|  |  |  |
| --- | --- | --- |
| **Tag** | **Purpose** | **Typical Use Case in Scraping** |
| <div> | Generic container (block-level) | Wrapping data blocks or sections |
| <span> | Generic container (inline) | Extract inline content like price, title, rating |
| <p> | Paragraph | Article text, reviews, descriptions |
| <a> | Anchor (link) | Extract URLs, navigation, article references |
| <img> | Image | Scrape image src, alt text |
| <ul> / <ol> | List container (unordered/ordered) | Review points, product features |
| <li> | List item | Items inside product features or menus |
| <table> | Table wrapper | Scrape tabular data |
| <thead> | Table header section | Extract column headers |
| <tbody> | Table body section | Extract rows of structured data |
| <tfoot> | Table footer section | Optional totals or summaries |
| <tr> | Table row | Scrape each row |
| <td> | Table data cell | Extract cell values |
| <th> | Table header cell | Get field names or categories |
| <form> | Web form container | Rare; may help simulate login/search (advanced use) |
| <input> | Form input field | Useful if scraping requires simulated form submission |
| <button> | Button (e.g., Load More) | For dynamic scraping with tools like Selenium |
| <section> | Semantic content block | Similar to <div>, used for grouping content |
| <article> | Full content item (e.g., a blog post or review) | Full scrape of structured post or entry |
| <label> | Describes input field | Identify or locate form inputs |
| <h1>–<h6> | Headings (hierarchy of content) | Titles, headings, blog/article section headers |
| <nav> | Navigation container | Sometimes used to skip menus in scraping |
| <iframe> | Embedded page | Advanced: may contain additional or hidden data |
| <meta> | Metadata (from <head>, e.g., description) | For SEO tags, OpenGraph data (only if needed) |

N8N HTML Extract Node - using HTML Tags:

**h3 > p > a**

N8N HTML Extract Node - using HTML & CSS Class Tags:

span.**sc-d541859f-1.imUuxf**

N8N HTML Extract Node - Using HTML & CSS Class and Attribute Tag**:**

span.rating-star[data-testid="star-rating"]