

Mesures descriptives pour décrire les variables QUANTITATIVES

Mihaela Iancu

Objectifs éducationnels

À la fin du TP, l'étudiant devra posséder les connaissances suivantes :

- Savoir calculer les statistiques descriptives pour décrire les valeurs d'une **seule variable quantitative** (discrète, continue) **en utilisant les fonctions prédéfinis (AVERAGE, MEDIAN, STDEV, QUARTILES)** de l'Excel
- Savoir calculer les statistiques descriptives pour décrire les valeurs d'une **seule variable quantitative** (discrète, continue) **en utilisant leurs formules de calcul (le coefficient de variation, l'erreur standard)**
- Savoir calculer les statistiques descriptives pour décrire les valeurs d'une **seule variable quantitative** (discrète, continue) **en utilisant l'option Data Analysis** de l'Excel
- Savoir interpréter les statistiques descriptives
- Savoir réaliser les types de graphiques appropriés pour décrire les valeurs d'une seule variable quantitative
- Savoir réaliser les types de graphiques appropriés pour la description de la relation entre **deux variables quantitatives**
- Savoir interpréter les graphiques obtenus.

Scenario

Une étude a été menée pour décrire le type dominant de l'indice nasal (IN) sur un échantillon aléatoire de 150 sujets adultes âgés de 18 à 45 ans qui ont été évalués entre le 1 février et le 1 décembre 2021 à Clinique de Chirurgie buccale-maxillo-faciale de Cluj Napoca.

L'indice nasal (IN) est l'un des paramètres anthropométriques cliniques très utilisé en médecine nasale ^[1]. Il est défini comme le rapport de la largeur maximale de la narine à la hauteur du nez squelette multiplié par 100 ^[2, 3].



Figure 1. Mesure de la largeur nasale et la hauteur du nez

Source : Dhulqarnain AO, Mokhtari T, Rastegar T, Mohammed I, Ijaz S, Hassanzadeh G. Comparison of Nasal Index Between Northwestern Nigeria and Northern Iranian Populations: An Anthropometric Study. J Maxillofac Oral Surg. 2020 Dec;19(4):596-602.

¹ Hansen B, Mygind N. How often do normal persons sneeze and blow the nose? Rhinol 2002; 40(1): 407-26.

² Porter JP, Olson KL. Analysis of the African American female nose. Plastic Reconstruct Surg 2003; 111(2): 620-26.

³ Sharma SK, Jehan M, Sharma RL, Saxena S, Trivedi A, Bhadkaria V. Anthropometric Comparison of Nasal Parameters between Male and Female of Gwalior Region. Journal of Dental and Medical Sciences 2014; 13(5): 57 – 62.

La forme du nez est divisé en cinq types :

- i) l'hyperleptorrhine est une nez très long et étroit avec un indice nasal allant de 40 à 54,9
- ii) la leptorrhine fait est un nez long et étroit avec un indice nasal de 55-69,9
- iii) la mésorrhine fait référence à un nez moyen avec un indice nasal de 70-84,9
- iv) la platyrrhine décrit un court et large avec un indice nasal de 85-99,9
- v) l'hyperplatyrrhine décrit un nez très large avec un indice nasal supérieur à 100.

Les données recueillies se trouvent dans le fichier Donnees_TP05.xlsx.

Demandes

1. Créez un nouveau dossier nommé **TP05NP** ou **N = votre nom, P = votre prénom** sur le Bureau (Desktop) de Windows.
2. Enregistrer le fichier Excel **Donnees_TP05.xlsx** dans le dossier **TP05NP**.
3. Dans le même fichier, créer plusieurs feuilles de calcul et renommez-les: **Histogramme_1_variable**, **Boite_moustaches_1_variable**, **Nouage de points** au lieu de Sheet1, Sheet2.

