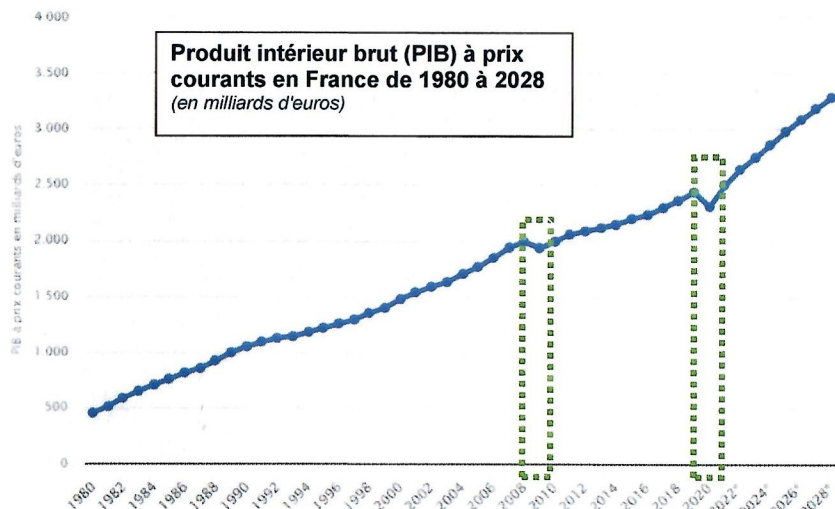


CONTRÔLE DE MI-SEMESTRE DE STATISTIQUES

Exercice 1

Cet exercice s'intéresse au produit intérieur brut (PIB) annuel de la France entre 1980 et 2023, avec des prévisions allant jusqu'en 2028, exprimé à prix courants et en milliards d'euros (source : Statistica). Le PIB à prix courants, ou PIB nominal, indique le PIB d'un pays en valeur de l'année prise. Il s'agit ainsi d'un PIB sans prise en compte de l'inflation et de la déflation. *prévisionnel



1.1. Commentez le graphique ci-dessus

Ce graphique représente l'évolution du PIB à prix courants, en milliards d'euros, en France de 1980 à 2028 (prévisionnels à partir de 2022). De façon générale, on observe une augmentation quasi constante du PIB entre 1980 et 2028. Une première diminution est présente en 2009 et correspond à la récession faisant suite à la crise économique mondiale de 2008. Une seconde diminution est présente en 2020 suite à la crise du coronavirus en 2019. Cette seconde récession (-8,3%) est bien plus importante que celle de 2009 (à titre informatif, c'est la plus importante récession mondiale depuis la Seconde Guerre Mondiale). Après la crise COVID, les données prévoient toutefois une accélération (une augmentation plus importante) du PIB en raison notamment des desserments progressifs des mesures de restriction et des différents plans de relance et soutien aux entreprises.

Voici les données des dernières années :

Date	PIB prix courants
2019	2 437,64
2020	2 309,61
2021	2 498,82
2022	2 641,69
2023	2 750,65

1.2. Calculer $\theta_{4|2}$ et interpréter

$$\theta_{4|2} = \frac{y_4}{y_2} = \frac{2\,641,69}{2\,309,61} = 1,14$$

Le PIB à prix courants en France a été multiplié par 1,14 entre 2020 et 2022, cela correspond à une augmentation de 14%.

* Points comptés avec la formule en X car pas encore dit que y ici

1.3. Remplir le tableau suivant. Vous détaillerez et interpréterez la case grisée

Date	t	Y_t	$\gamma(Y_t)$	En %
2019	1	2 437,64	/	/
2020	2	2 309,61	0,053	5,25%
2021	3	2 498,82	0,082	8,19%
2022	4	2 641,69	0,057	5,72%
2023	5	2 750,65	0,041	4,12%

$$\gamma(Y_t) = \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{2\,641,69 - 2\,498,82}{2\,498,82} = 0,0572 = +5,72\%$$

Le PIB à prix courants en France a augmenté de 5,72% entre 2021 et 2022.

1.4. Calculer le taux de croissance moyen de y sur la période et interpréter

$$\overline{\gamma(Y_t)} = \frac{1}{T-1} \times \sum_{t=2}^T \gamma(Y_t) = \frac{-0,053 + 0,082 + 0,057 + 0,041}{4} = \frac{0,128}{4} = 0,032 = +3,2\%$$

En moyenne, le PIB à prix courants en France a augmenté de 3,2% entre 2019 et 2023 (et ce, malgré la crise sanitaire de COVID-19).

