

Exploratory Data Analysis

J. Hamski

4/15/2017

```
library(dplyr)
library(ggplot2)
library(ggfortify)
library(magrittr)
library(xts)
library(broom)

load("wti_project.Rda")
```

Exploratory Data Analysis

Data Source

The data sources used were the West Texas Intermediate (WTI) and Brent nominal (i.e. not inflation adjusted) daily spot price record from the U.S. Energy Information Administration. The WTI series was used from January 2, 1986 through December 30, 2016, the Brent series from May 20 1987 through December 30, 2016.

```
summary(wti.xts) ##>% tidy()
```

##	Index	wti.xts
##	Min. :1986-01-03	Min. : 10.25
##	1st Qu.:1993-09-01	1st Qu.: 19.38
##	Median :2001-06-11	Median : 28.01
##	Mean :2001-06-19	Mean : 42.87
##	3rd Qu.:2009-03-31	3rd Qu.: 63.47
##	Max. :2016-12-30	Max. :145.31

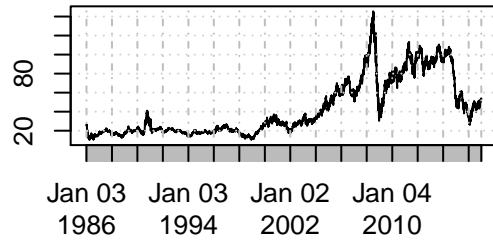
Returns and Volatility

A challenge in analyzing financial time series in general, and spot oil market prices specifically, is that the variance structure may be independent, but not identically distributed. Oil prices exhibit periods of low volatility (i.e. relatively constant prices) and periods of high volatility (i.e. changing prices). This is referred to as volatility clustering. This violates the assumption in the most frequently used time series model, the autoregressive integrated moving average (ARIMA) model.

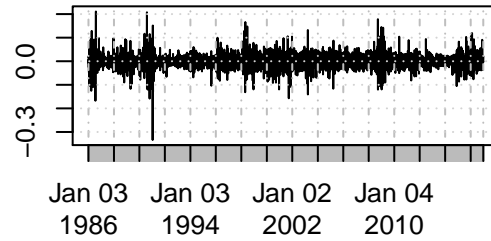
In this research paper, volatility is characterized two ways: (1) 5-day historic volatility and (2) 30-day historic volatility. In addition, the relationship between the returns themselves and price level is investigated.

```
par(mfrow = c(2, 2))
plot.xts(wti.xts, type = "l")
plot.xts(wti.return, type = "l")
plot.xts(wti.combined[,3], type = "l")
plot.xts(wti.combined[,4], type = "l")
```

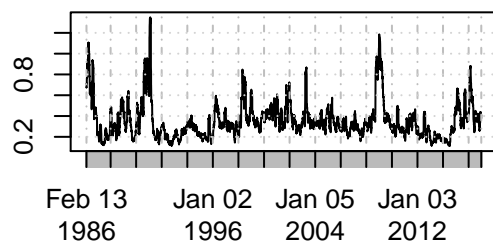
wti.xts



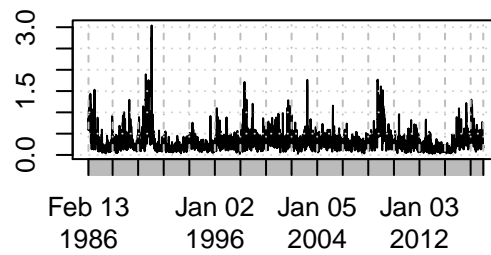
wti.return



wti.combined[, 3]



wti.combined[, 4]



```
par(mfrow = c(1, 1))
```

Times series exploration

```
acf(wti.return, lag.max = 50)
```

Series wti.return

