COLOMBIA POTENCIA DE LA VIDA

TIC

Misión 2

Lección 1: Criptografía Básica

Criptografía Básica

Tiempo de ejecución: 6 horas

Planteamiento de la sesión:

En la primera sesión se introducirán los conceptos de cifrado, por lo que es necesario que los estudiantes cuenten con un equipo de cómputo, un entorno de Python configurado para programar ejemplos y lápiz y papel para realizar algunas actividades de forma manual.

Materiales

- Computador con conexión a internet
- Lápiz y papel

Desarrollo de la sesión

La palabra criptografía proviene del griego kryptos (recubierto, oculto) y grafein (escritura). Es decir, significa ocultar o recubrir un mensaje o un escrito mediante un sistema o algoritmo. Lo interesante de un mensaje recubierto o cifrado es que el mensaje puede estar a la vista de cualquier persona, sin embargo, solo puede ser comprendido por el escritor del mensaje y el destinatario del mensaje. Es decir, existe una metodología que permite tanto al escritor como al lector crear una codificación y posteriormente decodificar el mensaje para su entendimiento.

Uno de los usos más antiguos de la criptografía se remonta a Egipto en donde se dibujaban jeroglíficos en paredes y en lugares añadiendo algo de misterio a un relato para que los lectores se sintieran retados a intentar leerlo y descifrarlo.

Se sabe también que los hebreos, y los griegos escondían mensajes bajo referencias (por ejemplo, se cree que el número 666 podía significar una referencia a un emperador o al imperio romano). Sin embargo, en Grecia se desarrollo un sistema basado en un dispositivo llamado la Escítala. Este sistema lo empleaban los espartanos para enviar mensajes secretos. La escítala está formada por dos varas de grosor variable (similares) y una tira de cuero. Para encriptar el mensaje se enrollaba el cuero sobre una de las dos varas y se escribía el mensaje (figura 1), posteriormente se desenrollaba el cuero y se enviaba al destinatario. De esta forma el mensaje era ilegible para cualquiera que no tuviese la misma vara. Una vez el mensaje en el cuero llegaba al receptor, este enrollaba la cinta sobre una vara gemela y podía leer el mensaje original sin problemas.

