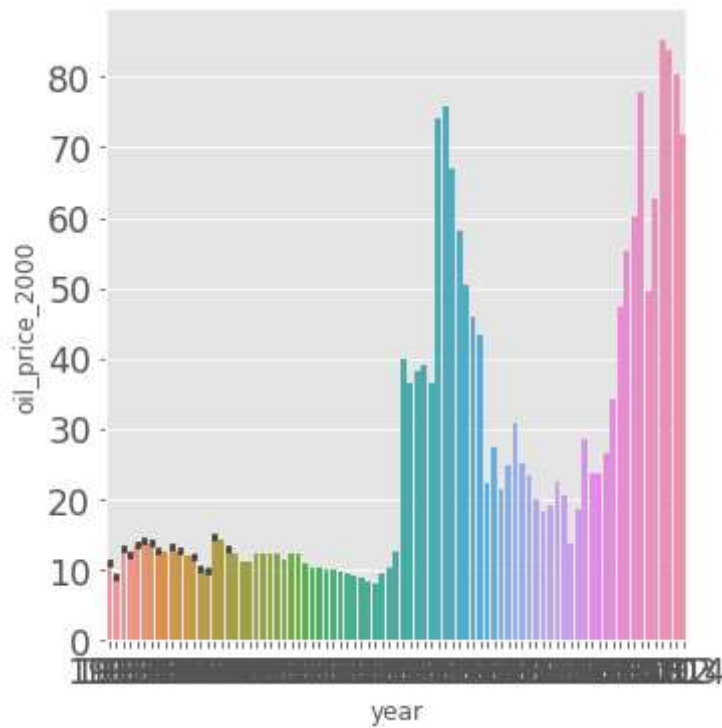
```
plt.title("OIL Prices from 1932-2014")  
plt.show()
```



```
In [146... df3["oil_price_2000"].hist(bins = 100)  
plt.show()
```



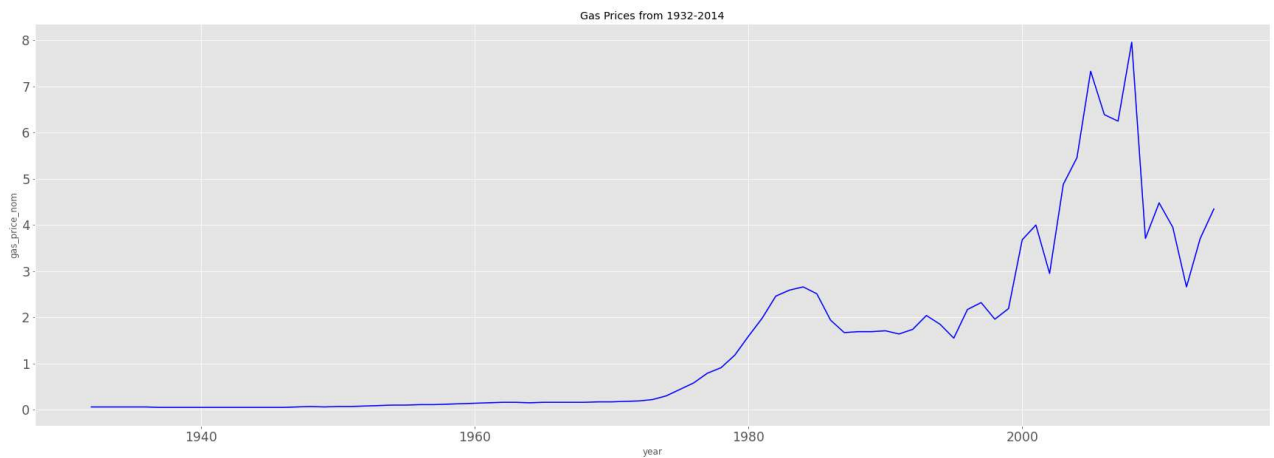
```
In [147... sns.catplot(data=df3, kind="bar", x="year",y="oil_price_2000")  
plt.show()
```



