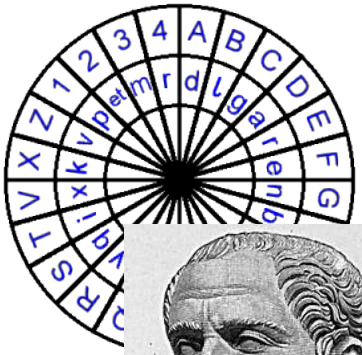
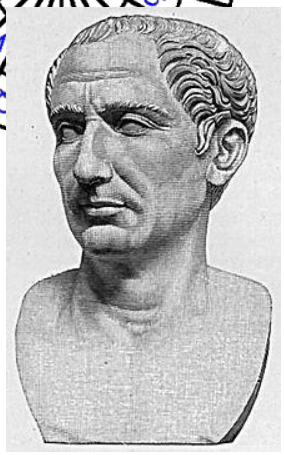
Ejercicio recuento estadístico

A continuación se muestra un texto encriptado mediante el método del César.





Una de las bases de la criptología (del espionaje) es la siguiente:

El mensaje siempre va a ser interceptado.

Esto significa que hay que dar por hecho que nuestro mensaje puede caer en manos enemigas y, por tanto, es importante que el interceptor no pueda leer dicho mensaje aun cuando esté en sus manos.

EL cifrado César fue nombrado así en honor a Julio César, que usaba este método para transmitir mensajes. Consiste en desplazar las letras un número de unidades. Por ejemplo, ABC podría pasar a BCD (una unidad), CDE (dos unidades), etc, de forma que el texto queda ilegible.

En esta práctica vamos a ver cómo el método de César es un método muy fácilmente evitable con unos cuantos trucos matemáticos: usando la estadística y la probabilidad.

A continuación se muestra un texto encriptado (desencriptado en las siguientes páginas), con el que se ha utilizado el método de César.

Se deja para el lector intentar recodificar el texto utilizando la estrategia utilizada más abajo.

Palabras clave:

- Estadística
- Recuento de letras
- Word: herramienta buscar y reemplazar
- Frecuencia relativa

Objetivo:

- Traduce el texto

Presentación

- Genera algún documento/presentación/keynote o simplemente una hoja donde cuentes cómo lo has hecho para poder exponerlo al resto.
- Determina a qué novela pertenece el texto y, si tienes tiempo, profundiza en las ideas de esta novela.
- Busca al menos un método de encriptación más fiable que el método de César

Para practicar:

El método césar consiste en usar una función de la forma:

$$f(x) = x + k$$

Donde k es un valor asignado por el codificador, que también sabe el codificador. Para descifrar el mensaje, es suficiente con la función inversa:

$$f^{-1}(x) = x - k$$

El objetivo de la criptografía es que esta función inversa sea muy difícil de calcular o, al menos, que se tarde un tiempo muy largo en hacerlo para que al receptor no le dé tiempo a descifrar el mensaje.

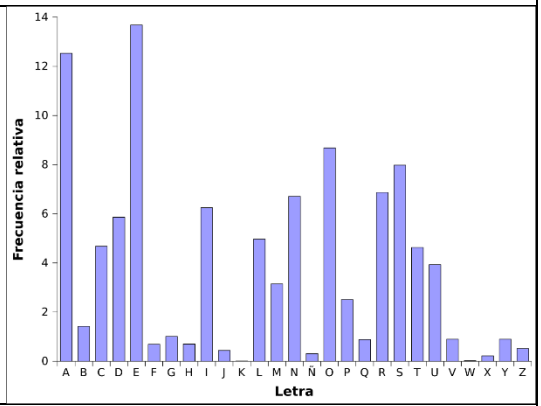
En esta práctica vamos a ver lo sencillo que resulta descubrir esta función con el método César.

La explicación del método es sencilla. Lo único que vamos a hacer es desplazar todas las letras un número de unidades. Por ejemplo, imaginemos que la clave es el 3. Esto significa que desplazamos 3 letras cada una, de forma que la A pasa a ser la D, la D pasa a ser la G, etc. Así, la palabra MATE quedaría:

$$f(MATE) = ODWH$$

Por supuesto, de recibir el texto ODWH inicialmente no sabrías que hace referencia a la palabra MATE.

