

Concours : _____

Nature : _____

Admissibilité : épreuve écrite

NOM patronymique (nom de naissance) : _____
(en MAJUSCULE)

NOM d'usage : _____

Prénoms : _____

Académie : PARIS



DRH – Service des concours - RH3
12 Rue de l'école de médecine
75270 PARIS CEDEX 06

Cadre réservé à l'anonymat

ACADEMIE DE PARIS – 1 Poste

Centre organisateur : Université Paris Descartes

Académie de Paris

**Concours externe Technicien de classe Supérieure
de recherche et de formation**

BAP D

Emploi type : Production & Analyse de données

Session 2014

Epreuve écrite d'admissibilité

Durée de l'épreuve : 3 heures – Coefficient 3

Date de l'épreuve : Mardi 3 juin 2014

Faculté de Médecine Site des Cordeliers – Salle : Pavillon 1

15 rue de l'école de médecine 75006 PARIS

Le sujet que vous devez traiter comporte en plus de cette page, 20 pages numérotées de 2 à 21.

Assurez-vous que cet exemplaire est complet.

S'il est incomplet, demandez un autre exemplaire au surveillant.

Important : Ce sujet ne doit pas être détaché.

L'usage de tout document est strictement interdit.

La calculatrice est autorisée.

L'usage du téléphone portable est interdit. Il doit être éteint et laissé dans vos sacs.

Vos sacs seront déposés sur le côté de la salle hormis le matériel dont vous avez besoin.

Répondre directement sur ce document (18 pages)

Cadre réservé à l'anonymat

Barème indicatif :

Exercice n°1	2
Exercice n°2	4
Exercice n°3	4
Exercice n°4	5
Exercice n°5	3
Exercice n°6	2

Exercice 1 :

Quelles sont les grandes étapes d'une enquête ?

Qu'est-ce que la CNIL? A quoi sert-elle?

Exercice 2 :

Partie A :

Une enquête sur les pratiques culturelles a été conduite sur un échantillon de 800 personnes actives (âgées entre 18 et 65 ans). Parmi les différents enseignements de cette enquête, on a notamment relevé les faits suivants :

Pour s'informer, 30% des personnes préfèrent consulter internet, 38% regarder la télévision et 32% un autre média.

67% des personnes disposent d'une connexion internet à leur domicile.

Il a aussi été demandé lors du sondage le niveau d'études (« Sans diplôme », « Bep », « Bac » ou « Études supérieures ») et le temps hebdomadaire (exprimé en heures) passé devant la télévision.

1. Identifier l'individu et l'échantillon visés par l'enquête.

2. Quelles variables (observées sur ces individus) peut-on identifier à partir de l'énoncé ci-dessus?

3. Quelle est la nature de chacune de ces variables? Indiquer également les modalités pour les variables qualitatives uniquement.

4. Combien de personnes (dans l'échantillon) ne disposent pas d'une connexion internet à leur domicile?

5. Parmi les personnes ayant fait des « études supérieures », on a noté la répartition suivante concernant le média préféré pour s'informer : 35% préfèrent consulter internet, 31% regarder la télévision et 34% un autre média.

A l'aide de variables aléatoires identifiées à la question 2, indiquer précisément quelle distribution est représentée ici.

6. Sachant que l'échantillon comprend 300 personnes ayant fait des « études supérieures », calculer les effectifs de la distribution représentée à la question précédente.

7. Dans cette enquête, on a également observé la variable « Nombre de livres lus au cours des 12 derniers mois ». On a alors calculé le coefficient de corrélation linéaire entre cette variable et la variable « Temps hebdomadaire passé devant la télévision » et on a obtenu un coefficient égal à -0,9.

Quelle interprétation pouvez-vous faire de cette valeur?