Exercice 2

Nous allons maintenant nous intéresser à la variable Population Urbaine, notée x_t . La population urbaine d'un pays correspond à la population nationale habitant dans des villes de plus de 50 000 habitants. Voici les données de 2020 issues de la Banque Mondiale, en milliers, dans l'ordre croissant :

Nom du pays	Abréviation du pays (code CIO)	Population urbaine 2020	x_t
Islande	ISL	344,10	.
Érythrée	ERI	1 470,25	.
Oman	OMN	3 919,86	.
Koweït	KWT	4 360,44	.
Haïti	HTI	6 454,83	.
Bélarus	BLR	7 455,47	.
République tchèque	CZE	7 922,94	.
Émirats arabes unis	ARE	8 084,40	.
Grèce	GRC	8 528,39	.
Israël	ISR	8 531,99	.
République dominicaine	DOM	9 079,12	.
Kazakhstan	KAZ	10 816,58	.
Pologne	POL	22 755,74	.
Maroc	MAR	23 309,11	.
Malaisie	MYS	25 617,12	.
Pérou	PER	26 076,63	.
Iraq	IRQ	30 169,92	.
Canada	CAN	30 999,41	.
Algérie	DZA	32 038,22	.
Espagne	ESP	38 276,19	.
Colombie	COL	41 470,29	.
Argentine	ARG	41 796,99	.
Corée, République de	KOR	42 201,96	.
Italie	ITA	42 224,77	.
Égypte, République arabe d'	EGY	45 976,81	.
Philippines	PHL	53 187,50	.
France	FRA	54 715,70	.
Royaume-Uni	GBR	56 283,17	.
Allemagne	DEU	64 410,59	.
Mexique	MEX	101 719,69	.
Japon	JPN	115 884,87	.
Indonésie	IDN	153 983,07	.

2.1. Remplir le tableau suivant

Population	Les pays du monde
Echantillon	32 pays du monde
Unité statistique	1 pays
Variable	Population urbaine (X, en milliers d'habitants)
Type de variable	Quantitatif continue

2.2 Remplir le tableau suivant et détailler les calculs pour $i = 3$

	Fréquences absolues (= effectifs)	Fréquences relatives (en %)	Fréquences relatives cumulées (en %)
[0 ; 40 000)	20	62.5%	62.5%
[40 000 ; 80 000)	9	28.1%	90.6%
[80 000 ; 120 000)	2	6.3%	96.9%
[120 000 ; 160 000)	1	3.1%	100.0%
TOTAL	32	100%	

$i = 3$ correspond à la troisième ligne du tableau soit la classe [80 000 ; 120 000)

$$n_3 = 2$$

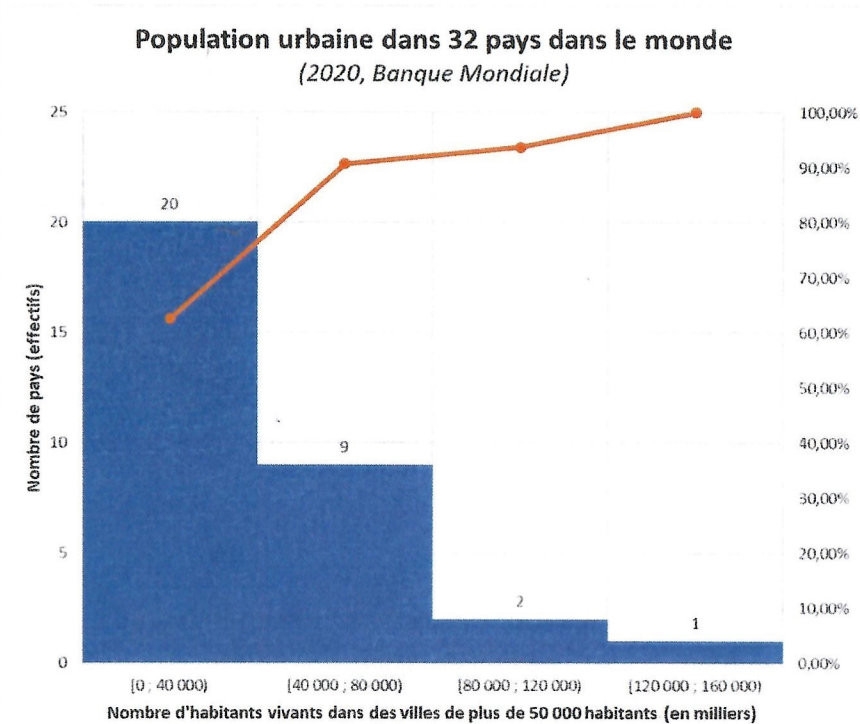
$$f_3 = \frac{n_3}{N} = \frac{2}{32} = 6.3\%$$

