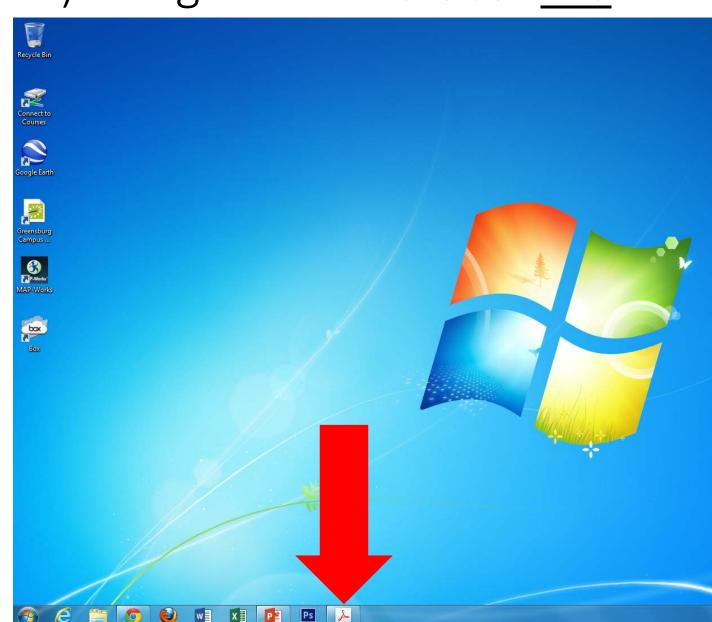
DH/DS OCR Homework Assignment

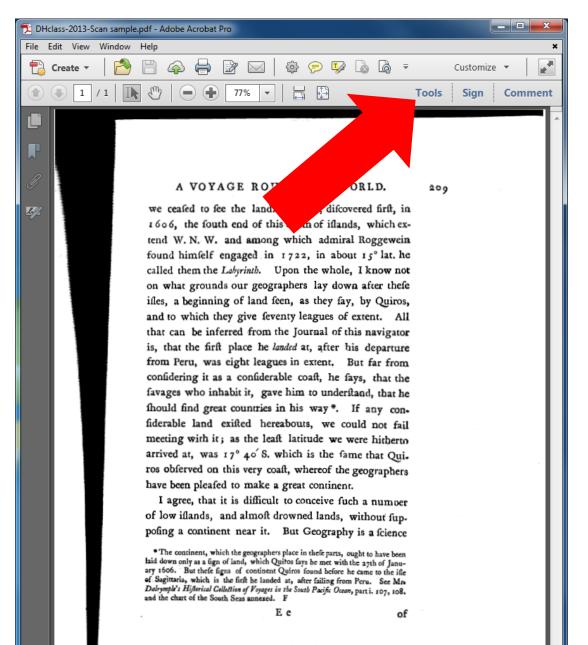
Due Monday Sept 9

1) Download and Open both of the PDF files on the Syllabus web page (or in Box) using Adobe Acrobat Pro

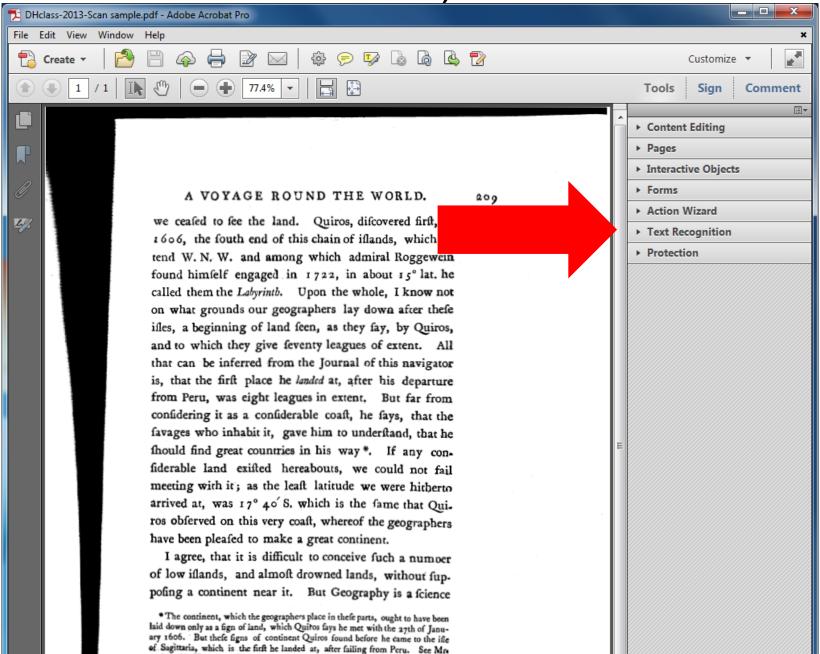
(not Reader)