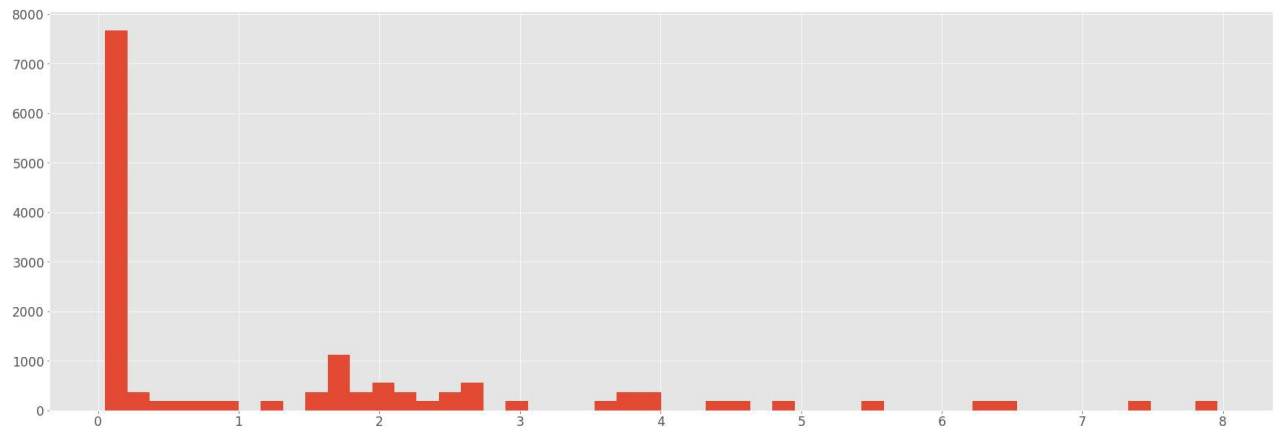
Visualization of Gas data set

```
In [143...  ## Visualising GASDataset

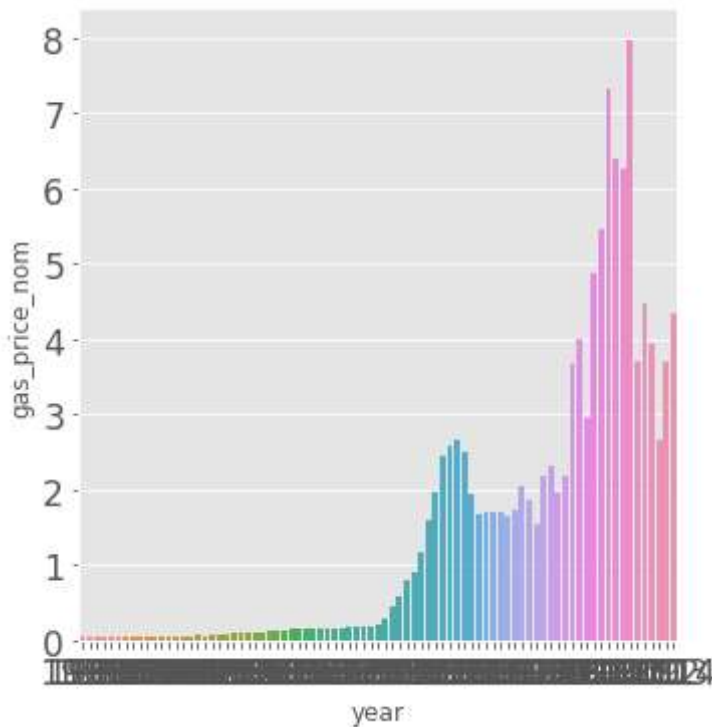
sns.lineplot(x="year",y='gas_price_nom',data=df5,color="blue")
plt.title("Gas Prices from 1932-2014")
plt.show()
```



```
In [144... df5["gas_price_nom"].hist(bins = 50)
plt.show()
```



```
In [148... sns.catplot(data=df5, kind="bar", x="year", y="gas_price_nom")
plt.show()
```



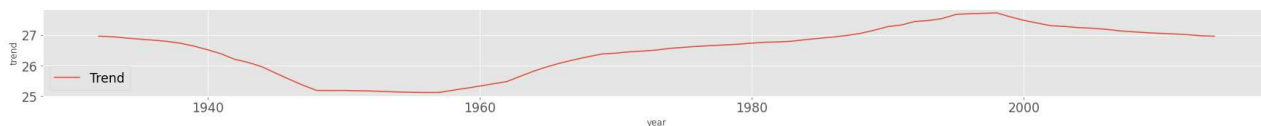
```
In [149... # Trend of oildata globally
from statsmodels.tsa.seasonal import seasonal_decompose
import seaborn as sns

ts= df3["oil_price_2000"]

decomposition = seasonal_decompose(ts, period = 365)

trend = decomposition.trend

#plt.subplot(411)
#sns.lineplot(y=ts, x='year', data=df3, label='Original')
#plt.legend(loc='best')
plt.subplot(412)
sns.lineplot(y=trend, x='year', data=df3, label='Trend')
plt.legend(loc='best')
plt.show()
```



Trend of the Oil Prices Globally

The trend of the Oil prices Globally shows that the

1. The oil prices decrease from 1932 to 1950 and remains constant upto 1960.
2. After 1960 the Oil prices are increasing reaches high increases price at and end of 2000 (1950 (lowest) to 2000 (highest) increases 600%)
3. In the beginning of 21st century (upto 2014) the oil prices are decreasing slightly.

