

SÉANCE TABLEUR (2H30)

Objectifs : être capable de

- créer et modifier un fichier tableur local ou en ligne, en distinguant, pour enregistrer, les formats éditables et finaux,
 - manipuler les feuilles de calcul (création, renommage...),
 - choisir et appliquer des formats de cellules appropriés aux données et maîtriser la mise en forme (y compris conditionnelle) des données et de la feuille,
 - utiliser les fonctionnalités de tri et de filtre des données,
 - saisir des formules de calcul en utilisant les références relatives, mixtes et absolues et les fonctions prédéfinies,
 - utiliser des fonctions mathématiques, statistiques et logiques (connaître en particulier SOMMEPROD, NBVAL, SI, SI.CONDITIONS, NB.SI, SOMME.SI),
 - tracer divers types de graphiques à partir d'un ensemble de données, en choisissant un type approprié :
 - diagrammes XY : $y = f(x)$,
 - diagrammes bâtons (colonnes ou barres horizontales)
 - camemberts
 - histogrammes de fréquences
- et en affichant dans les cas pertinents les barres d'erreur sur les données.
- Légénder et ajouter des éléments (texte, symboles) sur ces graphiques,
- effectuer une régression (i.e. ajustement, courbe de tendance, *fit* en anglais) sur une loi linéaire en affichant la droite de régression, son équation et son R^2 , ses paramètres avec DROITEREG, et connaître l'existence d'autres types de régressions (logarithmique, exponentielle, régressions multiples ...),
 - sélectionner une partie choisie d'une feuille de calcul en vue d'une impression papier ou pdf en contrôlant le nombre de pages et le format.

PIX : compétence 1.3. Traiter des données (seul le niveau 5 est à viser : surtout tableur)

Appliquer des traitements à des données pour les analyser et les interpréter (avec un tableur, un programme, un logiciel de traitement d'enquête, une requête calcul dans une base de données, etc.).

THÉMATIQUES ASSOCIÉES

Données quantitatives, type et format de données ; Calcul, traitement statistique et représentation graphique ; Flux de données ; Collecte et exploitation de données massives ; Pensée algorithmique et informatique ; Vie privée et confidentialité ; Interopérabilité

Important pour le Pix 1.3 :

Les fichiers supports des questions du Pix 1.3 sont souvent issus de données publiques ouvertes (*Open Data*). Parfois elles sont à chercher sur le web sur divers sites officiels, ou directement sur : <https://data.gouv.fr>
Il s'agit de données dites structurées, aux formats csv, json etc. Ce point est abordé dans l'exercice 6.

Logiciel tableur : vous pouvez utiliser (cf document annexe *Suite bureautique pour plus de détails*)

- **LibreOffice Calc** (équivalent un peu simplifié de Excel) pour Windows, Linux et Mac : logiciel libre et gratuit, qui sera utilisé en TD pour les corrections des exercices.

- **Microsoft Office Excel** installé sur les ordinateurs du SIF pour Windows

- **Microsoft 365 (dont une version allégée d'Excel)** en ligne à laquelle vous avez accès gratuitement en tant qu'étudiants et que vous pouvez aussi utiliser depuis chez vous ; il permet aussi de télécharger une version locale.

La fiche de cours, l'énoncé du TD, les fichiers-supports du TD (mis dans un dossier zippé) sont sur ecampus > Formation numérique PCST > Section Pix_Tableur

Les corrigés seront placés sur ecampus, en temps utile.

ATTENTION ! Avant de commencer l'exercice 1 :

Dans votre dossier PIX local, créer un **sous-dossier Tableur** (à la fin, comme vous avez des comptes invités, pour le conserver il faudra le copier sur clé ou le zipper et se l'envoyer par mail)

Depuis eCampus, télécharger le fichier .zip des fichiers supports du TD tableur sur votre ordinateur, idéalement directement dans le sous-dossier Tableur, sinon l'y déplacer.

NE **PAS** DOUBLE CLIQUER sur le zip ! Par un clic droit sur le .zip, extraire les fichiers de l'archive de format zip en choisissant comme dossier de destination, le sous-dossier Tableur de votre dossier PIX.

