Voilà, comme je l'avais promis, un petit tut sur une manière de traiter une crypto par substitution de façon artisanale avec Excel. Bien évidemment, ce que je vais montrer par la manière « manuelle » peut être sensiblement amélioré avec de petites macros VBA Excel. Et à partir de là, on peut aussi utiliser d'autres langages plus sympathiques que VBA. Mais une feuille Excel permet souvent de tester et visualiser facilement beaucoup de choses. Comme je ne suis pas là pour faire un cours sur Excel, je n'expliquerai pas le pourquoi du comment des formules. Juste un rappel : le $ permet de rendre absolues les références d'une cellule (utiliser la touche de fonction F4).

Le texte à décoder provient d'une épreuve de feu Espionet.
LRJEGCBICTJRBERGPRBKDSRLDSRJEQSJFCBJSBRPCTJQBUQSGURGJSCFRGDSRIRSXDSTJQBKF
RNQGJBRLRJQBKUCJ (le cryptogramme original comportait les espaces entre les mots : trop facile !).
Première étape, on colle le texte chiffré en A1.
En A2, histoire de voir la longueur du texte, on entre =NBCAR(A1) qui donne 89.

Un peu plus bas, ici à partir de A7, on entre la suite de 1 à 89 (je vous passe les deux cents méthodes possibles pour faire ça). Ici, jusqu'en A95. Et en B7, on entre =STXT($A$1;A7;1) pour afficher le premier caractère de la chaîne cryptée (non, pas Canal Plus). Recopie vers le bas par un clic sur le coin inférieur droit, pour avoir les 89 caractères de B7 à B95. Si le texte est très long, on peut opérer avec plusieurs lignes ou plusieurs colonnes pour tout avoir sous les yeux.

Un peu plus à gauche, on entre l'alphabet de F7 à F32 par exemple. Et de G7 à G32, on va compter les caractères du cryptogramme avec la formule =NB.SI($B$7:$B$95;F7). Les valeurs sont 0, 9, 6, 4, 3 etc. Afin de mieux visualiser, on va faire apparaître les pourcentages par rapport au nombre de caractères total (89). Donc d'abord calculer la somme en bas de liste, en G34, avec la formule =SOMME(G7:G33), qu'on peut obtenir avec la combinaison Alt= pour les paresseux comme moi. Ensuite, la formule de calcul des pourcentages à entrer en H7, =G7/$G$34, à recopier pour les vingt-six lettres. Un petit coup d'œil à cette analyse de fréquences permet de voir que la lettre cryptée R est présente 15,7 fois sur cent. Vous trouverez, avec quelques différences, la répartition des lettres françaises sur plusieurs sites comme
http://www.apprendre-en-ligne.net/crypto/stat/francais.html
mais aussi avec Google : fréquence lettres françaises =>
http://pedroiy.free.fr/alphabets/frequence.htm
ou encore
http://gala.univ-perp.fr/~cnegre/Cryptographie/frequence-lettre-bigramme-francais.pdf
et bien d'autres…

Retenez en gros ESANTIRULO, avec plus de 17% pour le E, une petite moitié pour le S et le A suivis par un tir groupé du N, du T et du I à 7% et des bricoles. Donc si le texte n'a pas été traficoté (voir les méthodes de la double ou triple représentation du E par exemple), une lettre qui représente 15,7% des caractères du texte a une bonne probabilité de correspondre au E. Par conséquent, ici R crypté = E clair.

Pour visualiser, l'avancement du travail, il suffit de commencer par mettre par exemple des * de I7 à I32, chaque étoile devant être remplacée au fur et à mesure par la bonne lettre. En C7 (à recopier jusqu'à C95 évidemment), on va entrer la formule donnant la « traduction » crypté/clair :
=RECHERCHEV(B7;$F$7:$I$32;4;0)

Une fois que les 89 cellules sont remplies d'étoiles, on va créer la chaîne qui permettra de visualiser la solution.
En A3, par exemple, on entre =C7&C8&C9&C10&C11&C12 …jusqu'à C95.
Si vraiment ça vous barbe, sélectionnez la zone C7:C95 et utilisez une macro genre :

```
Sub concatSelection()
concat = ""
        For Each c In Selection
                concat = concat & c.Address(RowAbsolute:=False, ColumnAbsolute:=False) & "&"
        Next
Range("A3").Formula = "=" &  Left(concat, Len(concat) - 1)
End Sub
```

C'est basique et ça peut être amélioré, je sais :)