In [120]...

```
## Percentage Changes in oil Data
df3["pc_oil"] = df3["oil_price_2000"].pct_change()

df3
```

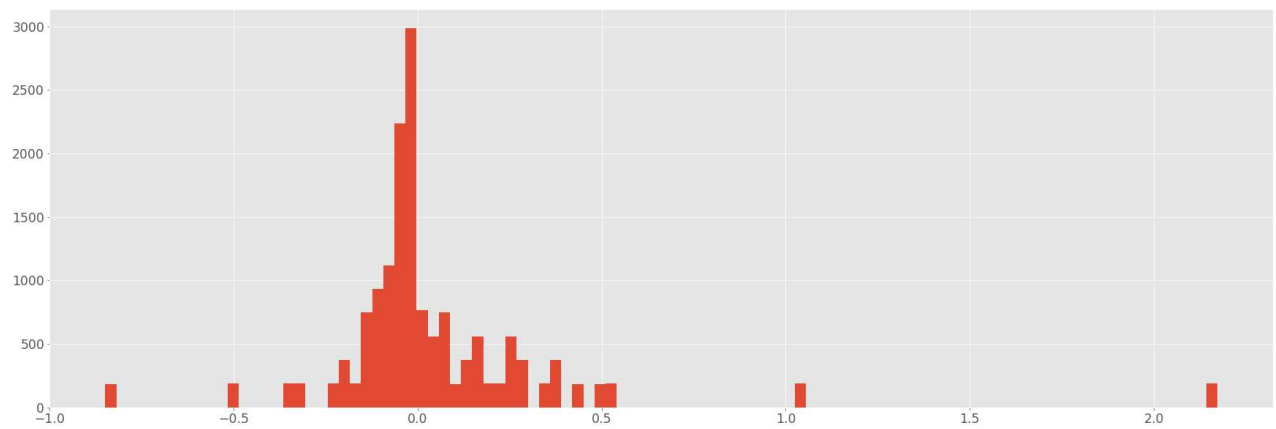
Out[120]...

	year	oil_price_2000	pc_oil
0	1932	10.860860	NaN
1	1933	8.815239	-0.188348
2	1934	12.715790	0.442478
3	1935	12.039690	-0.053170
4	1936	13.383210	0.111591
...
15516	2010	62.778052	0.268218
15517	2011	85.170708	0.356696
15518	2012	83.754465	-0.016628
15519	2013	80.319530	-0.041012
15520	2014	71.972588	-0.103922

15521 rows × 3 columns

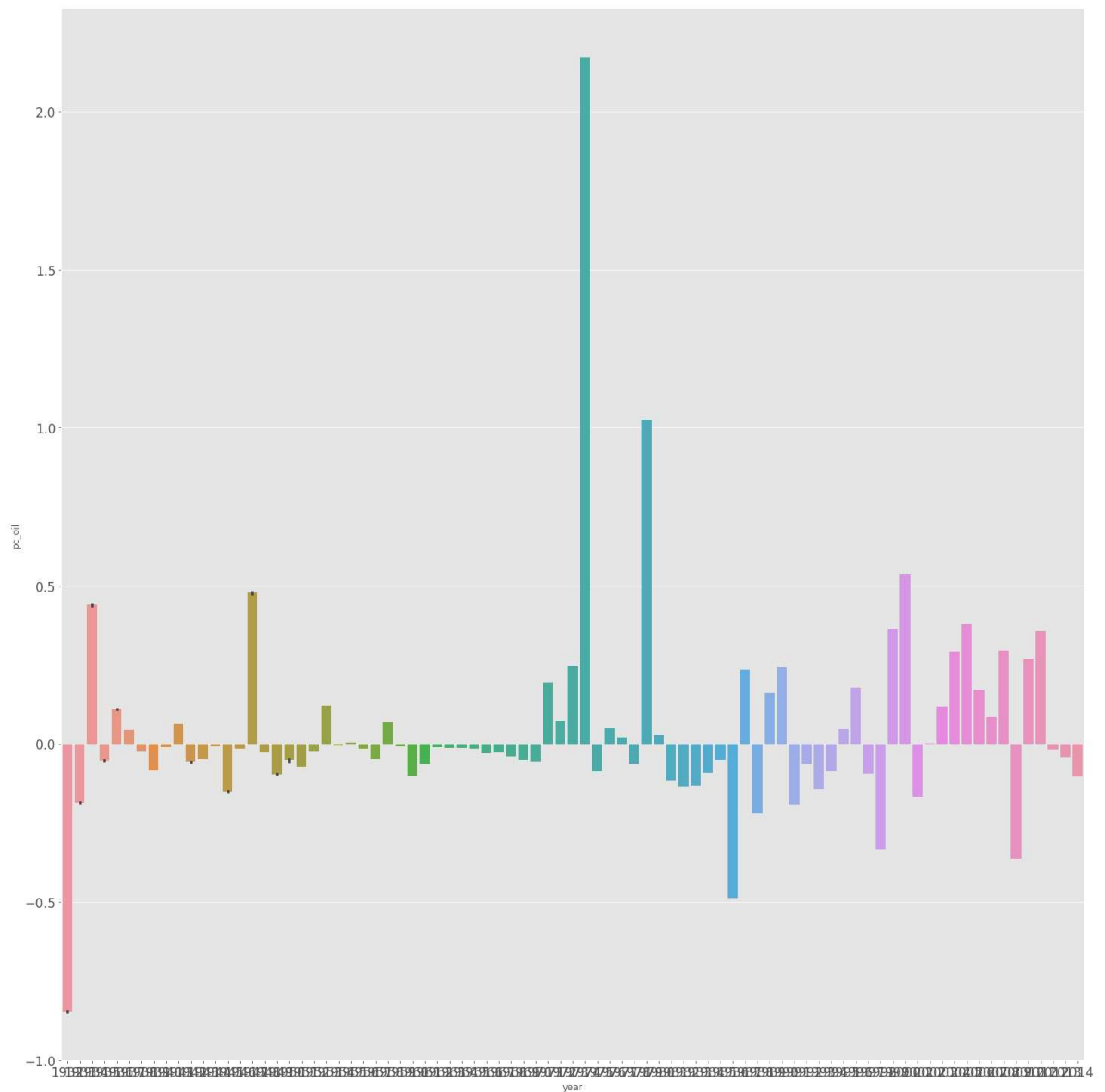
In [150]...

```
df3['pc_oil'].hist(bins=100)
plt.show()
```



