Homework 1

We load data on monthly stock returns from 1926 to 2021 (source).

In [38]:
```
import numpy as np
import pandas as pd

stocks = pd.read_csv("stocks.csv")
ret = stocks["Mkt-RF"].values
```

A *density histogram* is a histogram that is normalized so that the total area under the bars sums to unity. Argue that the stock return is not normally distributed as follows:

- Plot a density histogram of the returns **`ret`** with 100 bins.
- Over this histogram, plot the PDF of a normal distribution with mean equal to the sample mean of the returns and standard deviation equal to the sample standard deviation of the returns.

In [39]:
```
%matplotlib inline
import matplotlib.pyplot as plt

# Histogram
plt.hist(ret,bins=100,density=True,label="Density Histogram of Returns")

# mean of the returns
mean = np.sum(ret)/len(ret)

# standard deviation of the returns
# add ddof=1 since np.std is for population standard deviation
# ddof=1 gives the calculation of sample standard deviation which is N-1 in the divisor
std = np.std(ret,ddof=1)

# plot the pdf
from scipy.stats import norm as gaussian
x = np.linspace(-30,40,len(ret))
plt.plot(x, gaussian.pdf(x,mean,std),label="PDF of Normal Distribution")
plt.legend()
plt.show()
```