Breaking the Vigenère Cipher

Unknown Key Length

Duke
UNIVERSITY

Breaking Vigenère

```
Qqd fv avd iaght aqg Eknnplag lkwqgy km avd kxp'c rwqd
von rna ugupvo?  Oqy elwvbaklb, ygvynl voxwnqv cjl epyjla dju
dpiaghtciug kgjjwzn pc zngtnf qcym ax mrpk von rna ugupvo.
Drvoxwa c lqtywant, eonerrpn gcnp j txflbv wwtkgy qm mlh
snpncjz tlzwpagz gelgzbkcn lohvav.
```

- Step 2:
  - Unknown Key Length

Try Key Length = 1?

```
Qqd fv avd iaght aqg Eknnplag lkwqgy km avd kxp'c rwqd
von rna ugupvo?  Oqy elwvbaklb, ygvynl voxwnqv cjl epyjla dju
dpiaghtciug kgjjwzn pc zngtnf qcym ax mrpk von rna ugupvo.
Drvoxwa c lqtywant, eonerrpn gcnp j txflbv wwtkgy qm mlh
snpncjz tlzwpagz gelgzbkcn lohvav.
```

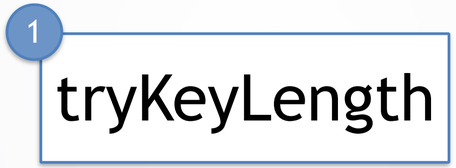
Key length = 1

1 tryKeyLength

Try Key Length = 1?

```
Hhu wm rmu zrxyk rhx Vbeegcrx cbnhxp bd rmu bog't inhu
mfe ier lxlgmf?  Fhp vcnmsrbcs, pxmpec mfonehm tac vgpacr ual
ugzrxyktzlx bxaanqe gt qexkew htpd ro digb mfe ier lxlgmf.
Uimfonr t chkpnrek, vfeviige xteg a kowcsm nnkbxp hd dcy
jegetaq kcqngrxq xvcxqsbte cfymrm.
```

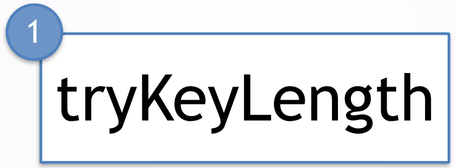
Key length = 1



Try Key Length = 1?

```
Hhu wm rmu zrxyk rhx Vbeegcrx cbnhxp bd rmu bog't inhu
mfe ier lxlgmf?  Fhp vcnmsrbcs, pxmpec mfonehm tac vgpacr ual
ugzrxyktzlx bxaanqe gt qexkew htpd ro digb mfe ier lxlgmf.
Uimfonr t chkpnrek, vfeviige xteg a kowcsm nnkbxp hd dcy
jegetaq kcqngrxq xvcxqsbte cfymrm.
```

Key length = 1



Probably not right

Try Key Length = 2?

```
Qqd fv avd iaght aqg Eknnplag lkwqgy km avd kxp'c rwqd
von rna ugupvo?  Oqy elwvbaklb, ygvynl voxwnqv cjl epyjla dju
dpiaghtciug kgjjwzn pc zngtnf qcym ax mrpk von rna ugupvo.
Drvoxwa c lqtywant, eonerrpn gcnp j txflbv wwtkgy qm mlh
snpncjz tlzwpagz gelgzbkcn lohvav.
```

Key length = 2

2 tryKeyLength

Try Key Length = 2?

```
Jow yt yob gtear yje Cdlgneyz eipozw if ttw dvi'v kujb
thl pgy szsith?  Hor xjptuydju, wztrle omqugoo vhe xnrhey bcs
bigtearvgne izhcusl nv slzrgd ovwf tv kknd omg klt nennom.
Wpomqut v eomwpygr, xmgckpil evli c mvyjut uprder jk kef
qgngacx rexpntes zceeszdag emattt.
```

Key length = 2

2 tryKeyLength

Probably not right

Try Key Length = 3?

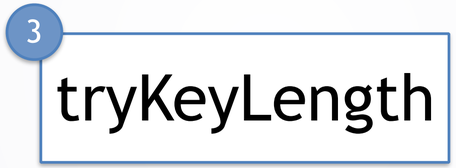
```
Qqd fv avd iaght aqg Eknnplag lkwqgy km avd kxp'c rwqd
von rna ugupvo?  Oqy elwvbaklb, ygvynl voxwnqv cjl epyjla dju
dpiaghtciug kgjjwzn pc zngtnf qcym ax mrpk von rna ugupvo.
Drvoxwa c lqtywant, eonerrpn gcnp j txflbv wwtkgy qm mlh
snpncjz tlzwpagz gelgzbkcn lohvav.
```

Key length = 3

Try Key Length = 3?