In [151...

```
sns.catplot(data=df3, kind="bar", x="year", y="pc_oil", height=20)
plt.show()
```



In [152...

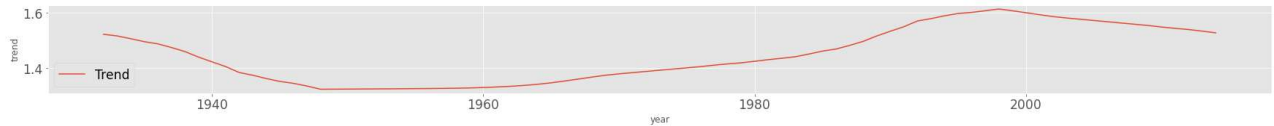
```
## Trend of Gasdata globally
from statsmodels.tsa.seasonal import seasonal_decompose
import seaborn as sns
```

```
ts1= df5["gas_price_nom"]

decomposition = seasonal_decompose(ts1, period = 365)

trend = decomposition.trend

plt.subplot(412)
sns.lineplot(y=trend,x='year', data=df5, label='Trend')
plt.legend(loc='best')
plt.show()
```



Trend of Gas prices Globally

The trend of the Gas prices Globally shows that the

- 1.The Gas prices decrease from 1932 to 1950 and remains constant upto 1965.
- 2.After 1965 the Gas prices are increaseing reaches hight increse price at and end of 2000(1965(lowest) t0 2000(heighest) increases 240%)
- 3.In the begining of 21st century(upto 2014) the oil prices are decreaseing slightly.

```
In [124... ## Percentage change in GasData
df5["pc_gas"]=df5["gas_price_nom"].pct_change()
df5
```

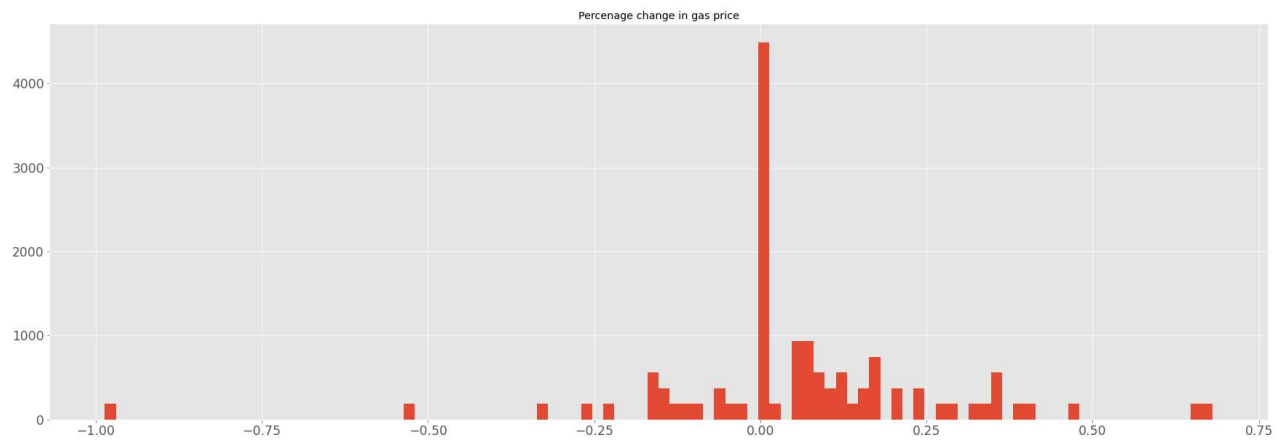
```
Out[124...
   year  gas_price_nom  pc_gas
0  1932      0.060000    NaN
1  1933      0.060000  0.000000
2  1934      0.060000  0.000000
3  1935      0.060000  0.000000
4  1936      0.060000  0.000000
...    ...          ...    ...
15516  2010      4.480000  0.207547
15517  2011      3.950000 -0.118304
15518  2012      2.660000 -0.326582
15519  2013      3.707290  0.393718
15520  2014      4.347742  0.172755
```

