**PROJECT-IDEA – Planning:**

* FAST WAY:
  + Go through Kaggle or Google datasets and skip the data collection stage.
* Collect own data: Data Scrapping
  + Think of a project problem
  + Brainstorming process
    - Glassdoor
      * Potential skills required
      * Career
      * Resume
      * Types of positions
        + Positions are different across the country

Salary varies based on

Job title

Company

Geography.

**DATA COLLECTION:**

* How to collect data and create git-hub repository.
* **Data Source** 
  + linkedIn
  + Glassdoor
* Data can be seen using the inspect tool
  + Class ‘span’ has the relevant data.
* **STEP – 1 :**
  + Initialize the git-repository
    - Go to git-bash command prompt and create a new project folder
    - Then inside the folder type, ‘git init’
      * Empty github repository will be created.
    - Now go to github and create a repository wit the folders name.
    - Later we will push our code to the git hub.
* **STEP – 2:** 
  + Data scrapping
    - Beautiful SOUP or Selenium
      * As glass door is JavaScript code, it is better to go with the selenium code as we can go to specific modules rather than grabbing the data of the complete webpage.
  + Data Cleaning
    - Create a separate branch for data cleaning
      * git checkout -b data\_cleaning