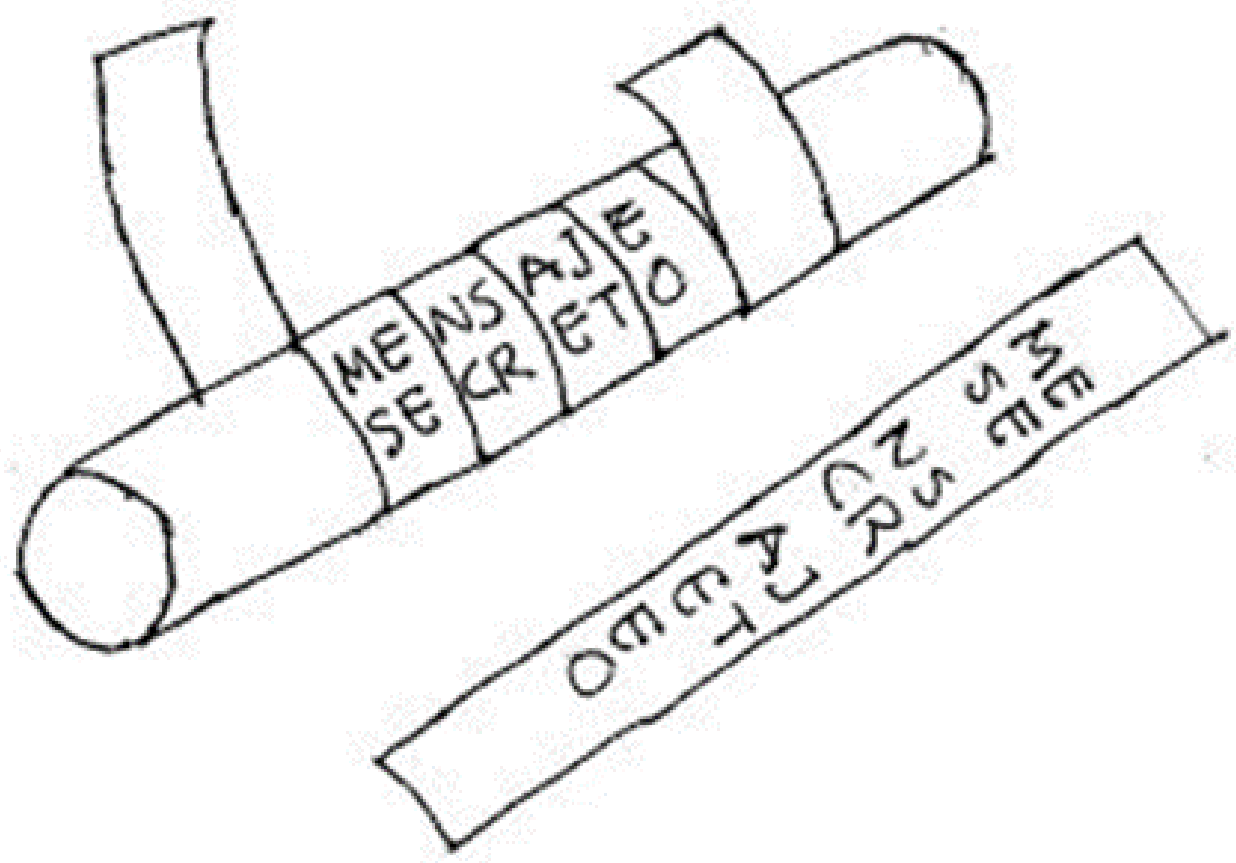

Figura 1: Representación del funcionamiento de una escítala

A este método se le conoce como cifrado por transposición y no es más que alterar el orden de los elementos en un mensaje para hacerlo ilegible.
Para entender mejor este sistema, y comprendiendo que hace parte de la familia de cifrado por trasposición, se puede intentar hacer un cifrado similar escribiendo un mensaje en una matriz. Supongamos que ud es un emperador griego y quiere advertir a uno de sus generales en el frente de batalla sobre qué acciones tomar.

El mensaje a codificar será:

diríjase al valle del norte y espere dos días, se encontrará con general Claus.

Para escribir el mensaje se puede poner en forma de matriz. Arbitrariamente hemos decidido que la matriz tendrá 10 columnas.

D	i	r	í	j	a	s	e	a	l
v	a	l	l	e	d	e	l	n	o
r	t	e	y	e	s	p	e	r	e
d	o	s	d	í	a	s	s	e	e
e	n	c	o	n	t	r	a	r	á
c	o	n	g	e	n	e	r	a	l
			C	l	a	u	s		

Tal como sucede con el cuero en la escítala, en nuestro caso podemos trasponer los caracteres simplemente intercambiando las filas con las columnas. De esta manera tendremos el mensaje organizado como: (matriz traspuesta)

D v r d e c C
i a t o n o l
r l e s c n a
í l y d o t u
j e e í t g s
a d s í r e
s e p a r e
e l e s a r
ar e r a
le á l
y u
s

nótese que el mensaje aún puede ser leído si se empieza en la primera columna y se intenta leer de arriba hacia abajo. Para entregar el mensaje codificado, basta con construir una oración mediante la sucesión de las filas de la matriz transpuesta. El mensaje a enviar será:

DvrdeCciatonolrlescnílydotujeítrsasdíreseparelsera rel eá y s

De esta forma hemos construido un mensaje que puede ser visto por cualquier persona, pero que está encriptado. Como plan militar, evidentemente el general que recibe el mensaje debe saber cómo reorganizar los caracteres. En este caso deberá colocar las letras de arriba hacia abajo hasta completar grupos de siete columnas (es decir, colocar los caracteres en grupos de siete será la clave para descifrar el mensaje) recuperando la matriz transpuesta. Finalmente, si se traspone tendrá el mensaje como se ve en la primera matriz.

Sin embargo, ese sistema puede tener ciertas dificultades, sobre todo cuando el mensaje puede ser más grande de lo esperado. Esto hará que el receptor no sepa cómo reordenar los caracteres para reconstruir la información.

