### **FEUILLE DE ROUTE**

### 

ROAD MAP en plusieurs étapes :  
✅ **Exploration des données & KPI**✅ **Analyse des corrélations & clustering**✅ **Modélisation (classification, régression)**✅ **Visualisation et interprétation des résultats**

### **Indicateurs Clés de Performance (KPI)**

#### **Démographie des Patients**

✅ Répartition de l'âge des patients 📊  
✅ Répartition par genre et ethnicité 🏥  
✅ Niveau d'éducation moyen 📚  
✅ Répartition du **BMI (Indice de Masse Corporelle)**

#### **Facteurs de Risque pour Alzheimer**

✅ Pourcentage de patients avec **antécédents familiaux d'Alzheimer** 🧠  
✅ Taux de **tabagisme et consommation d’alcool** 🚬🍷  
✅ Niveau moyen d'**activité physique** 🏃‍♂️  
✅ Moyenne des scores **MMSE (Mini Mental State Exam)**

#### **Facteurs Cardiovasculaires & Santé Générale**

✅ Moyenne de la **pression artérielle (systolique/diastolique)** 🩸  
✅ Niveau moyen de **cholestérol total, LDL et HDL**✅ Nombre de patients souffrant de **diabète, hypertension, dépression**

#### **Symptômes Cognitifs et Comportementaux**

✅ Pourcentage de patients signalant des **problèmes de mémoire, confusion, désorientation**✅ Score moyen des **activités de la vie quotidienne (ADL)**

## **2️⃣ Algorithmes à Appliquer 🤖**

Maintenant, voici quelques analyses que tu peux effectuer :

### **📌 Prédiction du Risque d'Alzheimer (Classification)**

➡ **Objectif** : Prédire si un patient est diagnostiqué avec **Alzheimer** en fonction de ses facteurs de risque.

📌 **Algorithmes possibles** :

* **Régression Logistique**
* **Random Forest**
* **Réseaux de neurones (MLP)**

✅ **Variable cible (diagnosis)** : Alzheimer ou Non-Alzheimer  
✅ **Variables explicatives** : âge, antécédents familiaux, MMSE, cholestérol, hypertension, etc.  
✅ **Évaluation** : Précision, rappel, matrice de confusion 🏥

### **📌 Détection des Facteurs de Risque (Analyse de Corrélation)**

➡ **Objectif** : Identifier les facteurs les plus liés à un diagnostic d'Alzheimer.  
📌 **Techniques** :

* **Heatmap des corrélations** 🔥
* **Test statistique (Chi², ANOVA)**
* **Feature Importance (Random Forest, XGBoost)**

✅ **Analyse intéressante** : Quels sont les **facteurs les plus influents** dans le développement d'Alzheimer ?

### **📌 Régression : Prédiction du Score Cognitif (MMSE)**

➡ **Objectif** : Prédire le score MMSE d’un patient à partir de ses caractéristiques.  
📌 **Algorithmes possibles** :

* **Régression Linéaire Multiple** 📈
* **Random Forest Regressor**
* **Gradient Boosting (XGBoost, LightGBM)**

✅ **Variable cible (mini\_mental\_state\_exam )**✅ **Prédicteurs** : âge, cholestérol, hypertension, mémoire, confusion…

## **3️⃣ Visualisation & Dashboard 📊**

Une fois les analyses faites, voici quelques idées de visualisations intéressantes :  
✅ **Heatmap des corrélations** entre les facteurs de risque et Alzheimer  
✅ **Courbes de tendance** des scores MMSE par tranche d'âge  
✅ **Distribution des patients** par niveau de diagnostic  
✅ **Carte interactive** des facteurs de risque (si tu as des données géographiques)