Partie 1 : fonctionnalités de base d'un tableur manipulation des données et fonctions prédéfinies

Exercice 1 Références absolues, relatives et mixtes : énergies E_n de l'atome d'Hydrogène et des ions hydrogénoïdes

Objectif : comprendre l'utilité des références absolues, relatives et mixtes et les utiliser à bon escient

Le but de l'exercice est de calculer l'énergie E_n de l'atome d'hydrogène H et des ions hydrogénoïdes pour des niveaux d'énergies successifs, via l'expression $E_n = -R_y \frac{Z^2}{n^2}$.

Ouvrir le fichier **NRJ_atomes.ods** fourni, qui est déjà en partie pré-rempli.

Ajouter dans des cellules bien choisies le paramètre manquant nécessaire dans la formule.

En saisissant une unique formule, remplir le tableau avec le calcul des énergies des 5 premiers niveaux ($n = 1$ à 5) de l'atome d'hydrogène H et des ions hydrogénoïdes pour Z variant de 1 à 10.

Choisir le format de cellule approprié et faire afficher 2 chiffres après la virgule.

Exercice 2 Utiliser des fonctions prédéfinies (outil fx : vérification de la loi normale ou loi de Gauss). Tracé d'une courbe simple $y = f(x)$

Objectif : saisir une formule mathématique combinant plusieurs fonctions prédéfinies — effectuer des calculs (simples) – créer et formater un graphique $y = f(x)$

En statistique, la loi normale¹ est une loi de probabilité très souvent utilisée. On peut d'ailleurs démontrer rigoureusement, par le biais du *théorème de la limite centrale* (ou théorème central limite), que cette loi émerge dans le cas d'un processus issu d'un grand nombre d'événements aléatoires.

La loi normale s'écrit :

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

où μ est la moyenne (ou espérance) de la variable aléatoire x et σ est son écart-type. Cette fonction est également appelé « Gaussienne » du nom du mathématicien et physicien allemand *Carl Friedrich Gauss* (1777-1855). Dans cet exercice, on cherche à vérifier l'implémentation de cette fonction « Gaussienne ».

- Créer un nouveau fichier tableur.** Dans la colonne A, saisir "x" dans la cellule A1, puis à partir de la cellule A2, créer une liste de nombres variant de manière régulière de -3 à +3 par incrément de 0,1. Pour ce faire vous pouvez soit saisir les premiers chiffres et étendre ce début de série à toutes les cellules, soit utiliser les séries, soit utiliser une formule avec la valeur précédente

1. Voir plus d'explications : https://fr.wikipedia.org/wiki/Loi_normale

additionn e de l'incr ment 0,1.

2. Dans les cellules isol es F1 et F2, donner des valeurs de la moyenne μ et de l' cart-type σ : par exemple $\mu = 0$ et $\sigma = 1$. *Attention, les cellules F1 et F2 ne doivent contenir que les nombres !*
3. Nous utilisons dans cette question la fonction pr d finie LOI.NORMALE. Dans la colonne B, cellule B1, saisir comme en-t te de colonne « Loi Normale » ; dans la cellule B2, utiliser   l'aide de l'assistant fonction fx (), la fonction pr d finie LOI.NORMALE [sous Excel, LOI.NORMALE.N] qui correspond   la fonction P(x).
Indication : il y a 2 variantes de cette loi, choisir celle qui n'est pas cumulative, pour cela donner au dernier argument de la fonction LOI.NORMALE, la valeur de 0 (ou FAUX).
4. Nous recalculons maintenant la fonction P(x). Dans la colonne C, cellule C1, saisir comme en-t te de colonne P(x) ; dans la cellule C2, saisir la formule de la fonction P(x) fournie en utilisant les fonctions pr d finies EXP(), PUISSANCE(), RACINE().
5. + Calculer dans la colonne D, la diff rence entre la colonne B et la colonne C.
Conclusion : la loi normale est-elle correctement impl ment e dans ce logiciel tableur par la fonction pr d finie LOI.NORMALE ?
6. Repr senter graphiquement la fonction P(x). Pr ciser le titre du graphique et les axes.
Menu : [Calc Insertion > Graphique > Nuage de points (reli s)]
Menu : [Excel Insertion > Graphique > Nuage de points (reli s)]

Exercice 3 Manipulation de donn es. Fonctionnalit s de tri et filtres. Fonctions statistiques : volume de lign es de grains de ma s

Objectif : savoir trier des donn es suivant la colonne choisie et filtrer des donn es avec les autofiltres, utiliser les fonctions statistiques et une fonction conditionnelle (SI)

Manipulations d'un fichier long, tri et filtre

Ouvrir le fichier **grains-mais-1.ods**, qui contient des donn es sur les dimensions des grains dans diverses lign es de ma s.