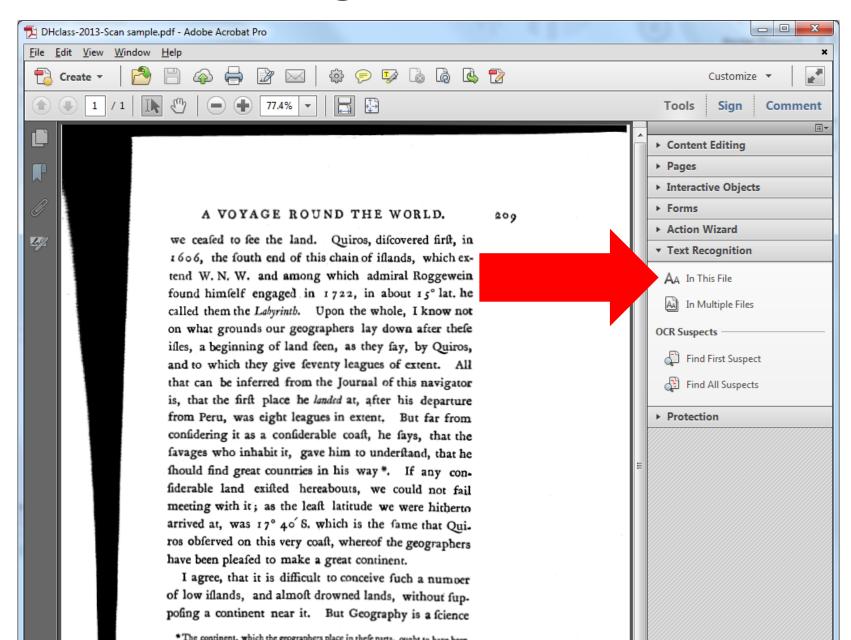
2) In Adobe Acrobat Pro, click on the Tools menu



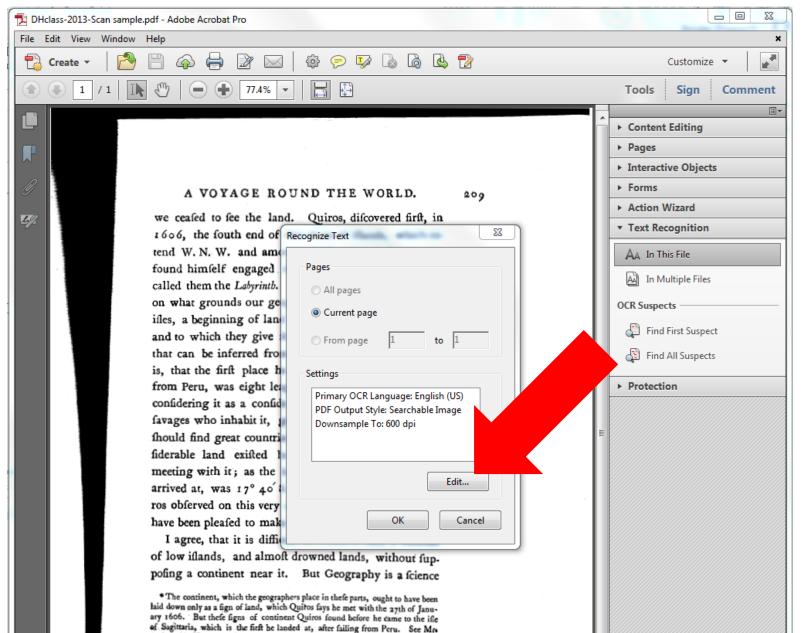
3) Under the Tools menu, select Text Recognition



