

# **Projet StressLess**

Analyse, visualisation et suivi du niveau de stress

**Préparé par :**  
AKHRAIS Hasnae

## Table des matières

|   |          |
|---|----------|
| <b>1 Le Problème (Business Case)</b>          | <b>2</b> |
| 1.1 Problématique centrale . . . . .          | 2        |
| <b>2 Objectif du projet</b>                   | <b>2</b> |
| <b>3 Description des variables</b>            | <b>2</b> |
| <b>4 Nature des données</b>                   | <b>2</b> |
| <b>5 Problématique des données manquantes</b> | <b>3</b> |
| <b>6 Stratégie de nettoyage adoptée</b>       | <b>3</b> |
| <b>7 Normalisation des variables</b>          | <b>3</b> |
| <b>8 Data Leakage</b>                         | <b>3</b> |
| <b>9 Distribution du niveau de stress</b>     | <b>3</b> |
| <b>10 Évolution temporelle du stress</b>      | <b>4</b> |
| <b>11 Analyse des corrélations</b>            | <b>4</b> |
| <b>12 Analyse des scénarios de test</b>       | <b>4</b> |
| <b>13 Indicateurs clés de suivi</b>           | <b>4</b> |
| <b>14 Fonctions d'évaluation</b>              | <b>4</b> |
| 14.1 Évolution temporelle du stress . . . . . | 4        |
| 14.2 Matrice de corrélation . . . . .         | 5        |
| 14.3 Diagramme de Pareto . . . . .            | 5        |
| 14.4 Heatmap activité / stress . . . . .      | 5        |

## 1 Le Problème (Business Case)

Le stress chronique est un enjeu majeur de santé publique et de performance professionnelle. Une mauvaise gestion du stress peut entraîner :

- baisse de productivité,
- troubles du sommeil,
- risques psychosociaux,
- dégradation du bien-être général.

Le stress est une variable **subjective, évolutive dans le temps**, et influencée par plusieurs facteurs simultanés (activité, sommeil, charge de travail).

### 1.1 Problématique centrale

Comment analyser, suivre et anticiper l'évolution du niveau de stress d'un individu à partir de données temporelles et comportementales ?

## 2 Objectif du projet

Le projet **StressLess** vise à construire une chaîne complète d'analyse de données permettant :

- d'explorer les niveaux de stress observés,
- d'identifier les tendances temporelles,
- de comprendre les relations entre stress et facteurs explicatifs,
- de fournir des indicateurs exploitables par des professionnels du bien-être.

L'objectif est descriptif et préventif, afin d'anticiper les périodes de stress élevé.

Les Données (Input du Système)

## 3 Description des variables

| Variable     | Description                             |
|--------------|---|
| date         | Date ou horodatage de l'observation     |
| stress_level | Niveau de stress (échelle 0–100 ou 1–5) |
| activity     | Niveau ou intensité d'activité          |
| sleep        | Qualité ou durée du sommeil             |
| workload     | Charge de travail                       |
| user_id      | Identifiant utilisateur                 |

TABLE 1 – Description des variables du projet StressLess

## 4 Nature des données

- Données **temporelles** (séries chronologiques),

- Données **quantitatives continues**,
- Données potentiellement **bruitées ou incomplètes**,
- Possibilité de **variations inter-individuelles**.

Analyse Approfondie : Nettoyage et Préparation des Données (Data Wrangling)

## 5 Problématique des données manquantes

Les algorithmes statistiques et analytiques ne peuvent pas fonctionner correctement en présence de valeurs manquantes (NaN). Une valeur absente peut fausser :

- les moyennes,
- les corrélations,
- les visualisations temporelles.

## 6 Stratégie de nettoyage adoptée

- Imputation des valeurs manquantes par la moyenne ou la médiane,
- Suppression des valeurs aberrantes (ex : stress négatif, stress supérieur au seuil maximal),
- Uniformisation des formats temporels :

Exemple de code Python

```
df['date'] = pd.to_datetime(df['date']) df = df.sort_values('date')
```

## 7 Normalisation des variables

Pour homogénéiser les échelles et faciliter l'analyse, une normalisation Min-Max a été appliquée :

Exemple de code Python

```
from sklearn.preprocessing import MinMaxScaler df['stress_norm'] = MinMaxScaler().fit_transform(df[['stresslevel']])
```

## 8 Data Leakage

Dans une approche rigoureuse, le nettoyage doit être calculé uniquement sur le jeu d'entraînement pour éviter toute fuite d'information vers le test.

Analyse Exploratoire des Données (EDA)

## 9 Distribution du niveau de stress

Cette étude permet de répondre aux questions suivantes :

- Le stress est-il majoritairement modéré ou élevé ?
- Existe-t-il des valeurs extrêmes ?
- La distribution est-elle symétrique ou biaisée ?

## 10 Évolution temporelle du stress

Graphique (remplacé)

Ici serait la courbe d'évolution du stress. Le graphique montre la tendance générale et la diminution progressive du stress.

## 11 Analyse des corrélations

Graphique (remplacé)

Ici serait la matrice de corrélation entre les facteurs, recommandations et score global.

## 12 Analyse des scénarios de test

Scénarios simulés pour tester le système :

- Surcharge prolongée de travail,
- Repos et amélioration du sommeil,
- Variations rapides du stress.

Analyse Méthodologique : Indicateurs et Évaluation

## 13 Indicateurs clés de suivi

- Niveau moyen de stress,
- Volatilité du stress,
- Durée en zone de stress élevé,
- Amplitude des variations journalières.

## 14 Fonctions d'évaluation

Exemple de fonctions Python

```
evaluate_stress_level() visualize_stress_evolution() visualize_correlation_matrix()
```

Visualisations et Interprétation

### 14.1 Évolution temporelle du stress

Graphique (remplacé)

Courbe montrant la tendance à la baisse du stress, indiquant une amélioration progressive.

## 14.2 Matrice de corrélation

Graphique (remplacé)

Matrice de corrélation entre facteurs, recommandations et score global.

## 14.3 Diagramme de Pareto

Graphique (remplacé)

Diagramme de Pareto mettant en évidence les facteurs de stress les plus critiques.

## 14.4 Heatmap activité / stress

Graphique (remplacé)

Heatmap représentant l'impact des différents scénarios sur le score global.

Conclusion et Perspectives

Le projet **StressLess** montre que l'analyse de données repose sur :

- compréhension métier,
- nettoyage rigoureux des données,
- exploration statistique approfondie,
- visualisation claire et interprétable.

Ce projet constitue une base solide pour :

- tableaux de bord de suivi du stress,
- systèmes de recommandation personnalisés,
- outils de détection précoce,
- assistants intelligents de bien-être.