

# CSS Selectors and XPath

---



**Clarke Bishop**

BIG DATA ENGINEER

@ClarkeBishop [www.clarkebishop.com](http://www.clarkebishop.com)

# Overview

**HTML & CSS selectors**

**XPath**

**Chrome developer tools**

# HTML & CSS Selectors

---

## HTML Example

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li>... </li>
      <li>... </li>
    </ul>
  </body>
</html>
```

# Used Tesla Cars in Your Area

## 2017 Tesla Model 3

\$34,586

A wonderful car and a great deal.

## 2019 Tesla Model 3

\$37,938

Great deal. Low mileage.

## 2018 Tesla Model 3

\$36,263

Bright red. Attract attention.

## HTML Example

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li>... </li>
      <li>... </li>
    </ul>
  </body>
</html>
```

## Select the Title Tag

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
      </div>
    </li>
    <li> ... </li>
    <li> ... </li>
  </ul>
```

css => 'title'

## Select an H1 Tag

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
      </div>
    </li>
    <li> ... </li>
    <li> ... </li>
  </ul>
```

css => 'h1'



## Select Multiple <li> Elements

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li> ... </li>
      <li> ... </li>
    </ul>
```

CSS => 'li'

## Select with an HTML ID

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li> ... </li>
      <li> ... </li>
    </ul>
```

css => '#vin3827'

## Select with a CSS Class

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li> ... </li>
      <li> ... </li>
    </ul>
```

css => `' .auto-listing'`

## Select by Parent-Child

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li> ... </li>
      <li> ... </li>
    </ul>
```

css => 'ul li'

## Combined Selectors

```
<html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1 id="main-page-title">Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <div>
        <h2 class="model">2017 Tesla Model 3</h2>
        <p class="price">$34,586</p>
        <p class="description">
          A wonderful car and a great deal.</p>
        </div>
      </li>
      <li> ... </li>
      <li> ... </li>
    </ul>
```

css => 'ul.listings li#vin3827'

```
example = open("example.html", "r")  
html = example.read()  
# html = requests.get(url).text  
example.close()
```

```
from bs4 import BeautifulSoup  
soup = BeautifulSoup(html)  
print(soup.prettify())
```

# Beautiful Soup

```
example = open("example.html", "r")  
html = example.read()  
# html = requests.get(url).text  
example.close()
```

```
from bs4 import BeautifulSoup  
soup = BeautifulSoup(html)  
print(soup.prettify())
```

# Beautiful Soup

```
soup.title
```

```
<title>Car Buying Web Page</title>
```



```
soup.title
```

```
<title>Car Buying Web Page</title>
```

```
soup.li
```

```
<li id="vin3827" class="auto-listing">  
  <div>  
    <h2 class="model">2017 Tesla  
      Model 3</h2>  
    <p class="price">$34,586</p>  
    <p class="description">A wonderful  
      car and a great deal.</p>  
  </div>  
</li>
```

```
soup.title
```

```
<title>Car Buying Web Page</title>
```

```
soup.li
```

```
<li id="vin3827" class="auto-listing">  
  <div>  
    <h2 class="model">2017 Tesla  
      Model 3</h2>  
    <p class="price">$34,586</p>  
    <p class="description">A wonderful  
      car and a great deal.</p>  
  </div>  
</li>
```

```
soup.find_all('li')
```

```
[ <li> ... </li>, <li> ... </li>, <li> ... </li> ]
```

