

WSJ

June 29, 2022

```
[1]: wsj_url = "https://www.wsj.com/search?  
      ↪query=cryptocurrency&mod=searchresults_viewallresults"
```

```
[4]: !pip install selenium
```

Collecting selenium

Downloading selenium-4.3.0-py3-none-any.whl (981 kB)

981.4/981.4

KB 4.9 MB/s eta 0:00:0000:0100:01

Collecting trio~=0.17

Downloading trio-0.21.0-py3-none-any.whl (358 kB)

359.0/359.0

KB 3.4 MB/s eta 0:00:0000:0100:01

Collecting trio-websocket~=0.9

Downloading trio_websocket-0.9.2-py3-none-any.whl (16 kB)

Requirement already satisfied: urllib3[secure,socks]~=1.26 in
/usr/local/lib/python3.9/site-packages (from selenium) (1.26.6)

Requirement already satisfied: idna in /usr/local/lib/python3.9/site-packages
(from trio~=0.17->selenium) (2.10)

Requirement already satisfied: sniffio in /usr/local/lib/python3.9/site-packages
(from trio~=0.17->selenium) (1.2.0)

Collecting outcome

Downloading outcome-1.2.0-py2.py3-none-any.whl (9.7 kB)

Collecting sortedcontainers

Downloading sortedcontainers-2.4.0-py2.py3-none-any.whl (29 kB)

Requirement already satisfied: attrs>=19.2.0 in /usr/local/lib/python3.9/site-
packages (from trio~=0.17->selenium) (21.4.0)

Collecting async-generator>=1.9

Downloading async_generator-1.10-py3-none-any.whl (18 kB)

Collecting wsproto>=0.14

Downloading wsproto-1.1.0-py3-none-any.whl (24 kB)

Collecting PySocks!=1.5.7,<2.0,>=1.5.6

Downloading PySocks-1.7.1-py3-none-any.whl (16 kB)

Collecting cryptography>=1.3.4

Downloading cryptography-37.0.2-cp36-abi3-macosx_10_10_x86_64.whl (2.8 MB)

2.8/2.8 MB

5.2 MB/s eta 0:00:0000:0100:01

```

Collecting pyOpenSSL>=0.14
  Downloading pyOpenSSL-22.0.0-py2.py3-none-any.whl (55 kB)
    55.8/55.8 KB
1.7 MB/s eta 0:00:00
Requirement already satisfied: certifi in /usr/local/lib/python3.9/site-
packages (from urllib3[secure,socks]~=1.26->selenium) (2021.5.30)
Requirement already satisfied: cffi>=1.12 in /usr/local/lib/python3.9/site-
packages (from cryptography>=1.3.4->urllib3[secure,socks]~=1.26->selenium)
(1.15.0)
Collecting h11<1,>=0.9.0
  Downloading h11-0.13.0-py3-none-any.whl (58 kB)
    58.2/58.2 KB
2.1 MB/s eta 0:00:00
Requirement already satisfied: pycparser in /usr/local/lib/python3.9/site-
packages (from
cffi>=1.12->cryptography>=1.3.4->urllib3[secure,socks]~=1.26->selenium) (2.21)
Installing collected packages: sortedcontainers, PySocks, outcome, h11, async-
generator, wsproto, trio, cryptography, trio-websocket, pyOpenSSL, selenium
Successfully installed PySocks-1.7.1 async-generator-1.10 cryptography-37.0.2
h11-0.13.0 outcome-1.2.0 pyOpenSSL-22.0.0 selenium-4.3.0 sortedcontainers-2.4.0
trio-0.21.0 trio-websocket-0.9.2 wsproto-1.1.0
WARNING: You are using pip version 22.0.3; however, version 22.1.2 is
available.

You should consider upgrading via the '/Users/anshulrao/ipy-jupyter-
venv3/bin/python -m pip install --upgrade pip' command.

