

*Agence Nationale de la Statistique et de la démographie (ANSD)*

*Ecole nationale de la statistique et de l’analyse économique*

Pierre Ndiaye (ENSAE)

**Dakar, Juin 2023**

**Résumé- Synthèse succ\_Exposé\_2023**

Groupe 2 Professeur:

Pathé diagne-Fanmara Sadio-Fallou Badji Mr Hema Aboubacar.

Elèves ingénieurs Statisticien Économiste en première année ISE.

[pathediagne25@gmail.com](mailto:pathediagne25@gmail.com)

# Objectifs generales

L'objectif général de ce document est de fournir un résumé synthétique complet et synthétique de l'ensemble des exposés présentés, mettant en évidence les principaux points discutés, les concepts clés abordés et les conclusions importantes tirées. Cela permettra a tout un chacun de bénéficier d'une vision d'ensemble de tous les sujets traités lors de l'exposé et de comprendre les principaux enseignements et messages transmis par les différents exposants.

Thème 3 : RMARKDOWN

Nous avons suivi avec beaucoup d’intérêt le troisième exposé qui a porté sur rmarkdown.

Dans cette présentation, les présentateurs ont exploré les fonctionnalités et l'importance de R Markdown, un langage de balisage puissant pour la génération de documents dynamiques.

Leurs développements et précisions nous ont permis de mieux comprendre l'importance de R Markdown dans la création de rapports interactifs, de présentations et d'autres types de documents. Cela nous a permis d'apprécier son utilité dans le domaine de l'analyse de données.

Ainsi, les points ci- dessous ont été développés tout au long de cet exposé.

1. **Fonctionnalités de R Markdown**.

S’agissant des fonctionnalités de rmarkdown les points importants ont été relevés.

* La syntaxe de base de R Markdown permet de combiner facilement du texte et du code dans un même document, facilitant ainsi la création de documents bien structurés.
* L'intégration de code R dans les documents R Markdown offre la possibilité d'exécuter des analyses, de générer des graphiques et de présenter les résultats de manière interactive.
* Avec R Markdown, les graphiques et les résultats peuvent être affichés de manière dynamique, permettant une exploration interactive des données.
* R Markdown prend en charge plusieurs formats de sortie tels que HTML, PDF et Word, offrant ainsi une flexibilité pour générer des documents dans le format souhaité.

1. **Gestion des tires et sous titres**

Nous avons appris comment gérer les titres et sous-titres en utilisant la syntaxe Markdown.

* Pour spécifier le titre principal de votre document, on utilise le symbole "#" suivi du texte du titre. Par exemple, "# Titre Principal".
* Sous-titres thématiques : On utilise le symbole "#" suivi du texte du sous-titre pour créer des sous-titres thématiques.
* Il est possible d’utiliser plusieurs niveaux de sous-titres en ajoutant des "#" supplémentaires. Par exemple, "## Sous-titre 1", "### Sous-titre 2".

1. **Utilisation et manipulation avancée de R Markdown A.**

* En utilisant R Markdown, il est possible d'inclure des images, des tableaux et des formules mathématiques pour enrichir les documents.
* La gestion des références bibliographiques dans R Markdown facilite l'intégration des sources et des citations dans les documents académiques.
* L'automatisation de la génération de rapports avec R Markdown permet de gagner du temps en générant automatiquement les documents à partir des données et du code.

*Résumé-synthèse-exposés-2023*

1. **Avantages de l'utilisation de R Markdown**

A travers l’exposé nous avons noté que l’utilisation de rmarkdown procure plusieurs avantages notamment:

* La reproductibilité des résultats qui est l'un des principaux avantages de l'utilisation de R Markdown, permettant de partager et de reproduire facilement les analyses.
* Grâce à R Markdown, le partage du code et des résultats devient simple, la collaboration et la transparence dans les projets d'analyse de données devient plus facile.
* L'automatisation de la génération de rapports avec R Markdown permet d'économiser du temps et de mettre en oeuvre des rapports d’analyse très présentables.

1. **Exemples d'utilisation de R Markdown**

* Exemples concrets de création de rapports interactifs avec R Markdown, mettant en évidence la flexibilité et l'interactivité offertes par cette approche.
* Nous avons vu comment R Markdown peut être utilisé pour générer des présentations Beamer, permettant une intégration harmonieuse du contenu et des résultats de l'analyse.
* Les exposants ont également examiné des exemples de production de documents scientifiques avec R Markdown, montrant comment il peut faciliter la création de documents bien structurés et professionnels.

