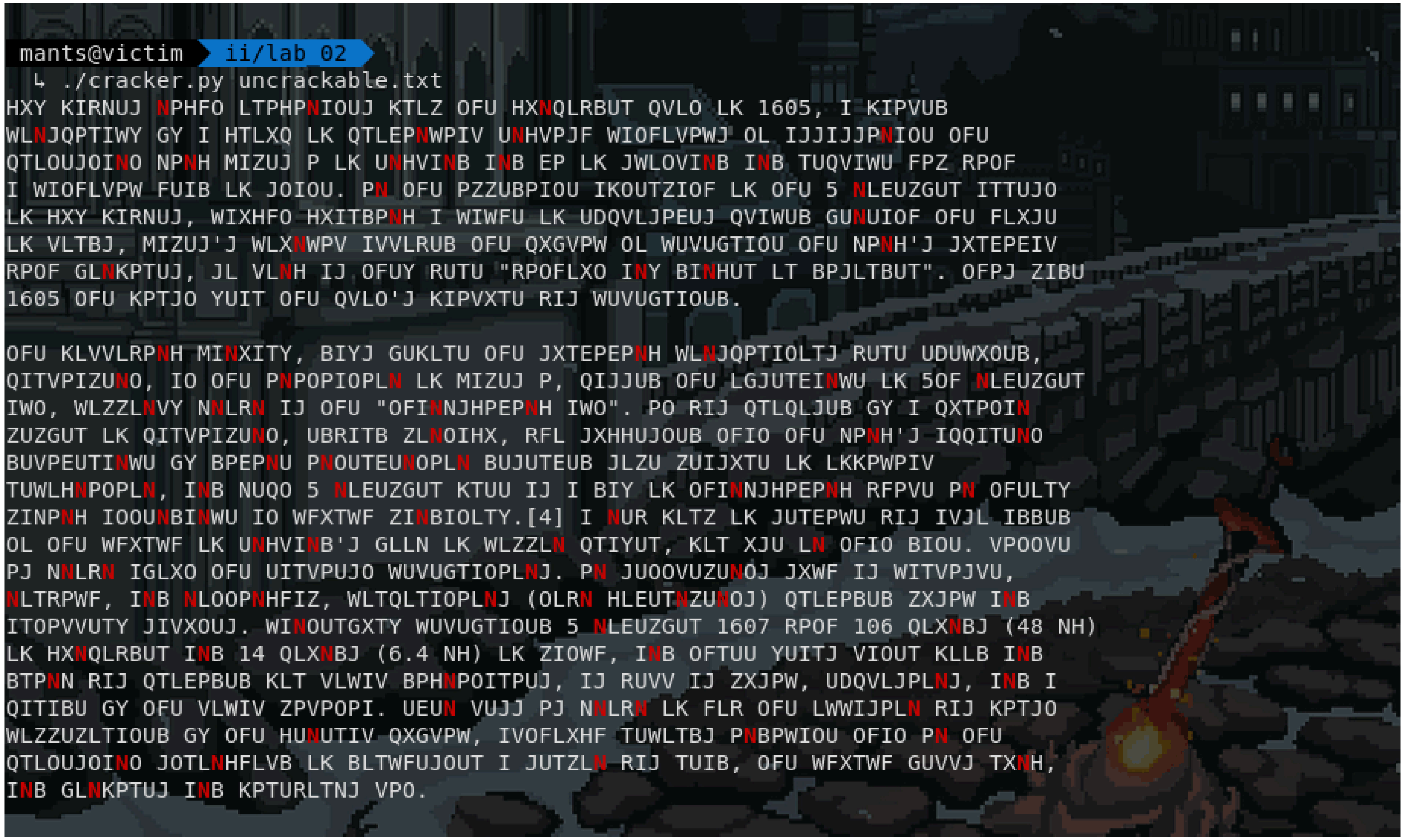
03. [50p] Solving a substitution cipher



```
mants@victim  ii/lab 02
  ↳ ./cracker.py uncrackable.txt
HXY KIRNUJ NPHFO LTPHPNIOUJ KTLZ OFU HXNQLRBUT QVLO LK 1605, I KIPVUB
WLNJQPTIWY GY I HTLXQ LK QTLEPNWPIV UNHVPJF WIOFLVPWJ OL IJJIJJPNIOU OFU
QTLOUJOINO NPNH MIZUJ P LK UNHVINB INB EP LK JWLOVINB INB TUQVIWU FPZ RPOF
I WIOFLVPW FUIB LK JOIOU. PN OFU PZZUBPIOU IKOUTZIOF LK OFU 5 NLEUZGUT ITTUJO
LK HXY KIRNUJ, WIXHFO HXITBPNH I WIWFU LK UDQVLJPEUJ QVIWUB GUNUIOF OFU FLXJU
LK VLTBJ, MIZUJ'J WLXNWPV IVVLRUB OFU QXGVPW OL WUVUGTIOU OFU NPNH'J JXTEPEIV
RPOF GLNKPTUJ, JL VLNH IJ OFUY RUTU "RPOFLXO INY BINHUT LT BPJLTBUT". OFPJ ZIBU
1605 OFU KPTJO YUIT OFU QVLO'J KIPVXTU RIJ WUVUGTIOUB.

OFU KLVVLRPNH MINXITY, BIYJ GUKLTU OFU JXTEPEPNH WLNJQPTIOLTJ RUTU UDUWXOUB,
QITVPIZUNO, IO OFU PNPOPIOPLN LK MIZUJ P, QIJJUB OFU LGJUTEINWU LK 5OF NLEUZGUT
IWO, WLZZLNVY NNLRN IJ OFU "OFINNJHPEPNH IWO". PO RIJ QTLQLJUB GY I QXTPOIN