Otra de las dificultades es que, si se cuentan los caracteres es posible determinar los submúltiplos del número. Como se sabe que la matriz tiene un número entero de filas y columnas, no tomará mucho tiempo descubrir cómo reorganizarlos para que tenga sentido el mensaje, con lo cual no es seguro el envío de información por este medio

Ejercicio en clase:

Crear un programa que calcule los submúltiplos del mensaje **(El mensaje debe incluir los espacios, puntos y comas).** E intente calcular el tiempo que le toma a un computador descifrarlo. Al guiar a los estudiantes a hacer esta actividad encontrarán que toma menos de un segundo descubrir el mensaje.

Otro de los cifrados que se desarrollaron en la antigüedad fue el cifrado César. Es una técnica de cifrado simple y pertenece a los cifrados por sustitución. Es decir, cada letra de un texto original es sustituida por otra letra que se encuentra un número fíjo de posiciones más adelante en el alfabeto.

Por ejemplo, una sustitución puede hacerse con el número 4 (cuatro desplazamientos hacia adelante en el texto) con lo que, los nuevos caracteres van a quedar representados de la siguiente manera:

A -> E B -> FC -> G D -> H

Y así sucesivamente, rotando el alfabeto cuatro posiciones.

Este tipo de cifrado lo empleaba el emperador Julio César (romano) para enviar mensajes a sus generales. Sin embargo, empleando un computador, toma menos de un segundo romper este tipo de cifrados.

Actividad en clase, cifrado césar.

En esta actividad programaremos el cifrado de mensajes utilizando el algoritmo de césar (sustitución simple). Para eso seguiremos el código del cuaderno de jupyter llamado **01_cifrado_cesar.ipynb**

En este cuaderno se tiene una función llamada rotar alfabeto que permite crear un diccionario con las sustituciones a aplicar:

Posteriormente se define la función encriptar_cesar que permite crear el texto encriptado.

Como el objetivo de la actividad es determinar cómo se encripta y qué tan seguro es un algoritmo de encriptación, se define la función desencriptar_fb que ejecuta todas las combinaciones posibles de clave para intentar descifrar el texto codificado y mide el tiempo que tarda un programa en resolver el cifrado.

En el cuaderno se hace el cálculo del tiempo y vemos que, probando las 27 combinaciones y , sin emplear ninguna técnica sofisticada para reducir el proceso, resolvemos el cifrado en menos de 10 mili segundos. Es decir es extremadamente fácil de descifrar.

Actividad en clase, descifrado automático:

Como actividad se propone que los estudiantes descarguen un diccionario de palabras en español, puede ser en el link :

https://raw.githubusercontent.com/words/an-array-of-spanish-words/master/index.json

que tiene cerca de 160 mil palabras en español. Con ayuda de la biblioteca JSON de Python, cargar este archivo y comparar cuantas coincidencias de palabras hay en el archivo descifrado. En caso de que hayan mas de 10 palabras que estén en el diccionario, se imprimirá solamente esa clave y el texto descifrado. El tiempo que tarda un programa en recorrer muchas veces todo el diccionario debe ser menor a 20 milisegundos, por lo que la respuesta se creará en menos de medio segundo aproximadamente y sin tener que revisar cada una de las 27 combinaciones de texto de salida.

Cuadrado de Polibio:

Este algoritmo reemplaza los caracteres por sus coordenadas en un cuadrado. Por ejemplo, se pueden ubicar las letras del alfabeto en un cuadrado de 5x5, encasillando dos letras en un mismo recuadro para hacer caber el alfabeto (no se emplean signos diacríticos como la ñ ni las tildes). En este ejemplo se junta la I con la J.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I, J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Para codificar un mensaje basta con sustituir cada letra con un número de dos cífras que indicarán la fila y la columna de cada letra en el cuadrado. Por ejemplo:

“mensaje secreto” se codificaría como:

32 15 33 43 11 24 15 43 15 13 42 15 44 34

Otros cifrados mono alfabéticos:

El algoritmo cesar tiene una sola regla de reemplazo para todas las letras en un mensaje (desplazar caracteres N posiciones). También existen sustituciones simples como es el caso del cifrado Atbash en donde el alfabeto de reemplazo se invierte (la A corresponde a la Z, la B corresponde por la Y y así sucesivamente). Esto hace más difícil descifrar el contenido del mensaje.