15521 rows × 3 columns

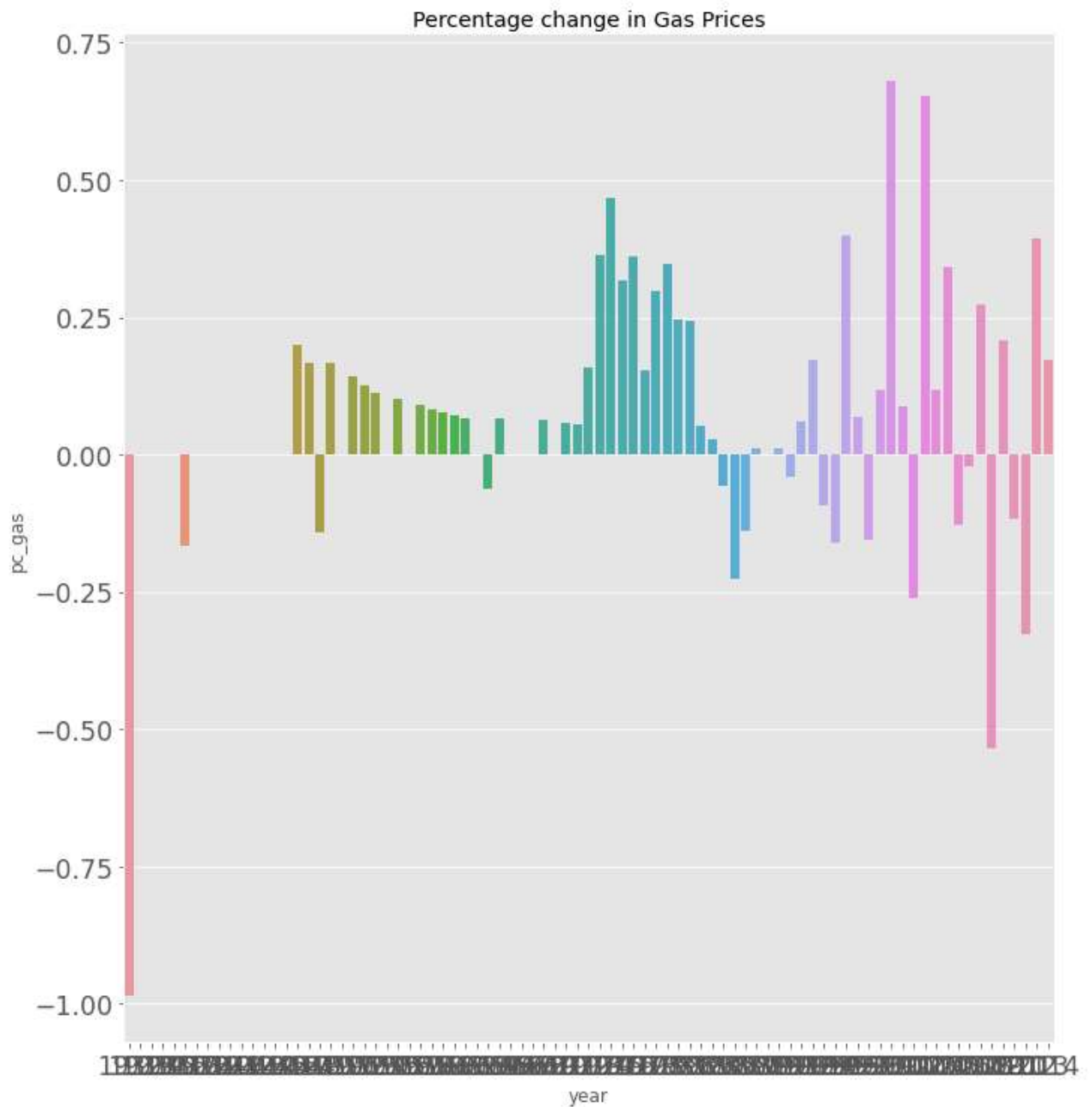
```
In [153... df5["pc_gas"].hist(bins=100)
```



