Assignment Title: PDF Content Segmenter with REST API in Java

Objective:

Develop a Java application that programmatically segments a system-generated PDF into distinct sections based on the whitespace between blocks of text, making exactly X cuts. The application should expose a REST API to accept PDF files, segment them, retrieve metadata, update segmentation details, and delete processed PDFs. The goal is to identify logical sections such as headings, paragraphs, and distinct blocks that are visually separated by increased whitespace, without using image processing techniques.

Requirements:

Input Specifications:

- The application should accept a PDF file as input via a REST API.
- The user should specify the number of cuts (X) to be made.

Processing Details:

- Analyze the PDF to identify sudden changes in whitespaces along the Y-axis that significantly separate text blocks.
- Define "significant whitespace" as vertical spaces that are noticeably larger than the typical line spacing within paragraphs.
- The application should make exactly X cuts based on the largest Y whitespace gaps.

Output Specifications:

- The output should be multiple PDF files, each containing one of the segments created by the cuts, returned via the REST API.
- Ensure that the segmentation does not cut through the middle of text blocks or paragraphs.

Constraints:

- Do not use image processing libraries or convert PDF pages to images for processing.
- Leverage Java PDF manipulation libraries such as Apache PDFBox or iText to analyze and manipulate the PDF content.

Technologies to Use:

- Java 11 or above
- Spring Boot for REST API development
- Maven or Gradle for build management
- Apache PDFBox or iText for PDF processing