Report for midterm

This midterm uses what we have learned since the beginning of the semester in regards to spark. This assignment is very similar to the last homework which did some of the same functions. It takes in a text file (user is prompted for the file name), creates a flat map of each line in the file, then makes a map of all the characters in those lines, reduces by key to consolidate the character frequency count, and sorts that frequency list in order. This uses the first entry in the list as the most common letter in the encrypted text.

The nest step is to find out the most common letter of the alphabet and check for its instances. The most common letter in the English alphabet is the letter E so the code will look for e then everything else afterwards.

Next section uses lambda in the sense that we must have every letter in the encrypted text to shift accordingly while special characters and numbers are left out in the sort. The shifted text is what we are after so it is printed out afterwards. Using a sample of 5 percent of words from the shifted text are checked to see if they are valid English words. The code checks for a accuracy rate of 75 percent then senses that it is completely decrypted then the percentage is printed. The user has an option to bypass the decryption if the cipher is deemed already decrypted. Lastly The code also will check google for the sources after the cipher is decrypted and saved. We can check our output in the saved file labeled decrypted.txt.

Sources:

Encrypted-1 - https://studfile.net/preview/393909/page:78/

Encrypted-2 - https://books.google.com/books?id=7x9yw5DQIOUC&pg=PA92&lpg=PA92&dq=times+of+peace,+are+of+every-day+occurrence+between+contending+armies&source=bl&ots=DkxSINdPlO&sig=ACfU3U2CWQf9gBcVwRpBMFrSjgq3QEDIHg&hl=en&sa=X&ved=2ahUKEwiFkbK53-7nAhXySt8KHSXZBa4Q6AEwAHoECAQQAQ#v=onepage&q=times%20of%20peace%2C%20are%20of%20every-day%20occurrence%20between%20contending%20armies&f=false