Statistiques descriptives & graphiques pour décrire une SEULE **variable QUANTITATIVE**

4. Dans la feuille nommée **Données**, calculer pour chaque patient son indice nasal (IN) en utilisant la formule de calcul suivante :

$$IN = \frac{\text{Largeur nasale (cm)}}{\text{Hauteur nasale (cm)}} \times 100$$

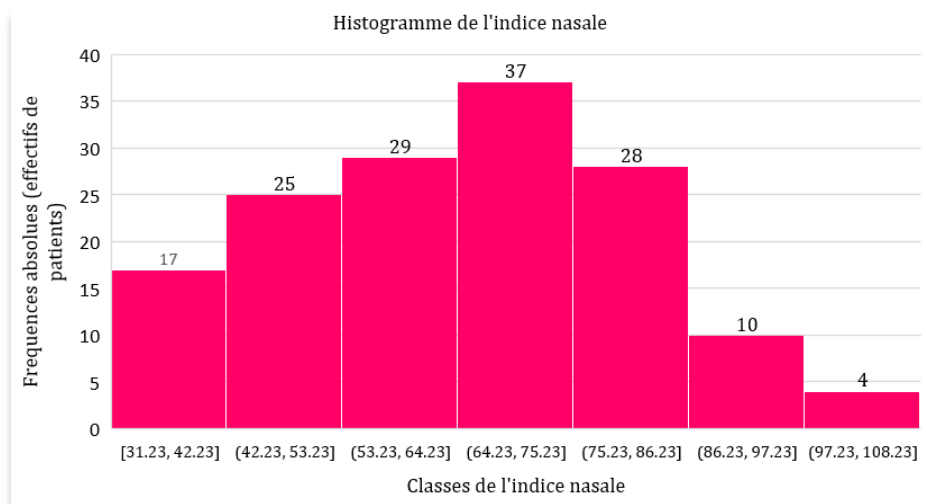
5. Formatez la colonne IN (Indice nasale) avec le format numérique ayant **2 décimales**.
6. Dans la même feuille **Données**, à côté droite de la base de données, créer une table de la forme suivante **puis à l'aide des notions du cours, remplir le tableau** :

Variables	Type de la variable (qualitative dichotomique, qualitative ordinale, qualitative nominale, quantitative discrète, quantitative continue)
Largeur nasale	
Hauteur nasale	
Indice nasal (IN)	
Age (ans)	

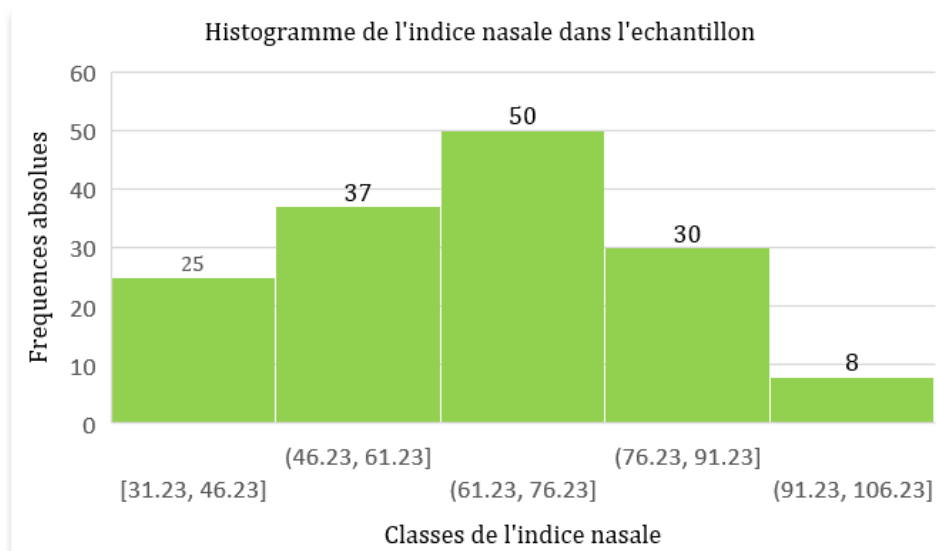
7. Dans la feuille nommée **Statistiques descriptives**, copier **les valeurs calculées de la variable Indice nasale (IN) et les valeurs de l'Age** en utilisant l'option Home→Copy et puis, dans la page Statistiques descriptives on va utiliser l'option Home→Paste → **Paste Values**.
8. Dans la feuille nommée **Statistiques descriptives**, **remplir** le tableau contenant les statistiques descriptives (*mesures de tendance centrale, mesures de dispersion, mesures de*

position, mesures d'asymétrie et mesures d'aplatissement) appropriés aux **VARIABLES QUANTITATIVES**.