También existen alfabetos reordenados, similar a como sucede con el cifrado Atbash, en donde los caracteres son arbitrariamente reordenados y luego se hace un reemplazo, por ejemplo:

Alfabeto normal: abcdefghijklmnñopqrstuvwxyz
Alfabeto reordenado: zebrascdfghijklmnñopqtuvwxy

En este caso la A debe ser reemplazada por la Z para codificar, la B por la E, la C por la B, la D por la R y así sucesivamente, dificultando su descifrado.

Algoritmos de sustitución polialfabéticos

Son métodos de cifrado por sustitución similar al algoritmo de cesar, sin embrago emplean múltiples alfabetos en vez de uno solo. Es decir, cada carácter está desplazado diferentes unidades o es cambiado dependiendo de una clave. De esta forma se resuelve el problema de seguridad del algoritmo cesar, ya que el cifrado cesar, como se vio en el ejercicio anterior, es fácil de descifrar con un computador.

Uno de los primeros cifrados poli-alfabéticos en aparecer fue el cifrado de Alberti. Este consiste en cambios de alfabeto no periódicos empleando un artilugio inventado por Alberti (los discos de Alberti). Como el de la figura 2.

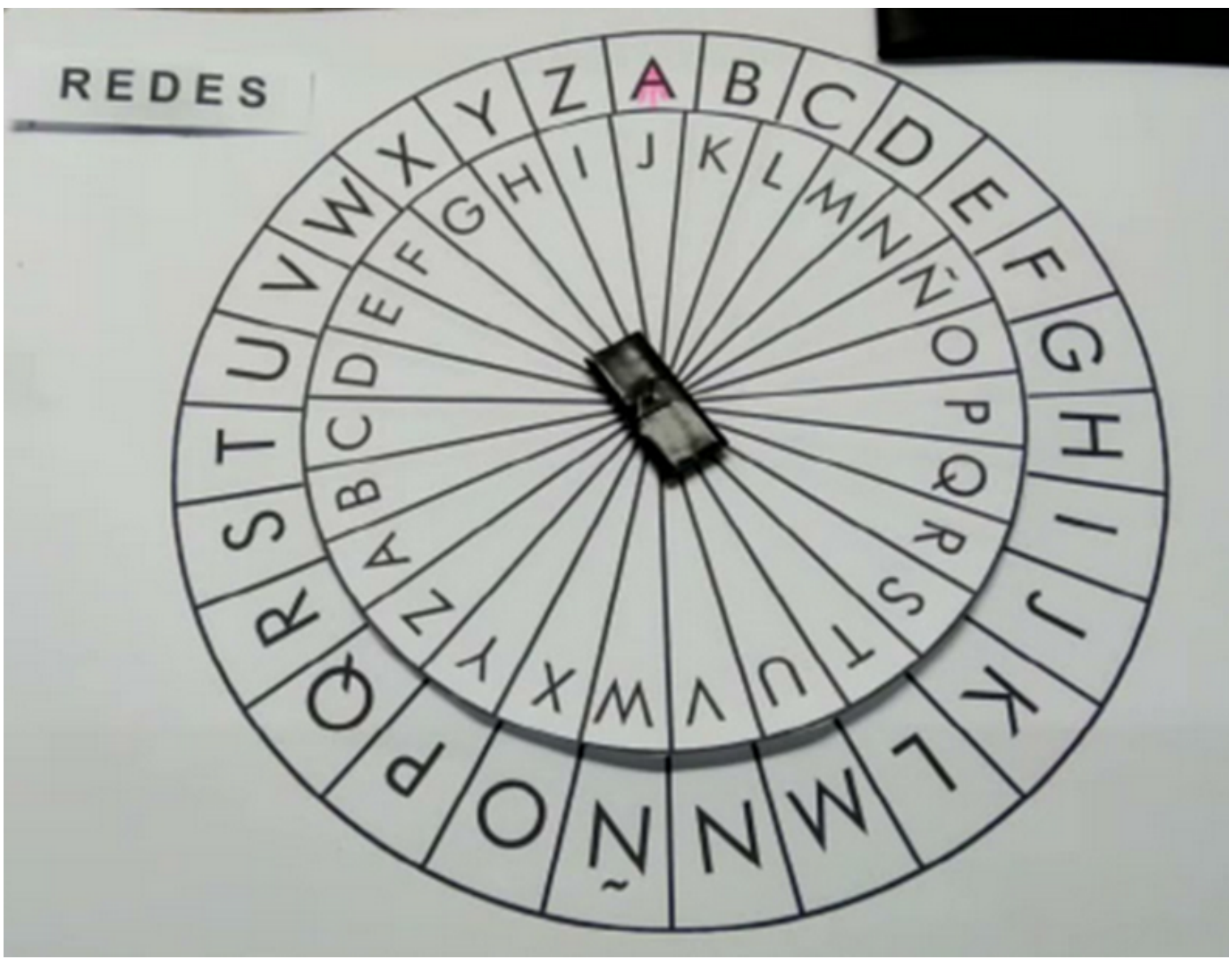

Figura 2. Disco de Alberti.

