

# Exercises: Vigenère Cipher and Kasiski Attacks

## Spreadsheet Exercise: Vigenère

To begin with, download the Excel spreadsheet: `02a-vigenere_kasiski.xlsx`. We will use some functions from last week to create an encryption tool for the Vigenère cipher and later use a premade tool to aid Kasiski attacks on encrypted text.

1. Make sure you are in the worksheet named 'Vigenere Cipher'.
2. In Cell B1, enter a plaintext string in **lower case** and **without spaces** or leave the text in that is already there.
3. In Cell B2 there is a passage of plaintext. Leave this for now.
4. In Row 4 we will use the MID function as before to split the plaintext characters out one by one. Enter `=MID($B$1,B3,1)` into Cell B4 to retrieve the first letter. Then copy or drag this formula across to the end of the row. (Remember that copying the formula and then holding **Ctrl-Shift-Right Arrow** will highlight the rest of the row before pasting.)
5. (Optional) Row 5 was already prefilled. Can you see what the formula from Cell C5 onwards is doing?
6. In Row 6 use the MID function again to split out the characters of the key using the positions in Row 5 above it. Remember to copy this to the end of the row so that it repeats the key word enough times.
7. We will use the CODE function to turn the letters in Row 4 and 6 into numbers. Since `=CODE("a")=97` we will need to adjust things a bit here. Enter `=CODE(B4)-97` into Cell B7 and copy this across the row. Do the same for the key letters in Cell B8 and copy this across too - you will need to point the CODE function at the key letters in Row 6, not the plaintext letters in Row 4.
8. In the 'Shifted' row, Row 9, we will use modular arithmetic to shift the plaintext letters by each key letter value. In Cell B9 enter `=MOD(B7+B8,26)+97`, remembering that we need the +97 to shift the value back into the right range.

9. We'll now turn our numbers back into ciphertext letters using the UPPER and CHAR functions. Enter `'=IFERROR(UPPER(CHAR(B9)), "")'` in Cell B10. The UPPER function forces our ciphertext to be upper case.
10. Use the CONCAT function to stick the ciphertext letters together in Cell B11.
11. Check that your spreadsheet works by seeing if the plaintext message 'introductiontoCyberSecurityMathematicsandCryptography' with key 'shannon' encrypts to 'AUTEB RHUAI BAHBU FBREG RUBRV GMZSA HRZOG AJSNA RPJFP GBUES WHL' (without spaces).
12. (Optional) What would happen if somebody entered the plaintext or the keyword with upper case letters? Can you use the LOWER function to account for this?

## Exercises

1. Within the spreadsheet linked above there is a second worksheet called 'Kasiski Analysis'.
2. You do not need to create anything in this spreadsheet and can just use it as a tool.
3. There is an encrypted message in Cell B1, the full ciphertext is copied below. Inspect the rest of the spreadsheet to try and get an idea of what it will do.