1. Trouver la fonctionnalit  permettant de rendre la 1 re ligne fixe. (2 solutions)
Cette ligne constitue l'en-t te du tableau et sert   d crire les donn es de chaque colonne. Si l'on fait d filer l'ascenseur vertical, cette ligne reste fixe et toujours visible. Si l'ordre des donn es est modifi , cet en-t te ne le sera pas.
Comment fixe-t-on la premi re colonne ?
Il est  galement possible de rendre plusieurs lignes fixes, et de m me pour les colonnes.
2. Combien de colonnes le fichier contient-il ? Combien de lignes ? (Trouver l' quivalent clavier pour arriver en bas d'un fichier long). Combien de cellules ? Consulter les statistiques du classeur pour v rifier. [Calc : Donn es>Statistiques>Statistiques descriptives]
[Excel : R vision>Statistiques du classeur]
3. En utilisant la fonctionnalit  de **tri**, classer les donn es par volume d croissant
4. Utiliser la fonctionnalit  de **filtre** pour n'afficher que la lign e n 3.
5. R afficher tout. **Masquer** les 10 premi res lignes (i.e. celles de la lign e n 1.) *Bien distinguer du filtrage.* Les r afficher. (On peut aussi masquer des colonnes).

Fonctions logiques et fonctions statistiques

Pour les questions 1   3, travailler sur le fichier **grains-mais-2.ods**

1. Utilisation de la fonction logique SI
[Outil fx > cat gorie logique]
Dans la colonne G,   l'aide de la fonction SI, faire afficher automatiquement la mention "maximum" dans la ligne correspondant au volume de grain le plus grand. Les autres cellules de la colonne doivent rester vides.

- Faire afficher automatiquement en rouge le mot "maximum", en utilisant un formatage conditionnel adapt . (On peut aussi faire afficher un fond color  pour cette cellule).
- Calcul de grandeurs statistiques en utilisant les fonctions statistiques pr d finies [Outil fx > cat gorie Statistiques]
On effectuera ces calculs uniquement pour les volumes de la lign e 1 puis 2 de grains de ma s.
MOYENNE (  calculer en I2 et I3)
MEDIANE (  calculer en J2 et J3)
ECART-TYPE (  calculer en K2 et K3)
VARIANCE (  calculer en L2 et L3).
Les autres fonctions   conna tre sont SOMMEPROD, NBVAL, SI, NB.SI, SOMME.SI, SI.CONDITIONS. L'assistant fonction fx en donne la syntaxe et une aide   la saisie.
 - + (Pour en savoir plus) Calculer   l'aide de la fonction de moyenne conditionnelle MOYENNE.SI et via une formule unique, dans la plage I16   I157, les moyennes des volumes des grains de chaque lign e indiqu e dans la colonne H. Utiliser l'aide pour comprendre la fonction MOYENNE.SI (difficile).
 - + (Entra nement) Dans le fichier **grains-mais-3.ods**, calculer dans les cellules I2   N2, par les fonctions statistiques pr d finies les grandeurs suivantes pour l'enti ret  de la colonne des volumes : la moyenne, la m diane, la variance de la population, la variance de l' chantillon, l' cart-type de la population et l' cart-type de l' chantillon.

Partie 2 :

les graphiques

Exercice 4 Diagramme « Nuage de points ». Courbe $y = f(x)$ avec barres d'incertitudes

Objectif : cr er un graphique (cas d'un graphique de type courbe, XY : $y = f(x)$) avec des barres d'incertitude) – analyser des donn es exp rimentales (d pendance d'une grandeur par rapport   une autre)

Calc : Menu Insertion > Diagramme > Nuage de points

Excel : Menu Insertion > Graphiques > Nuage de points

(ou outil de la barre d'outils)

1) Fond diffus cosmologique ("rayonnement fossile")

La feuille de calcul devra pr senter les r sultats de mani re claire.