ZUZGUT LK QITVPIZUNO, UBRITB ZLNOIHX, RFL JXHHUJOUB OFIO OFU NPNH'J IQQITUNO
BUVPEUTINWU GY BPEPNU PNOUTEUNOPLN BUJUTEUB JLZU ZUIJXTU LK LKKPWPIV
TUWLHNPOPLN, INB NUQO 5 NLEUZGUT KTUU IJ I BIY LK OFINNJHPEPNH RFPVU PN OFULTY
ZINPNH IOOUNBINWU IO WFXTWF ZINBIOLTY.[4] I NUR KLTZ LK JUTEPWU RIJ IVJL IBBUB
OL OFU WFXTWF LK UNHVINB'J GLLN LK WLZZLN QTIYUT, KLT XJU LN OFIO BIOU. VPOOVU
PJ NNLRN IGLXO OFU UITVPUJO WUVUGTIOPLNJ. PN JUOOVUZUNOJ JXWF IJ WITVPJVU,
NLTRPWF, INB NLOOPNHFIZ, WLTQLTIOPLNJ (OLRN HLEUTNZUNOJ) QTLEPBUB ZXJPW INB
ITOPVVUTY JIVXOUJ. WINOUTGXTY WUVUGTIOUB 5 NLEUZGUT 1607 RPOF 106 QLXNBJ (48 NH)
LK HXNQLRBUT INB 14 QLXNBJ (6.4 NH) LK ZIOWF, INB OFTUU YUITJ VIOUT KLLB INB
BTPNN RIJ QTLEPBUB KLT VLWIV BPHNPOITPUJ, IJ RUVV IJ ZXJPW, UDQVLJPLNJ, INB I
QITIBU GY OFU VLWIV ZPVPOPI. UEUN VUJJ PJ NNLRN LK FLR OFU LWWIJPLN RIJ KPTJO
WLZZUZLTIOUB GY OFU HUNUTIV QXGVPW, IVOFLXHF TUWLTBJ PNBPWIOU OFIO PN OFU
QTLOUJOINO JOTLNHFLVB LK BLTWFUJOUT I JUTZLN RIJ TUIB, OFU WFXTWF GUVVJ TXNH,
INB GLNKPTUJ INB KPTURLTNJ VPO.
```