En Conclusion, R Markdown offre des fonctionnalités puissantes pour la génération de documents dynamiques et reproductibles, combinant du texte, du code et des résultats dans un seul document.

Thème 4 : R - Quarto

R Quarto est un package R puissant et polyvalent pour la création de rapports dynamiques et interactifs. L'exposé sur R Quarto a mis en lumière ses fonctionnalités essentielles et son importance dans la génération de rapports.

1. **Fonctionnalités de R Quarto :**

Syntaxe claire et simple : R Quarto utilise une syntaxe intuitive qui permet de combiner facilement du texte et du code pour créer des rapports bien structurés.

Intégration de code R : R Quarto permet d'inclure du code R dans les rapports, ce qui facilite l'analyse des données, la création de graphiques et la génération de résultats interactifs.

1. **Personnalisation avancée**

R Quarto offre une grande flexibilité en termes de personnalisation des rapports, avec des options pour modifier l'apparence, le style et le formatage du contenu.

Prise en charge de multiples formats de sortie : R Quarto prend en charge plusieurs formats de sortie tels que HTML, PDF et Word, offrant ainsi une adaptabilité pour générer des rapports dans le format souhaité.

**II. Gestion des titres et sous-titres :**

R Quarto utilise une syntaxe similaire à R Markdown pour gérer les titres et les sous-titres. Les symboles "#" sont utilisés pour spécifier les niveaux de titres, permettant ainsi une hiérarchisation claire du contenu.

En utilisant cette syntaxe, il est facile de créer des titres principaux, des sous-titres thématiques et des sous-sous-titres pour organiser efficacement le rapport.

**III. Utilisation et manipulation avancée de R Quarto** :

L’exposant montre que R Quarto offre des fonctionnalités avancées pour enrichir les rapports, y compris l'inclusion d'images, de tableaux et de formules mathématiques.

La gestion des références bibliographiques dans R Quarto facilite l'intégration des sources et des citations dans les documents académiques.

L'automatisation de la génération de rapports avec R Quarto permet de gagner du temps en générant automatiquement les rapports à partir des données et du code.

**IV. Avantages de l'utilisation de R Quarto :**

La principale force de R Quarto réside dans sa capacité à créer des rapports interactifs et dynamiques, permettant aux utilisateurs d'explorer les données et d'interagir avec les résultats.

R Quarto favorise la reproductibilité des résultats, facilitant le partage et la reproduction des analyses.

Grâce à R Quarto, le partage du code, des résultats et des visualisations devient simple, favorisant la collaboration et la transparence dans les projets d'analyse de données.

En résumé, l'exposé sur R Quarto a souligné son utilité dans la création de rapports dynamiques et interactifs. Ses fonctionnalités, la gestion des titres et sous-titres, ainsi que l'utilisation avancée avec inclusion d'images et de références bibliographiques ont été discutées. Les avantages de R Quarto incluent la reproductibilité des résultats, le partage facile du code et des résultats, ainsi que la facilité d'automatisation de la génération de rapports.

Thème 6 :  **R\_vers\_python-Reticulate**

Nous avons assisté avec beaucoup d'intérêt à la présentation du groupe 9 qui portait sur la transition de R vers Python en utilisant le package Reticulate. Voici un résumé synthétique des points saillants de l'exposé

**Introduction à Reticulate :**

Les présentateurs ont commencé par expliquer ce qu'est Reticulate, un package de R permettant d'utiliser des fonctions et des scripts R dans un environnement Python. Cela permet aux utilisateurs de tirer parti des fonctionnalités de R tout en bénéficiant de l'écosystème Python. A travers l’expose nous sommes capable de realiser les manipulations en amont, nécessaires pour pouvoir installer le package reticulate, les extensions python nécessaires, etc…

**Avantages de la transition :** Les présentateurs ont souligné les avantages de passer de R à Python en utilisant Reticulate. Cela inclut la possibilité de combiner les capacités de R et de Python dans un même projet, d'accéder à une plus grande variété de packages et de bénéficier d'une plus grande flexibilité dans le développement et le déploiement de projets.