Le fond diffus cosmologique (en anglais, *Cosmic Microwave Background* ou CMB), est un rayonnement initialement tr s chaud  mis vers 380000 ans apr s le Big Bang dans les mod les cosmologiques actuels, et devenu au cours de l'expansion de l'univers un rayonnement froid   2,7 K.

Dans les ann es 1990, le satellite COBE a enregistr  de fa on pr cise ce rayonnement   l'aide de l'instrument FIRAS. Les donn es sont fournies dans le fichier **CMB-COBE.ods**.

fr�quence ν (cm ⁻¹)	longueur d'onde λ (mm)	densit� du flux (MJy/sr)	incertitude (kJy/str) (1σ)
2,27	4,405	200,723	14
2,72	3,676	249,508	19
3,18	3,145	293,024	25

Tableau 1: Fond diffus cosmologique (CMB) : extrait des donn es observationnelles de COBE/FIRAS

- Ouvrir le fichier **CMB-COBE.ods**. Renommer la feuille 2 avec le nom *mesures*. Reproduire la mise en forme du tableau ci-dessus (apparence, chiffres significatifs...).
- Tracer le graphe densité du flux = $f(\lambda)$ sur la même feuille. Indiquez les noms des axes.
+ Mettez un fond (paroi du diagramme) en gris 30% et un arrière-plan en gris plus foncé.
- Afficher sur le graphe des barres d'incertitude (de valeur σ) symétriques en Y, affectant la densité de flux, en utilisant la colonne d'incertitudes fournies.
Calc : clic droit sur le graphe, insérer des Barres d'erreur Y, plage de cellules, valeurs identiques en + et -, sélectionner la plage où se trouvent les valeurs des incertitudes.
Excel : *suivant la version d'Excel* cliquer sur le + en haut à droite du graphe, Barres d'erreur, plus d'options, le volet droit s'ouvre : en bas, personnaliser, fenêtre de dialogue, puis pour chaque valeur (côté positif et côté négatif de l'erreur), aller chercher par clic sur la flèche, la ligne des écarts-types.
Ou Outil Format ou Clic droit sur les colonnes, format série 1, activer le bouton barres d'erreurs : par défaut c'est l'erreur standard qui est utilisée.
Ou par un clic sur le diagramme, menu Outils graphiques, Ajouter un élément graphique, Barres d'erreur (plusieurs types : erreur standard, écart-type...).
- +Sur la courbe, grâce à la barre d'outils dessin, insérer un symbole sur le maximum de la courbe.
- Sélectionner pour un export pdf ou une impression la partie de la feuille 2 contenant les données et la courbe ("zone d'impression"). Choisir un format de page approprié (portrait ou paysage) dans Format > Page. Vérifier que tout tient sur une page. Exporter la zone sélectionnée en pdf et ouvrir le pdf pour vérifier la mise en page.

2) (+ Entraînement) Détermination du volume équivalent de soude lors d'un dosage acido-basique

On cherche à doser 50 mL d'acide chlorhydrique HCl à la concentration de $0,1 \text{ mol.L}^{-1}$ par une solution de soude NaOH de même concentration. Le dosage est suivi par pH-métrie. On va grâce à la courbe $\text{pH} = f(v(\text{NaOH}))$ déterminer le volume équivalent de soude. Ouvrir le fichier **courbepH.ods**.

V_{NaOH} en ml	pH
0,000	1
40,910	2
49,010	3
49,900	4
49,990	5
49,999	6
50,000	7
50,001	8
50,010	9
50,100	10
51,010	11
61,110	12

Tableau 2: pH de la solution en fonction du volume de la solution de soude versée

- Dans le fichier **courbepH.ods**, renommer la feuille en lui donnant le nom *mesures*. Reproduire la mise en forme du tableau ci-dessus : apparence, chiffres significatifs...

2. Tracer le graphe approprié $\text{pH} = f(V_{\text{NaOH}})$ sur la même feuille. Mettez des noms pour les axes, un fond (paroi du diagramme) en gris 30% et un arrière-plan en gris plus foncé.
3. Afficher sur le graphe des barres d'incertitude symétriques en X, affectant la mesure du volume de soude versée. Faire de même pour les barres d'incertitude symétriques en Y, affectant la mesure du pH. Les valeurs des incertitudes sont données dans le fichier.