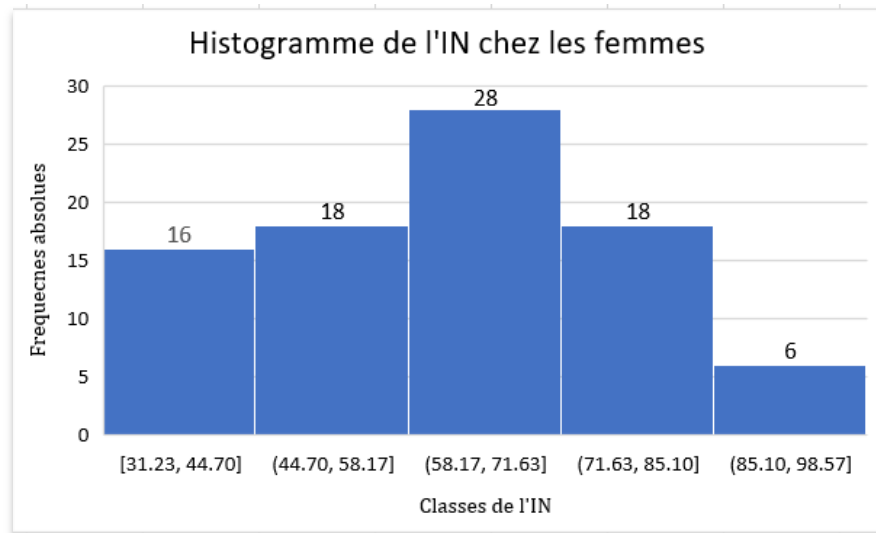
9. Dans une nouvelle page, calculer les **mesures descriptives de l'IN** en utilisant l'option **DATA ANALYSIS**.
10. Dans la feuille nommée **Histogramme_1_variable**, copier les variables **Indice nasale (IN)** et la variable **Sexe** et puis :
 - a. réalisez un **HISTOGRAMME** pour représenter les données de **l'indice nasal (IN)** dans **l'échantillon d'étude**. Le graphique créé devrait ressembler à celui ci-dessous:



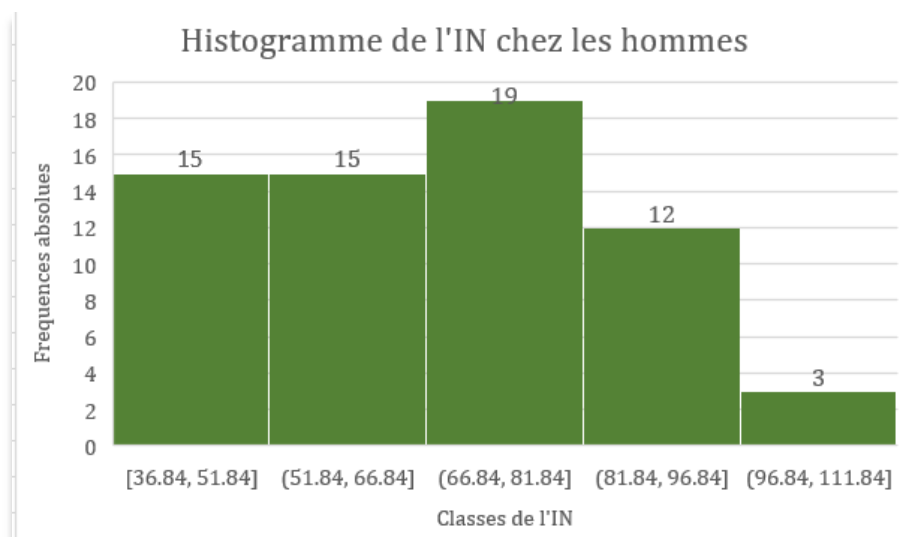
- b. réalisez un **HISTOGRAMME** pour représenter les données de **l'indice nasal (IN)** dans **l'échantillon d'étude avec des classes de longueur =15**. Le graphique créé devrait ressembler à celui ci-dessous:



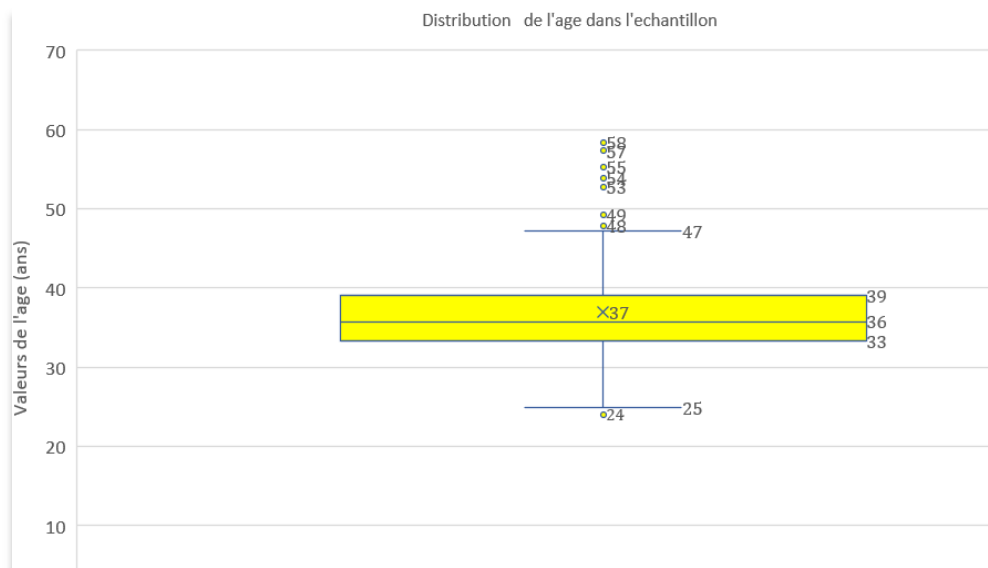
- c. réalisez un **HISTOGRAMME** pour représenter les données de l'**indice nasal (IN)** chez les **femmes ayant 5 classes de fréquences**. Le graphique créé devrait ressembler à celui ci-dessous:



- d. réalisez un **HISTOGRAMME** pour représenter les données de l'**indice nasal (IN)** chez les **hommes ayant 5 classes de fréquences**. Le graphique créé devrait ressembler à celui ci-dessous:



- e. Dans la feuille **Boite_moustaches**, copier la variable **Age** et réaliser un **graphique boîte a moustaches** (engl. **box-whisker plot**) pour les valeurs de l'Age dans l'échantillon d'étude. Le graphique créé devrait ressembler à celui ci-dessous:



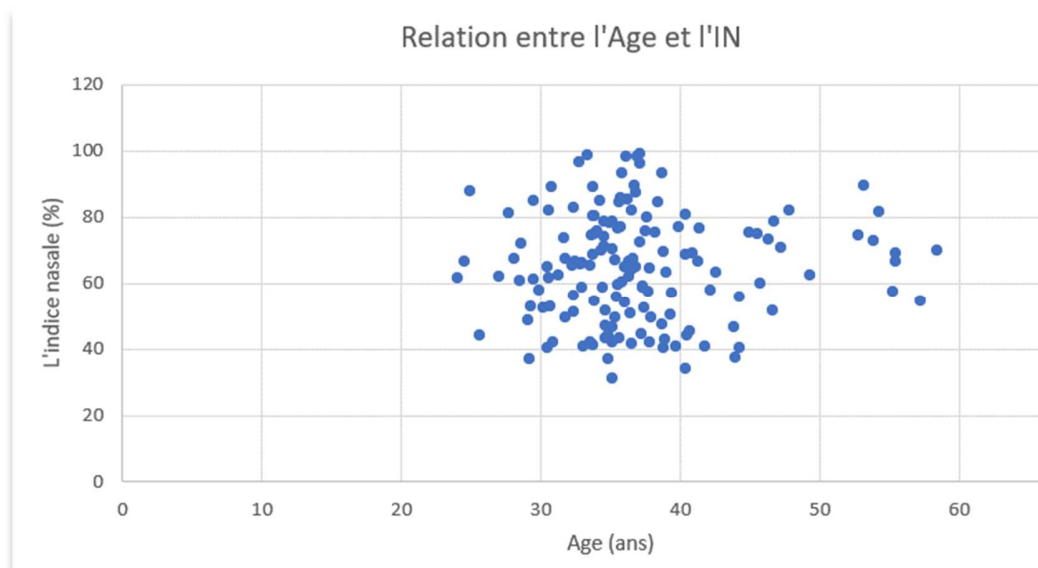
11. Dans la même feuille **Boite_moustaches**, créer et remplir (en utilisant le graphique) le tableau suivant :

Statistiques descriptives	Valeurs
Médiane	
Moyenne	
Q1	
Q3	
Intervalle interquartile (IQR) : [Q1 ; Q3]	
Valeurs aberrantes	

Représentation graphique pour la relation entre **DEUX variables QUANTITATIVES**

12. Dans la feuille **Nouage de points** de votre fichier Excel, copier les variables **Age** et **l'Indice nasale** de sorte que sur la première colonne de votre page se trouve l'âge puis l'index nasale.

