

# Rapport du projet en programmation R

Trabelsi Foued & Abderrazek Rania



**Année universitaire :** 2023 /2024

**Sous la supervision de :** Mme Rihab Mersni

**Membres du groupe :**

- Trabelsi Foued
- Abderrazek Rania

**Classe :** IGL3

---

---

## SOMMAIRE :

<b>1-Introduction</b>	page 2
<b>2-Processus du développement de projet</b>	page 3
<b>3-Analyse , Visualisation &amp; interpretation des resultats</b>	page 8
• Le volumes des précipitations et le remplissage des barrages	page 8
– volumes des précipitations	page 8
– remplissage des barrages	page13
• Rapport entre les précipitations et le remplissage des barrages	page 14
• Repartition des utilisations principales de l'eau	page 17
• Tarification de la SONEDE	page 19
• Production des céréales	page 22
<b>4 -Des prédictions</b>	page 23
<b>5- Solutions</b>	page 27
<b>6- Conclusion</b>	page 29
<b>7- Bibliographie</b>	page 30

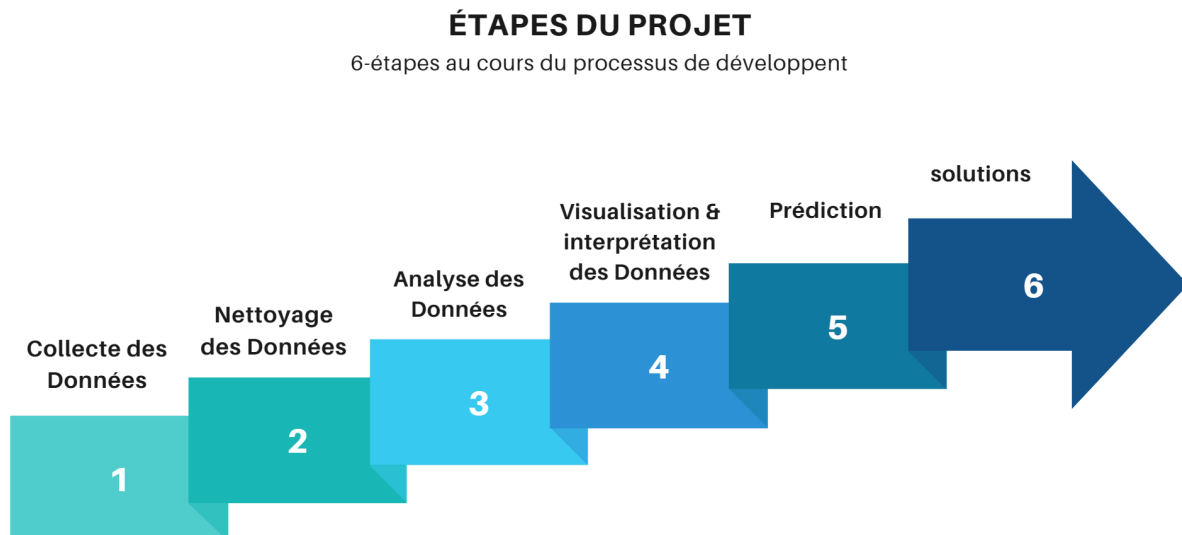
---

## 1-Introduction :

L'eau est une ressource vitale pour la vie et le développement socio-économique. En Tunisie, un pays caractérisé par un climat méditerranéen semi-aride, la gestion efficace de l'eau revêt une importance cruciale pour assurer la sécurité alimentaire, garantir l'accès à l'eau potable et promouvoir le développement durable. Dans ce contexte, notre projet vise à analyser l'impact des précipitations sur les ressources en eau en Tunisie, en mettant en évidence le rôle critique du remplissage des barrages et ses implications dans divers secteurs. Nous examinerons également les répercussions de ces conditions hydrologiques sur la consommation quotidienne en eau, l'agriculture et l'économie nationale.

Cette analyse approfondie nous permettra de mieux comprendre les défis liés à la disponibilité de l'eau dans le pays et d'identifier des stratégies pour une gestion durable des ressources en eau.

## 2-Processus du développement de projet :



### Collecte des Données :

Dans cette étape, on a recherché des sources fiables et pertinentes pour les données économiques qui repondent a notre sujet .

*( Toutes les sources utilisées sont mentionnées dans la partie bibliographie).*

On a sélectionné les ensembles de données correspondant à notre domaine d'intérêt et à notre objectif d'analyse économique. Ensuite, on a téléchargé ces données dans un format approprié pour l'analyse dans R, tel que CSV, Excel.