FUIKBYMACFPJSLGSTGBUEKHRFPGUIHKCBVRGAUHCJTJEWKFKTVIDGHZEMIEIW  
FYGLCQKARGMVOPWRCFKGFXIPIJJWFGVHVHWJSVATVHZAVIKQUUTRVFHMGTY  
CHAHCDSGHLETSKQDASVAIVSVURAEFTDUPGKJSTQKZPHGTJEWKFWPNATGHZEX  
OCWAWSVRLEYLHGBRTFWDIRRVSSNFFVNZGPGNFPHZEJERTHZIPTYGRTARGLZIVL  
IHIYCRFOVCRWUZTCGCKAHSGOWVWOREEKBYWQRUUWFTJEJGQGNFATVWKHQUCF  
BWVGRYCJWFKNZUVWDKTSWHSSKSKQCHEFTFRWUKKTLRAQAEJDSUAOEIKJWTGD  
KQHZEQPVPDSSGAEFKATJATMGFVEITCJOTPTVTVSPUIKYOKOHJFAGGPQIXPOFT  
VHRVWVUHFVTSVIPMGFQNGRMGWKNCTTJSVTJEKJWFGQUKQTLHGCFZKAPDTTS  
HTUHRMWFGVODAPWDTOFKZETEZTSSDKTRPRJETERFWLAPDNGDLAPDCCYHGD  
PRLRGMSNSVWKTYCVGRTOIYVACJAKVWEEUJUOALUMVASLTJIKGLHGTYKBYTJ  
AKVFGUDLVUAWFQRZEOFNQWQFYEVCRITQGSWYGFBNATMGLATSYCBIPTYGVW  
AXEEUKZETEKJSKHCDFYGGFOEEUHZOWGYVGDEPGKJSFIPTYGOXTGREQCFWJEEV  
VWTYIEUIFSUIEMWFTQTYGZSKGOWJODICNUOMEIPDNKZDBGAIHCJEXEIVVWMGM  
FTMGFVHVRDLKDDCGCIRRRAUGDYICNQMRUEKJSORKTVTOKTJENTWLETHRUQMR  
UEUVVWWQRCFKATJTYKGTECUKKTMLUTLRSFDQUJEFWAVIFPHWRTISNSANKTJUW  
EPNITKHQITRVUWKTKBCGWFIVSKTILHCWFTZVWJITJBGWVRVOPDEUBVHCJEVHV  
MWFKNPGZDOYWYGBLHGFIBUHIOMGFFMGNKUSAZGDKJSLRCNJNOLEFCFRWWSY  
HZEZVAFJLUHSRTIMGRANRAIKGDOPDFPCXCQUIUSTEEADGSSGGRKQFWAFIKKHA  
SYECNYFOYNYQKLHGBFQYKPTERFZAKGAEBXEETZQIKDKSVCGWFTODEWLYVOTK  
HQFTODECFTKNVPHLOEOEVWFEPSTCFJEFOLVVWRGCFPTASEAKGRLHGRVFSFOWN  
TGRTYRRVUGSNFPLNDATEEEUIJEFEMGBTYVHVOCKTCDMCBUEFOWNWLETAIAOFA  
TCYKGLSPOUGTANKTVRFANEIGNSKHCDSGSFVKOCCHWDKNKJCKEYITMSVPCGVUB  
GDQCKTWFERFOIDGCTVFBGCQNMKQLIQNJQILRCGVFWLCCUCFBGTDEAWRYEFBP  
CBQKPONPGLAPDRTRQEVACVVGUIHZVKSSCCBPCOLGDXGRLHCTKJSKURRVOSFOV

EFHOJTJAUDSWNUTIWQCIPTYGYANIEASDLQWRNZXENTKJOLHWMRPBSTWRVECM  
 LFNFPWATTYGGRLRCIEPCJTRZXSGNYOIFGANYHZEVLHGEJUSFCGOWRIJEUTGQ  
 WKOPLLTYWDVHVXSJYDAECZATAAEFWFNQCV PQWOHTYGTARUTREHGNNYRNZGWGD  
 KJSTLQWKQTSLNWVSJWCRUYWLHOOIGOOFWLVHTWCV

- When you have had a look around the sheet, use the following table to start implementing a Kasiski attack on the ciphertext.

sequence	interval	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
TJEW	126	X	X			X	X		X					X			
KZET	168	X	X	X		X	X	X				X		X			
IPTY	434		X				X										
QMRU	28	X		X			X							X			
EWKF	126	X	X			X	X		X					X			
WFGV	406	X					X							X			

- First guess what the most probable key length is from studying the table (it is no bigger than 10). Enter your guess in Cell C2. You should find that the spreadsheet splits the ciphertext out into as many groups as your key length guess.
- Starting in Cell A18, the frequency of the letters in each group are counted and compared with the standard frequencies. This first group of letters is shifted by a Caesar cipher. When you think you have guessed how far the group is shifted, enter this guess in Cell AC22. For the first two groups there are bar charts to help you spot the patterns, and also a chart for the standard frequencies. Remember that if you think that shifts are one out from what you expect, so if C is the most common letter in this group, then the shift is 2, not 3. There is a reference for the shift values in Row 17.
- Continue to do this for each of the groups. As you enter guesses for how each group is shifted, the spreadsheet will update below Row 88. Keep an eye on this to see if your guesses start to reveal anything sensible.
- Once you have deciphered the text, try to identify the source.
- Repeat the task for the ciphertext below, using the table to help start building a Kasiski attack. Downloading this as a text file makes this easier as it gives you the text without any spaces. (This turns out to be much harder if you copy the text below with spaces in it, as it throws off the frequencies of each letter.) Again, try to identify the source of the plaintext.