Après ça, vous vous retrouvez avec une chaîne de 89 étoiles en A3.
Maintenant, miracle de la technologie, en I24, en face du R, entrez E (en minuscule si vous préférez).
A3 devient :

*e*********e**e**e****e***e**********e**********e*****e***e*e**********e*****e*e*******

La deuxième lettre cryptée la plus fréquente est le J avec 12,4%. Donc optons pour le S qu'on entre en I16 et voyons ce que ça donne :
*es*******se**e**e****e***es***s***s**e***s******e*s***e***e*e*****s****e***s*e*es*****s
Pas idiot pour le moment. Si bien sûr on aboutit à une ineptie ou quelque chose qui semble curieux, on essaie autre chose : si on pense que ce n'est pas le S, on essaie le A, le N et on regarde si ce qui se forme sous nos yeux ressemble ou non à quelque chose qui peut aboutir à un texte intelligible ou si au contraire les assemblages de lettres sont vraiment trop bizarres (trop de voyelles ou trop de consonnes). Je n'aborde pas le sujet, mais si nécessaire, on peut chercher à utiliser une analyse de fréquence sur les bigrammes.

Continuons notre approche. Deux lettres suivent, le B et le S avec 10,1%. Si tout allait pour le mieux, le B représenterait le A et le S représenterait le N ou réciproquement… Essayons avec B=>A.
*es***a***sea*e**ea***e***es***s**as*ae***s*a*****e*s***e***e*e*****s*a**e***sae*es*a***s
Le bigramme ea (*sea*, *ea*) et le bigramme ae (*ae*, sae**es) sont possibles mais guère enthousiasmants, même si pour ea on peut supposer que le e termine un mot et que le a en commence un autre. Remettons l'étoile en face du B et essayons S=>A. On obtient :
*es*******se**e**e***ae**aes**as***sa*e***s****a**e*sa**e**ae*ea**a*s****e***s*e*es*****s
Encore du ae et du ea.
Le tout étant d'essayer d'accrocher un mot visuellement, si rien ne vous « parle », passez à autre chose, comme B=>N. Tout en pensant qu'il est important de placer les voyelles qui représentent plus de 40% d'un texte. Donc si B ou S => N, il est fort probable que l'autre sera une voyelle, sans doute le U si on continue d'écarter le A.
Autre chose : chercher à identifier soit des mots courts (articles, conjonctions de coordination, pronoms…), des trigrammes classiques (ent, ant…), ou des mots attendus (message, challenge, mot de passe…)
Avec B=>N, on arrive à :
*es***n***sen*e**en***e***es***s**ns*ne***s*n*****e*s***e***e*e*****s*n**e***sne*es*n***s
Pas idiot, même si le *sne* vers la fin peut surprendre, si on oublie que le s peut terminer un mot, le suivant commençant par ne. Notre nouvelle chaîne ne paraît pas stupide. Si on réessaie S=>A on voit qu'on a toujours des choses peu sympathiques.
Allez, on se lance… S=>U !
*es***n***sen*e**en**ue**ues**us**nsune***s*n**u**e*su**e**ue*eu**u*s*n**e***sne*es*n***s
Mouais.
Et si on verse le N et le U ?
*es***u***seu*e**eu**ne**nes**ns**usnue***s*u**n**e*sn**e**ne*en**n*s*u**e***sue*es*u***s
Mouais, pas mal non plus… Le premier fait apparaître *nsune* qui pourrait donc comprendre l'article une. Donc, je continue sur le premier. Mais c'est purement intuitif, et non d'une totale rationalité.

On essaie de regarder ce que donnent les voyelles. Pour l'instant, on a placé que le E et le U en clair. Reste, dans l'ordre de fréquence, le A, le I et le O. Parmi celles qui restent, les deux lettres cryptées les plus fréquentes sont, avec 6,7%, le C et le Q. Il y a de grandes chances de trouver le A dans l'une des deux. On essaie le premier venu, C=>A.
*es**an*a*sen*e**en**ue**ues**us*ansune*a*s*n**u**e*sua*e**ue*eu**u*s*n**e***sne*es*n**as
Le *sua* n'est pas trop plaisant, mais le *ansune* pourrait être « dans une ».