Il y a 2 pays dans notre échantillon dont la population urbaine est comprise entre 80 000 000 et 120 000 000 habitants : le Mexique et le Japon, ce qui correspond à 6.3% des 32 pays de l'échantillon.

$$F_3 = \sum_{s=1}^3 f_s = f_1 + f_2 + f_3 = 62.5 + 28.1 + 6.3 = 96.9\%$$

Sur les 32 pays de l'échantillon, 96.9% ont une population urbaine strictement inférieure à 120 000 000 habitants.

2.3. Représenter la distribution de ces données dans le format adéquat et interpréter



En 2020, plus de la majorité des pays de l'échantillon (62.5%) ont une population urbaine entre 0 et 40 000 000 habitants (20/32) et seuls 3 pays (9.4%) ont plus de 80 000 000 personnes résidant dans des villes de plus de 50 000 habitants.

*Note pour plus tard : On voit donc ici que la distribution semble oblique à gauche et étalée à droite : il y a donc une forte concentration des pays avec un faible nombre d'habitants dans des villes de plus de 50 000 habitants. Si on devait calculer b1, on s'attend à ce qu'il soit positif.

Exercice 3

Dans cet exercice, nous nous focalisons sur un sous-échantillon de 9 pays issu du tableau précédent.

Nom du pays	Abréviation du pays (code CIO)	Population urbaine 2020 x_i
Islande	ISL	344,10
Érythrée	ERI	1 470,25
Maroc	MAR	23 309,11
Malaisie	MYS	25 617,12
Colombie	COL	41 470,29
Égypte, République arabe d'	EGY	45 976,81
Philippines	PHL	53 187,50
France	FRA	54 715,70
Royaume-Uni	GBR	56 283,17

3.1. – 3.7. Calculer et interpréter la moyenne empirique, la variance empirique, l'écart-type, la médiane, Q_1 et Q_3 , l'étendue, l'écart interquartile et le coefficient de variation

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{344,10 + 1470,25 + 23309,11 + 25617,12 + 41470,29 + 45976,81 + 53187,50 + 54715,70 + 56283,17}{9} = \frac{302374,05}{9} = 33597,12$$

$$s_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{1}{N-1} \left(\sum_{i=1}^N x_i^2 - N\bar{x}^2 \right)$$

$$= \frac{(344,10^2 + 1470,25^2 + 23309,11^2 + 25617,12^2 + 41470,29^2 + 45976,81^2 + 53187,50^2 + 54715,70^2 + 56283,17^2) - 9 \times 33597,12^2}{8}$$

$$= \frac{14\,025\,996\,704,41 - 10\,158\,896\,234,82}{8} = \frac{3\,867\,099\,469,59}{8} = 483\,387\,558,70$$

La variance s'exprime en (milliers d'habitants)²

$$s_x = \sqrt{s_x^2} = \sqrt{483\,387\,558,70} = 21\,986,08$$

Les données sont déjà organisées dans l'ordre croissant

$$m_x = Q_2 = x_{(1+(N-1)/2)} = x_{(1+(9-1)/2)} = x_{(1+4)} = x_{(5)} = 41\,470,29$$

Autre calcul de médiane possible (car N est impair) :

$$m_x = Q_2 = x_{(N+1)/2} = x_{(9+1)/2} = x_{(10/2)} = x_{(5)} = 41\,470,29$$

$$Q_1(x) = x_{(1+(N-1)/4)} = x_{(1+(9-1)/4)} = x_{(1+2)} = x_{(3)} = 23\,309,11$$

$$Q_3(x) = x_{(1+3(N-1)/4)} = x_{(1+3(9-1)/4)} = x_{(1+6)} = x_{(7)} = 53\,187,50$$

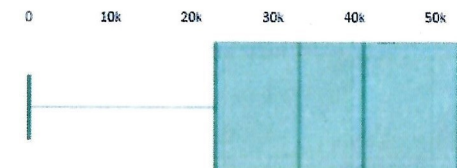
$$ETE(x) = \text{Max}(x) - \text{Min}(x) = 56\,283,17 - 344,10 = 55\,939,07$$

$$EIQ(x) = Q_3(x) - Q_1(x) = 53\,187,50 - 23\,309,11 = 29\,878,39$$

$$CV(x) = \frac{s_x}{\bar{x}} = \frac{21\,986,08}{33\,597,12} = 0,6544 = 65,44\%$$