Calc : clic droit sur le graphe, insérer des Barres d'erreurs (choisir en X ou en Y), valeur constante à saisir.

Excel : cliquer sur le + en haut à droite du graphe, Barres d'erreur, plus d'options, le volet droit s'ouvre : en bas, personnaliser, dans la fenêtre de dialogue, puis pour chaque valeur (côté positif et côté négatif de l'erreur) saisir la valeur de l'erreur. Autres versions : menu Outils graphiques, Ajouter un élément graphique, Barres d'erreur.

4. En utilisant la fonction SI.CONDITIONS, dans la colonne E, faire afficher automatiquement les mots "acide", "neutre", "basique" en fonction de la valeur du pH située en colonne B. Puis faire colorer *automatiquement* les cellules "acide" en rouge, la cellule "neutre" en bleu, la cellule "basique" en jaune.
5. +Le volume équivalent est le volume pour lequel la dérivée de la fonction $\text{pH} = f(V_{\text{NaOH}})$ est maximale. Dans la colonne C sont fournies les valeurs de cette dérivée calculées à l'aide de l'approximation du taux d'accroissement (*cf. explications en fin d'exercice*), pour les volumes de solution basique introduits de 40,910 à 51,010 mL. Tracer le graphe de la dérivée avec une échelle des abscisses appropriée pour mettre en évidence le pic.
+Tracer ensuite le pH et la dérivée sur le même graphe en veillant aux échelles des ordonnées (afficher 2 axes Y et modifier après coup leur échelle par clic droit).
6. + (*Entraînement*) Dans la colonne D, à l'aide de la fonction SI, faire afficher automatiquement la mention "équivalence" dans la ligne correspondant au volume équivalent. Les autres cellules de la colonne doivent rester vides.
7. + (*Entraînement*) Faire afficher avec un fond orange la cellule contenant le mot "équivalence", en utilisant un formatage conditionnel adapté. Créer pour cela un style de cellule "orange".
8. +Sur la courbe, insérer un symbole grâce à la barre d'outils dessin et placer ce symbole sur le point équivalent. Ajouter à côté à l'aide de l'outil Texte, le mot Equivalence.
9. Sélectionner pour un export pdf ou une impression la partie de la feuille contenant les données et la courbe ("zone d'impression"). Choisir un format de page approprié (portrait ou paysage) dans Format > Page. Vérifier que tout tient sur une page. Exporter en pdf et ouvrir le pdf pour vérifier la mise en page.

Comment calculer une dérivée numériquement :

Mathématiquement la dérivée en x est définie par la limite du taux d'accroissement de la fonction :

$$f'(x) = \frac{df}{dx} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

Numériquement, cela revient à approximer la dérivée $f'(x_i)$ pour chaque point (i), par le rapport de l'écart entre les abscisses et les ordonnées des points précédent (i-1) et suivant (i+1), selon l'expression :

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1}))}{x_{i+1} - x_{i-1}}$$

Exercice 5 Diagrammes « nuages de points » superposés : évolution des émissions de CO₂ de quelques pays d'Europe au cours du temps

Objectif : représenter et comparer une grandeur (ici les émissions de CO₂) pour différentes catégories (ici les pays) en fonction d'une même variable (ici le temps)

Calc : Menu Insertion > Diagrammes > XY points reliés

Excel : Menu Insertion > Graphiques > Nuage de points reliés

Ouvrir le fichier **emissions-CO2.ods**.

Il contient les valeurs annuelles des émissions de CO₂, dues aux combustibles fossiles et à l'industrie, de 5 pays d'Europe de 1860 à 2021. Source des données : Our World in data <https://ourworldindata.org/co2-emissions>

Tracer sur un même graphe pour les 5 pays indiqués, les valeurs de CO₂ émis par année (en GtCO₂/an) par ces pays de 1860 à 2021. Les points peuvent être reliés pour un même pays. Commenter.

Exercice 6 Diagramme secteurs ("camembert") : part de chaque filière énergétique dans la production d'électricité en France en 2023

Objectif : créer un diagramme secteur ("camembert") adapté à représenter la répartition de valeurs continues, par exemple en %.