```

```
[6]: !pip install webdriver-manager
```

```

Collecting webdriver-manager
  Downloading webdriver_manager-3.7.1-py2.py3-none-any.whl (25 kB)
Collecting python-dotenv
  Downloading python_dotenv-0.20.0-py3-none-any.whl (17 kB)
Requirement already satisfied: requests in /usr/local/lib/python3.9/site-
packages (from webdriver-manager) (2.25.1)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.9/site-packages (from requests->webdriver-manager)
(2021.5.30)
Requirement already satisfied: chardet<5,>=3.0.2 in
/usr/local/lib/python3.9/site-packages (from requests->webdriver-manager)
(4.0.0)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/usr/local/lib/python3.9/site-packages (from requests->webdriver-manager)
(1.26.6)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.9/site-
packages (from requests->webdriver-manager) (2.10)

```

Installing collected packages: python-dotenv, webdriver-manager
Successfully installed python-dotenv-0.20.0 webdriver-manager-3.7.1
WARNING: You are using pip version 22.0.3; however, version 22.1.2 is
available.

You should consider upgrading via the '/Users/anshulrao/ipy-jupyter-venv3/bin/python -m pip install --upgrade pip' command.

```
[9]: from selenium import webdriver
from webdriver_manager.chrome import ChromeDriverManager

driver = webdriver.Chrome(ChromeDriverManager().install())
```

```
[WDM] - ===== WebDriver manager =====
[WDM] - Current google-chrome version is 103.0.5060
[WDM] - Get LATEST chromedriver version for 103.0.5060 google-chrome
[WDM] - Driver
[/Users/anshulrao/.wdm/drivers/chromedriver/mac64/103.0.5060.53/chromedriver]
found in cache
/var/folders/rd/80b2d3pj44n9nwhd83335__40000gn/T/ipykernel_1337/2160715986.py:4:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
    driver = webdriver.Chrome(ChromeDriverManager().install())
```

```
[10]: driver.get(wsj_url)
```

```
[12]: from selenium.webdriver.common.by import By
```

```
[30]: title_xpath = '//*[@id="main"]/div/div[1]/article[{}]/div/div[1]/div[1]/div[2]/
↳h3/a'
```

```
[32]: body_xpath = '//*[@id="main"]/div/div[1]/article[{}]/div/div[1]/div[1]/p'
```

```
[35]: articles = {
    'title': [],
    'body': [],
    'link': []
}
```

```
[36]: # reading the 20 articles present on the page
limit = 20
for i in range(1, limit + 1):
    title_elem = driver.find_element(By.XPATH, title_xpath.format(i))
    articles['title'].append(title_elem.text)
    articles['link'].append(title_elem.get_attribute('href'))
```

```
body_elem = driver.find_element(By.XPATH, body_xpath.format(i))
articles['body'].append(body_elem.text)
```

```
[37]: articles
```

```
[37]: {'title': ['Stocks Waver as Central Bankers Hold Panel',
'Crypto Hedge Fund Three Arrows Ordered to Liquidate by Court',
'DeFi's Existential Problem: It Only Lends Money to Itself',
'Behind the Celsius Sales Pitch Was a Crypto Firm Built on Risk',
'What Happens If a Crypto Platform Goes Bankrupt?',
'NFT Artists Get First Major Test Since Crypto's Collapse',
'Opinion: 'Irrational Exuberance' on a Plate',
'Do Kwon Tries Again After $40 Billion Crypto Crash',
'Hedge Funds Bet Against Tether, as Crypto Melts Down',
'How Margin Calls Can Transform Market Downturns Into Market Crashes',
'Risk of Lockdowns in China Delays Return to Business as Usual',
'Do Kwon Tries Again After $40 Billion Crypto Crash',
'Robinhood Shares Soar on Takeover Hopes',
'Crypto Hedge Fund Three Arrows Defaulted on Loan, Says Broker',
'Studios Battle for Players of 'Hypercasual' Videogames, the Latest Mobile
Distraction',
'Marketers at NFT.NYC Upbeat on Brand-Building in the Metaverse',
'More Hedge Funds Bet Against Tether as Crypto Melts Down',
'VR Tech to Combat Social Isolation Among Seniors',
'Dow Jumps as Rate-Hike Expectations Cool',
'Banks Face Risks From Geopolitics, Staff Turnover, OCC Says'],
'body': ['The panel of major central bank officials is expected to provide
insight into their views on the economy, inflation and the path of monetary
policy.',
'A British Virgin Islands court ordered the liquidation of the cryptocurrency
hedge fund after creditors sued it for failure to repay debts.',
'Crypto lenders' exclusive focus on other crypto projects suggests that their
problems run much deeper than a Lehman-style liquidity crisis.',
'Celsius CEO Alex Mashinsky built his cryptocurrency lender on a pitch that it
was better than a bank. But investor documents show it had far more risk than a
traditional lender. ',
'When cryptocurrency lending platform Celsius froze user accounts amid a
plunge in valuations, it sent ripples across the industry and raised questions
about what happens to user assets if a crypto platform files for bankruptcy.
WSJ's Vicky Ge Huang explains. Photo illustration: Jordan Kransse',
'Last year Christie's sold nearly $150 million worth of NFTs. This year? So
far, it's only sold $4.6 million in NFT art.',
'A bursting of the fake-burger bubble.',
'Algorithmic stablecoin TerraUSD was supposed to maintain its value at $1 with
the help of its sister-crypto, Luna. But TerraUSD's price collapsed to pennies
in May, while Luna plunged by more than 99%, wiping out around $40 billion in
value. Their creator Do Kwon has been attempting a comeback by relaunching Luna.
```