**Syntaxe et utilisation de Reticulate :**

Les présentateurs ont présenté la syntaxe de base de Reticulate et ont montré des exemples concrets d'utilisation. Ils ont souligné la simplicité de l'intégration de scripts R dans du code Python existant et ont expliqué comment utiliser des fonctions R dans Python.

**Conversion de données :** Un aspect important de la transition de R à Python est la conversion des données entre les deux langages. Les présentateurs ont discuté des différentes méthodes de conversion de données, y compris l'utilisation de bibliothèques telles que pandas pour manipuler les données dans Python.

**Cas d'utilisation** : Les présentateurs ont partagé des exemples de cas d'utilisation où la combinaison de R et de Python avec Reticulate a été bénéfique. Cela inclut l'analyse de données, la visualisation, l'apprentissage automatique et d'autres domaines de l'analyse quantitative.

En conclusion, cet exposé sur la transition de R vers Python avec Reticulate nous a permis de comprendre les avantages et les possibilités offertes par l'utilisation de ce package. Nous avons appris comment intégrer facilement des scripts R dans du code Python, convertir des données entre les deux langages et exploiter les fonctionnalités de chaque langage dans nos projets d'analyse de données.

Thème 9 : R - Cartographie

Voici un résumé synthétique des points saillants de la cartographie avec R presente par le groupe 9 : les exposants ont expliciter les points essentiels qui suivent.

**Importation des données géographiques** :

D’abord ils ont précisé l’importance de la cartographie dans le domaine de l’analyse des données aujourd’hui. Ainsi, la méthodologie de l’importation des donnée a été mis en exergue. Pour cela, des packages comme "sf" ou "sp" pour importer des données géographiques dans R. Nous pouvons charger des fichiers shapefile, des données au format GeoJSON ou utiliser des données en ligne via des API.

**Préparation des données** :

Avant de créer une carte, il est nécessaire de préparer vos données. Cela peut inclure la projection des coordonnées, la fusion de différentes sources de données ou la manipulation des attributs.

**Création de la carte de base** :

Nous pouvons utiliser des packages tels que "ggplot2" pour créer une carte de base. Vous pouvez spécifier les coordonnées géographiques, ajouter des fonds de carte, des axes et des légendes.

**Ajout des données :**

Nous pouvons également utiliser les données préparées pour ajouter des éléments géographiques à votre carte. Par exemple, nous pouvons ajouter des points, des lignes ou des polygones correspondant à vos données. Les attributs peuvent être utilisés pour définir les couleurs, les tailles ou les symboles des éléments.

**Personnalisation de la carte :**

Ils nous ont suggéré des fonctionnalités de personnalisation de R pour ajouter des titres, des étiquettes, des échelles, des légendes ou d'autres éléments de design à votre carte. Il est possible également d’appliquer des techniques de mise en forme graphique pour améliorer l'esthétique et la lisibilité.

**Exportation et partage :**

Une fois votre carte créée, nous pouvons l'exporter au format souhaité, comme une image (PNG, JPEG) ou un document PDF. nous pouvons également intégrer vos cartes dans des rapports ou des applications interactives.

Il est important de noter que la cartographie avec R est un domaine vaste et complexe, et qu'il existe de nombreux packages et fonctionnalités avancées pour répondre à des besoins spécifiques. Leur présentation synthétique vise à fournir une vue d'ensemble générale du processus de cartographie avec R.

窗体顶端

En conclusion, nous retenons que la cartographie via le logiciel R offre de nombreuses fonctionnalités pour créer des représentations visuelles de données géographiques. À travers les étapes présentées, nous pouvons importer et préparer vos données, créer une carte de base, ajouter des éléments géographiques et personnaliser votre carte selon vos besoins.

R offre également des possibilités d'exportation et de partage de nos cartes.

La cartographie avec R peut être utilisée dans de nombreux domaines tels que l'analyse de données, la géographie, l'épidémiologie, l'environnement, l'urbanisme et bien d'autres. Elle permet de visualiser et d'analyser des informations spatiales de manière efficace et attrayante.

Thème 8: R - Shiny

Nous avons assisté avec grand intérêt à l'exposé sur RShiny, un package de R dédié à la création d'applications web interactives. Voici un résumé synthétique des points saillants de l'exposé :

**Introduction à RShiny** :

Les présentateurs ont expliqué que RShiny est une bibliothèque qui permet de développer des applications web interactives directement à partir de R. Cela permet aux utilisateurs de partager leurs analyses et leurs visualisations de données de manière dynamique et conviviale.