Calc : Menu Insertion > Diagrammes > Secteur

Excel : Menu Insertion > Graphiques > Secteur

Se rendre sur le **site des données publiques ouvertes françaises** <https://data.gouv.fr>

Chercher la production d'électricité annuelle par filière à la maille nationale.

URL directe sinon : <https://www.data.gouv.fr/fr/datasets/production-nationale-annuelle-par-filiere-2012-a-2021/>

La page permet de télécharger les données à différents **formats structurés de données** :

csv (*Coma Separated Values*), JSON (*JavaScript Object Notation*).

Télécharger le premier fichier au format csv. L'ouvrir avec LibreOffice Calc. Dans la fenêtre de dialogue de l'import des valeurs, ne rien changer et valider. Sauver ensuite le fichier au format .ods : energie.ods

On cherche à représenter la part de chaque filière (exprimée en %) dans la production d'électricité pour l'année 2023. *Remarquer que les valeurs des TWh produits ne sont pas ordonnées dans le fichier fourni.*

1. Dans la ligne 14, calculer le % de production d'électricité assurée par chaque filière en 2023.
2. Tracer un diagramme secteurs ("camembert") représentant ce % : pour cela, ne sélectionner que les cellules indiquant la filière et celles contenant les %.

Le diagramme secteur obtenu est-il lisible ?

3. On va maintenant ordonner les valeurs des % par ordre décroissant.
Créer une nouvelle feuille. Pour plus de clarté, on va placer dans un tableau uniquement les filières et les % : pour cela, copier-coller les cellules contenant les filières et les % (attention, faire un collage spécial : nombres et formats). On va ensuite disposer le tableau en colonnes : par copier-coller avec transposition (inversion lignes/colonnes). On peut alors plus facilement trier les % par ordre décroissant. Tracer alors le diagramme secteurs des % en lui donnant un titre.
Comparer à celui tracé question 2 et conclure.

Exercice 7 Diagramme à bâtons (colonnes : « histogramme », ou barres horizontales) : comparaison de données. Rats, nécropole.

Objectif : créer un diagramme bâton (i.e. vertical en colonnes, parfois nommé « histogramme », ou horizontal à barres) adapté à représenter et comparer des données dépendant de catégories discrètes

Calc : Menu Insertion > Diagrammes > Colonnes ou Barres

Excel : Menu Insertion > Graphiques > Histogramme

1) Comparaison des poids de catégories de rats

Ouvrir le fichier **stats_bio.ods** et aller dans l'onglet **rats**.

Le fichier contient un tableau de mesures de poids de 3 catégories de rats : témoins , traités anabolisants, traités placebo.

1. Calculer la moyenne de chaque catégorie et l'écart-type.
2. Tracer un diagramme bâton (« histogramme »), adapté à la représentation des données des moyennes par catégories de rats.
3. Placer les "barres d'erreur" (écart-type) sur ce diagramme.

Calc : clic droit, insérer des barres d'erreurs Y, plage de cellules, sélectionner la ligne des écart-types

Excel : Outil Format ou Clic droit sur les colonnes, format série 1, activer le bouton barres d'erreurs : par défaut c'est l'erreur standard qui est utilisée.

Ou par un clic sur le diagramme, menu Outils graphiques, Ajouter un élément graphique, Barres d'erreur (plusieurs types : erreur standard, écart-type...).

2) (Entraînement) Etude d'une nécropole

Ouvrir le fichier **stats_bio.ods** et aller dans l'onglet **nécropole**.

Le tableau fournit pour 10 sites archéologiques (les catégories discrètes), le nombre d'ossements trouvés (la variable -ici discrète, entière).

1. Calculer la moyenne, (+médiane), (+variance) et écart-type du nombre d'ossements sur les 10 sites puis sur 9 sites (n°1 à 9). Commenter.
2. Représenter sous forme de diagramme bâton (« histogramme ») la distribution du nombre d'ossements :
 - des 10 sites. Quel site a le plus grand nombre d'ossements ? En quoi le 10e site gêne-t-il l'analyse de la distribution des 9 autres ?
 - des 9 sites (n° 1 à 9) en excluant le site 10.