Partie B :

Lors de la même enquête, on a étudié le temps hebdomadaire, exprimé en heures, passé devant la télévision en fonction du niveau d'études. L'échantillon est constitué de $n = 800$ personnes actives. Notant $X =$ « Niveau d'études » et $Y =$ « Temps hebdomadaire passé devant la TV », le tableau de contingence en effectifs est le suivant :

$X \setminus Y$	$[0, 10[$	$[10, 20[$	$[20, 25[$	$[25, 30[$
Sans diplôme	15	30	n_{13}	n_{14}
Bep	45	80	68	57
Bac	33	58	35	36
Études supérieures	72	103	72	53

1. Quelle est la distribution représentée dans ce tableau?

2. Parmi les personnes « sans diplôme », il y en 25% qui passent plus de 25 heures devant la télévision. Calculer alors les 2 effectifs n_{13} et n_{14} .

3. Donner la distribution conditionnelle (en fréquences) du « Temps hebdomadaire passé devant la TV » chez les personnes ayant fait des études supérieures. Indiquer également la distribution cumulée (en fréquences) associée.

4. Calculer alors le 1er quartile pour la variable « Temps hebdomadaire passé devant la TV » chez les personnes ayant fait des études supérieures.

5. Indiquer (en justifiant) la proportion des personnes ayant le niveau « Bac » parmi celles qui passent entre 10h et 20h devant la TV.

6. Calculer le temps hebdomadaire moyen passé devant la TV chez les personnes ayant le niveau « Bep ». Calculer alors la variance de la variable Y chez ces mêmes personnes.

7. Le temps hebdomadaire moyen passé devant la TV chez les personnes « sans diplôme » est $\bar{y}_1 = 18,21$. Pour les niveaux d'études « Bac » et « Études supérieures », on a respectivement les moyennes suivantes $\bar{y}_3 = 17,36$ et $\bar{y}_4 = 16,61$. En utilisant la formule de la décomposition de la moyenne, en déduire le temps hebdomadaire moyen passé devant la TV dans l'ensemble de l'échantillon.

Exercice 3 :

L'évolution de la population municipale au recensement de la région Île-de-France entre 1962 et 2011 a permis de construire le tableau suivant :

Année (x_i)	1962	1968	1975	1982	1990	1999	2006	2011
Population en millions (y_i)	8,47	9,25	9,88	10,07	10,66	10,95	11,53	11,85

(Source : INSEE)

1. Calculer le pourcentage d'évolution de la population de cette région entre 1962 et 2011.

2. Représenter le nuage de points associé à la série statistique ($x_i ; y_i$).

3. Le nuage de points montre qu'un ajustement affine est justifié.
Donner une équation de la droite d'ajustement affine D de y en x , obtenue par la méthode des moindres carrés. (*Les résultats obtenus à la calculatrice ne seront pas justifiés*).
Représenter D dans le repère précédent.