**Structure d'une application Shiny :**

Les présentateurs ont expliqué la structure de base d'une application Shiny, qui comprend deux éléments principaux : le fichier UI (User Interface) et le fichier Server. Le fichier UI définit l'apparence et les éléments de l'interface utilisateur, tandis que le fichier Server contient le code R pour le traitement des données et la génération des sorties.

**Création de l'interface utilisateur :**

Les présentateurs ont montré comment créer une interface utilisateur attrayante et conviviale à l'aide des différents éléments disponibles dans Shiny, tels que les boutons, les cases à cocher, les menus déroulants, les graphiques interactifs, etc. Ils ont également souligné l'importance de la mise en page et de l'organisation des éléments pour une meilleure expérience utilisateur.

**Programmation réactive :**

Un aspect clé de Shiny est la programmation réactive, qui permet de lier les entrées de l'utilisateur aux sorties correspondantes de manière dynamique. Les présentateurs ont expliqué comment définir les réactions entre les différents éléments de l'interface et les opérations de calcul dans le fichier Server, assurant ainsi une mise à jour automatique des sorties en fonction des actions de l'utilisateur.

**Personnalisation et déploiement de l'application :**

Les présentateurs ont montré comment personnaliser l'apparence de l'application Shiny en utilisant des thèmes prédéfinis ou en personnalisant les styles CSS. Ils ont également discuté des options de déploiement, notamment l'exécution de l'application localement ou son déploiement sur un serveur web pour un accès public.

**Cas d'utilisation :**

Les présentateurs ont partagé des exemples de cas d'utilisation où l'utilisation de RShiny a été bénéfique, tels que la création de tableaux de bord interactifs, la visualisation de données dynamiques, la création d'outils d'aide à la décision, etc.

En conclusion, cet exposé sur RShiny nous a permis de comprendre comment créer des applications web interactives à l'aide de R. Grâce à sa facilité d'utilisation et à sa puissance, RShiny offre la possibilité de partager des analyses de données de manière conviviale et dynamique. L'utilisation de RShiny ouvre de nouvelles perspectives pour la diffusion et l'interaction avec les résultats de nos analyses, offrant ainsi une expérience utilisateur enrichissante et engageante.

窗体底端

Theme 7: résolution d’ équations non linéaires

Nous avons assisté à une présentation intéressante sur la résolution d'équations non linéaires avec R. Voici un résumé synthétique des points clés abordés lors de cette présentation :

**Introduction aux équations non linéaires :**

Les présentateurs ont expliqué ce qu'est une équation non linéaire, qui est une équation où les termes contiennent des variables élevées à des puissances autres que 1. Les équations non linéaires peuvent être plus complexes à résoudre que les équations linéaires.

**Méthodes de résolution d'équations non linéaires :**

Les présentateurs ont présenté différentes méthodes couramment utilisées pour résoudre les équations non linéaires. Cela comprend des méthodes itératives telles que la méthode de Newton-Raphson, la méthode de la sécante, la méthode de la dichotomie, la méthode de Brent, etc.

**Utilisation des fonctions de R pour résoudre des équations non linéaires :**

R propose des fonctions prédéfinies pour résoudre des équations non linéaires. Les présentateurs ont montré comment utiliser ces fonctions, telles que "uniroot()", "optimize()", "nleqslv()", etc., en fournissant des exemples concrets.

En résolution d'équations non linéaires, on distingue généralement deux grandes catégories de méthodes : les méthodes directes (ou explicites) et les méthodes indirectes (ou itératives). Voici une brève explication de ces deux types de méthodes :

**Méthodes directes (ou explicites) :**

Les méthodes directes cherchent à trouver une solution exacte à l'équation non linéaire en utilisant des formules mathématiques ou des algorithmes spécifiques. Elles sont basées sur des propriétés analytiques ou géométriques des équations. Les méthodes directes peuvent être utilisées lorsque l'équation a une forme particulière ou lorsque des simplifications peuvent être faites pour résoudre directement l'équation. Cependant, ces méthodes ne sont généralement applicables qu'à des cas spécifiques et ne sont pas toujours disponibles pour toutes les équations non linéaires.

