

DATE :21/12/2020-5/1/2021.....
 EPREUVE: Statistiques et informatique appliquées aux Sciences sociales
 ENSEIGNANT: ...Renaud Orain <renaud.orain@univ-paris1.fr>
 DIPLOME : Licence de Sciences sociales, 3^e année
 Étudiant-es de tous parcours, inscrit-es en contrôle continu...
 DUREE conseillée: ...2. h
 DOCUMENT AUTORISE Devoir à la maison (tous documents autorisés)

Sujet obligatoire pour les étudiant-es inscrit-es en contrôle continu

Ce devoir devra être rendu sous forme manuscrite, en main propre ou en version scannée ou photographiée sur le devoir. En cas d'impossibilité, prière de me contacter au moyen de l'adresse de courriel indiquée ci-dessus et de contacter votre chargé de TD en expliquant bien la situation.

Il vous est demandé de traiter les questions qui suivent, en faisant les calculs « à la main » (calculatrices autorisées, y compris sur ordinateur) en indiquant si possible le temps nécessaire pour réaliser le devoir. Cette mention ne sera pas prise en compte dans l'évaluation. Elle vise à mieux calibrer les sujets futurs.

Distribution d'âge et sorties culturelles

Les étudiants en Master Démographie en 2019-2021 ont construit un questionnaire portant sur les sorties et pratiques culturelles des jeunes mais ce questionnaire est accessible en ligne à quiconque. L'échantillon n'est donc pas tiré aléatoirement à partir d'une base de sondage dans la population cible (jeunes résidant en Ile de France, mais aussi dans d'autres régions), mais d'un appel à participation à partir de plusieurs réseaux sociaux au sens large. De plus, l'enquête est encore en cours au moment des calculs. Dans un premier temps, nous voulons créer une variable d'âge exploitable pour ensuite regarder comment varient les pratiques en fonction de cette variable. La question posée est « Quelle est votre année de naissance ? » avec la modalité et la répartition ci-dessous parmi les personnes ayant répondu au questionnaire avant le 16/12/2020 :

A02Année de naissance (codes)	Année de Naissance (libellés)	Classe d'âge (au 31/12)	Âge moyen dans chaque classe	Effectifs	Somme des âges dans chaque classe	Somme des arrés des âges dans chaque classe
01	2002 ou après	[17,5;19[18,25	11	200,75	3663,6875
02	2001	[19;20[19,5	10	195	3802,5
03	2000	[20;21[20,5	12	246	5043
04	1999	[21;22[21,5	21	451,5	9707,25
05	1998	[22;23[22,5	45	1012,5	22781,25
06	1997	[23;24[23,5	32	752	17672
07	1996	[24;25[24,5	21	514,5	12605,25
08	1995	[25;26[25,5	16	408	10404
09	1994	[26;27[26,5	12	318	8427
10	1993	[27;28[27,5	6	165	4537,5
11	1992	[28;29[28,5	7	199,5	5685,75
12	1991	[29;30[29,5	8	236	6962
13	1990	[30;31[30,5	9	274,5	8372,25
15	1989 ou avant	[31;40[35,5	43	1526,5	54190,75
Total	Ensemble	[17,5;40]	/	253	6499,75	173854,1875
Z	Ne sait pas	Inconnue	/	2	NB : pas de refus	

- Tracez l'histogramme correspondant à ces données d'âge au 31 décembre et commentez le cas échéant le choix des bornes de 17,5 ans et 40 ans.
- Quelle est la médiane, quels sont les premiers et troisième quartiles de cette distribution ?
- En faisant comme si chaque individu avait l'âge moyen de sa classe, calculez la moyenne d'âge au 31/12/2020 (de l'ordre de 25-26 ans) , l'écart-type (de l'ordre de 5 ans) et le coefficient de variation (écart-type divisé par la moyenne) de cette distribution. Attention à bien tenir compte de l'effectif de chaque classe.

On se demande comment varie le nombre de sorties en 2019 en fonction de l'âge en milieu d'année. Ceci revient à décaler les âges de 1,5 ans dans le passé. On se restreint aux personnes ayant répondu sur leur année de naissance et on évalue les corrélations de l'âge avec un décompte de différent type de sortie faites dans l'année.

Coefficients de corrélation de Pearson, N = 253 Proba > r sous H0: Rho=0							Mo- yennes	Ec- type	Coefficients de variation
Âge Nb sorties	Age mi 2019	1:cinéma	2 :Concert	3 :Musee	4 :Monuments	5 : Cirque, théâtre			
Age mi 2019	1,00000	-0,03598 0,5689	0,12422 0,0484	-0,06295 0,3186	-0,02928 0,6430	0,15574 0,0131	24,19	5,21	0,22
1:cinéma	-0,03598 0,5689	1,00000	0,21627 0,0005	0,31075 <,0001	0,23937 0,0001	0,29990 <,0001	7,15	6,60	0,92
2 :Concert	0,12422 0,0484	0,21627 0,0005	1,00000	0,15306 0,0148	0,18197 0,0037	0,33887 <,0001	3,40	4,96	1,46
3 :Musee	-0,06295 0,3186	0,31075 <,0001	0,15306 0,0148	1,00000	0,60333 <,0001	0,42451 <,0001	6,80	7,03	1,03
4 :Monu- ments	-0,02928 0,6430	0,23937 0,0001	0,18197 0,0037	0,60333 <,0001	1,00000	0,29076 <,0001	5,49	5,76	1,05
5 : Cirque, théâtre	0,15574 0,0131	0,29990 <,0001	0,33887 <,0001	0,42451 <,0001	0,29076 <,0001	1,00000	2,33	3,60	1,54

- À l'aide des sommes de produits ci-dessous, expliquer les étapes du calcul de la corrélation entre l'âge et le nombre de sorties au cirque ou au théâtre dans l'année.

Somme des produits	Âges mi 2019	Cirque, théâtre	Somme	Effectif
Âges mi 2019	154 924,2	15 010,8	6120,25	253
Cirque, théâtre	15 010,8	4 646,0	590	253

- Déterminer les coefficients et le coefficient de détermination dans la régression linéaire du nombre de sorties au cirque et au théâtre sur l'âge.
- Commentez ces résultats. Vous pourrez utiliser le graphique ci-dessous pour visualiser l'ajustement linéaire présenté à la question précédente.