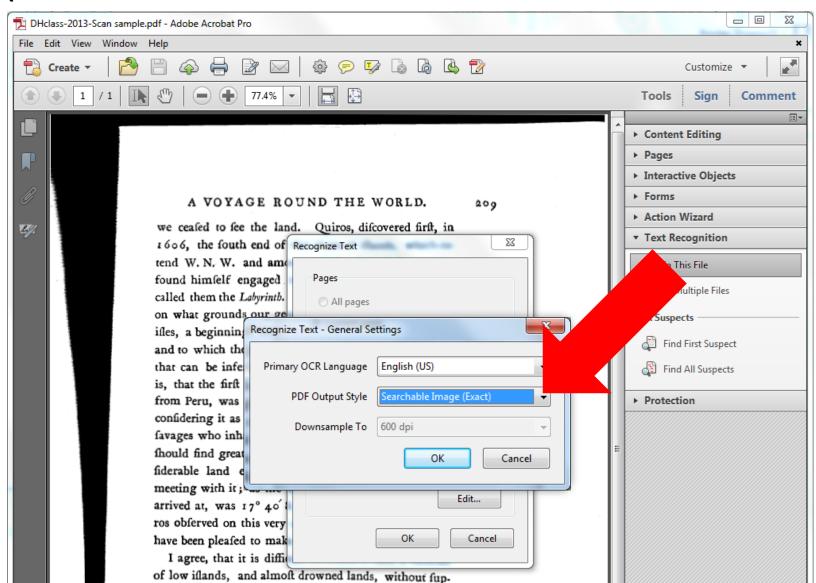
4) Under Text Recognition, select In This File



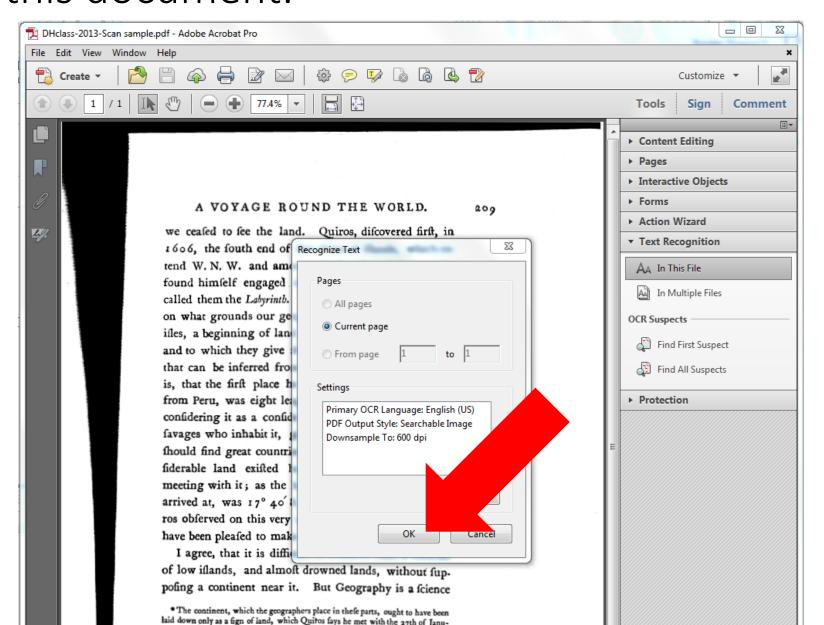
5) In the Recognize Text dialog box, click on Edit



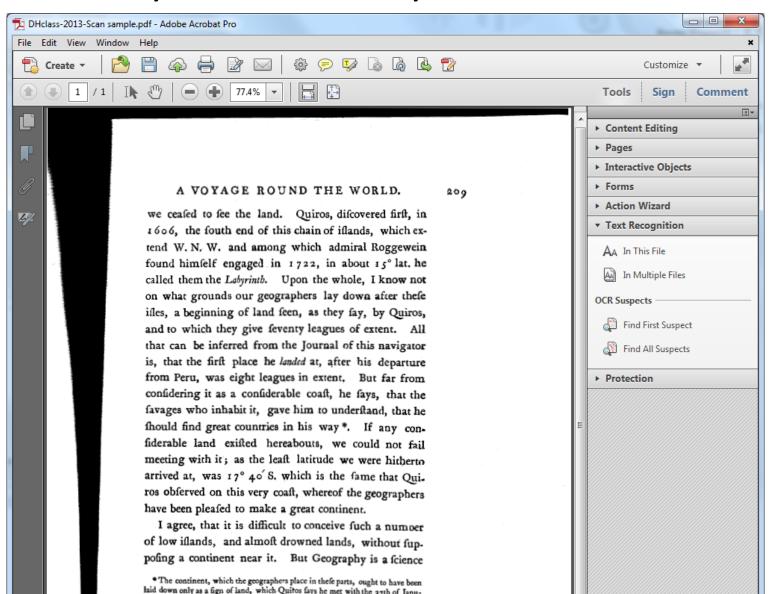
- 6) In the Recognize Text General Settings dialog box:
 - Change PDF Output Style to Searchable Image (Exact)
 - Click OK



7) Click OK on the Recognize Text dialog box to perform OCR on this document.