Les 9 pays de ce sous-échantillon ont, en moyenne, 33 597 120 ± 21 986 080 habitants dans des villes de plus de 50 000 habitants, ce qui correspond à un coefficient de variation de 65,44%. La dispersion autour de la moyenne est donc très forte. La différence entre les valeurs extrêmes est de 55 939 070 habitants. 25% des pays ont une population urbaine strictement inférieure à 23 309 110 habitants. 50% ont de moins de 41 470 290 habitants et 25% une population urbaine de strictement plus de 53 187 500 habitants. Les 50% médian se situe donc dans une tranche de 29 878 390 habitants.

*Si on avait demandé une boîte à moustache :



Exercice 4

Nous cherchons maintenant s'il peut exister une relation entre les émissions de CO₂ (y_i , en kilotonnes) et la population urbaine (x_i , en milliers d'habitants). Nous complétons donc les informations détenues sur notre échantillon dans le tableau de l'exercice 2. Les pays sont présentés dans le même ordre que dans l'exercice 2 (croissant selon x_i) :

Nom du pays	Abréviation du pays (code CIO)	Population urbaine 2020 x_i	Emission CO ₂ 2020 y_i
Islande	ISL	344,10	1 446,60
Érythrée	ERI	1 470,25	706,40
Oman	OMN	3 919,86	71 041,50
Koweït	KWT	4 360,44	92 308,90
Haiti	HTI	6 454,83	3 209,04
Bélarus	BLR	7 455,47	54 801,50
République tchèque	CZE	7 922,94	88 835,20
Émirats arabes unis	ARE	8 084,40	188 088,70
Grèce	GRC	8 528,39	51 002,20
Israël	ISR	8 531,99	58 471,80
République dominicaine	DOM	9 079,12	22 881,30
Kazakhstan	KAZ	10 816,58	211 896,70
Pologne	POL	22 755,74	279 223,80
Maroc	MAR	23 309,11	66 719,50
Malaisie	MYS	25 617,12	245 139,30
Pérou	PER	26 076,63	46 578,90
Iraq	IRQ	30 169,92	163 511,50
Canada	CAN	30 999,41	516 873,70
Algérie	DZA	32 038,22	161 563,00
Espagne	ESP	38 276,19	202 705,80
Colombie	COL	41 470,29	79 057,60
Argentine	ARG	41 796,99	154 535,90
Corée, République de	KOR	42 201,96	569 681,80
Italie	ITA	42 224,77	281 286,80
Égypte, République arabe d'	EGY	45 976,81	210 752,30
Philippines	PHL	53 187,50	133 471,30
France	FRA	54 715,70	267 154,70
Royaume-Uni	GBR	56 283,17	308 650,30
Allemagne	DEU	64 410,59	603 350,50
Mexique	MEX	101 719,69	383 131,40
Japon	JPN	115 884,87	1 014 064,70
Indonésie	IDN	153 983,07	563 197,00

4.1. Remplir le tableau suivant à l'aide des effectifs concernés

X/Y	[0 ; 300 000]	[300 000 ; 600 000]	[600 000 ; 900 000]	[900 000 ; 1 200 000]	TOTAL
[0 ; 40 000]	19	1	0	0	20
[40 000 ; 80 000]	6	2	1	0	9
[80 000 ; 120 000]	0	1	0	1	2
[120 000 ; 160 000]	0	1	0	0	1
TOTAL	25	5	1	1	32

4.2. Comment s'appelle le tableau de la question précédente ?

Un tableau de contingence

Un graphique représente ces données page suivante.

