This is a PDF file to extract the text