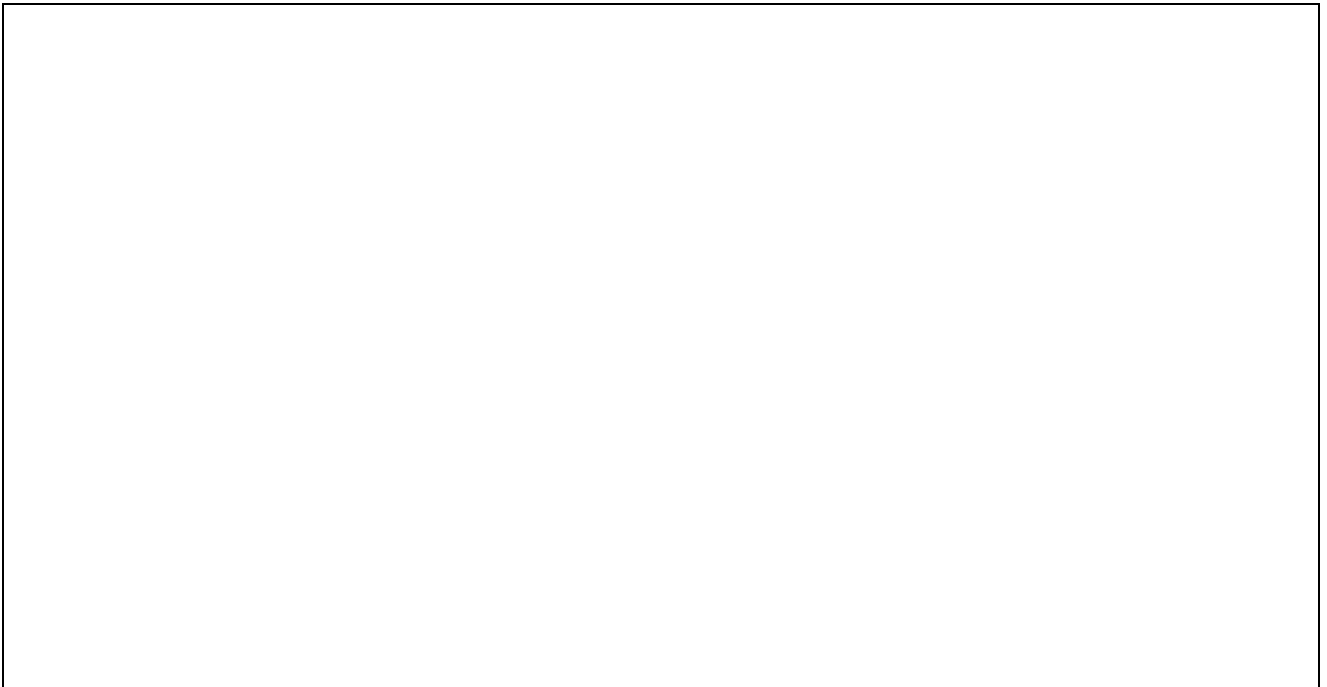
4. Quelle prévision ferait-on avec cette approximation pour la population de la région en l'an 2020 ?

5. On envisage un autre modèle pour prévoir l'augmentation de la population de la région Île-de-France. On suppose qu'à partir de 2011, la population augmentera de 1,5% par an. On désigne par u_n le nombre d'habitants, en millions, de l'année (2011 + n). On a donc $u_0 = 11,85$.

a. Calculer le nombre d'habitants en 2020.



b. En quelle année la population de cette région dépassera 15 millions d'habitants ?



Exercice 4 :

Le tableau ci-dessous est le tableau de contingence des deux variables SANG et SEXE (O_{jk}).

SANG	SEXE		Total
	Masculin	Féminin	
A	12	.	73
B	6	16	.
AB	1	.	9
O	.	52	77
Total	44	137	.

Partie A :

1. Compléter les cellules vides du tableau de contingence et les marges.
2. Quelle est la proportion :
 - De femmes dans cet échantillon ?

- D'hommes de groupe sanguin « A » dans cet échantillon ?

- D'hommes de groupe sanguin « A » parmi l'ensemble des hommes de cet échantillon ?

3. Les individus dont le groupe sanguin est « O » sont-ils relativement plus nombreux lorsqu'il s'agit d'individus de sexe masculin ou féminin ?

Partie B :

Nous souhaitons vérifier si la répartition du groupe sanguin entre les 4 classes dépend du sexe des individus. À cet effet, considérons les 181 données de notre échantillon.

1. Quel test devons-nous utiliser pour répondre à cette question sur la base de notre échantillon ci-dessous.

2. Qu'appelle-t-on effectif théorique (ou effectif d'indépendance) d'une modalité conjointe ?

3. Déterminez le tableau des effectifs théoriques (E_{jk}) :

SANG	SEXE		Total
	Masculin	Féminin	
A			73
B			22
AB			9
O			77
Total	44	137	181

4. Déterminez le tableau des carrés des écarts entre les effectifs observés et théoriques divisés par les effectifs attendus $((O_{jk}-E_{jk})^2 / E_{jk})$:

