

Week 1 – 12/25/2017 – 01/01/2018

- Basic web scraping using python APIs.
 - Finding the best possible API to use for the data extraction.
 - The ones that I know as of now are
 - a.) BeautifulSoup Library
 - b.) Scrapy (Collaborative framework built on python for extracting data from websites)
 - With the help of chrome developer tools.
- Will first export the basic columns (the ones only with html tags without any stylesheets)
- Dataset_v1 should be ready by the end of the week (with the following columns - Brewery 1, Brewery 2, Brewery 3, Beer Name, Brewed by (brewer), Brewed in (location)).

Week 2 – 01/02/2018 – 01/09/2018

- Will export the rest of the columns – including the ones with CSS stylesheets (Website Rating, User Rating etc.).
- The next challenge would be to get the most recent review dates in ratebeer.com. Will try to sort the dates in the JavaScript itself and then get the least(oldest) one.
- Once we have a basic csv (or output) file ready, we can discuss further on the adding additional columns if needed.
- Dataset_v2 should be ready by the end of the week with all the columns that we have discussed.

Week 3 – 01/10/2018 – 01/17/2018

- Wrapping it up – we should be ready with the final exported file by the end of the week.
- Code review and the data correctness.