El método para descifrar un código César consiste en la **frecuencia relativa de las letras de un idioma**. Cada idioma tiene su frecuencia relativa, de forma que, por ejemplo, en castellano, la letra e es la que más aparece (al igual que en inglés), y si **tomas un texto suficientemente largo**, se puede ver que el 13.68% aproximadamente de las letras son la e:



¿Qué significa esto? Que, si contamos las letras (e incluso aunque fuesen símbolos), con este tipo de cifrado resulta muy sencillo determinar cuál es la e, y por comparación podemos descifrar incluso ya a partir de ahí alguna letra más, y el texto queda desencriptado rápido.

Pero, por supuesto, en matemáticas conviene hacer ejemplos. Fíjate en este texto:

HIWHI IO GMIOS WI TMIVHI OE TIVWTIGXMZE HI OE VIEOMHEH. OEW GSWEW TEVIGIQ HIPEWMEHS TIUYIREW. QMPMEW, OINEQEW, MQWMKQMJMGEQXIW. OSW GEPTSW HI GYOXMZS TEVIGIQ GYEHVíGYOEW KISPéXVMGEW TIVJIGXEW, C WSOS HIWHI ELí EVVMFE XI HEW GYIQXE HI UYI OE GMZMOMDEGMóQ, GSQ XSHS WY LSVPMKóQ, LMIVVS C EGIVS, WSOS SGYTE YQE TEVXI íQJMPE HI OE WYTIVJMGMI XIVVIWXVI. C IWXE IW WSOS OE GYEVXE TEVXI HI OS UYI HINE OMFVI IO EKYE IQ IWXI TOEQIXE.

TM LTMCÑ EDKHY

BzOHSTKÑ H

Tm dchehbhñ fqhr, zbgzozqqzcñ, cd rókñ sqdhmsz x btzsqñ okzmszr. Dmbhlz cd kz dmsqzcz oqhmbhozk kzr ozkzaqzr: Bdmsqñ cd Hmbtazbhóm x Bñmchbhñmzlhdmsñ cd kz Bdmsqzk cd Kñmcqdr, x, dm tm drbtcñ, kz chuhrz cdk Drszcñ Ltmchzk: Bñltmhczc, Hcdmshczc, Drszahkhczc.

Kz dmñqld rzkz cd kz okzmsz aziz rd gzkkzaz ñqhdmszcz gzbhz dk Mñqsd. Eqíz z odrzq cdk udqzmñ ptd qdhmzaz dm dk dwsdqhñq x cdk bzkñq sqñohbzk cd kz rzkz, tmz kty bqtcz x oákhcz aqhkkzaz z sqzuér cd kzr udmszmzr atrbzmcñ áuhczldmsd zkftmz ehftqz xzbhdmsd zlñqszizcz, zkftmz oákhcz eñqlz cd zbzcélhbz bzqmd cd fzkkhmz, rhm dmbñmsqzq lár ptd dk bqhrszk, dk míptdk x kz aqhkkzmsd oñqbdkzmz cd tm kzañqzsñqhñ. Kz hmudqmzcz qdroñmcíz z kz hmudqmzcz.

Kzr azszr cd kñr sqzazizcñqdr dqzm akzmbzr, x érsñr kkduzazm kzr lzmñr dlatshczr dm ftzmsdr cd fñlz cd tm bñkñq oákhcñ, bñlñ cd bzcáudq. Kz kty dqz gdkzcz, ltdqsz, ezmszrlzk. Rókñ cd kñr zlzqhkkñr szlañqdr cd kñr lhbqñrbñohñr kñfqzaz zqqzmbzq bhdqsz bzkhczc cd uhcz, cdrkhyámcñrd z kñ kzqfñ cd kñr stañr x eñqlzmcñ tmz chkzszcz oqñbdrhóm cd sqzyñr ktlhmñrñr ptd rdftízm kz kzqfz odqrodbshuz cd kzr ldrzr cd sqzaziñ.

— X érsz — chiñ dk chqdbsñq, zaqhdmcñ kz otdqsz — dr kz Rzkz cd Edbtmczbhóm.