**Méthodes indirectes (ou itératives) :**

Les méthodes indirectes sont les plus couramment utilisées pour résoudre des équations non linéaires de manière numérique. Elles se basent sur des approximations successives de la solution en utilisant des itérations. Les méthodes indirectes commencent souvent par une estimation initiale de la solution et effectuent des itérations successives en utilisant des formules itératives pour converger vers la solution. Les méthodes itératives populaires comprennent la méthode de Newton-Raphson, la méthode de la sécante, la méthode de la dichotomie, la méthode de Brent, etc. Ces méthodes sont souvent plus flexibles et applicables à un large éventail d'équations non linéaires, mais elles peuvent nécessiter plus d'itérations pour converger vers la solution et peuvent être sensibles au choix de l'estimation initiale.

Il est important de choisir la méthode la plus appropriée en fonction de la nature de l'équation non linéaire, des contraintes de calcul et des objectifs spécifiques de la résolution. Les méthodes directes peuvent être plus efficaces pour certaines équations particulières, tandis que les méthodes indirectes offrent généralement une approche plus générale et adaptable à diverses situations.

**Personnalisation et ajustement des méthodes :**

Les présentateurs ont souligné qu'il est souvent nécessaire de personnaliser et d'ajuster les méthodes de résolution en fonction de la nature spécifique de l'équation non linéaire à résoudre. Ils ont expliqué comment cela peut être réalisé en utilisant des options et des arguments spécifiques dans les fonctions de résolution de R.

**Cas d'utilisation et limitations :**

Les présentateurs ont partagé des exemples de cas d'utilisation où la résolution d'équations non linéaires avec R a été utile, tels que l'optimisation de fonctions, l'estimation de paramètres, la modélisation mathématique, etc. Ils ont également discuté des limitations potentielles des méthodes et des erreurs numériques associées.

En conclusion, la résolution d'équations non linéaires avec R offre une gamme de méthodes et de fonctions pour aborder ce type de problèmes mathématiques complexes. R propose des outils puissants et flexibles pour résoudre ces équations, ce qui facilite l'analyse et la modélisation dans divers domaines. Cependant, il est important de comprendre les caractéristiques des équations non linéaires et de personnaliser les méthodes de résolution en fonction des exigences spécifiques de chaque situation.

窗体顶端

Theme 8: R-Text\_Mining

Nous avons assisté à une présentation intéressante sur l'exploration de textes (text mining) avec R. Voici un résumé synthétique des points saillants de l'exposé :

**Introduction à l'exploration de textes :**

Les présentateurs ont expliqué que l'exploration de textes est un domaine qui vise à extraire des informations utiles et significatives à partir de textes non structurés tels que des documents, des articles, des tweets, etc. Cela peut inclure l'analyse des sentiments, la classification de texte, l'extraction d'entités, la détection de thèmes, etc.

**Importation et prétraitement des données textuelles :**

Les présentateurs ont montré comment importer des données textuelles dans R à partir de différents formats tels que des fichiers texte, des bases de données ou des sources en ligne. Ils ont également discuté des techniques de prétraitement des données textuelles, telles que la suppression des mots vides (stop words), la lemmatisation, la normalisation des termes, la suppression des ponctuations, etc.

**Analyse des sentiments :**

Les présentateurs ont présenté des techniques d'analyse des sentiments qui permettent de déterminer les sentiments associés à un texte. Ils ont expliqué comment utiliser des dictionnaires de sentiments ou des modèles d'apprentissage automatique pour attribuer des polarités positives, négatives ou neutres aux textes.

**Classification de texte :**

Les présentateurs ont abordé la classification de texte, qui consiste à attribuer des étiquettes ou des catégories prédéfinies aux documents. Ils ont montré comment utiliser des méthodes telles que la classification bayésienne naïve, les machines à vecteurs de support (SVM) ou les réseaux de neurones pour entraîner des modèles de classification et les appliquer aux nouveaux textes.

**Visualisation des résultats :**

Les présentateurs ont souligné l'importance de la visualisation pour comprendre et présenter les résultats de l'exploration de textes. Ils ont montré comment créer des nuages de mots, des graphiques de fréquence de termes, des diagrammes de réseaux de co-occurrence, etc., pour visualiser les informations extraites des textes.