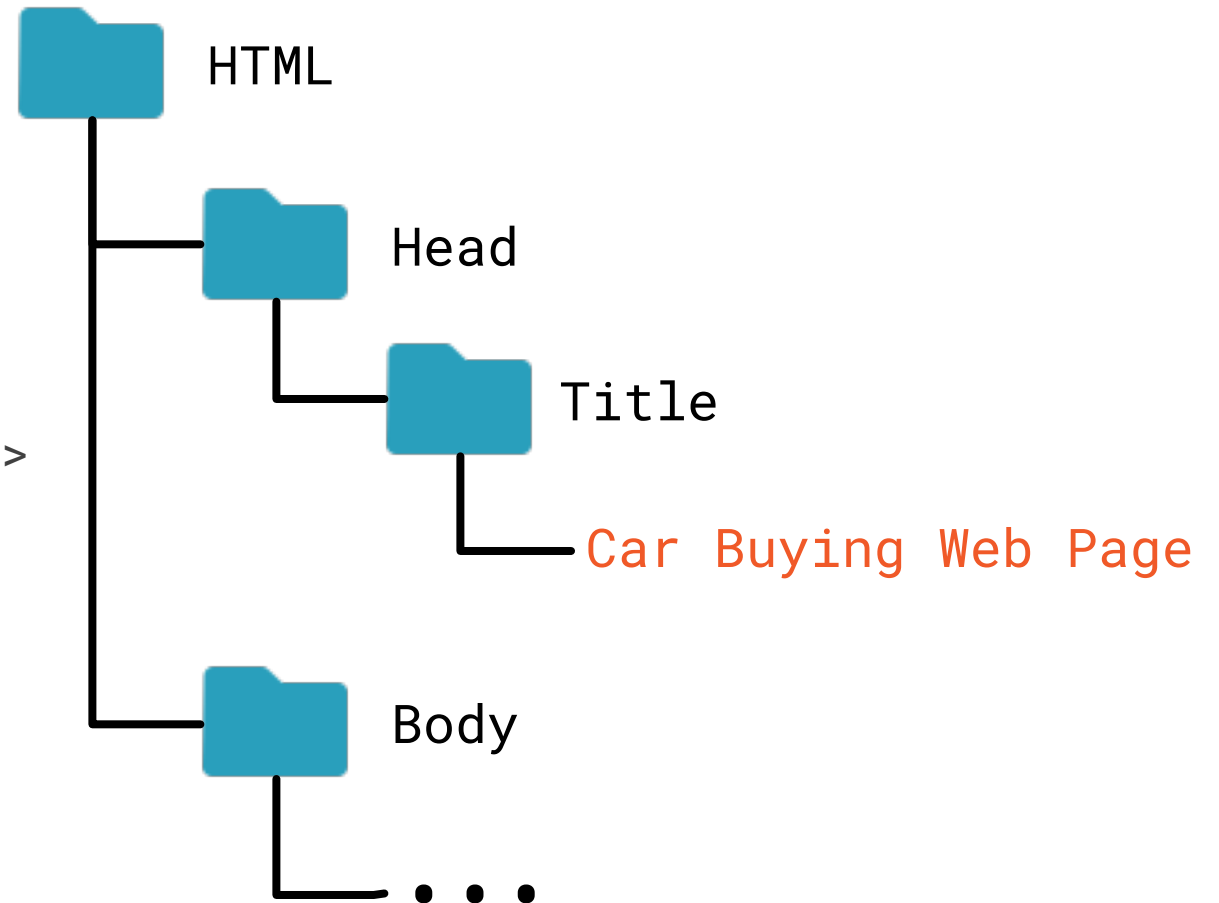
XPath

---

# XPath Document Tree

```
<!DOCTYPE html>
<head>
  <title>Car Buying Web Page</title>
</head>

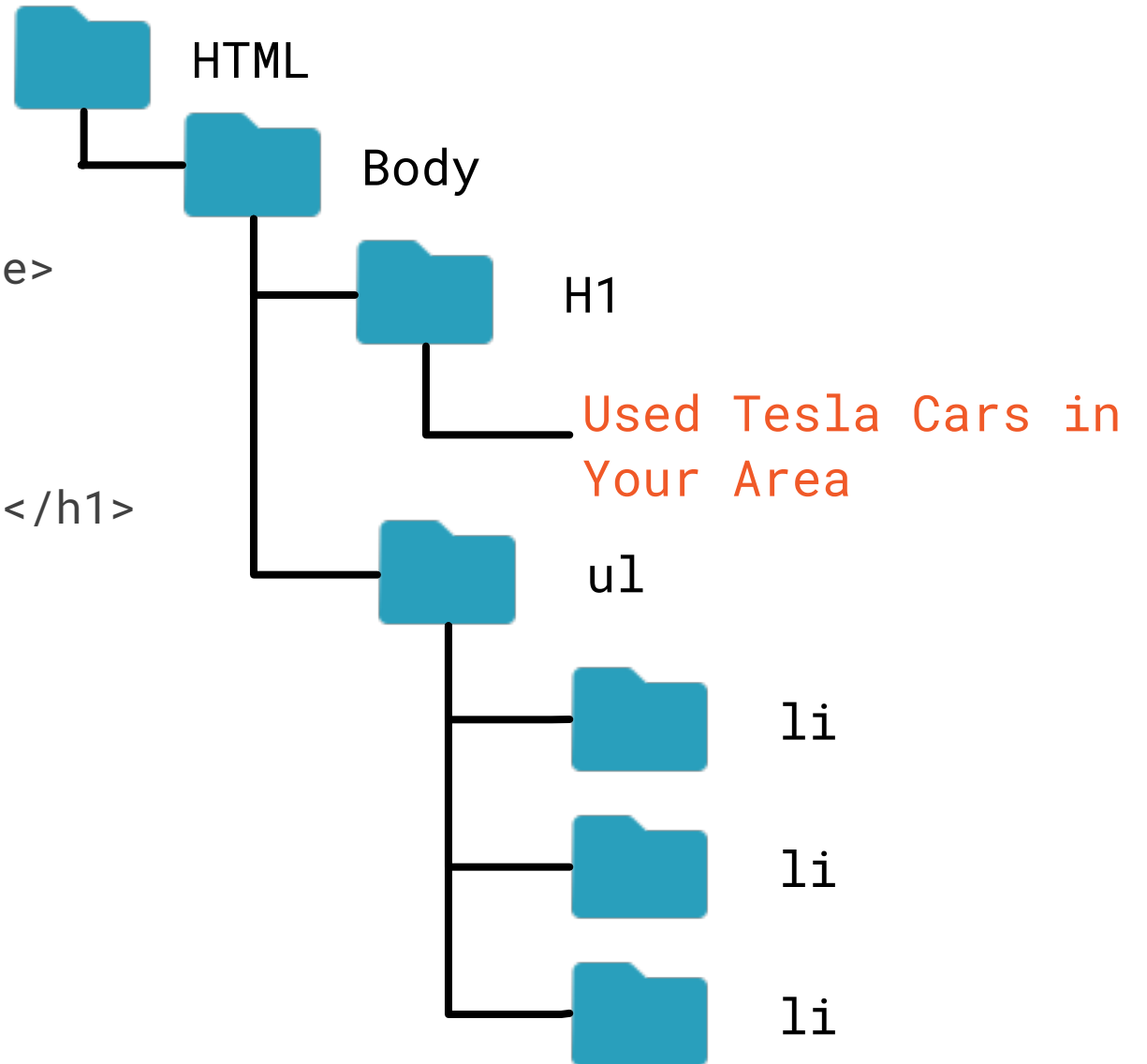
<body>
  <h1>Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li> ... </li>
    <li> ... </li>
    <li> ... </li>
  </ul>
</body>
</html>
```



# XPath Document Tree

```
<!DOCTYPE html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1>Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li> ... </li>
    <li> ... </li>
    <li> ... </li>
  </ul>
</body>
</html>
```



## XPath: Select Multiple <li>

```
<!DOCTYPE html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1>Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <h2 class="model">2015 Tesla Model S</h2>
      <p class="description"><span class="price">$34,586</span>
        A wonderful car and a great deal.</p>
    </li>
    <li> ... </li>
```

css => 'li'

xpath => '//li'

## XPath: Combined Selectors

```
<!DOCTYPE html>
<head>
  <title>Car Buying Web Page</title>
</head>

<body>
  <h1>Used Tesla Cars in Your Area</h1>
  <ul class="listings">
    <li id="vin3827" class="auto-listing">
      <h2 class="model">2015 Tesla Model S</h2>
      <p class="description"><span class="price">$34,586</span>
        A wonderful car and a great deal.</p>
    </li>
    <li> ... </li>
```

```
css => 'ul.listings li#vin3827'
```

```
xpath => '//ul[@class="listings"]/li[@id="vin3827"]'
```

# XPath or CSS Selector



**Which is Better?**



# Chrome Developer Tools

---

# Demo

**Accessing Chrome developer tools**

**Using developer tools panes**

**Editing styles to visualize selectors**

**Generating CSS selectors inside Chrome**

**Generating XPath selectors inside  
Chrome**

**Real world challenges**

# Summary

**HTML & CSS selectors**

**XPath selectors**

**Chrome developer tools**