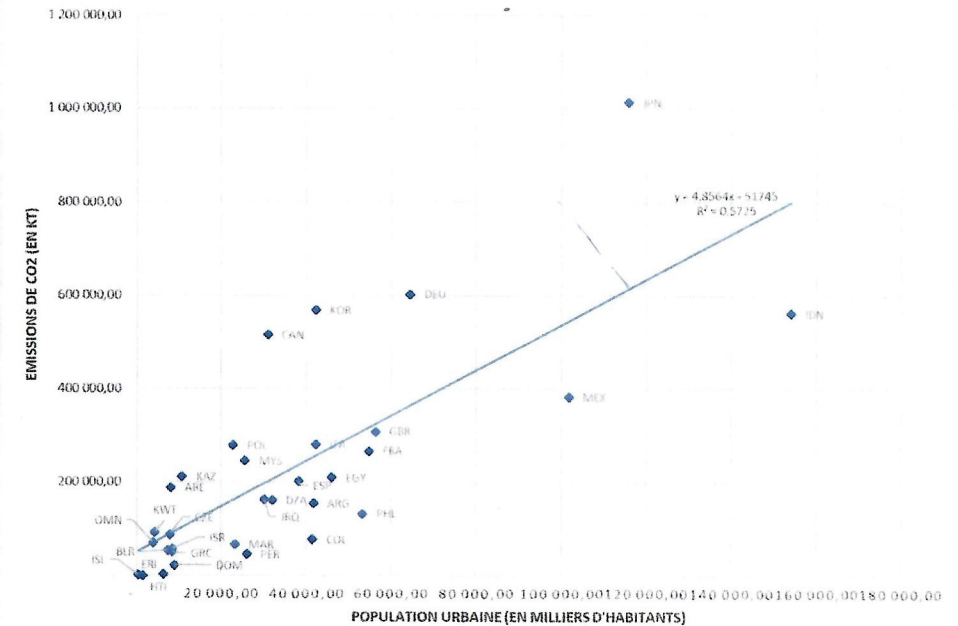
4.3. Comment s'appelle ce type de graphique ?

C'est un nuage de points

4.4. Quelle première interprétation pouvez-vous en faire ?

Nous voyons qu'il semble exister une relation linéaire entre les deux variables (droite de régression sur le graphique). Cette relation est positive : si la population urbaine augmente, le volume d'émission de CO₂ augmente aussi. Cette relation semble assez forte pour les pays avec une faible population urbaine, moins pour ceux avec une population urbaine supérieure à 60 000 000 habitants.

EMISSION DE CO₂ EN FONCTION DE LA POPULATION URBAINE DANS 32 PAYS DU MONDE (EN 2020)



4.5. Remplir le tableau suivant

	n_{13}	n_2
Type d'effectif	Effectif partiel	Effectif marginal
Interprétation	Il s'agit du nombre de pays avec une population urbaine entre 0 et 40 000 000 habitants ($n=1$) ET une émission de CO ₂ entre 600 000 et 900 000 kT ($n=3$)	Il s'agit du nombre de pays avec une population urbaine entre 40 000 000 et 80 000 000 habitants ($n=2$), quel que soit le volume d'émission de CO ₂ ($n=1$)
Pays concernés (abréviations)	Aucun	ITA : FRA : EGY : ARG : PHL : COL : KOR : GBR : DEU
Calcul de la fréquence relative associée	$f_{13} = \frac{n_{13}}{N} = \frac{0}{32} = 0 = 0\%$	$f_2 = \frac{n_2}{N} = \frac{9}{32} = 0,2813 = 28,13\%$
Interprétation de la fréquence relative associée	En 2020, sur les 32 pays de notre échantillon, 0% ont à la fois une population urbaine entre 0 et 40 000 000 habitants ET une émission de CO ₂ entre 600 000 et 900 000 kT.	En 2020, sur les 32 pays de notre échantillon, 28,13% ont une population urbaine entre 40 000 000 et 80 000 000 habitants.

4.6. Remplir le tableau suivant à l'aide des fréquences relatives conditionnelles de y

	[0 ; 300 000]	[300 000 ; 600 000]	[600 000 ; 900 000]	[900 000 ; 1 200 000]	TOTAL
[0 ; 40 000]	95,00%	5,00%	0,00%	0,00%	100,00%
[40 000 ; 80 000]	66,67%	22,22%	11,11%	0,00%	100,00%
[80 000 ; 120 000]	0,00%	50,00%	0,00%	50,00%	100,00%
[120 000 ; 160 000]	0,00%	100,00%	0,00%	0,00%	100,00%
TOTAL	78,13%	15,63%	3,13%	3,13%	100,00%

4.7. Détailler le calcul de la cellule encadrée en gras et interpréter

$$f_{42(y|x)} = \frac{n_{42}}{n_4} = \frac{1}{1} = 1 = 100\%$$

En 2020, parmi les 32 pays de notre échantillon dont la population urbaine est comprise entre 120 000 000 et 160 000 000 habitants, 100% présentent une émission de CO₂ entre 300 000 et 600 000 kT.