El cifrado se realizar exactamente igual que con el cifrado césar, ya que girando el disco se puede encontrar qué letra equivale con la cifrada, con la diferencia que para diferentes grupos de letras el disco se hace girar más posiciones. Esto cambia el cifrado por completo y hace que las técnicas de fuerza bruta sean ineficaces para romper el cifrado.

Ejercicio en clase:

Guíe a los estudiantes a consultar cómo fabricar un disco de cifrado de Alberti usando papel. Posteriormente ellos deben cifrar y descifrar mensajes para entender a profundidad el funcionamiento del cifrado y del disco de desplazamiento de alfabetos

Cifrado Vigenère.

Es uno de los cifrados poli alfabéticos más famosos ya que lo llamaron en francés le chiffre indéchiffrable (el cifrado indescifrable). Su fama se debe a que es fácil de entender e implementar y parece imposible de resolver, en escencia es similar a aplicar muchas veces diferentes cifrados césar.
Para codificar un mensaje se debe escoger una clave que sea más corta que el mensaje. Por ejemplo, arbitrariamente escogemos como clave la palabra limón.
Posteriormente necesitamos repetir la clave hasta completar todos los caracteres del texto a cifrar. Como ejemplo se codificará el mensaje: saludos desde el otro lado.
Ubicamos la frase a codificar y debajo situamos la clave repetida tantas veces como caracteres (por practicidad sin espacios):

s	a	l	u	d	o	s	d	e	s	d	e	e	l	o	t	r	o	l	a	d	o
l	i	m	o	n	l	i	m	o	n	l	i	m	o	n	l	i	m	o	n	l	i

Posteriormente requerimos realizar desplazamientos de cada letra dependiendo de la clave. Esto se hace con facilidad a mano empleando una tabla vigenere (que es una tabla con todas las rotaciones del alfabeto posibles, como la de la figura 3.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O
Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q
S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S
U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T
V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U
W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V
X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W
Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X
Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y

Figura 3. Tabla Vigenere

Para codificar el mensaje se debe ir a la fila de la primera letra de la clave (en este caso la L) y a la columna de la primera letra del mensaje. La intersección entre ambas será la letra que debe ir en el mensaje cifrado.

En nuestro caso la fila L y la columna S corresponden a la letra D

Este proceso se debe repetir hasta completar el mensaje. El mensaje codificado en este caso es:

Dixiqzapsfomqzbezaznow

Para descifrar el mensaje se sigue el proceso inverso, es decir, se encuentra la fila correspondiente a la primera letra de la clave (la letra L) y hay que desplazarse sobre esa fila hasta encontrar la primera letra del mensaje cifrado (la letra D). La salida es la letra que está en la primera columna (en este caso la S). Nota: es obligatorio conocer la clave para descifrar el texto.

Ejercicio en clase:

Programar el cifrado Vigenere usando Python. Crear una función para codificar en Vigenere dado un texto y una clave. Crear una función para decodificar un texto teniendo la clave y un texto cifrado. Utilizar como base las funciones vistas En el cifrado césar (normalizar textos, quitar tildes) y probar el funcionamiento. Guíe a los estudiantes a que tomen unos minutos para enviarse mensajes codificados entre ellos. Los mensajes pueden enviarse por correo a todos los remitentes o utilizando lápiz y papel de tal forma que todos los alumnos puedan ver el mensaje, pero solo el destinatario que sabe la clave lo pueda descifrar. Debata con los estudiantes si es posible descifrar el mensaje y cómo podría hacerse.

Si bien este es un cifrado mucho más complejo de descifrar hay metodologías que permiten romper el cifrado y encontrar la clave. Estas se practicarán en el taller del final de la unidad.