**Mini Project Ideas to Master BeautifulSoup (Without Selenium)**

Each of these projects uses only requests and BeautifulSoup. They are ideal for practicing real-world scraping techniques with static HTML content.

# 1. Top News Headlines Scraper

🔗 **Target Site:** <https://www.bbc.com/news> or <https://inshorts.com/en/read>

* **📌 Concepts Covered:**
* Tag selection
* Clean text
* Looping through headlines

📄 **Output Format:** | Headline | Category | URL |

# 2. Job Listings Aggregator

🔗 **Target Site:** https://remoteok.io or https://weworkremotely.com

* 📌 **Concepts Covered**:
* Extracting job title, company, tags
* Pagination
* Data cleanup and saving to CSV

# 3. Wikipedia Summary Extractor

🔗 **Tar**ge**t Site:** Wikipedia pages (e.g., <https://en.wikipedia.org/wiki/Python_(programming_language)>)

* **📌 Concepts Covered:**
* .find() and .find\_all()
* Navigating headings
* Selective tag parsing

📄 **Output Format:** Title and First Paragraph

# 4. IMDB Top 100 Movies Scraper

🔗 **Target Site:** <https://www.imdb.com/chart/top>

* **📌 Concepts Covered:**
* Tables
* Sorting data
* Tag attributes

📄 **Output Format:** | Rank | Movie Name | Year | Rating |

# 5. Crypto Price Tracker

🔗 Target Site: <https://www.coingecko.com/en> or <https://www.investing.com/crypto>

* **📌 Concepts Covered:**
* Nested tags
* Class-based filtering
* Currency formatting

# 6. Quotes Collector

🔗 **Target Site:** <http://quotes.toscrape.com/>

* 📌 Concepts Covered:
* Pagination
* Tag selection
* Attribute access

📄 **Output Format:** | Quote | Author | Tags |

# 7. Stock Prices Scraper

🔗 Target Site: <https://www.moneycontrol.com> or <https://www.nseindia.com>

* **📌 Concepts Covered:**
* Headers to avoid 403 errors
* Table scraping
* Data cleanup

# 8. Event Listing Scraper

🔗 **Target Site:** <https://www.eventbrite.com/d/online/all-events/>

* 📌 **Concepts Covered:**
* Titles, date, time, location
* Parsing datetime formats
* Pagination logic

# 9. E-book Deals Finder

🔗 **Target Site:** <https://www.packtpub.com/free-learning> or <https://manybooks.net>

* 📌 **Concepts Covered:**
* Titles and offer expiration
* Timers/countdowns
* Export to JSON

# 10. Sports Scoreboard Scraper

🔗 **Target Site:** <https://www.espncricinfo.com> or [https://www.livescore.com/en /](https://www.livescore.com/en%20/)

* **📌 Concepts Covered:**
* Real-time score data
* Extract team names and scores
* Use of .select() and .text