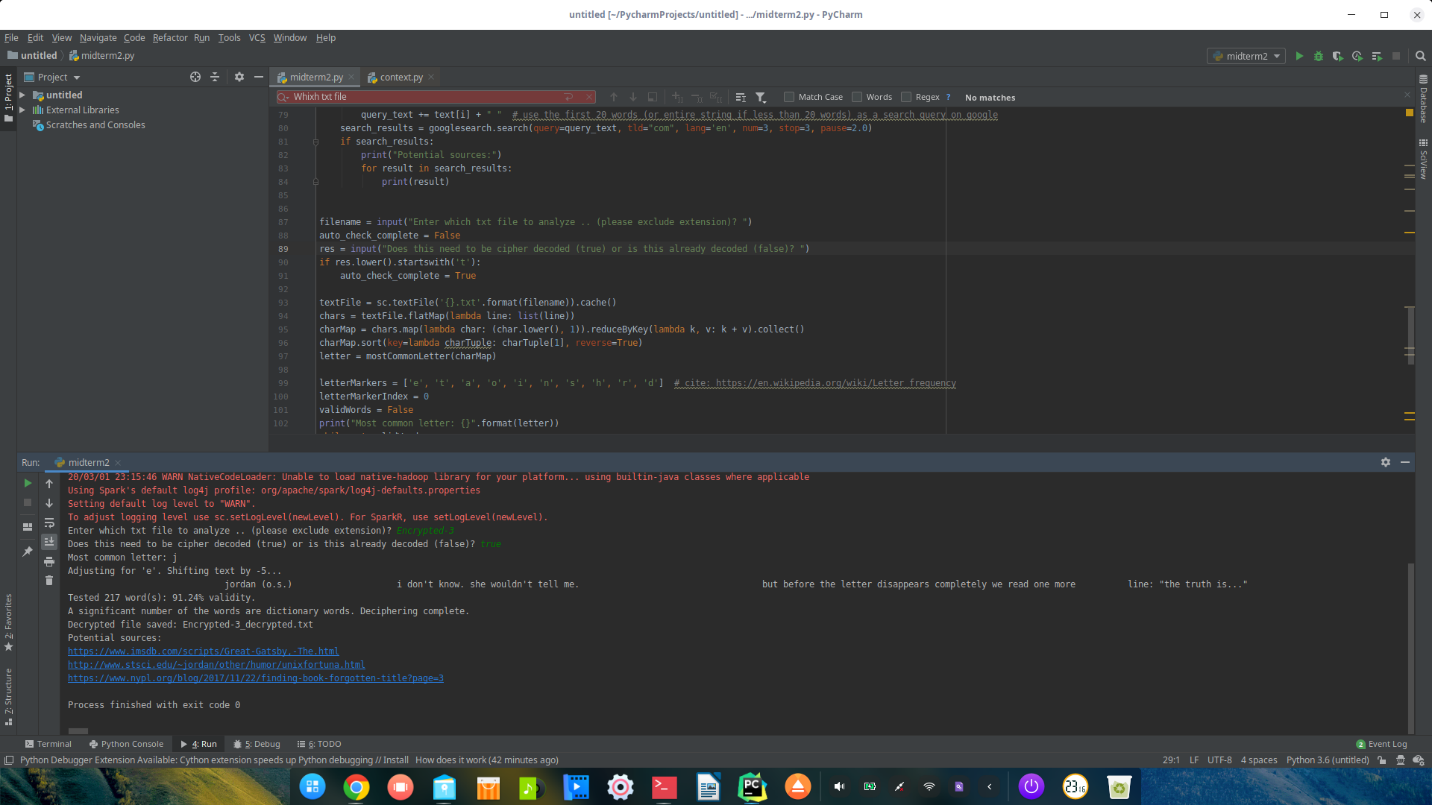
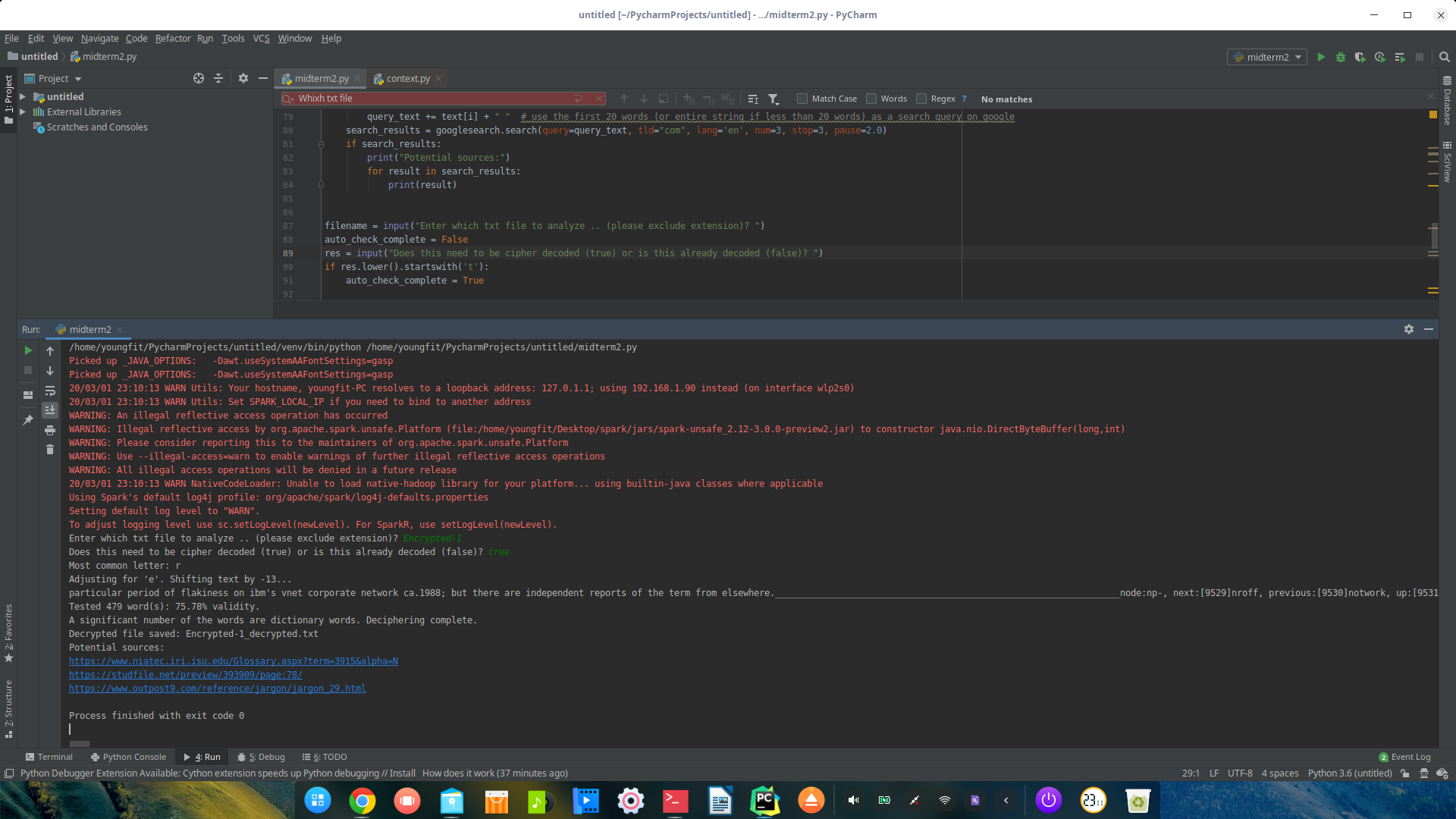
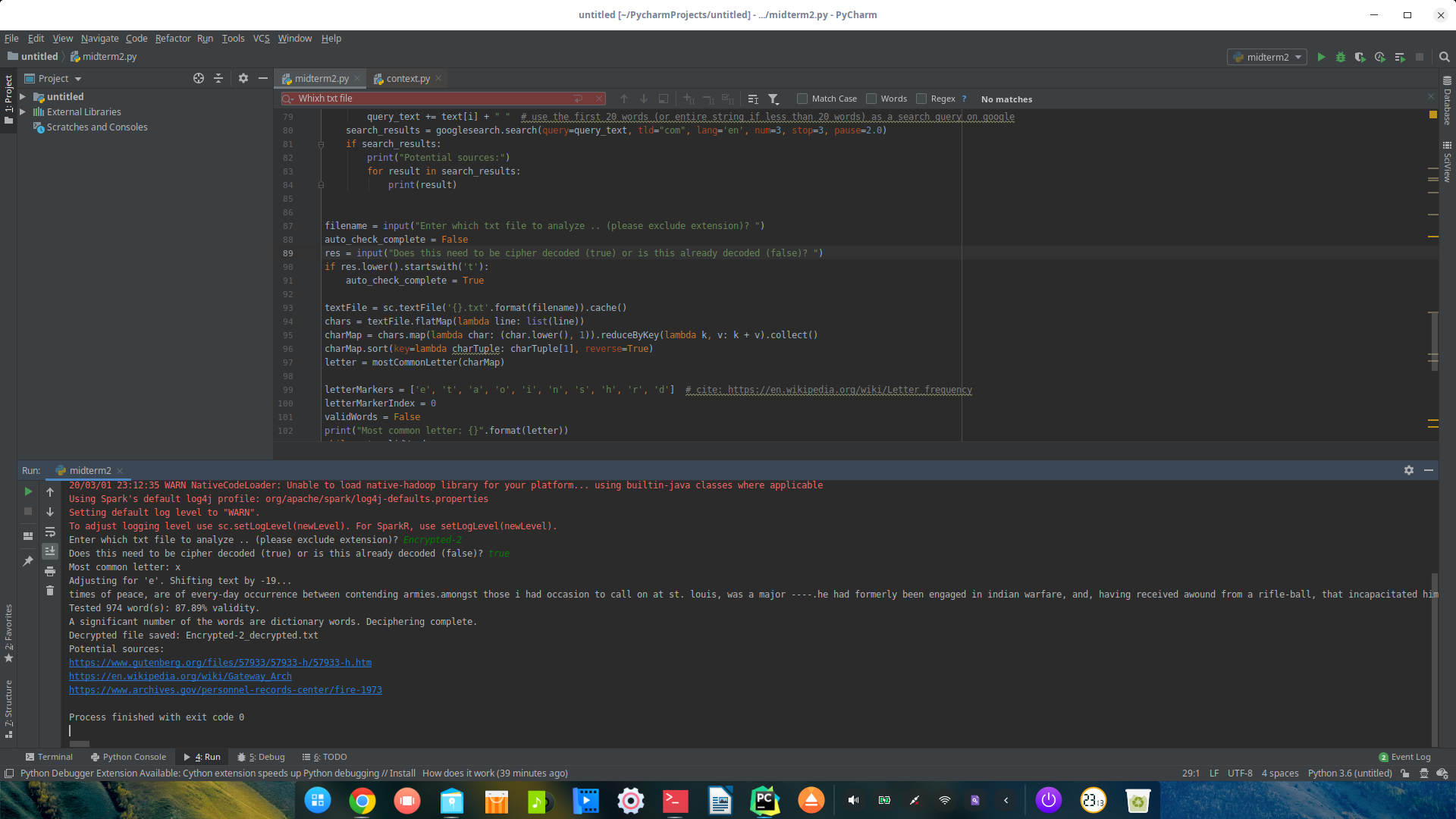
Encrypted-3 - https://www.imsdb.com/scripts/Great-Gatsby,-The.html

Screenshots of Code:





Screenshots of terminal:



Screenshots of Immediate results:

When a user is prompted if it is deciphered or not