Time to get your hands dirty! Below, you have a ciphertext [https://en.wikipedia.org/wiki/Ciphertext]. Specifically, this output is the result of a substitution cipher [https://en.wikipedia.org/wiki/Substitution_cipher], meaning that every letter in the English alphabet has been assigned a random, unique correspondent. As you may have noticed, digits and special characters remain unchanged.

```
HXY KIRNUJ SPHFO LTPHPSIOUJ KTLZ OFU HXSQLRBUT QVLO LK 1605, I KIPVUB
WLSJQPTIWY GY I HTLXQ LK QTLEPSWPIV USHVPJF WIOFLVPWJ OL IJJIJJPSIOU OFU
QTLOUJOISO NPSH MIZUJ P LK USHVISB ISB EP LK JWLOVISB ISB TUQVIWU FPZ RPOF
I WIOFLVPW FUIB LK JOIOU. PS OFU PZZUBPIOU IKOUTZIOF LK OFU 5 SLEUZGUT ITTUJO
LK HXY KIRNUJ, WIXHFO HXITBPSH I WIWFU LK UDQVLJPEUJ QVIWUB GUSUIOF OFU FLXJU
LK VLTBJ, MIZUJ'J WLXSWPV IVVLRUB OFU QXGVPW OL WUVUGTIOU OFU NPSH'J JXTEPEIV
RPOF GLSKPTUJ, JL VLSH IJ OFUY RUTU "RPOFLXO ISY BISHUT LT BPJLTBUT". OFPJ ZIBU
1605 OFU KPTJO YUIT OFU QVLO'J KIPVXTU RIJ WUVUGTIOUB.

OFU KLVVLRPSH MISXITY, BIYJ GUKLTU OFU JXTEPEPSH WLSJQPTIOLTJ RUTU UDUWXOUB,
QITVPIZUSO, IO OFU PSPOPIOPLS LK MIZUJ P, QIJJUB OFU LGJUTEISWU LK 5OF SLEUZGUT
```

```
IWO, WLZZLSVY NSLRS IJ OFU "OFISNJHPEPSH IWO". PO RIJ QTLQLJUB GY I QXTPOIS
ZUZGUT LK QITVPIZUSO, UBRITB ZLSOIHX, RFL JXHHUJOUB OFIO OFU NPSH'J IQQITUSO
BUVPEUTISWU GY BPEPSU PSOUTEUSOPLS BUJUTEUB JLZU ZUIJXTU LK LKKPWPIV
TUWLHSPOPLS, ISB NUQO 5 SLEUZGUT KTUU IJ I BIY LK OFISNJHPEPSH RFPVU PS OFULTY
ZINPSH IOOUSBISWU IO WFXTWF ZISBIOLTY.[4] I SUR KLTZ LK JUTEPWU RIJ IVJL IBBUB
OL OFU WFXTWF LK USHVISB'J GLLN LK WLZZLS QTIYUT, KLT XJU LS OFIO BIOU. VPOOVU
PJ NSLRS IGLXO OFU UITVPUJO WUVUGTIOPLSJ. PS JUOOVUZUSOJ JXWF IJ WITVPJVU,
SLTRPWF, ISB SLOOPSHFIZ, WLTQLTIOPLSJ (OLRS HLEUTSZUSOJ) QTLEPBUB ZXJPW ISB
ITOPVVUTY JIVXOUJ. WISOUTGXTY WUVUGTIOUB 5 SLEUZGUT 1607 RPOF 106 QLXSBJ (48 NH)
LK HXSQLRBUT ISB 14 QLXSBJ (6.4 NH) LK ZIOWF, ISB OFTUU YUITJ VIOUT KLLB ISB
BTPSN RIJ QTLEPBUB KLT VLWIV BPHSPOITPUJ, IJ RUVV IJ ZXJPW, UDQVLJPLSJ, ISB I
QITIBU GY OFU VLWIV ZPVPOPI. UEUS VUJJ PJ NSLRS LK FLR OFU LWWIJPLS RIJ KPTJO
WLZZUZLTIOUB GY OFU HUSUTIV QXGVPW, IVOFLXHF TUWLTBJ PSBPWIOU OFIO PS OFU
QTLOUJOISO JOTLSHFLVB LK BLTWFUJOUT I JUTZLS RIJ TUIB, OFU WFXTWF GUVVJ TXSH,
ISB GLSKPTUJ ISB KPTURLTNJ VPO.
```

Your task is to write a *Python* script that will help you break the cipher and decode the original text:

- read [https://www.pythontutorial.net/python-basics/python-read-text-file/] the ciphertext from a file specified as a command line argument
- use a dictionary to map each encoded character back to it's original value
- manually populate this dictionary as you progress in your attempt and reveal new characters
- whenever you run the script, it should print the text to the screen, with a few minor changes:
  - any character that exists as a key in the dictionary should be replaced with what you think the correspondent is.
  - any replaced character should be highlighted in bold red.

Remember ANSI codes?

```
$ echo "\033[1;34m I'm blue, da ba dee, dabba daa-ee, dabba dee-a dabba da \033[0m"
```

In breaking a short substitution cipher like this while also knowing the original language, you need to look at bigrams and trigrams. Small groups of letters that have a limited amount of possible values that make sense: *“to”*, *“and”*, *“the”*, etc. As you reveal more and more of the original text, words will begin to form, making everything progressively easier.

If you need an extra hint:

> *“5 SLEUZGUT”* looks like a date. Hmm… ***“SLEUZGUT”***…

Already done? Try this challenge [https://ctflearn.com/challenge/238] as well.

There are a lot more out there for you to find!