# Extracting and Structuring News Article Content from the Indian Express Homepage Using Web Scraping

**Problem Statement:**

**You are tasked with developing a Python script that automatically extracts news headlines and their corresponding article text from the Indian Express (https://indianexpress.com/) homepage. The goal is to collect and organize this data into a structured format (like a DataFrame) for further analysis, such as summarization, keyword extraction, or topic modeling.**

**The script should:**

- Scrape the homepage for Major Stories (Top News/Latest News) and the actual content of these articles
- Store the article title, link, and full text in a structured format like a pandas DataFrame.
- Include basic error handling for network failures or invalid pages.

**Constraints:**

- The script must use requests, BeautifulSoup, and pandas.
- Avoid scraping too aggressively — implement a polite and efficient scraping strategy.
- Should work on live pages and be easy to modify to handle saved HTML files later.

**Deliverable/ Submission Format:**

- Solution 1: A clean, well-commented Python script (.Py) Only.
- Solution 2: A CSV file containing the scraped article data in the below mentioned Google Drive link. Column headings of CSV be strictly as follows:

NEWS_TITLE_Interns Name_NEWS_LINK_FULL_SCRAPED_TEXT

**Google Drive link:** Scraping News Article_CSV Files

You are supposed to scrape 20 to 30 pieces of news.