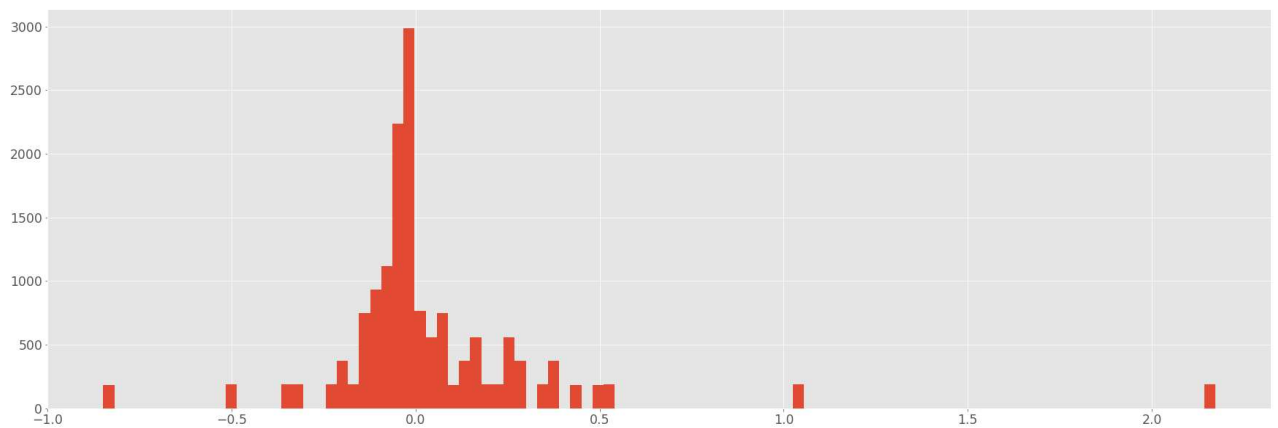
```
plt.title("Percentage change in gas price")  
plt.show()
```



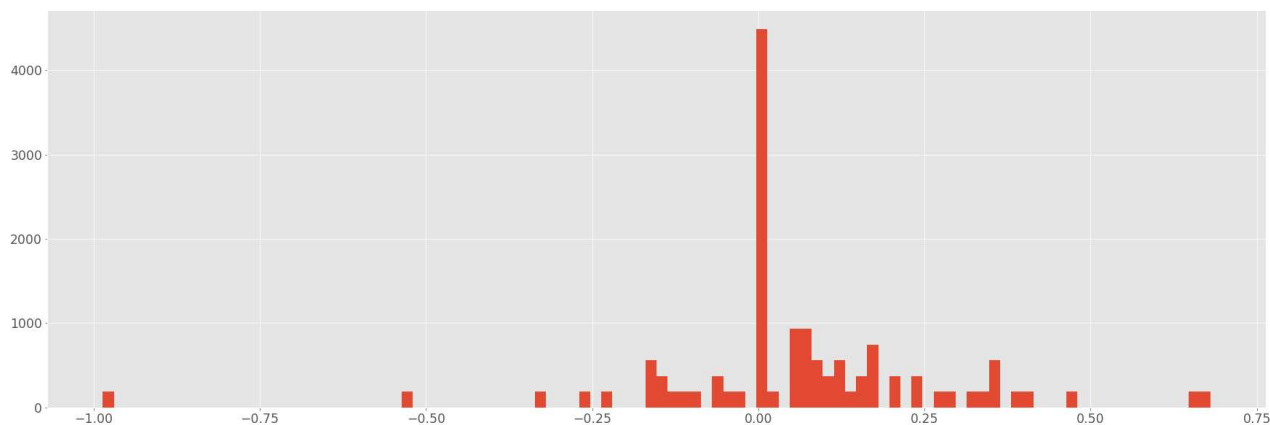
```
In [154... sns.catplot(data=df5, kind="bar", x="year",y="pc_gas",height=10)  
plt.title("Percentage change in Gas Prices")  
plt.show()
```



```
In [155... hist(df3['pc_oil'],bins=100)
plt.show()
```



```
In [156... hist(df5["pc_gas"],bins=100)
plt.show()
```



```
In [129... df3['pc_oil'].shape
```

```
Out[129... (15521,)
```

```
In [130... df3['pc_oil'].count
```

```
Out[130... <bound method Series.count of 0          NaN
1          -0.188348
2           0.442478
3          -0.053170
4           0.111591
...
15516       0.268218
15517       0.356696
15518      -0.016628
15519      -0.041012
15520      -0.103922
Name: pc_oil, Length: 15521, dtype: float64>
```

```
In [131... pc_oil=df3["pc_oil"]
```

```
In [132... df5["pc_gas"].shape
```

```
Out[132... (15521,)
```

```
In [133... df5["pc_gas"].count
```

```
Out[133... <bound method Series.count of 0          NaN
1           0.000000
2           0.000000
3           0.000000
4           0.000000
...
15516       0.207547
15517      -0.118304
15518      -0.326582
15519       0.393718
15520       0.172755
Name: pc_gas, Length: 15521, dtype: float64>
```

```
In [134... ##Normal and Extreme Conditions Probabilities
##Look at the probability of +/-1%; +/-3%; +/-5% change in oil prices.
##Out of %5 change in price can be investigated as an extreme condition

print("The Probability of Oil price changes between 1% and -1% is %1.2f%% " % (100*pc
```

```

        & (pc_oil<0.01)].shape[0]/pc_oil.shape[0]))

print("The Probability of Oil price changes between 3% and -3% is %1.2f%% " % (100*pc
        & (pc_oil<0.03)].shape[0]/pc_oil.shape[0]))

print("The Probability of Oil price changes between 5% and -5% is %1.2f%% " % (100*pc
        & (pc_oil<0.05)].shape[0]/pc_oil.shape[0]))

print#("The Probability of Oil price changes between 7% and -7% is %1.2f%% " % (100*p
        #      & (pc_oil<0.07)].shape[0]/pc_oil.shape[0]))

print("The Probability of Oil price change is morethan 5% is %1.2f%% " % (100*pc_oil[(
print("The Probability of Oil price change is lessthan -5% is %1.2f%% " % (100*pc_oil[