13. Dans la même feuille **Nouage de points**, réaliser un graphique par nouage de points pour mettre en évidence la possible relation entre l'Age et l'indice nasale. Dans votre graphique, on va considérer que **l'âge** est la **variable indépendante** et **l'index nasal** est la **variable dépendante**. Le graphique créé devrait ressembler à celui ci-dessous:



14. Peut-on dire que dans l'échantillon étudié, les valeurs de l'indice nasale augmentent avec l'âge des patients ? ? répondez à cette question dans votre page de calcul à l'aide du graphique (regarder s'il y a une tendance croissante ou décroissante du nuage de points).

15. Créer une présentation dans Microsoft **PowerPoint** concernant les interprétations des statistiques descriptives antérieurement calculées en suivant les exigences ci-dessous :

a. Ouvrir une nouvelle présentation dans PowerPoint et enregistrez votre présentation sous le nom **Interprétation de l'analyse statistique.pptx** dans le dossier **TP05NP**.

b. Choisissez un design (format) pour votre présentation.

c. La structure de la présentation doit être comme suit :

📌 **Diapositive 1 :**

÷ **Titre:** Statistiques descriptives pour décrire les variables quantitatives

÷ **Sous-titres:** Votre nom (comme la Faculté de médecine / Faculté de médecine dentaire) et votre affiliation (Iuliu Hațieganu Université de Médecine et Pharmacie de Cluj-Napoca)

📌 **Diapositive 2:** Copier le tableau contenant les **statistiques descriptives de tendance centrale** pour la variable IN (**exercice 8**) **ET** interpréter **la symétrie** des valeurs de l'indice nasale en regardant : i) aux statistiques de tendance centrale et ii) au coefficient d'asymétrie (engl.skewness).

📌 **Diapositive 4 :** Copier l'histogramme de l'IN (**exercice 10.a**) et interprétez-la (préciser si elle a une forme symétrique et quelle est la classe modale)

📌 **Diapositive 5** Copier le tableau contenant les statistiques descriptives **de dispersion** pour la variable IN (**exercice 8**) **ET** interpréter **le coefficient de variation**.

📌 **Diapositive 6.** Copier le graphique boîte à des moustaches de l'Age et le tableau associé (**exercices 10.e & exercice 11**) et interpréter **l'intervalle IQR**.

Probleme récapitulatif/révision

On considère la même base de données.

1. Insérez une nouvelle variable appelée **Forme du nez**. En utilisant une **fonction prédéfinie**, établir pour chaque sujet si la forme du nez est :

- vi) l'hyperleptorrhine SI l'indice nasal allant de 40 à 54,9
 - vii) la leptorrhine SI l'indice nasal est compris entre 55-69,9
 - viii) la mésorrhine SI l'indice nasal est compris entre 70-84,9
 - ix) la platyrrhine SI l'indice nasal est compris entre 85-99,9
 - x) l'hyperplatyrrhine SI l'indice nasal est supérieur à 100.
2. Réaliser le tableau de fréquence pour la variable **Forme du nez**.
 3. Créer/Réaliser ensuite un graphique qui montre la répartition des patients par rapport au **Sexe et Forme du nez**.
 4. Calculer les mesures de dispersion de l'Indice nasale (IN) chez les hommes.

Rappelez-vous que....

1. Le calcul de mesures descriptives associées à une variable QUANTITATIVE peuvent se calculer en utilisant les fonctions prédéfinies (AVERAGE, MEDIAN, STDEV, QUARTILES, etc.) ou par leurs formules (CV=coefficient de variation, ES=erreur standard) ou l'option DATA ANALYSIS (la dernière option ne calculera pas les quartiles ni le coefficient de variation).
2. Le graphique de l'HISTOGRAMME est approprié pour la description d'une SEULE variable QUANTITATIVE.
3. La relation entre DEUX variables QUANTITATIVES sera décrite par un graphique par nouage de points.

A la fin du TP....

Le fichier sera envoyé à l'enseignant du TP par la manière/méthode établie de celle-ci mais aussi à votre adresse e-mail (cela peut être celle sur Outlook Teams votre_nom_utilisateur@elearn.umfcluj.ro)