How do you break the Vigenere cipher if you don't know the key length? For centuries, people thought the cipher was unbreakable because it seemed hard to find the key length. Without a computer, checking even a modest number of key lengths requires excessive effort.

Key length = 3



Probably right!

Automating Trying Multiple Key Lengths

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each

Automating Trying Multiple Key Lengths

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?

Automating Trying Multiple Key Lengths

Jow yt yob gtear yje Cdlgneyz eipozw if ttw dvi'v kujb thl pgy szsith? Hor xjptuydju, wztrle omqugoo vhe xnrhey bcs bigtearvgne izhcusl nv slzrgd ovwf tv kknd omg klt nennom. Wpomqut v eomwpygr, xmgckpil evli c mvyjut uprder jk kef qgngacx rexpntes zceeszdag emattt.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?
- Can we automate “eyeball” algorithm?

Automating Trying Multiple Key Lengths

Jow yt yob gtear yje Cdlgneyz eipozw if ttw dvi'v kujb thl pgy szsith? Hor xjptuydju, wztrle omqugoo vhe xnrhey bcs bigtearvgne izhcusl nv slzrgd ovwf tv kknd omg klt nennom. Wpomqut v eomwpygr, xmgckpil evli c mvyjut uprder jk kef qgngacx rexpntes zceeszdag emattt.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?
- Can we automate “eyeball” algorithm?

Automating Trying Multiple Key Lengths

Jow yt yob gtear yje Cdlgneyz eipozw if ttw dvi'v kujb thl pgy szsith? Hor xjptuydju, wztrle omqugoo vhe xnrhey bcs bigtearvgne izhcusl nv slzrgd ovwf tv kknd omg klt nennom. Wpomqut v eomwpygr, xmgckpil evli c mvyjut uprder jk kef qgngacx rexpntes zceeszdag emattt.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?
- Can we automate “eyeball” algorithm?

Automating Trying Multiple Key Lengths

Jow yt yob gtear yje Cdlgneyz eipozw if ttw dvi'v kujb thl pgy szsith? Hor xjptuydju, wztrle omqugoo vhe xnrhey bcs bigtearvgne izhcusl nv slzrgd ovwf tv kknd omg klt nennom. Wpomqut v eomwpygr, xmgckpil evli c mvyjut uprder jk kef qgngacx rexpntes zceeszdag emattt.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?
- Can we automate “eyeball” algorithm?

Automating Trying Multiple Key Lengths

How do you break the Vigenere cipher if you don't know the key length? For centuries, people thought the cipher was unbreakable because it seemed hard to find the key length. Without a computer, checking even a modest number of key lengths requires excessive effort.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look (“eyeball” approach)?
- Can we automate “eyeball” algorithm?

Automating Trying Multiple Key Lengths

How do you break the Vigenere cipher if you don't know the key length? For centuries, people thought the cipher was unbreakable because it seemed hard to find the key length. Without a computer, checking even a modest number of key lengths requires excessive effort.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look ("eyeball" approach)?
- Can we automate "eyeball" algorithm?

Automating Trying Multiple Key Lengths

How do you break the Vigenere cipher if you don't know the key length? For centuries, people thought the cipher was unbreakable because it seemed hard to find the key length. Without a computer, checking even a modest number of key lengths requires excessive effort.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look ("eyeball" approach)?
- Can we automate "eyeball" algorithm?

Automating Trying Multiple Key Lengths

How do you break the Vigenere cipher if you don't know the key length? For centuries, people thought the cipher was unbreakable because it seemed hard to find the key length. Without a computer, checking even a modest number of key lengths requires excessive effort.

- Loop: try key length 1,2,3,4,5…
  - Fractions of a second each
- How to tell if right length?
  - Human look ("eyeball" approach)?
- Can we automate "eyeball" algorithm?

Idea: Count Real Words

- Read in list of English words
  - ArrayList?
  - Better: HashSet
- Try decryption
- See how many real words you have
  - Appears in English word list?
- Choose key length with most real words
  - Have seen “max” many times by now!

HashSet

- ArrayList: would work fine
  - Read file, .add() each word to list
  - Use .contains() to check if word in list
- Better: HashSet<String>
  - Use .add() and .contains()
  - .contains() will be *much* faster
  - Why/how faster? UCSD's *Java Programming: Object-Oriented Design of Data Structures Specialization*

Split String into Words

```
for (String word : decryptedMessage.split("\\W")) {

}
```

- Need to split String into words
  - Can use String method .split()
  - Pass in “\\W”
    - Divides between “non-word” characters