```

The Probability of Oil price changes between 1% and -1% is 7.33%
 The Probability of Oil price changes between 3% and -3% is 22.97%
 The Probability of Oil price changes between 5% and -5% is 33.79%
 The Probability of Oil price change is morethan 5% is 30.09%
 The Probability of Oil price change is lessthan -5% is 36.11%

This tells that the probability of change in Oil prices increasing Globally more than 5% is 30%

and decreasing morethat 5% is 36%.

In [135... `gas=df5["pc_gas"]`
`gas`

Out[135... `0` NaN
`1` 0.000000
`2` 0.000000
`3` 0.000000
`4` 0.000000
...
`15516` 0.207547
`15517` -0.118304
`15518` -0.326582
`15519` 0.393718
`15520` 0.172755
Name: pc_gas, Length: 15521, dtype: float64

In [136... *##Normal and Extreme Conditions Probabilities*
##Look at the probability of +/-1%; +/-3%; +/-5 change in gas prices.
##Out of %5 change in price can be investigated as an extreme condition

```

print("The Probability of gas price changes between 1% and -1% is %1.2f%% " % (100*ga
        & (gas<0.01)].shape[0]/gas.shape[0]))

print("The Probability of gas price changes between 3% and -3% is %1.2f%% " % (100*ga
        & (gas<0.03)].shape[0]/gas.shape[0]))

print("The Probability of gas price changes between 5% and -5% is %1.2f%% " % (100*ga
        & (gas<0.05)].shape[0]/gas.shape[0]))

#print("The Probability of gas price changes between 7% and -7% is %1.2f%% " % (100*g
        #      & (gas<0.07)].shape[0]/gas.shape[0]))

print("The Probability of gas price change morethan 5%is %1.2f%% " % (100*gas[gas>0.05
print("The Probability of gas price change lessthan -5%is %1.2f%% " % (100*gas[gas<-0.

```

The Probability of gas price changes between 1% and -1% is 26.51%
 The Probability of gas price changes between 3% and -3% is 31.33%
 The Probability of gas price changes between 5% and -5% is 32.53%
 The Probability of gas price change morethan 5%is 49.40%
 The Probability of gas price change lessthan -5%is 18.07%

This tells that the probability of change in gas prices increasing Globally more than 5% is 49%

and decreasing morethan 5% is 18%.

```
In [137...  ##Max and Min Oil & Gas Prices:
          ##Let's try to pick up when max increase and decrease in oil and gas prices happened.

print("Minmun value =" , df3['pc_oil'].min())
print("Index of the Minimum value =" ,df3['pc_oil'].idxmin())
print("MAximum value =" ,df3['pc_oil'].max())
print("Index of the Maximum value =" ,df3['pc_oil'].idxmax())

Minmun value  = -0.8490972699205265
Index of the Minimum value  = 83
MAximum value  = 2.1728650690422593
Index of the Maximum value  = 42
```

