

NAME: SHASHANK NAGUMANTRI

ROLL NO: 19B21A05C3

E-MAIL ID: shashankn7261@gmail.com

COURSE: FAST TRACK PYTHON

Fast Track Python (Assignment 6)

1. Explain why selenium is important in web scraping.

A. Selenium is a python library and tool used for automating web browsers to do a number of tasks. One of such is web-scraping to extract useful data and information that may be otherwise unavailable. Selenium is an excellent automation tool and scrapy is by far the most robust web scraping framework. While dealing with JavaScript based websites where we need to make AJAX/PJAX requests, selenium can work better.

2. what the difference between scraping images and scraping websites? Use an example to demonstrate your point.

A.

SCRAPING IMAGES	SCRAPING WEBSITES
1.Scraping image is a subset of the web scraping technology.	1. Scraping web is basically extracting data from websites in an automated manner.
2.A technology in which a computer program data from output generated from another program.	2. Different types of web scrapers:1. self-built web scrapers, 2. Brower extensions web scrapers, 3. Cloud web scrapers.
3. Image scraping only focus on the media side – images, videos, audio, and so on.	3. Web scraping has multiple applications across various industries. Like; price monitoring, market research,

	news monitoring, sentiment, email marketing.
4. Web scraping and crawling aren't illegal by themselves.	4. Price monitoring, enriching machine learning models, financial data aggregation, monitoring consumer sentiment, etc.
5. For example, Instagram scraping, Facebook scraping, google scraping, and so on.	5. For example, you might scrape product information from an ecommerce website onto an excel spreadsheet.

3. Explain how MongoDB indexes data.

A. MongoDB uses multikey indexes to index the content stored in arrays. If you index a field that holds an array value, MongoDB creates separate index entries for every element of the array. These multikey indexes allow queries to select documents that contain arrays by matching on element of the array.

4. What is the significance of the \$SET modifier?

A. Set expressions are used to define the scope of a calculation. The central part of the set expression is the \$SET modifier that specifies a selection, this is used to modify the user selection, or the selection in the set identifier, and the result defines a new scope for the calculation.

5. Explain the MongoDB aggregation framework.

A. The aggregation framework is a set of analytics tools within MongoDB that allows us to run various types of reports or analysis on documents in one or more collections. Based on the idea of a pipeline.