SANG	SEXE	
	Masculin	Féminin
A		
B		
AB		
O		

5. Que vaut la statistique de test $\chi^2_{\text{calculé}}$:

6. Combien de degrés de liberté (d.d.l.) compte la loi Khi-carré que nous utilisons ici ?

7. On fournit la table de Khi² ci-dessous :

Table de Khi-Deux χ^2

v / a	0.1	0.05	0.025	0.01	v / a	0.1	0.05	0.025	0.01
1	2.71	3.84	5.02	6.63	16	23.54	26.30	28.84	32.00
2	4.61	5.99	7.38	9.21	17	24.77	27.59	30.19	33.41
3	6.25	7.81	9.35	11.34	18	25.99	28.87	31.53	34.80
4	7.78	9.49	11.14	13.28	19	27.20	30.14	32.85	36.19
5	9.24	11.07	12.83	15.09	20	28.41	31.41	34.17	37.57
6	10.64	12.59	14.45	16.81	21	29.61	32.67	35.48	38.93
7	12.02	14.07	16.01	18.47	22	30.81	33.92	36.78	40.29
8	13.36	15.51	17.53	20.09	23	32.01	35.17	38.08	41.64
9	14.68	16.92	19.02	21.67	24	33.20	36.41	39.37	42.98
10	15.99	18.31	20.48	23.21	25	34.38	37.65	40.65	44.31
11	17.27	19.67	21.92	24.72	26	35.56	38.88	41.92	45.64
12	18.55	21.03	23.34	26.22	27	36.74	40.11	43.19	46.96
13	19.81	22.36	24.74	27.69	28	37.92	41.34	44.46	48.28
14	21.06	23.68	26.12	29.14	29	39.09	42.56	45.72	49.59
15	22.31	25.00	27.49	30.58	30	40.26	43.77	46.98	50.89

v : le nombre de degré de liberté. a : le risque d'erreur.

Si nous fixons le risque d'erreur à 5 %, à partir de quelle valeur de $\chi^2_{\text{théorique}}$ pourrions-nous rejeter l'hypothèse que les distributions du groupe sanguin et du sexe sont indépendantes ?

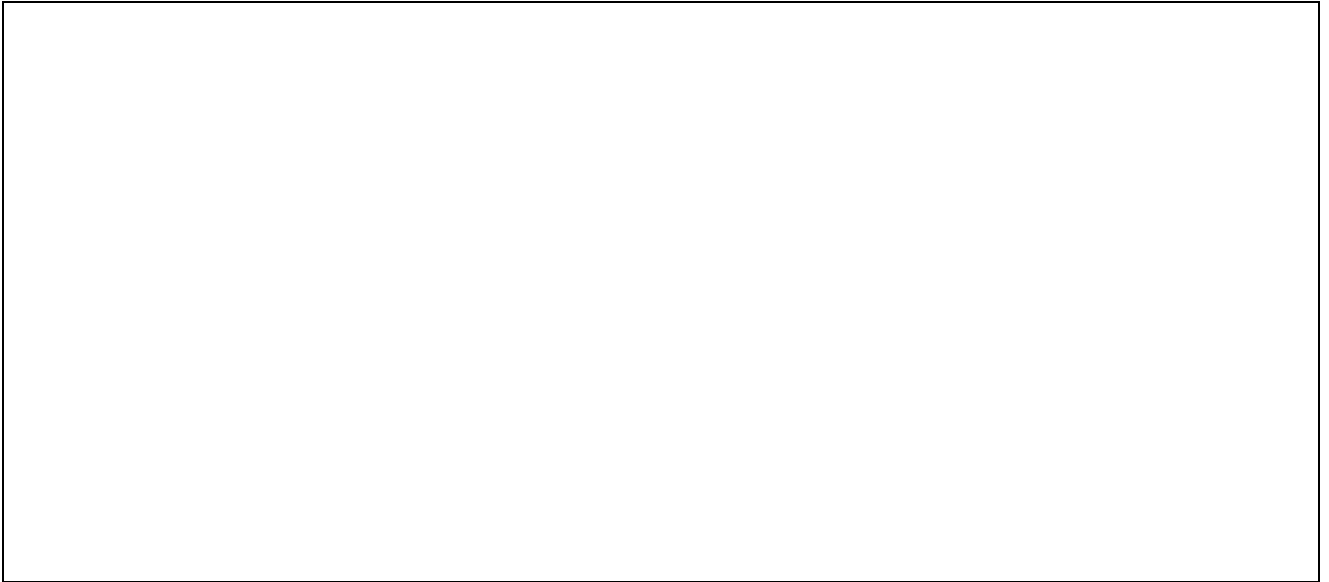
8. Que pouvez-vous conclure ?

