Digital Forensics Assignment

Chythanya Gudla

a) Chythanya Gudla

b)

Algorithm:

- The MD5 checksum value of the corrupted file is printed first.
- Using the file handling concepts from tuitorialspoint.com I have opened the correpted file and decoded using base64.base64decode standard function
- The decoded base64 string is now converted into byte array and passed to a standard function “binascii.b2a_hex” inorder to convert it into hex.
- Now the magic numbers comes into picture. Every jpeg,png,pdf,gif,docx file has a constant header and footer respectively. Using the headers and footers the required content is trimmed using “trimmed_file=deco_string[initialindex:finalindex+tail_length]”
- Now the obtained hex trimmed string is converted to ascii
- This is the final output which we require
- Now the MD5 checksum values of all the files are printed and verified to be same.
- The extra file which is not mentioned in the problem statement is the .docx file(“BINGO”)
- Now all the output files are created in the desired location using “output_file = open("C:\Users\Chythanya\Desktop\output.docx", "wb")
  output_file.write(ascii_string)”

c) I have downloaded Anaconda software which consists of all the required packages required for python. I have done my code in ipython notebook which is available in anaconda software. We just need to open the new python shell in ipython notebook, open the ipython code, provide the local corrupted file path in the code “file = open("C:\Users\Chythanya\Desktop\cc.docx", "r")”. Provide the path where output file have to be created “output_file = open("C:\Users\Chythanya\Desktop\output.png", "wb")” . Run the shell the MD5 hash values for the corrupted and the output files will be printed and the output files will be created in the provided path.