ZRPSLTZRNAZHYMTVKHCINFPASMYSZTZRQSIIOAROB EYHJSRVJSISIQRLLOHYL  
 MAHRROGHFZSFWICCYXIBTWSCKBHKIYPBFVASMYSZRZHQSURPEEZPRLTCWEKDY  
 QRVRWCGVVVPGNAVEEEDZRRFCWWZQRCEIRIAHCCCECDZRRVNDGMYKNHDCNLN  
 ASICHCIXTWFCDIGMFWKSCMZIKXPVRRKEATVBXEEQLQYEXFRFUSZVBBCCELVG  
 RRORBHMYKZCIILLQWJXTRPHCCTVRAVQMI EWKALWVBKLPFYSROOIPSDFPVNB  
 ILGUGVTLNVHVHJMAUVQMIEKISFKUHZXDKUCJXFTBBKLPJYCFVPETSIPJMJWJL  
 PHGVVQZVECNZLMAZPMSEQGFYRLGHFFZVECNJCSZAPFZS XGJYCGROJIZJFCIVZ

AFCIVZASCIXSIYCJXWIIACIIQSEHYICEESRROVNRZEYXZOHPRJVFQELROEKPP  
FBRQPPRBFVPRNAVPPWFVVVPJBFVZPVZCIIILRQHYIDMYVVRDEQIEGPVGQZRCYF  
HCMYKBTVENLCIITWIPIIIXLMAHYVTPYSUQPJVZCIOQRKZXSJNBKEDXVQKICVBF  
JRPZRFWIWXSOWSCIFCKLLXACNXZGWCPPELRPVEEMAUFJXCUSRVEMFHFISOVRDV  
EEMAUKMDWBAVZTWVHFVPRGFVEEAMUVREVNBTILXZMTLLQOSIHZSEGFQPPNHVZ  
TWVHFVPRGFVEEAMUVREVNBTILXZMTLLQOSIHZSEHYMDMGWJEYHACKLRTAFVP  
TESJIIYXMYDCDSHZXVPAFHISYKRFYIDMGOKMYKGVVRYSYCEKPVFWIWLQMWFVXE  
QODXCYYMPSFVSCIKTZRBVWDMVAGPZVRPLXELRTRGEMFWNEDRNDGMYKNBUWZKR  
BKPJCBITEXIEOGTTRTOEHDSSOZREPLMFYNEZSKEATVBXXLTCWEKLXZMTLLQOS  
IHZSEHYEEMFQRVNIJOJWFVRWYILVQMFYSIESZSAIASUATHRHYYIOSBFUECOASJ  
WELRFVEYHACKLRTAFVPHRSGMYXBHYEEHNFBRPWFVICMAUCSYKVGSZHGUVV  
PABBUICMAUWILVVBXHZYOHZRRHESRQTRTRIILQFBFQZVGOCIGIERRVPHGCUVP  
EZPVJZVRPLXELRGZPPRPSNEDYAPISVIAOEHELKGKMPASJWREISESESXSEEYH  
GVVSYPKLFVOXUSIIDTBYVRHEFHYYIHLVGGICIQKFVOPRBFVPXUWJMHVLGGICIQ  
OEHLRRQYSXYEALVPHOOTELRKFOVRBFVPQRFVPJXUWJEYHACKLRTAFVFPFNQ  
BMYXBHYINLNASICXHFEMYKNZCQJWBICATXUWEQPFHFEMYKCFRLKNWEMSINFU  
EEECDZRRWBABASEGZFYOIEHYEYFRTFVPWHFVPJWNWUMDYESCCELNHZWSZSKL  
TRTOKQJAVBUSHPNHKMNIYSKQPWRSKLPRJVRXELRFVEEMFOEHELVGDCDXRFP  
TYCIIWIGAPLPEEHSIDVZCEXSZSEXLRQHYMDQLGKICCRLGPZVRHZWELRKZROE  
ARESELVBXQZVRGCIYLRVFMQPHBXXSIFVLXEIEKYIYAVHYQLRLOWPTVGOEHQPH  
HKICMAHYICIFHVTAIQOJXLXRZPVLZRBFLJELRGRMYXYMUEJWBTPSCIACKXSIYS  