9. Après avoir listé les conditions d'application d'un test Khi-carré, vérifiez qu'elles sont bien remplies dans le cas de cet échantillon.

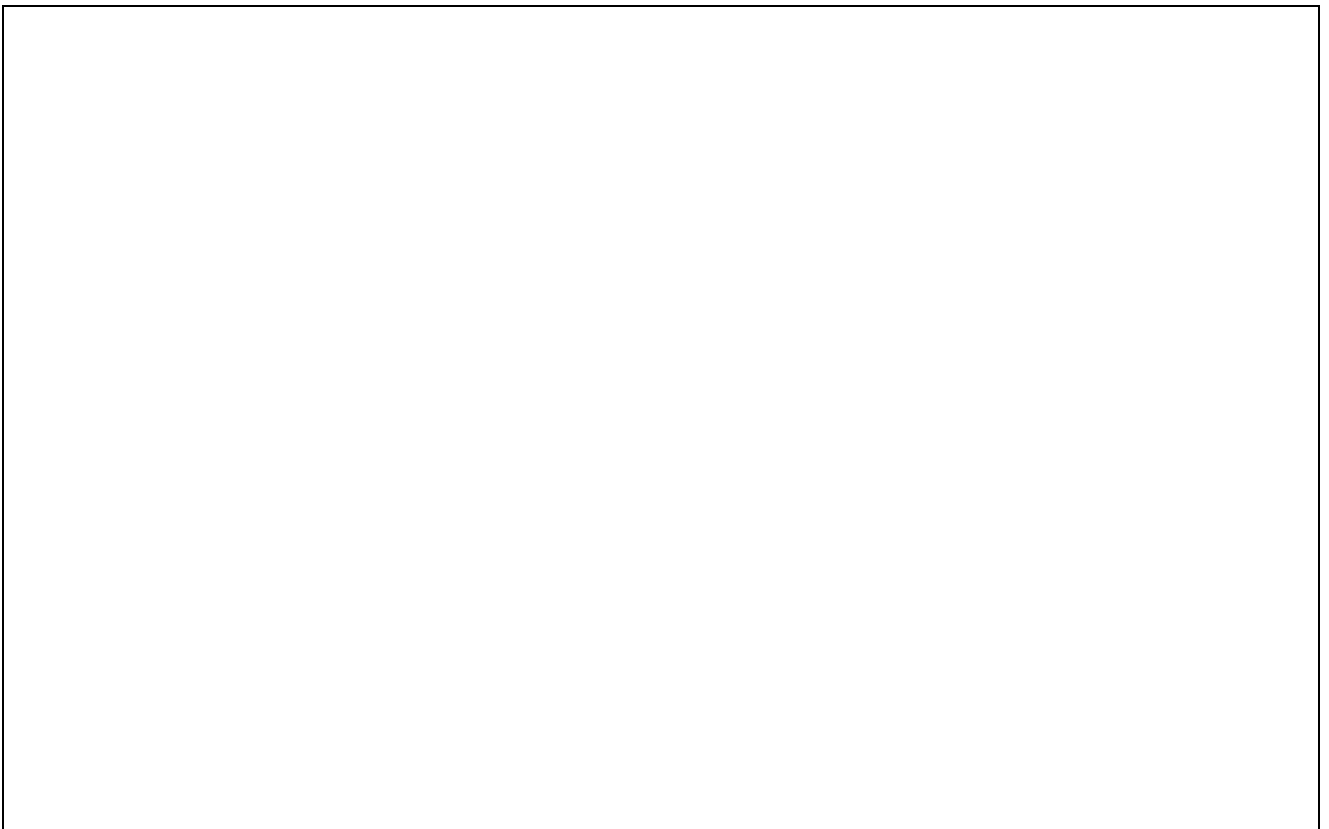
Exercice 5 :