Wall Street Journal markets reporter Alexander Osipovich joins WSJ Tech News Briefing host Julie Chang to explain the collapse of Do Kwon's crypto empire and what his second try looks like.',

'Short sellers have been ramping up their bets against tether, the world's largest stablecoin, amid a broad market selloff that has called into doubt the financial health of some crypto companies. Wall Street Journal markets reporter Caitlin McCabe joins WSJ What's News host Luke Vargas with more.',

'Investors rarely hear about margin calls except during market downturns. WSJ's Dion Rabouin explains what they are and the unique way they are contributing to the current cryptocurrency crash. Illustration: Adele Morgan',

'A.M. Edition for June 28. The city of Shanghai may be emerging from a monthslong Covid-induced lockdown, but for companies across China, concerns about abrupt future lockdowns loom large. WSJ China corporate correspondent Liza Lin explains how a recent cycle of pandemic restrictions could complicate future business decisions and risk denting consumer activity. Luke Vargas hosts.',

'Algorithmic stablecoin TerraUSD was supposed to maintain its value at \$1 with the help of its sister-crypto, Luna. But TerraUSD's price collapsed to pennies in May, while Luna plunged by more than 99%, wiping out around \$40 billion in value. Their creator Do Kwon has been attempting a comeback by relaunching Luna. WSJ markets reporter Alexander Osipovich joins host Julie Chang to explain the collapse of Do Kwon's crypto empire and what his second try looks like.',

'Robinhood Markets shares surged Monday on reports that cryptocurrency exchange FTX is exploring a potential acquisition of the brokerage.',

'Three Arrows has defaulted, failing to make loan payments tied to large bets in the digital-currency realm, crypto broker Voyager Digital said.',

'Mobile games that require little brainpower to play are becoming more sophisticated as their publishers try to cling on to fickle-and monetizable-players.',

'Companies' enthusiasm for the advertising potential of the virtual world was undimmed at NFT.NYC, a four-day event last week, in contrast to the recent downturn in cryptocurrencies and NFTs.',

'Short sellers have been ramping up their bets against tether, the world's largest stablecoin, amid a broad market selloff that has called into doubt the financial health of some crypto companies.',