**Applications et limites :**

Les présentateurs ont partagé des exemples d'applications de l'exploration de textes dans différents domaines tels que l'analyse des médias sociaux, l'analyse de sentiments pour les avis clients, l'extraction d'entités dans les articles de presse, etc. Ils ont également discuté des limites potentielles de l'exploration de textes, telles que la qualité des données, la représentativité des échantillons, etc.

En conclusion, l'exploration de textes avec R offre des outils puissants pour analyser et comprendre les données textuelles. Grâce à des techniques telles que l'analyse des sentiments et la classification de texte, on peut extraire des informations précieuses et des connaissances à partir de textes non structurés. La visualisation des résultats permet de communiquer de manière efficace les insights obtenus. Cependant, il est important de prendre en compte les défis spécifiques liés à la nature subjective et complexe des données textuelles lors de l'exploration de textes avec R.

Theme 1: R-Package\_Janitor

L'exposé portait sur le package "janitor" en mettant l'accent sur son utilité pour le nettoyage et la préparation des données. Voici un résumé synthétique des points clés de l'exposé :

**Introduction à "janitor" :**

Le présentateur a expliqué que "janitor" est un package R qui facilite le nettoyage et la préparation des données. Il a souligné l'importance d'avoir une base de données propre et bien préparée avant d'effectuer des analyses.

**Fonctions de nettoyage des noms de variables :**

L'exposé a mis en avant la fonction "clean\_names()" qui permet de changer les noms des variables dans une base de données pour les rendre plus compréhensibles et significatifs. Cela améliore la lisibilité du code et facilite l'interprétation des résultats.

**Gestion des valeurs manquantes :**

Le package "janitor" propose la fonction "remove\_empty()" qui permet de supprimer les valeurs manquantes dans une base de données. Cela contribue à améliorer la qualité des données et évite les biais dans les analyses.

**Élimination des colonnes constantes :**

L'exposé a abordé la fonction "remove\_constant()" qui permet de détecter et de supprimer les colonnes qui contiennent des valeurs constantes, c'est-à-dire des colonnes où toutes les observations ont la même valeur. Cela permet de réduire la redondance dans les données.

Identification et suppression des doublons : Le package "janitor" offre la fonction "get\_dupes()" qui permet d'identifier les observations en double dans une base de données. Cette fonction facilite le nettoyage des données en éliminant les duplications indésirables.

**Limites du package :** L'exposé a souligné que le package "janitor" ne fournit pas de fonctionnalités pour générer des tableaux. Pour formater les tableaux, il est nécessaire de faire appel à d'autres fonctions provenant d'autres packages.

En conclusion, le package "janitor" est un outil précieux pour le nettoyage et la préparation des données. Il propose des fonctions pratiques pour renommer les variables, gérer les valeurs manquantes, éliminer les colonnes constantes et détecter les doublons. Cependant, il ne fournit pas de fonctionnalités pour la génération de tableaux, nécessitant l'utilisation d'autres packages pour cette tâche.

Theme 2 le Package gt-summury

L'exposé du groupe 2 portait sur le package "gtsummary" qui permet de créer des tableaux statistiques descriptifs et de combiner plusieurs tableaux. Voici un résumé synthétique des points clés de l'exposé :

**Introduction à "gtsummary" :**

Les présentateurs ont expliqué que "gtsummary" est un package R qui facilite la création de tableaux statistiques descriptifs. Il offre des fonctionnalités pour générer des tableaux récapitulatifs et des tableaux croisés, ainsi que pour personnaliser l'apparence des tableaux.

**Fonctionnalités clés :**

L'exposé a mis en évidence plusieurs fonctions essentielles du package "gtsummary" :

Fonction tbl\_summary() : Cette fonction permet de créer des tableaux statistiques descriptifs en spécifiant les variables d'intérêt. Elle affiche des mesures statistiques telles que les moyennes, les médianes, les écarts-types, etc.

Fonction flextable() : Elle permet d'exporter les tableaux générés avec "gtsummary" dans différents formats tels que Word, PDF, etc. Cela facilite le partage des résultats de l'analyse.

Fonction digits() : Elle permet d'ajuster le nombre de chiffres après la virgule dans les résultats numériques affichés dans les tableaux.

Fonction tbl\_cross() : Cette fonction permet de créer des tableaux de contingence en spécifiant les variables à croiser. Elle facilite l'analyse des relations entre les variables catégorielles.