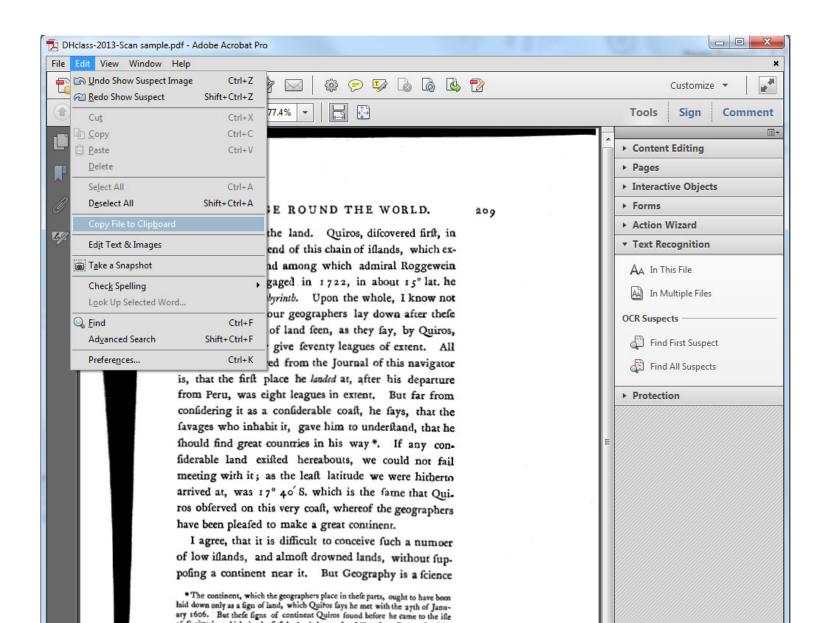
8) The document now includes formatted text, though probably with many errors.



- 9) Under Text Recognition, select Find All Suspects
 The words in red are the words that Acrobat SUSPECTS might
- be incorrect. - 0 DHclass-2013-Scan sample.pdf - Adobe Acrobat Pro File Edit View Window Help Customize Create * 77.4% Tools Comment ▶ Content Editing Pages ► Interactive Objects ▶ Forms A VOYAGE ROUND THE WORLD. 200 ▶ Action Wizard we ceased to see the land. Quiros, discovered first, in ▼ Text Recognition 1606, the fouth end of this chain of islands, which extend W. N. W. and among which admiral Roggewein AA In This File found himself engaged in 1722, in about 15° lat. he In Multiple Files called them the Labyrintb. Upon the whole, I know not on what grounds our geographers lay down after these **OCR Suspects** isles, a beginning of land seen, as they say, by Quiros, Find First Suspect and to which they give feventy leagues of extent. All that can be inferred from the Journal of this navigator Find All Suspects is, that the first place he landed at, after his departure from Peru, was eight leagues in extent. But far from ▶ Protection confidering it as a confiderable coast, he says, that the favages who inhabit it, gave him to understand, that he fhould find great countries in his way *. If any confiderable land existed hereabouts, we could not fail meeting with it; as the least latitude we were hitherto arrived at, was 17° 40' S. which is the fame that Quiros observed on this very coast, whereof the geographers

have been pleafed to make a great continent.

10) Under the Edit menu, select Copy File to Clipboard



- 11) ...and paste into a New Document in MS-Word to see how Adobe Acrobat Pro actually read the PDF image 12) Repeat for the second PDF image file that you downloaded
 - In a brief one-page response paper due before class on Monday, Sept 9:
 - Discuss at least FIVE characteristics of the original text or PDF image documents that produced errors in the OCR process.
 - Include the OCR'd text from your MS-Word files as a separate pages with your response paper.
 - Please always remember to include your name in your submitted filename.