"To tackle social isolation among seniors, some caregivers have turned to virtual reality. One company developing that technology is Rendeever. At The Wall Street Journal's recent Future of Everything Festival, personal tech columnist Nicole Nguyen spoke with Rendeever's co-founder and CEO about where this tech stands today and how it can be leveraged to transform the day-to-day lives of residents in senior communities. Julie Chang hosts.",

'U.S. stock indexes moved more than 2% higher, as investors shifted their views about central bank policy after weak economic data.',

'The Office of the Comptroller of the Currency said geopolitical events, inflation and higher interest rates are posing headwinds for banks.',

'link': ['<https://www.wsj.com/articles/global-stocks-markets-dow-update-06-29-2022-11656487667?page=1>'],

'<https://www.wsj.com/articles/crypto-fund-three-arrows-ordered-to-liquidate->

```

by-court-11656506404?page=1',
'https://www.wsj.com/articles/defis-existential-problem-it-only-lends-money-
to-itself-11656503840?page=1',
'https://www.wsj.com/articles/behind-the-celsius-sales-pitch-was-a-crypto-
firm-built-on-risk-11656498142?page=1',
'https://www.wsj.com/video/series/news-explainers/what-happens-if-a-crypto-
platform-goes-bankrupt/492D3C38-C15A-4654-AA1F-9CFDC52CEADB?page=1',
'https://www.wsj.com/articles/nft-artists-get-first-major-test-since-cryptos-
collapse-11656453852?page=1',
'https://www.wsj.com/articles/irrational-exuberance-on-a-
plate-11656435348?page=1',
'https://www.wsj.com/podcasts/google-news-update/do-kwon-tries-again-
after-40-billion-crypto-crash/A5C77DD5-6385-492C-AA5F-0FB353DF9DA8?page=1',
'https://www.wsj.com/podcasts/google-news-update/hedge-funds-bet-against-
tether-as-crypto-melts-down/44684BF4-3B44-456C-903F-20A86D3A24C1?page=1',
'https://www.wsj.com/video/series/dion-rabouin/how-margin-calls-can-transform-
market-downturns-into-market-
crashes/D4936064-C30C-4BA6-801D-13EB6703A5B1?page=1',
'https://www.wsj.com/podcasts/whats-news/risk-of-lockdowns-in-china-delays-
return-to-business-as-usual/3FE3FCBE-C20D-479E-A56A-E6520D29AFCE?page=1',
'https://www.wsj.com/podcasts/tech-news-briefing/do-kwon-tries-again-
after-40-billion-crypto-crash/E657E91A-15A7-4234-99DE-78F9DD5B3079?page=1',
'https://www.wsj.com/articles/robinhood-shares-soar-on-takeover-
hopes-11656362590?page=1',
'https://www.wsj.com/articles/crypto-hedge-fund-three-arrows-defaulted-on-
loan-says-broker-voyager-digital-11656336534?page=1',
'https://www.wsj.com/articles/studios-battle-for-players-of-hypercasual-
videogames-the-latest-mobile-distraction-11656324001?page=1',
'https://www.wsj.com/articles/marketers-at-nft-nyc-upbeat-on-brand-building-
in-the-metaverse-11656324000?page=1',
'https://www.wsj.com/articles/more-hedge-funds-are-betting-against-tether-as-
crypto-melts-down-11656322200?page=1',
'https://www.wsj.com/podcasts/tech-news-briefing/vr-tech-to-combat-social-
isolation-among-seniors/ABEBD6F8-EFBE-4314-BC49-617CF092D9CE?page=1',
'https://www.wsj.com/articles/global-stocks-markets-dow-
update-06-24-2022-11656055923?page=1',
'https://www.wsj.com/articles/banks-face-risks-from-geopolitics-staff-
turnover-occ-says-11656099709?page=1']}]

```

```
[38]: import pandas as pd
```

```
[41]: df = pd.DataFrame(articles)
```

```
[42]: df
```

```
[42]:
0                               title \
    Stocks Waver as Central Bankers Hold Panel
```