Exercice 8 Histogramme des fréquences de classes : glycémie

Objectif : représenter graphiquement une distribution de valeurs continues regroupées en classes, par un histogramme des fréquences, afin de visualiser la classe la plus fréquente et la dispersion.

Calc : pas encore de représentation directe Histogramme, donc il faut créer les classes et calculer les fréquences avant le tracé

Excel : Menu Insertion > icône des graphiques statistiques (icône de forme histogramme) > Histogramme

Ouvrir le fichier **stats_bio.ods** et aller dans l'onglet **glycémie**.

Le fichier contient des mesures de glycémie de différents patients classées par ordre croissant.

1. + (Entraînement) Calculer la moyenne, médiane, variance et écart-type des valeurs de glycémie.
2. Tracer l'histogramme des données de glycémie à partir du tableau 1.

Calc : il faut d'abord définir les classes regroupant les données, et calculer la fréquence des données de chaque classe par la fonction matricielle [FREQUENCE(Plage des données ; Plage des Classes)] (tableau à 2 colonnes à remplir). Puis tracer l'histogramme comme diagramme bâton.

Excel : directement à l'aide de la représentation Histogramme d'Excel.

Si besoin, changer la largeur des classes à une valeur que vous jugerez plus appropriée pour la

lisibilit  de l'histogramme.

Calc : changer le tableau des classes

Excel : clic droit, format des donn es. Attention, le volet droit doit parfois  tre  largi pour voir le champ de saisie de l'intervalle (largeur de la classe).

- Quelle est la classe de valeurs de glyc mie la plus fr quente dans l' chantillon de patients ?
Pour comparaison, une glyc mie   jeun normale est entre 0,74 et 1,06 g/L, avec une moyenne de 0,85 g/L.

Exercice 9 Nuage de points et r gressions² :  talonnage, Loi de Hubble

Objectif : tracer un graphique du type *Nuage de points* puis effectuer un ajustement de donn es sur une loi de *r gression lin aire*, la tracer sur le graphe exp rimental, et d terminer ses param tres ( quation et R^2)

Calc et Excel : utilisation de la fonction matricielle DROITEREG.

1) Droite d' talonnage en spectrophotom trie

La technique de spectrophotom trie peut  tre utilis e pour analyser la teneur d'une esp ce color e en solution, ou d'esp ces incolores transform es par r action en esp ces color es. Par exemple on peut l'utiliser pour doser des ions ou des mol cules dans les eaux (ici on dose les ions nitrates NO_3^- qui par r action conduisent   un colorant : d'o  la technique de spectrophotom trie employ e).

Les donn es du fichier **spectrophotometrie.ods** repr sentent l'absorbance A relative   des solutions de concentrations connues,   la longueur d'onde du maximum d'absorption de l'esp ce color e absorbante. Le fichier contient  galement la valeur de la mesure de A sur un  chantillon dont on cherche   conna tre la concentration en ions nitrates.

- Tracer le graphe des points servant   l' talonnage ($A = f(c)$).
- Effectuer sur ce graphe une **r gression lin aire** pour tracer la **droite de tendance**, en laissant l'ordonn e   l'origine libre de varier ($y = ax+b$), et faire afficher son  quation et R^2 .
Clic droit sur les points pour faire appara tre le menu contextuel avec les courbes de tendances.
- Faire afficher, avec la fonction matricielle DROITEREG, tous les param tres de la r gression lin aire : rep rer l'incertitude sur la pente, l'incertitude sur l'ordonn e   l'origine (cf explications dans le fichier). V rifier avec cette incertitude que la droite dans tous les cas contient l'origine.
S'aider de l'assistant fonctions fx pour DROITEREG.
- D duire de la mesure de A sur l' chantillon, la valeur de concentration cherch e,   l'aide de l' quation trouv e ainsi que son incertitude.

2) (Entra nement) Loi de Hubble

La loi de Hubble  tablit une relation de proportionnalit  entre la vitesse de r cession des galaxies et leur distance   l'observateur : $v = H_0 d$ o  H_0 est la constante de Hubble.

Le fichier **Loi_de_Hubble.ods** contient des observations de Hubble de 1929   partir desquelles on va v rifier cette loi et calculer H_0 .

