

Data Scientist (NLP)

1. A question and answers dataset (JSON file) can be downloaded from the link below. It contains 307,373 questions and answers.
 - a. Click [Here](#) to download.
2. Clean the dataset and prepare a CSV file that contains `document_title`, `question_text`, `short_answer`, `yes_no_answer` columns.
 - a. If you find two or more values for `short_answer` column, you can name them as `short_answer1`, `short_answer2` and so on.
3. Build a Gradio or Streamlit app to explore the cleaned dataset with some basic filter options.
4. Maintain all source files in any VCS.