

**Muriel ZOUZZOU**

**NOTE METHODOLOGIQUE**

-

**ANALYSE EN COMPOSANTE PRINCIPALE**

**Antoine ROLLAND**

## **1. Objectif du projet :**

L'objectif de ce projet est d'explorer les habitudes de sommeil, l'activité physique, le stress et les signes vitaux d'une population de 373 individus en utilisant une **Analyse en Composantes Principales (ACP)**. L'objectif principal est d'identifier des profils types de sommeil et de santé, tout en explorant les liens entre ces profils et des facteurs sociodémographiques tels que le **genre** et la **catégorie d'IMC** (indice de masse corporelle).

## **2. Sélection des variables :**

Les variables sélectionnées pour l'analyse sont de deux types :

- Variables quantitatives continues :**

Nous avons choisi sept variables liées au sommeil et à la santé physique :

- Durée de sommeil (Sleep.Duration)
- Qualité du sommeil (Quality.of.Sleep)
- Niveau d'activité physique (Physical.Activity.Level)
- Niveau de stress (Stress.Level)
- Pression artérielle (Blood.Pressure)
- Fréquence cardiaque (Heart.Rate)
- Nombre de pas quotidiens (Daily.Steps)

Ces variables sont toutes des indicateurs pertinents pour évaluer l'hygiène de vie et la santé d'un individu.

- Variables qualitatives :**

Deux variables qualitatives ont été intégrées dans l'ACP pour étudier les influences sociodémographiques sur les comportements de sommeil :

- **Genre** (homme/femme)
- **Catégorie d'IMC** (normal, surpoids, obèse...)

Le genre et la catégorie d'IMC ont été intégrés dans l'ACP pour analyser l'impact des facteurs sociodémographiques sur les comportements de sommeil et la santé. Le genre peut influencer la qualité du sommeil et les niveaux de stress, tandis que l'IMC est lié à des problèmes de sommeil (comme l'apnée) et à l'activité physique. En ajoutant ces variables, on explore les différences entre les groupes (hommes/femmes, normal/surpoids/obésité) pour mieux comprendre leurs effets sur la santé physique et mentale. Cela permet d'identifier des profils de comportement spécifiques à chaque groupe.

## **3. Préparation des données :**

Avant de commencer l'ACP, les données ont été préparées comme suit :

- Centrage et réduction des variables :**

L'ACP nécessite que les variables soient **centrées et réduites**, ce qui signifie que chaque variable a été ajustée pour avoir une moyenne de 0 et un écart-type de 1. Cela permet d'éviter que des variables avec des échelles de mesure différentes influencent de manière disproportionnée les résultats.

- **Transformation des variables qualitatives :**

Les variables **Genre** et **Catégorie d'IMC** ont été transformées en variables factorielles, ce qui permet de les intégrer correctement dans l'ACP sans perturber l'analyse des données numériques.

#### **4. Application de l'ACP :**

L'ACP a été réalisée sur les sept variables quantitatives et ensuite sur un sous-ensemble des données incluant des variables qualitatives (genre et IMC). Les principales étapes ont été :

- **Calcul des composantes principales :**

L'ACP permet de transformer les variables originales en quelques "composantes principales" qui résument la majorité de l'information. Les deux premières composantes principales expliquent près de 74 % de la variance totale des données. La première composante (47 %) met en évidence une opposition entre la qualité du sommeil et le stress élevé, tandis que la deuxième composante (27 %) distingue le niveau d'activité physique du mode de vie plus sédentaire.

- **Visualisation des résultats :**

Des graphiques ont été générés pour visualiser les relations entre les variables et les profils individuels dans l'espace factoriel (par exemple, avec les graphiques de variables et d'individus réalisés à l'aide de la fonction `fviz_pca_var` de la bibliothèque `factoextra`).

#### **5. Analyse des facteurs sociodémographiques (Genre et IMC) :**

En intégrant des variables qualitatives dans l'ACP, nous avons exploré les différences entre les hommes et les femmes, ainsi qu'entre les individus en fonction de leur catégorie d'IMC. Les résultats montrent, par exemple, que les personnes en surpoids ou obèses sont généralement associées à un stress plus élevé, une fréquence cardiaque plus importante et une activité physique plus faible, ce qui influence leur profil de sommeil.

#### **6. Résultats et interprétation :**

Les résultats de l'ACP révèlent deux grands profils :

- **Profil "reposé/actif"** : individus qui dorment bien, sont peu stressés, physiquement actifs et en bonne santé.
- **Profil "sous tension"** : individus avec un sommeil de moins bonne qualité, un stress élevé, une fréquence cardiaque plus élevée et une faible activité physique.

L'ACP montre des liens clairs entre **l'hygiène de vie (sommeil et activité physique)**, le **stress** et **l'état de santé**. Ces résultats confirment l'importance d'un sommeil réparateur et d'une activité physique régulière pour limiter les effets négatifs du stress et améliorer les paramètres physiologiques.

## **CONCLUSION**

Cette analyse en composantes principales (ACP) a permis d'identifier des profils distincts de sommeil et de santé au sein d'une population de 373 individus, en mettant en lumière l'interaction entre le sommeil, l'activité physique, le stress et les signes vitaux. Les deux grands profils identifiés, "**reposé/actif**" et "**sous tension**", illustrent clairement l'impact d'une hygiène

de vie saine (bonne qualité de sommeil et activité physique régulière) sur le bien-être physique et mental. En outre, l'intégration de variables sociodémographiques telles que le genre et l'IMC a permis d'approfondir l'analyse en identifiant des différences importantes entre les groupes d'individus. Par exemple, les personnes en surpoids ou obèses présentent des comportements de sommeil et des indicateurs de santé moins favorables, ce qui souligne la nécessité d'adapter les recommandations en fonction des caractéristiques individuelles. En somme, cette étude confirme l'importance de maintenir un équilibre entre sommeil, activité physique et gestion du stress pour améliorer la qualité de vie et les paramètres physiologiques, et met en évidence la pertinence d'une approche individualisée dans l'analyse de la santé publique.

## **INFORMATIONS :**

### **Pourquoi du vert ?**

J'ai choisi le vert pour la couleur des titres de ce rapport car il évoque la santé, l'équilibre et le bien-être, en harmonie avec le thème de l'étude sur les habitudes de sommeil et la santé physique.