On effectue des tirages sans remise dans une urne contenant initialement 3 boules rouges et 3 boules noires jusqu'à obtenir une boule noire. On appelle X le numéro du tirage de cette boule noire (ainsi, $X = 1$ si la première boule tirée est noire).

1. Quelles valeurs peut prendre la variable aléatoire X ? Avec quelles probabilités?



2. Représenter sa fonction de répartition F .



3. Calculer l'espérance et la variance de X .

Exercice 6 : Questionnaire à choix multiple (QCM)

Il peut y avoir une ou plusieurs bonnes réponses. Pour chaque question, les réponses correctes sont comptées positivement, les réponses fausses sont comptées négativement. Si le bilan de la question est négatif il est ramené à 0. Ne rien cocher vaut 0.

Que désigne Intranet ?

- Un réseau informatique utilisé à l'intérieur d'une entreprise utilisant les techniques de communication d'internet
- Une technique de sauvegarde de données dans une entreprise, sur un serveur spécialisé
- Un réseau de fibres optiques installé dans un immeuble
- Un code de bonne conduite au sein d'une entreprise
- Un ensemble de logiciels libres

Quelle est la caractéristique d'un CD-R ?

- Il ne peut être gravé qu'une seule fois
- Il ne peut contenir que de la musique
- Il peut être gravé plusieurs fois
- Il ne peut être que lu
- Il ne peut être que gravé

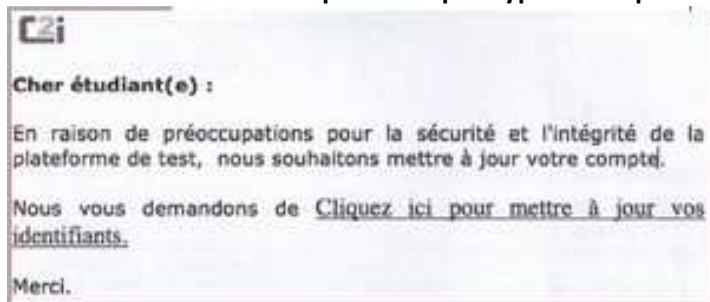
Quelle(s) affirmation(s) concernant les systèmes d'exploitation est/sont vraie(s)?

- Chaque fabricant d'ordinateurs a son propre système d'exploitation.
- Le système d'exploitation est le programme informatique de base indispensable au bon fonctionnement d'un ordinateur.
- On peut installer plusieurs systèmes d'exploitation dans des partitions différenciées du disque dur.
- Tous les systèmes d'exploitation sont des logiciels propriétaires.

Vous travaillez sur un document et le message suivant apparaît: « Espace disque Insuffisant pour enregistrer ce document; veuillez libérer de l'espace en supprimant des fichiers ». Quels risques encourrez-vous?

- Le disque dur risque d'être physiquement endommagé en raison de la trop lourde charge.
- Sans libérer de place, les fichiers les plus anciens présents sur votre disque dur risquent d'être supprimés automatiquement pour libérer de la place.
- Sans libérer de place, les dernières modifications des fichiers actuellement utilisés risquent d'être perdues.
- Les dernières modifications des fichiers en cours d'utilisation risquent d'écraser des fichiers système.
- La situation ne peut pas se produire, le système d'exploitation prévoit toujours une solution de secours.

Le courriel ci-dessous fait penser à quel type d'attaque ?



- Un hameçonnage (phishing).
- Un courriel normal.
- Un canular (hoax).
- Une diffamation.

Pour se protéger d'éventuelles intrusions d'un pirate sur son ordinateur, que doit-on installer ?