RWESOSZWLRSDEOIUSESEEZWEYEIFHFTAIQICWEELSULPFHHNMELZWVRZJYCI  
HZVYOUCAIEQYIOEOMIXCPVRQMIERFSCTRTFTLPHHDFRLFHGKSQTNZCEDNHGKE  
MSISDCNLNASICHBCITPVPVHLRQGRXLRQBFXSMAUDSCIGVVRELGVFZRLPZVO  
FRULMWMAUDCDEQTRRNCVBKSDQVZZRRFLHYIRVNJVEYHFHVYHRQFVFBTKLPG  
BIEXPRNBTITXJCIIELBIXLELLQIIDXOSJLZVAOEHDNLJVRLELBIZWLMQOIXDYE  
SESNNJVRRNLGKPKJEWDEYHNBTMPRGFRZPRJOEHPVVBXJCSZHYIYMTVKPJWUC  
IIEIYZDIHLNHLKJPBFUPJRNAVMDSAHYIYMTVKWAPHHFRTEAGYSCIDIFXSXUSI  
EGIAVBZPVZCIIIXPVZQLVISCPHGVZWFRTOZRWCSCNPESUSRVOMFQFYCWRGFT  
WEVBCCELBIXLTXFOEWHIEZZXEPRAVEYMAUCMEXYSIIWIIOEGJFBFVJZVJSTEY  
RBXYIWTNUIIPMAUKLLXACCMGMAUYXEAAPVMYKRJVJVGKRWMPRGJIOAVHYWPI  
VBXFVTVQSSGIUWJGSEZPVVOSBFSMCHBFSILWIGISYXUSJGFPCHLVPHOIJXLFB  
JVLTPVVRQMIERFSCAVHYWFGUBRQPEFBVZPVZCIIIMYGHYICEISEWTXGWEKWSAS  
CCZRGVVVTWEPWUFFWGGGSVIBBCCELNHFRPABFUEDMSVZWSHZZRELNHFRPABFU  
LPHVRFYETBIIRZXUWEKQEEHYICXUSELPYGHVVPHACKEQINHYICXUSELPJYIKX  
PVRKMWMPVGTCEGRZPQZVRHYEYQHKKICIQCKLPVSFZFYHFVRZPJYCNRMISCIIZ  
RGVVQZVECNLPAVZCPPEISDILWZMYSIAIFVRZPJYCNRMISCIIELRBKLPFVFWLM  
QBVZPVZCIIIDXNFKPPHNHKLPGWGWCPYIFGSVZORBSCCICZPWZECHCCDTBYVROSH  
PKPPWFGRMOMJVRXTXHHKICWVGZXSZAZPWESPYRROWGCIINEHUYXQVBAJSXIHB  
YEATLARWEIEKYSXYAAVNNMSICTWNGKICJBZCSHIQTRWEEARWSWPBKVHQEFHV  
VEMYZYMDWBBXWZRRPLVOIAPFVPXVZCXSIQWIKPWBTYMDLBDVXSEGAVPLRPVFP  
JFHFUIYFBFVSQRRJVVYIISIQZVRPLXELRFRZPRFHZPWFRULMWMAURPWQLTRN  
CVBKSDQVZZRRWGFRMRLGWNLPISUENYFVZSYIQGVEEMATISYXBTSMCHNBUFFW  
GOEHOSBFKLPRHDFRELJVPGIGGZRVMAUZFPXBCBQJWRZWXZPVBBMYKSOEGJYA  
HFJLRLPMKLTRXWEKHLNHLTWBAZRYFPZVOSSMFVPAUOKXSMFUIMXYAURMYPLU  
YEDXYMXEFRGOEHZQVBFYDFVFUSQCBFVQPEAHZRNVBOMYKASMICQBFVXSMFWJ  
EEIAURKPHVBXYPWFWEKMYGBFWJPYOSPPKDIIDWVBXXZXUSWSHPJVFVPJVSIC  
PCRGESHFHFEIOMAHFQJFBGFQDGBFVXSMFOEHXSESZWLXQWMMYMAUNMELZMYIL  
HNHVEDIESTPTRVBXSXYUSTYDLVCEWGIYJVXWMAWEKELNHLPPNAGPTKUHPZE  
GSUSPVOIKASSFSMIWZRHMZPRHCMYMAUNMELGVVPLQCZZKSXTZFEEMAUFICWU

