Problem Statement

Executive Summary:  
Preliminary research has shown little impact between large cocaine seizures and undercover purchases by law enforcement. This is contrary to what is expected and can have multiple explanations including i) supply chain robustness or ii) bad data.   
The primary goal of this research project is to apply text mining/web scraping to extract cocaine listings, location, price, and other useful information in order to study the impact of large seizures on the price.   
We started off by i) analyzing various research papers to understand the approach used by previous authors. ii) extracting appropriate files based on keywords and iii) applying state-of-the-art web scraping methods to extract the data.  
  
Data Description:  
Dark Net Markets (DNM) are online markets typically hosted as Tor hidden services providing escrow services between buyers & sellers transacting in Bitcoin or other crypto coins, usually for drugs or other illegal/regulated goods; the most famous DNM was Silk Road 1, which pioneered the business model in 2011. The author scraped all existing English-language DNMs between 2013-2015. It amounts to 50GB (~1.6TB uncompressed) collection covering 89 DNMs & 37+ related forums.  
The following project focuses on Silk Road 2, which was allegedly run by former administrators of Silk Road.

­­­­