- Tracer le graphe $v = f(d)$ puis effectuer une r gression lin aire pour tracer la droite de tendance et afficher son  quation et son coefficient de d termination R^2 . La loi est-elle lin aire ? En d duire la valeur de la constante de Hubble H_0 (estimation de l' poque).
- Faire afficher avec la fonction matricielle DROITEREG tous les param tres de la r gression lin aire : rep rer l'incertitude sur la pente, l'incertitude sur l'ordonn e   l'origine.

2. Autres termes utilis s pour r gression : ajustement, courbe de tendance, *fit* (en anglais)

Exercice 10 + Pour en savoir plus Diagramme à bâtons (« histogrammes ») groupés : comparaison des volumes des grains de maïs de différentes lignées

Objectif : créer un diagramme à bâtons groupés adapté pour représenter et comparer des ensemble de données (ici les volumes des 10 grains) dépendant de catégories discrètes (ici les lignées)

Ouvrir le fichier **grains-mais-4.xlsx**. (explications pour Excel).

On souhaite tracer un même diagramme regroupant plusieurs diagrammes chacun correspondant aux volumes de tous les grains d'une lignée, en vue de les comparer. On souhaite aussi ajouter le volume moyen de chaque lignée. On se limitera ici à comparer les 3 premières lignées.

Disposition des données : il faut pour cela disposer d'un tableau à double entrée avec les numéros des 3 lignées en colonnes, les numéros des 10 grains en lignes, et les volumes des grains dans les cellules au croisement. Les volumes moyens des grains par lignée ont également été calculés dans le fichier.

	Lignée 1	Lignée 2	Lignée 3
Grain 1	393,47	325,60	540,00
Grain 2	378,00	414,54	480,00
Grain 3	336,96	622,08	507,65
Grain 4	313,24	390,00	311,90
Grain 5	273,31	413,10	437,78
Grain 6	336,53	345,00	607,6
Grain 7	299,52	315,25	396,864
Grain 8	427,20	482,16	427,284
Grain 9	306,00	433,88	385,776
Grain 10	288,00	389,94	410,13
Volumes moyens	335,22	413,15	450,50

1. Sélectionner le tableau puis insérer un graphique à bâtons verticaux groupés, appelé dans Excel « Histogramme groupé » (soit via [Insertion > Graphiques recommandés], soit via [Insertion > icône nommée « Insérer un histogramme ou un graphique à barres » > Histogramme 2D > Histogramme groupé]).
2. Cliquer sur le graphique pour activer la catégorie de menus supplémentaires intitulée [Outils de graphique]. Cliquer sur le menu [Intervertir ligne/colonne] pour obtenir un graphique avec 3 ensembles de diagrammes associés chacun à une lignée.
3. On souhaite faire afficher sur le graphe les valeurs numériques des volumes moyens. Pour cela, faire un clic droit sur la colonne de droite de chaque lignée (i.e. la colonne des volumes moyens), puis dans le menu contextuel qui apparaît, choisir « Ajouter des étiquettes de données » : la valeur de chaque volume moyen apparaît alors sur le graphe.
4. On peut améliorer la présentation en choisissant des couleurs identiques pour les grains d'une même lignée. La procédure n'est pas automatique.

Dans la fenêtre latérale intitulée « Format de la zone graphique », choisir dans une flèche du menu « Série « grain n°1 » puis choisir la couleur de *remplissage* (double-cliquer sur une barre pour sélectionner la barre spécifique de cette lignée et pour mettre une couleur différente dans chaque lignée). Procéder ainsi pour chaque série de grains.

On peut de même mettre les colonnes des *moyennes* des volumes dans une autre couleur, ou une couleur assortie à la série mais plus intense, pour les distinguer des autres colonnes. (choisir aussi la couleur dans le volet de droite, *remplissage*).

Entraînement autonome : création du tableau à double entrée (On cherche ici à reproduire le tableau fourni)

1. On va se placer à partir de la colonne M. A partir de la cellule M1, créer le tableau à double entrée vide en intitulant les colonnes « Lignée n°... » et les lignes « Grain n°... »
2. Remplir ce tableau par copie-collage spécial des valeurs seules à l'aide des données de volume colonne F.