**Notes:**

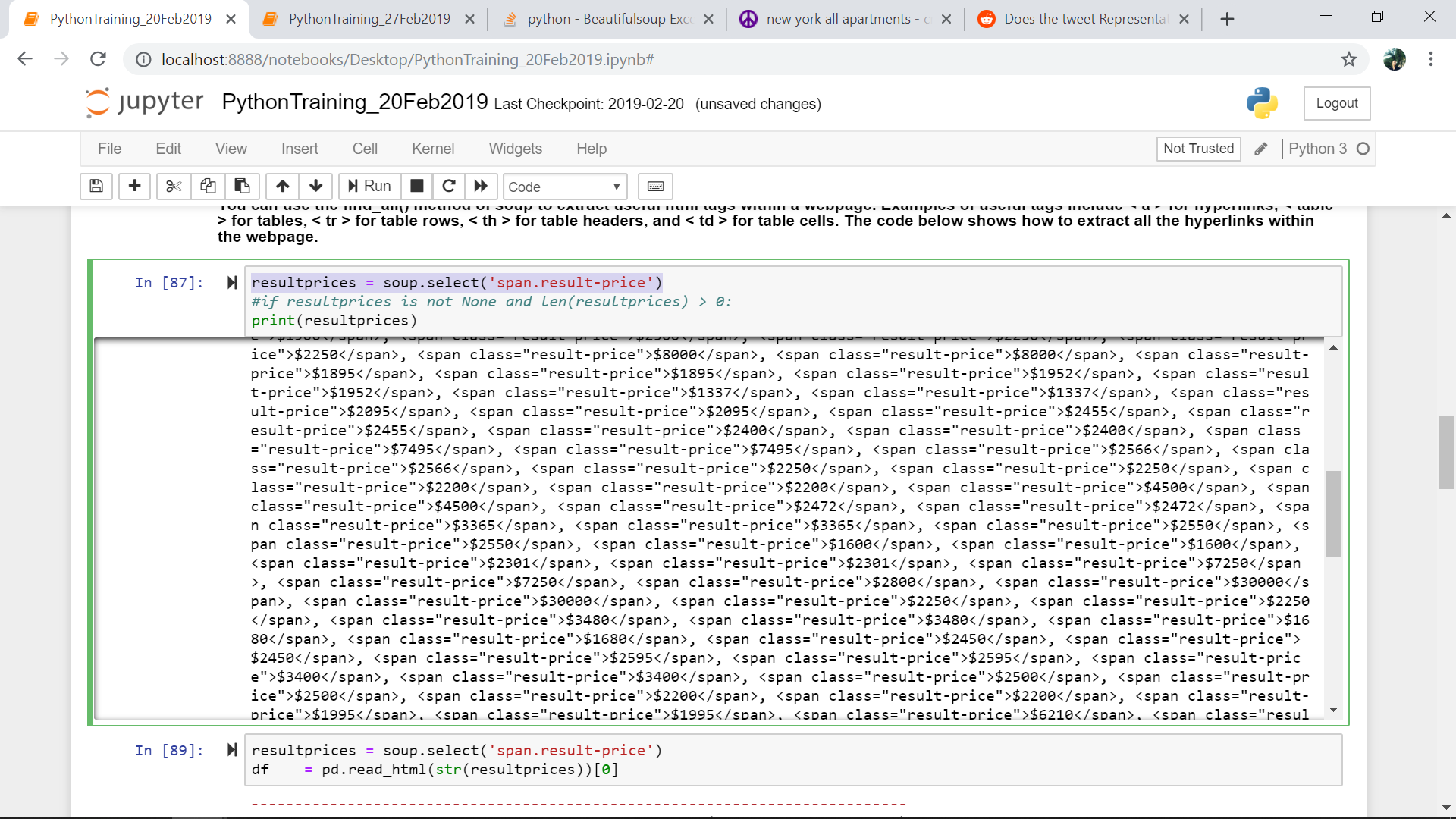
* Recap – Last week we scraped, interpreted and then visualized data from a bitcoin website.
* This week we are going to go to craigslist and collect data on houses: find the highest, lowest, and average price: <http://newyork.craigslist.org/search/aap>
* The first thing is identifying your data. Where is it in the page? Where does it show in the HTML code? This makes sure we know what we want to get from the website. Looking at the code of a page we can find the information

**´´Class is divided between between first-timers and those who went through last lesson´´**

1. We are going to go to <http://newyork.craigslist.org/search/aap>
2. Insert the new address in the code from last lesson (20th feb)
3. Run the information request.
4. Looked for the price data in the page code
   1. <span class="result-price">$4200</span>
   2. This is where the price is encoded in the page
5. Tried running the line in soup to find the price data: print(soup.find\_all('result-price')[0])
6. I get this error: IndexError: list index out of range
7. Stuck!

**\*I asked for advice\***

* soup.find\_all('result-price')
  + This was not the right soup function to find the price data, since this time we are not dealing with tables, but instead prices attached to house pictures.
* resultprices = soup.select('span.result-price')
  + here we define **variable** “resultprices” as what **soup.select function** finds on “span.result-price”
* print(resultprices)
  + This line of code then tells the computer to print (write out) the result of variable “resultprices”

**Here’s the result:**

* Now we can see our data, and we’ve taught the computer how to find the relevant lines on craigslist to find out data prices.
* Next step is ordering the data and doing some statistics! How can we can find highest, lowest and average? Have a look in Saulo’s document **PythonTraining\_27Feb2019...**