

DMDD CLASS NOTES

SEPTEMBER 29th, 2022

- Discussion on class quiz :
 - Showed a practice quiz and how we should attempt it.
 - Questions will be from the respective videos given in Canvas.
 - Students are free to share videos to TAs, if they find informative videos on Youtube.
 - Speed is important : 1 question has 1 minute.
 - Total 10 questions for 10 minutes.
- Data Scrapping
 - Abstract: Communication of Application and Database.
 - Opened a link and changed numbers on the link.. Which is implemented in changing the page numbers.
 - Parse a page and check. If you can parse, then you can parse multiple pages.
 - Logic is to parse a page, and then we want to turn the page from 20 to 30 till it stops.
 - A loop without end is a while loop.
 - BeautifulSoup → parses a web page. Parses **Document Object Model(DOM)**
 - Diff between internet and Web : SSH works on Internet.
 - The Internet is connecting packages from one end to another.
 - If I download something using university wifi.. It will go through Northeastern, then Engineers and then will come to my laptop.
 - HTTP :
 - Why did DOM and HTML become popular is because anyone could create it.
 - Discussed a professional webpage and checked the html on **NU HTML Checker** which showed all the errors and warnings.
 - ` .. ` it is a link tag.

- Remove all the tags and fetch the links : DOM is hierarchy of tags
- Also search for specific annotations like for example: how do we know from where it is coming.. So the link will start with the website name. For example : *www.indeed.com/bicau98hwh*
- Project → Job Database. :
 - If someone is facing issues with Python, contact the people who wanna offer help in the Jobs database.
- Showed a video on how to use BeautifulSoup using Python:
 - “Get all links from a website with Python and BeautifulSoup”
 - <https://www.youtube.com/watch?v=Jf3qinibzyg>
- Robots will scrap according to date.
- Suggestion for next Assignment :
 - Don't do double work.. If there is an assignment to scrap data then you can scrap the same data which you will be using for your project.
- Discussed about APIs
- Showed Another video :
 - Python MySQL Tutorial - Setup & Basic Queries (w/ MySQL Connector)
 - <https://www.youtube.com/watch?v=3vsC05rxZ8c>

*****End of Class*****