Fonctions de fusion et de division de tableaux : Le groupe a présenté les fonctions tbl\_stack(), tbl\_merge() et tbl\_split() qui permettent de combiner ou de diviser des tableaux pour obtenir des regroupements ou des sous-ensembles spécifiques.

Fonction tbl\_strata() : Elle permet de comparer des groupes spécifiques à l'intérieur d'autres groupes dans un tableau croisé.

Présentation de gtsummary-compact : Les présentateurs ont mentionné le style "gtsummary-compact" qui offre une apparence condensée et concise aux tableaux, facilitant ainsi la visualisation et l'interprétation des résultats.

En conclusion, le package "gtsummary" offre des fonctionnalités pratiques pour créer des tableaux statistiques descriptifs et croisés. Il permet de combiner, diviser et personnaliser les tableaux pour répondre aux besoins spécifiques de l'analyse. Grâce à des fonctions telles que tbl\_summary() et tbl\_cross(), il est possible de générer rapidement des tableaux informatifs et de les exporter dans différents formats avec flextable(). L'ajustement du nombre de chiffres décimaux avec digits() permet de contrôler la précision des résultats affichés. Le style "gtsummary-compact" offre une présentation condensée pour une visualisation plus claire des tableaux.

窗体顶端

R

Theme 10 La Programmation Parallèleavec R

L'exposé portait sur le calcul parallèle en utilisant R. Voici un résumé synthétique des points clés de l'exposé :

**Introduction au calcul parallèle :**

Les présentateurs ont souligné l'importance du calcul parallèle pour accélérer les calculs et traiter des volumes de données plus importants. Ils ont expliqué que le calcul parallèle consiste à diviser une tâche en plusieurs sous-tâches qui sont exécutées simultanément sur plusieurs processeurs ou cœurs.

**Packages pour le calcul parallèle :**

L'exposé a mis en évidence les packages R les plus couramment utilisés pour le calcul parallèle, tels que "parallel" et "foreach". Ces packages fournissent des fonctionnalités pour exécuter des boucles parallèles et effectuer des calculs sur plusieurs cœurs ou machines.

**Boucles parallèles avec "foreach" :**

Les présentateurs ont expliqué comment utiliser la fonction "foreach" pour créer des boucles parallèles. Cette approche permet d'effectuer des calculs sur des itérations indépendantes de manière simultanée, réduisant ainsi le temps d'exécution global.

**Partage des données :**

L'exposé a abordé les différentes méthodes de partage des données entre les processus parallèles, telles que le partage explicite des données ou l'utilisation de mémoires partagées.

Gestion des résultats : Les présentateurs ont également souligné l'importance de la gestion des résultats lors de l'utilisation du calcul parallèle. Ils ont expliqué comment collecter et combiner les résultats des différentes tâches parallèles pour obtenir le résultat final.

**Limites du calcul parallèle :**

L'exposé a mentionné certaines limitations du calcul parallèle, telles que la nécessité d'avoir du matériel adapté avec plusieurs cœurs de processeur et une quantité suffisante de mémoire. Il a également souligné que certaines tâches ne peuvent pas être facilement parallélisées en raison de leur nature séquentielle ou de dépendances entre les étapes.

En conclusion, le calcul parallèle est une approche puissante pour accélérer les calculs et traiter des volumes importants de données en utilisant R. Les packages tels que "parallel" et "foreach" offrent des fonctionnalités pour créer des boucles parallèles et exécuter des calculs sur plusieurs cœurs ou machines. Cependant, il est important de prendre en compte les limitations matérielles et les contraintes de dépendances lors de l'utilisation du calcul parallèle.

# Conclusions et remarques

Nous avons assisté au différents exposés que nous jugeons très intéressants sur différents sujets, notamment R Markdown et le package Reticulate, R vers Excel, RCartographie……..pour l’interopérabilité entre R et divers outils de traitement, d’analyse et d’exploration des données. Ces exposés ont mis en avant les fonctionnalités et l’importance de ces outils dans le domaine de l’analyse de données, en permettant la création de documents dynamiques et interactifs, ainsi que leur utilisation conjointe avec le langage R.

Enfin, ces exposés ont fourni des informations précieuses sur des outils puissants pour l’analyse de données, en mettant en avant leur utilité et leurs avantages pour les utilisateurs. Ils nous ont également permis en tant que participants de mieux comprendre comment ces outils peuvent être utilisés dans nos propres projets d’analyse de données et être plus efficient plus performants.