Si on cherche à placer nos mots courts, on pense naturellement à l'article les ou des au début de la phrase. Si on opte pour « dans une », le F crypté donnerait D et l'article du début ne serait pas « des », mais « les », le L crypté correspondant au L clair ! Je place le L (je n'ai pas encore mis le D clair) et je vois :
les**an*a*sen*e**en**uel*ues**us*ansune*a*s*n**u**e*sua*e**ue*eu**u*s*n**e***sneles*n**as
Le *uel*ues* attire l'œil, d'autant que les deux lettres manquantes correspondent au D crypté, et qu'il n'y en a que quatre dans le texte. Ca vous parle ? Moi ça me fait immanquablement penser à « quelques ». On teste, et on voit que les deux autres Q se placent bien avant un U, donnant que pour le premier, et donc qui pour le second (plutôt que quo car quos… ne parle pas vraiment). Allez, on fonce : T=> I
les**an*aisen*e**en*quelques**us*ansune*ais*n**u**e*sua*e*que*eu*quis*n**e***sneles*n**as
On n'a pas encore placé le D clair (« dans une » ?) et il reste le O clair également (les voyelles !).
On a deux fois s*n, où on glisserait bien un O. Est-ce la même lettre cryptée à chaque fois ? Oui, Q crypté !
Yop, Q=> O :)
les**an*aisen*e**en*quelques*ous*ansune*aison*ou**e*sua*e*que*eu*quison**e*o*sneleson**as

Dans les lettres fréquentes, il nous manque le T clair qu'on voit bien après le quison* (qui sont) voire aussi à la fin de la phrase avec un « ne le sont » qui se dessine, laissant même entrevoir un « ne le sont pas ».
Le T d'abord, K=>T et U=>P dans l'excitation du moment…
les**an*aisen*e**entquelques*ous*ansune*aisonpou*pe*sua*e*que*eu*quisont*e*o*snelesontpas

Ca commence à ressembler quelque chose, et il nous reste des lettres fréquentes à placer comme le R. Ah, notre D de « dans une maison », pour voir. F=>D.
les**an*aisen*e**entquelques*ousdansune*aisonpou*pe*suade*que*eu*quisontde*o*snelesontpas

Le « pou*pe*suade* » avec le G crypté à chaque fois ne laisse pas de doute : G=>R va donner « pour persuader ».
les*ran*aisen*er*entquelques*ousdansune*aisonpourpersuaderque*eu*quisontde*orsnelesontpas

A ce stade, comme pour le R clair, on peut chercher à placer des lettres courantes au pifomètre, ou a deviner des mots. La fin de la phrase « quisontde*orsnelesontpas » montre que le N crypté équivaut au H clair (malheureusement il n'y en a pas d'autre dans le texte…). Et petit à petit, ce qui manque se dessine intuitivement.
« que*eu*quisontdehors… » ne peut être, compte tenu des lettres restantes, que « que ceux qui… ».
I=>C et X=>X ;
les*rancaisen*er*entquelques*ousdansune*aisonpourpersuaderqueceuxquisontdehorsnelesontpas
Le reste vient tout seul avec le M de maison qu'on avait déjà en tête et le F.
P=>M et E=>F.
lesfrancaisenfermentquelquesfousdansunemaisonpourpersuaderqueceuxquisontdehorsnelesontpas
Les français enferment quelques fous dans une maison pour persuader que ceux qui sont dehors ne le sont pas.
C'est ce brave Montesquieu qui aurait écrit ça :)

LRJEGCBICTJRBERGPRBKDSRLDSRJEQSJFCBJSBRPCTJQBUQSGURGJSCFRGDSRIRSXDSTJQBKFRNQGJBI										
89										
**********										
					CRYPTE			CLAIR		Solution
1	L	*			A	0	0,0%	*		*
2	R	*			B	3	10,3%	*		n
3	J	*			C	2	6,9%	*		a
4	E	*			D	2	6,9%	*		q
5	G	*			E	3	10,3%	*		f
6	C	*			F	0	0,0%	*		d
7	B	*			G	2	6,9%	*		r
8	I	*			H	0	0,0%	*		*
9	C	*			I	1	3,4%	*		c
10	T	*			J	3	10,3%	*		s
11	J	*			K	1	3,4%	*		t
12	R	*			L	2	6,9%	*		l
13	B	*			M	0	0,0%	*		*
14	E	*			N	0	0,0%	*		h
15	R	*			O	0	0,0%	*		*
16	G	*			P	1	3,4%	*		m
17	P	*			Q	0	0,0%	*		o
18	R	*			R	6	20,7%	*		e
19	B	*			S	2	6,9%	*		u
20	K	*			T	1	3,4%	*		i
21	D	*			U	0	0,0%	*		p
22	S	*			V	0	0,0%	*		*
23	R	*			W	0	0,0%	*		*
24	L	*			X	0	0,0%	*		x
25	D	*			Y	0	0,0%	*		*
26	S	*			Z	0	0,0%	*		*
27	R	*								
28	J	*				**29**				
29	E	*								