- Un antivirus
- Un logiciel de contrôle parental
- Un espion
- Un plan vigipirate
- Un pare-feu

Quelles sont les informations personnelles considérées comme sensibles et qui requièrent une dérogation pour figurer dans un fichier informatisé ?

- L'appartenance syndicale
- Le prénom
- L'origine ethnique
- Le numéro de carte bleue
- L'opinion religieuse

Quels sont les dispositifs qui sont susceptibles de conserver des traces de la navigation d'un internaute après qu'il ait éteint son ordinateur ?

- L'historique de navigation du navigateur.
- Le cache du navigateur.
- Les cookies.
- La mémoire vive de l'ordinateur.

Lorsque le contenu d'un document est confidentiel, que peut-on faire pour éviter toute indiscretion?

- Mettre le fichier en lecture seule.
- Protéger le fichier en écriture par un mot de passe.
- Protéger le fichier en lecture par un mot de passe.
- Convertir le fichier au format PDF.

Parmi les supports suivants, lesquels sont des supports de sauvegarde amovibles?

- Le CD-ROM.
- La clé USB.
- L'espace de stockage de l'ENT.
- Le disque dur externe.
- La RAM.

Si Pierre met un commentaire sur le blog de Jean, qui est responsable des propos de Pierre?

- L'hébergeur du blog de Jean.
- Pierre, l'auteur du commentaire.
- Jean, l'éditeur du blog.
- Le fournisseur d'accès à Internet de Pierre.

Comment appelle-t-on un règlement intérieur définissant les règles d'utilisation des ressources numériques?

- La Netiquette.
- La charte d'établissement.
- Les règles d'accessibilité.
- Le hoax.

Dans un document texte, que faut-il faire pour pouvoir générer automatiquement une table des matières?

- Il faut insérer un saut de page avant chaque titre.
- Il faut numéroter les pages.
- Il faut justifier les paragraphes.
- Il faut structurer le document à l'aide de styles, de niveaux de plan ou d'entrées marquées.

Dans un tableur, si on recopie la formule « =A1+A\$2 » d'une cellule vers le bas, quelle formule obtient-on?

- = A1+A\$3
- = A2+A\$3
- = A2+A\$2
- = B1+B\$2
- = A1+A\$2

Dans un tableur, quelles sont les formules qui réalisent la moyenne des nombres contenus dans les cellules allant de A1 à A4?

- =(A1+A2+A3+A4)/4
- =MOYENNE(A1;A4)
- =MOYENNE(A4:A1)
- = A1+A2+A3+A4/4
- =MOYENNE(A1-A4)

Dans un tableur, que se passe-t-il si le résultat d'une formule affiche « ##### »?

- Le format de nombre n'est pas compatible avec la donnée.
- La colonne est trop étroite pour afficher le résultat tout entier.
- La formule contient une erreur.
- La formule n'a pas de sens du point de vue mathématique.
- Il n'y a pas assez de mémoire pour que le logiciel fasse le calcul.

Dans une présentation sur ordinateur, vers quels types d'objets peut pointer un lien hypertexte?

- Vers une adresse de messagerie électronique.
- Vers une page Web.
- Vers une vidéo.
- Vers une autre diapositive de la présentation.

Dans l'URL <http://catalogue.bnf.fr/>, quel est le nom de domaine?

- catalogue
- bnf
- fr
- catalogue.bnf.fr
- bnf.fr

Dans une page web, il y a un lien vers un fichier au format PDF. Comment faire pour télécharger ce fichier sur son poste?

- On enregistre l'URL.
- On enregistre la page html seule.
- On enregistre la cible du lien.
- On enregistre la page web complète.

Comment procéder pour envoyer le même courriel à trois personnes sans qu'elles connaissent l'existence des autres destinataires?

- On place une adresse différente dans chaque champ : A, Cc et Cci.
- On place toutes les adresses dans le champ A.
- On place toutes les adresses dans le champ Cc.
- On place toutes les adresses dans le champ Cci.