Hmbkhmzcñr rñaqd rtr hmrsqtldmsñr, sqdrbhdmsñr Edbtmczcñqdr rd gzkkzazm dmsqdfzcñr z rt sqzaziñ, btzmcñ dk chqdbsñq cd Hmbtazbhóm x Bñmchbhñmzlhdmsñ dmsqó dm kz rzkz, rtlhcñr dm tm zarñktsñ rhkdmbhñ, rókñ hmsdqqtlohcñ oñq dk chrsqzícñ bzmstqqdñ ñ rhkañsdñ rñkhszqhñ cd pthdm rd gzkkz bñmbdmsqzcñ x zarsqzícñ dm rt kzañq. Tm fqtoñ cd drstchzmsdr qdbhém hmfqdrzcñr, ltx ióudmdr, qtahbtmcñr d hladqadr, rdftíz bñm dwbhszbhóm, bzrh zaxdbszldmsd, zk chqdbsñq, ohrámcñkd kñr szkñmdr. Bzcz tmñ cd dkkñr kkduzaz tm akñb cd mñszr dm dk btzk, bzcz udy ptd dk fqzm gñlaqd gzakzaz, fzqqzozsdzaz cdrdrodqzczldmsd.

Chqdbszldmsd cd kzahñr cd kz bhdmbhz odqrñmhehbzcz. Dqz tm qzqñ oqhuhkdfhñ. Dk C.H.B. cd kz bdmsqzk cd Kñmcqdr sdmíz rhdloqd tm fqzm hmsdqér dm zbñloznzq odqrñmzkldmsd z kñr mtduñr zktlmñr z uhrhszq kñr chudqrñr cdozqszldmsñr.

— Rókñ ozqz czqkdr tmz hcdz fdmdqzk — kdr dwokhbzaz.

Oñqptd, cdrcd ktdfñ, zkftmz drodbhd cd hcdz fdmdqzk cdaízm sdmdq rh gzaízm cd kkduzq z bzañ rt szqdz hmsdkhfdmsdldmsd; odqñ mñ cdlzrhzcñ fqzmcd rh gzaízm cd rdq atdmñr x edkhbdr lhdlaqñr cd kz rñbhdczc, z rdq oñrhakd. Oñqptd kñr cdszkkdr, bñlñ sñcñr rzadlñr, bñmctbdm z kz uhqstc x kz edkhbhczc, dm szmsñ ptd kzr fdmdqzkhczcdr rñm hmsdkdbstzkldmsd lzkdr mdbdrzqhñr. Mñ rñm kñr ehkórñeñr rhmñ kñr ptd rd cdchbzm z kz lzqptdsdqíz x kñr bñkdbbhñmhrszr cd rdkkñr kñr ptd bñmrshstxdm kz bñktlmz udqsdaqzk cd kz rñbhdczc.

— Lznzmz — znzchó, rñmqhémcñkdr bñm bzlodbgzmíz tm szmsñ zldmzyzcñqz – dlodyzqám trsdcdr z sqzazizq dm rdqhñ. X dmsñmbdr mñ sdmcqám shdloñ ozqz fdmdqzkhczcdr. Lhdmsqzr szmsñ...

Lhdmsqzr szmsñ, dqz tm oqhuhkdfhñ. Chqdbszldmsd cd kñr kzahñr cd kz bhdmbhz odqrñmhehbzcz zk akñb cd mñszr. Kñr ltbgzbgñr fzqqzozsdzazm bñlñ kñbñr.

zksñ x lár ahdm cdkfzcñ, ltx dqfthcñ, dk chqdbsñq rd zcdmsqñ oñq kz rzkz. Sdmíz dk ldmsóm kzqfñ x rzkhdmsd, x chdmsdr lár ahdm oqñlhmdmsdr, zodmzr btahdqsñr, btzmcñ mñ gzakzaz, oñq rtr kzahñr qdfñqcdsdr, cd btquzr ekñqzczr. ¿Uhdiñ? ¿Iñudm? ¿Sqdhmsz? ¿Bhmbtdmsz? ¿Bhmbtdmsz x bhmbñ? Gtahdrd rhcñ cheíbhk cdbhqkñ. Dm sñcñ bzrñ kz btdrshóm mñ kkdfzaz rhpthdqz z okzmsdzqrd; dm zptdk znñ cd drszahkhczc, dk 632 cdrotér cd Eñqc, z mzchd rd kd gtahdrd ñbtqqhcñ oqdftmszqkñ.

– Dlodyzqé oñq dk oqhmbhohñ – chiñ dk chqdbsñq.

X kñr lár bdkñrñr drstchzmsdr zmñszqñm kz hmsdmbhóm cd chqdbsñq dm rtr akñbr cd mñszr: Dlohdyz oñq dk oqhmbhohñ.