1 Crypto Hedge Fund Three Arrows Ordered to Liqu...
 2 DeFi's Existential Problem: It Only Lends Mone...
 3 Behind the Celsius Sales Pitch Was a Crypto Fi...
 4 What Happens If a Crypto Platform Goes Bankrupt?
 5 NFT Artists Get First Major Test Since Crypto'...
 6 Opinion: 'Irrational Exuberance' on a Plate
 7 Do Kwon Tries Again After \$40 Billion Crypto C...
 8 Hedge Funds Bet Against Tether, as Crypto Melt...
 9 How Margin Calls Can Transform Market Downturn...
 10 Risk of Lockdowns in China Delays Return to Bu...
 11 Do Kwon Tries Again After \$40 Billion Crypto C...
 12 Robinhood Shares Soar on Takeover Hopes
 13 Crypto Hedge Fund Three Arrows Defaulted on Lo...
 14 Studios Battle for Players of 'Hypercasual' Vi...
 15 Marketers at NFT.NYC Upbeat on Brand-Building ...
 16 More Hedge Funds Bet Against Tether as Crypto ...
 17 VR Tech to Combat Social Isolation Among Seniors
 18 Dow Jumps as Rate-Hike Expectations Cool
 19 Banks Face Risks From Geopolitics, Staff Turno...

body \

0 The panel of major central bank officials is e...
 1 A British Virgin Islands court ordered the liq...
 2 Crypto lenders' exclusive focus on other crypt...
 3 Celsius CEO Alex Mashinsky built his cryptocur...
 4 When cryptocurrency lending platform Celsius f...
 5 Last year Christie's sold nearly \$150 million ...
 6 A bursting of the fake-burger bubble.
 7 Algorithmic stablecoin TerraUSD was supposed t...
 8 Short sellers have been ramping up their bets ...
 9 Investors rarely hear about margin calls excep...
 10 A.M. Edition for June 28. The city of Shanghai...
 11 Algorithmic stablecoin TerraUSD was supposed t...
 12 Robinhood Markets shares surged Monday on repo...
 13 Three Arrows has defaulted, failing to make lo...
 14 Mobile games that require little brainpower to...
 15 Companies' enthusiasm for the advertising pote...
 16 Short sellers have been ramping up their bets ...
 17 To tackle social isolation among seniors, some...
 18 U.S. stock indexes moved more than 2% higher, ...
 19 The Office of the Comptroller of the Currency ...

link

0 <https://www.wsj.com/articles/global-stocks-mar...>
 1 <https://www.wsj.com/articles/crypto-fund-three...>
 2 <https://www.wsj.com/articles/defis-existential...>
 3 <https://www.wsj.com/articles/behind-the-celsiu...>

```
4 https://www.wsj.com/video/series/news-explaine...
5 https://www.wsj.com/articles/nft-artists-get-f...
6 https://www.wsj.com/articles/irrational-exuber...
7 https://www.wsj.com/podcasts/google-news-updat...
8 https://www.wsj.com/podcasts/google-news-updat...
9 https://www.wsj.com/video/series/dion-rabouin/...
10 https://www.wsj.com/podcasts/whats-news/risk-o...
11 https://www.wsj.com/podcasts/tech-news-briefin...
12 https://www.wsj.com/articles/robinhood-shares-...
13 https://www.wsj.com/articles/crypto-hedge-fund...
14 https://www.wsj.com/articles/studios-battle-fo...
15 https://www.wsj.com/articles/marketers-at-nft-...
16 https://www.wsj.com/articles/more-hedge-funds-...
17 https://www.wsj.com/podcasts/tech-news-briefin...
18 https://www.wsj.com/articles/global-stocks-mar...
19 https://www.wsj.com/articles/banks-face-risks-...
```

```
[44]: df.to_csv("wsj_articles.csv", index=False)
```