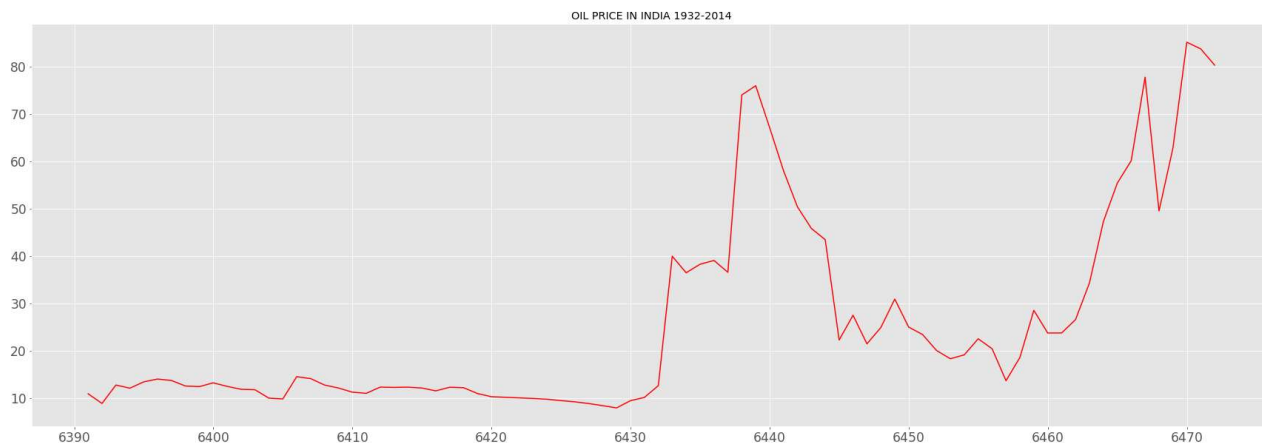
```
In [138... df3["oil_price_2000"].iloc[[82,83,84,85,86]] ## decrease from 72 to 12 600% decrease
```

```
Out[138... 82    71.972588
            83    10.860860
            84     8.815239
            85    12.715790
            86    12.039690
            Name: oil_price_2000, dtype: float64
```

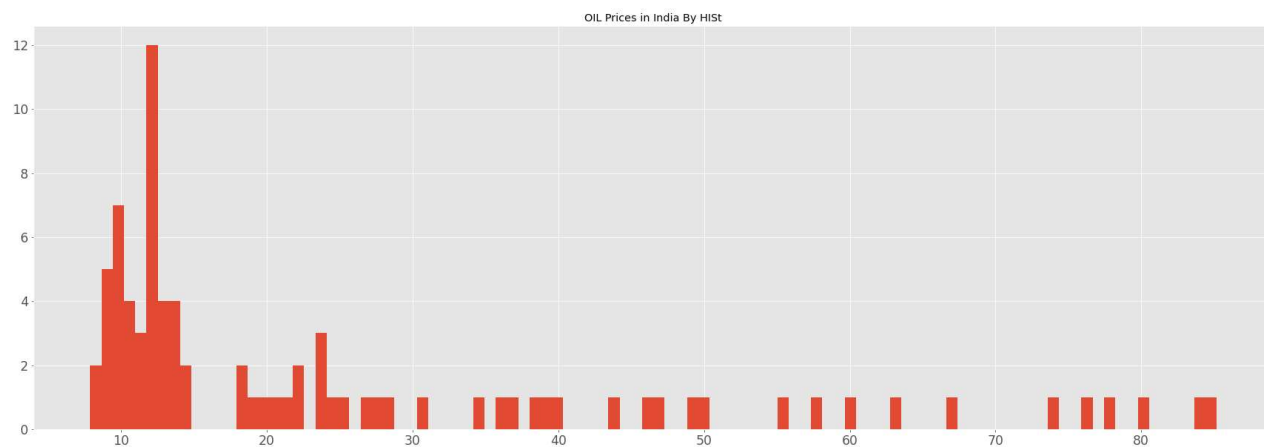
```
In [139... gas.iloc[[41,42,43,44,45]] ## increase from 0.15 to 0.36 240% increase
```

```
Out[139... 41    0.157895
            42    0.363636
            43    0.466667
            44    0.318182
            45    0.362069
            Name: pc_gas, dtype: float64
```

```
In [140...  ## oil prices in India
          plot(df3['oil_price_2000'].iloc[6391:6473], color="red")
          plt.title("OIL PRICE IN INDIA 1932-2014")
          plt.show()
```



```
In [141... df3['oil_price_2000'].iloc[6391:6473].hist(bins=100)
plt.title("OIL Prices in India By HIsT")
plt.show()
```



```
In [157... ## Gas prices in India
plot(df5["gas_price_nom"].iloc[6391:6473], color="red")
plt.title("GAS PRICES IN INDIA 1932-2014")
plt.show()
```