*(Vous trouvez ci jointes toutes les bases de données en leurs états brutes : barrage.csv , cereales.csv , sonede.csv , usage.xlsx , pluviometrie.csv )*

### **Nettoyage des Données :**

Dans cette étape, nous avons travaillé sur le nettoyage et la préparation des données pour l'analyse ultérieure , et on a procédé comme suit :

1. **Importation des données :** Les données ont été importées des fichiers CSV et Excel
2. **Conversion et enregistrement :** Les données provenant d'un fichier Excel ont été converties en format CSV pour assurer une uniformité .
3. **Nettoyage des données :**
  - Les colonnes non pertinentes ont été supprimées
  - Les lignes contenant des valeurs manquantes ont été éliminées .
  - Les valeurs qui ne respectent pas le format demandé ou écrites en langue arabe dans certaines colonnes ont été supprimées .
4. **Transformation des données :**
  - Les colonnes contenant des dates ont été converties en types de données de date.
  - Normalisation des valeurs des colonnes en minuscules
5. **Création d'un nouveau jeu de données :** des nouveaux data frame issus de l'opération du nettoyage avec les pourcentages correspondants.

\*\*\*\*\* code commenté correspondant a cette étape \*\*\*\*\*

```
library(readr)
#install.packages("readxl")
library(readxl)
# Importer les données de barrages
barrage <- read.csv("C:/Users/21629/OneDrive/Bureau/projet programmation R/barrages.csv")
# Importer les données de barrages
pluviometrie <- read.csv("C:/Users/21629/OneDrive/Bureau/projet programmation R/pluviometrie.csv")
# Importer les données de cereale
cereale <- read.csv("C:/Users/21629/OneDrive/Bureau/projet programmation R/cereale.csv")
head(cereale)
```

		X	X2014	X2015	X2016	X2017	X2018	X2019	X2020	X2021	X2022
1	Blé dur	596.1	540.0	515.4	580	534.0	562.0	548.0	544.0	515.4	
2	Blé tendre	124.8	110.7	95.0	95	85.3	81.0	62.0	65.0	64.4	
3	Orge	574.7	513.2	495.1	544	525.0	568.0	546.0	510.0	429.9	
4	Triticale	15.0	24.5	21.3	11	10.0	13.3	12.5	9.5	9.6	

```
# Charger le fichier de usage
donnees <- read_excel("C:/Users/21629/OneDrive/Bureau/projet programmation R/usage.xlsx")
# Charger le fichier de sonede
sonede <- read_excel("C:/Users/21629/OneDrive/Bureau/projet programmation R/Sonede.xlsx")

# Convertir et enregistrer en format CSV
write.csv(sonede, "C:/Users/21629/OneDrive/Bureau/projet programmation R/Sonede.csv", row.names=FALSE)
head(sonede)
```

```
# A tibble: 6 x 4
  id `tranche-metre-cube-par-trimestre` `2023 (en dt)` `2024 (en dt)`
  <dbl> <chr>                                <dbl>          <dbl>
1     1 0-20                                0.2            0.2
2     2 20-40                               0.665          0.74
3     3 40-70                               0.93           1.04
4     4 70-100                              1.31           1.49
5     5 100-150                             1.54           1.77
6     6 150 et plus                         1.99           2.31
```

```
#class(sonede)
head(pluviometrie)
```

	Date	station	Pluvio_du_jour	Cumul_du_mois	Cumul_moy_du_mois
1	2024-03-04		0.0	2.5	49.90800
2	2024-03-04		1.5	5.5	45.33111
3	2024-03-04		1.3	3.3	50.75532
4	2024-03-04		8.0	25.0	161.05800
5	2024-03-04		2.2	10.2	93.65800
6	2024-03-04		3.3	12.9	68.59783

	Cumul_periode	Cumul_moy_periode	Cumul_mois_precedent	Cumul_periode_precedente
1	247.9	304.2431	1.4	169.0
2	287.5	294.6855	2.2	167.6
3	235.1	287.4481	3.5	168.3
4	1145.0	1148.8457	55.0	683.0
5	565.0	728.6094	56.4	588.7
6	499.3	521.0805	11.1	347.8

```
#class(pluviometrie)
#pour verifier que les donnees sont bien sous format frame
#class(barrage)
# Lire les premières lignes
head(barrage)
```