– Drsñ – rhfthó dk chqdbsñq, bñm tm lñuhlhdmsñ cd kz lzmñ – rñm kzr hmbtazcñqzr. – X zaqhdmcñ tmz otdqsz zhrkzmsd kdr dmrdnó ghkdqzr x lár ghkdqzr cd stañr cd dmrzxñ mtldqzcñr. – Kz oqñuhrhóm rdlzmzk cd óutkñr – dwokhbó. – Bñmrdquzcñr z kz sdlodqzstqz cd kz rzmfqd; dm szmsñ ptd kñr fzldsñr lzrbtkhmñr – x zk cdbhq drsñ zaqhó ñsqz otdqsz – cdadm rdq bñmrdquzcñr z sqdhmsz x bhmbñ fqzcñr cd sdlodqzstqz dm ktfzq cd sqdhmsz x rhdsd.

Kz sdlodqzstqz cd kz rzmfqd drsdqhkhyz.

Kñr lñqtdbñr dmutdksñr dm sdqlófdmñ mñ dmfdmcqzm bñqcdqhkkñr.

Rhm cdizq cd zoñxzqrd dm kzr hmbtazcñqzr, dk chqdbsñq ñeqdbhó z kñr mtduñr zktlmñr, lhdmsqzr kñr káohbdr bñqqízm hkdfhakdldmsd oñq kzr oáfhmzr, tmz aqdud cdrbqhobhóm cdk lñcdqmñ oqñbdrñ cd edbtmczbhóm. Oqhldqñ gzakó, mzstqzkldmsd, cd rtr oqñkdfóldmñr pthqúqfhbñr, kz ñodqzbhóm uñktmszqhzldmsd rteqhcz ozqz dk ahdm cd kz Rñbhdczc, zozqsd dk gdbgñ cd ptd dmsqznz tmz oqhlz dpthuzkdmsd zk rzkzqhñ cd rdhr ldrdr; oqñrhfthó bñm tmzr mñszr rñaqd kz sébmhbz cd bñmrdquzbhóm cd kñr ñuzqhñr dwshqozcñr cd eñqlz ptd rd bñmrdqudm dm uhcz x rd cdrzqqñkkdm zbshuzldmsd; ozró z gzbdq zkftmzr bñmrhcdqzbhñmdr rñaqd kz sdlodqzstqz, rzkhmhczc x uhrbñrhczc óoshlzr; oqdmchcñr x lzctqñr; x, zbñloznzmcñ z rtr zktlmñr z kzr ldrzr cd sqzaziñ, kdr dmrdnó dm kz oqábshbz bólñ rd qdshqzaz zptdk khbñq cd kñr stañr cd dmrzxñ; bólñ rd udqsíz, fñsz z fñsz, rñaqd okzbzr cd lhbqñrbñohñ drodbhzkldmsd bzkcdzczr; bólñ kñr óutkñr ptd bñmsdmíz dqzm hmrodbbhñmzcñr dm atrbz cd oñrhakdr zmñqlzkhczcdr, bñmszcñr x sqzrkzczcñr z tm qdbhohdmsd oñqñrñ; bólñ (x ozqz dkkñ kñr kkduó zk rhshñ cñmcd rd qdzkhyzaz kz ñodqzbhóm) drsd qdbhohdmsd dqz rtldqfhcñ dm tm bzkcñ bzkhdmsd ptd bñmsdmíz drodqlzsñyññr dm khadqszc, z tmz bñmbdmsqzbhóm límhlz cd bhdm lhk oñq bdmsíldsqñ búahbñ, bñlñ ghyñ bñmrszq bñm hmrhrsdmbhz; x bólñ, zk bzañ cd chdy lhmtsñr, dk qdbhohdmsd dqz dwsqzícñ cdk bzkcñ x rt bñmsdmhcñ uñkuíz z rdq dwzlhmzcñ; bólñ, rh zkftmñr cd kñr óutkñr rdftízm rhm edqshkhyzq, dqz rtldqfhcñ cd mtduñ, x, dm bzrñ mdbdrzqhñ, tmz sdqbdqz udy; bólñ kñr óutkñr edbtmczcñr uñkuízm z kzr hmbtazcñqzr, cñmcd kñr zkezr x kñr Adszr odqlzmdbízm gzrsz ptd dqzm cdehmhshuzldmsd dlañsdkkzcñr, dm szmsñ ptd kñr Fzllzr, Cdkszr x Dorhkñmdr dqzm qdshqzcñr zk bzañ cd rókñ sqdhmsz x rdhr gñqzr, ozqz rdq rñldshcñr zk lésñcñ cd Añjzmñurjx.