

# Warmup: web scraping

## Group members:

Here is the source HTML for a simple web page

```
<h1> The Great British Bake Off </h1>
```

```
<p>
```

```
  <b>The Great British Bake Off</b> is a British television baking competition.
</p>
```

```
<h2> Background </h2>
```

```
<p> The baking competition was conceived by producer Anna Beattie.
```

Here is a table of executive producers:

```
  <table class='wikitable'>
    <tr>
      <td>Anne Beattie</td>
      <td>2010 - present</td>
    </tr>
    <tr>
      <td>Richard McKerrow</td>
      <td>2010</td>
    </tr>
    <tr>
      <td>Sarah Thompson-Woolley</td>
      <td>2016-2017</td>
    </tr>
  </table>
</p>
```

```
<p> Here is a table of judges:
```

```
  <table class='wikitable2'>
    <tr>
      <td>Paul Hollywood</td>
      <td>2010 - present</td>
    </tr>
    <tr>
      <td>Mary Berry</td>
```

```
      <td>2010 - 2016</td>
    </tr>
    <tr>
      <td>Prue Leith</td>
      <td>2017 - present</td>
    </tr>
  </table>
</p>
```

## Questions

1. Suppose that the url for this page is `https://fake_web_page.html`. Write code that will import the table of judges into R, as a data frame.