SJLLPYDIIDWNVEIGIEAFVPXUSEQPXUCLKSXGVVETVTFVAOIAGVVAIETLQPHSF  
FQLRHB JIPRPSEWPVFKLRRFLGVVLTUWDASSFSWSZXSOC PDXVBBPPHBBKLPXHTK  
IOJYCFVHVRHTLTGEWVHELLUFHSEGVCIYXGVVIMCGVVWPEAUVPDLRVRXSWRBKX  
SIRFVWAMGSIIDTVHVEYHASGIYXUSWVZQGVPPQBFZIDSSZVRZVRELEQJBVHYL  
JSHYMDOVBURPTRBKLPEARWSCRHKLTWYCJXWIACIIBYBHYXSIEOMIYRRJVVXS  
ESGVZTUSKWL MQWKLTRTCWIGMYDISALRHJXTPYWFTVQCIHPZVZNLPXUSIXPQC  
HVVDIAHFVHLRHYICXRAGIDXGCJWPHGVVISIESRWSESUIDSYOKIJIGOC PFRQO  
LREIQCEXSMFRVWPVGZRROIAQY EYXRRFRELVGYSXIOMYSCVBFEYFRGSUXPPYAV  
XCYYMZMXTYCIITWGVVPMFHYICIOOCQTRTWCILHGSCPXIGSCPXIVWDTWSESHY  
ZXUHYICEISERPZRFDSCICFFTSIGGRMOMGVZRRSSMMWTECGLPXFHZP WMSPVZO  
SERVZTPOMKLLXUSRZPRGVRXMIARJEMSISLWMC GVRXRSQKVFZXUOUSCIGSCPEL  
VGJSFPJWKLDSEFFAWEQSEM QAVHYMYXUSUMDXNBKETHRBEMEUOCNPNNGGEDEV  
BKIOQNWUIYAUCDXSINBXIWWAODIWIACIINPNNGGECEESROVNRZEYXZOZHPRJV  
FQELROEKPPFBRQPPRBFVPUHCKLELRFRZPRASMICQBVFVFXUOKAZVQCLVDMTBF  
JAEHZZRRFVFUSCJVSEHTWUFZIVIQIGWEEHZZRRKRHKLP IOOTOTRGCKLPXRAGI  
DXNBUXSIAWXLEWCZLXZRVOEWSSESCILZRBFFWEPYGP FQROJEESXSESQXUOKPT  
IGVPWZYYVRXSWCCBIYPROMIXCYCEIWMASJWFROFFOPRDIZXELRPLWEEOCMIXC  
QCFVEEXSKLJFROBJCSZCLXXCUSRVEEARKEVIGVPJZVZTISXSSTDCOSBFHYZXU  
HYICEISERPZRFDSCINBUXSIEOMIYRRJVVQP VHKMYKFH ZPWFZGZEMA UJXTPYW  
JWTXGWEKZRGVVTLPYWUFFWGCWTLPYOJNFWGOSSGIZMTLLQOSIHZSEOEHSMFSP  
IDLNJVEWPGVWP IZWEKZJNRVQZRFHYEEMFRIILQVBXEYHGVVPLQCZZKSXBSIL  
TQFHIILQVBXXSVBKJLTW FVRHZABBKLPJYCFVLRQAPWZYYTISXSHHKLLXFVRHZ  
AGVRXWMRGWPZEGWEKZRGVVJWSBFJLLPYPVPTJGSURPZRFDSCI

sequence	interval	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ASMY	72	X	X	X		X		X	X			X					
ECDZ	24																
ZVEC	24																
FCIV	6																
VEEM	24																
EEMA	48																