	Nom_Fr	Nom_Ar	apports	id_barrage	lachers	stock	
1	mellegue	0.024128200		1	0.000000	8.81804	
2	sarrat	0.000713428		36	0.076415	6.75642	
3	benmetir	0.036453400		2	0.034560	41.92950	
4	kasseb	0.052349600		3	0.028512	22.97700	
5	barbara	0.060773900		4	0.095794	59.71730	
6	sidisalem	0.401331000		5	0.043200	245.67000	
	cumul_mensuel_apports	cumul_annuel_apports	cumul_mensuel_lachers				
1	1.123480		7.65798			0.000000	
2	0.131568		1.96184			0.498819	
3	2.172250		27.48690			0.552960	
4	1.850020		22.71430			0.471240	
5	4.856180		65.35770			4.542170	
6	22.250100		110.42800			0.759766	
	cumul_annuel_lachers	moy_mois	cumul_annee_prec	moy_periode	stock_annee_prec		
1	11.15080	12.62160	14.80270	114.1830		14.56180	
2	1.83912	1.90300	2.79072	11.2456		2.99817	
3	6.84288	7.20793	7.56689	33.7292		20.14490	
4	6.56965	7.47992	8.64542	34.8313		13.64430	
5	58.40000	11.90270	18.65860	67.6031		62.16960	
6	49.79470	72.75120	31.66980	286.5290		97.08830	
	Cap_tot_act	Cote	Cap_tot_init	fonctionnel	Annee_prod	Latitude	Longitude
1	51.130	265.00	267.670	1	1954	36.3144	8.70273
2	21.000	546.00	21.000	1	2016	49.9167	26.58330
3	60.393	436.50	61.630	1	1954	36.7471	8.74188
4	76.900	292.04	81.875	1	1968	36.7614	9.00230
5	64.700	182.00	74.820	1	1998	36.7341	8.53173
6	580.350	115.00	814.000	1	1981	36.5907	9.39706
	Bassin_versant			Date			
1	10300	2024-03-16	00:00:00				
2	1850	2024-03-16	00:00:00				
3	103	2024-03-16	00:00:00				
4	101	2024-03-16	00:00:00				
5	177	2024-03-16	00:00:00				
6	7950	2024-03-16	00:00:00				

```

# Supprimer les colonnes Nom_Ar et Annee_prod
barrage <- subset(barrage, select = -c(Nom_Ar, Annee_prod))

# Supprimer des lignes avec des valeurs manquantes
barrage <- na.omit(barrage)
pluviometrie <- na.omit(pluviometrie)

# Supprimer des lignes avec des valeurs négatives dans la colonne apports
barrage <- barrage[barrage$apports >= 0, ]
pluviometrie <- pluviometrie[pluviometrie$Cumul_periode >= 0, ]
# Convertir les colonnes contenant des dates en types de données de date
barrage$Date <- as.Date(barrage$Date)
pluviometrie$Date <- as.Date(pluviometrie$Date)

# Normaliser les valeurs de la colonne Nom_Fr en minuscules
barrage$Nom_Fr <- tolower(barrage$Nom_Fr)
head(barrage)

```

	Nom_Fr	apports	id_barrage	lachers	stock	cumul_mensuel_apports
1	mellegue	0.024128200	1	0.000000	8.81804	1.123480
2	sarrat	0.000713428	36	0.076415	6.75642	0.131568
3	benmetir	0.036453400	2	0.034560	41.92950	2.172250
4	kasseb	0.052349600	3	0.028512	22.97700	1.850020
5	barbara	0.060773900	4	0.095794	59.71730	4.856180
6	sidisalem	0.401331000	5	0.043200	245.67000	22.250100

	cumul_annuel_apports	cumul_mensuel_lachers	cumul_annuel_lachers	moy_mois
1	7.65798	0.000000	11.15080	12.62160
2	1.96184	0.498819	1.83912	1.90300
3	27.48690	0.552960	6.84288	7.20793
4	22.71430	0.471240	6.56965	7.47992
5	65.35770	4.542170	58.40000	11.90270
6	110.42800	0.759766	49.79470	72.75120

	cumul_annee_prec	moy_periode	stock_annee_prec	Cap_tot_act	Cote	Cap_tot_init
1	14.80270	114.1830	14.56180	51.130	265.00	267.670
2	2.79072	11.2456	2.99817	21.000	546.00	21.000
3	7.56689	33.7292	20.14490	60.393	436.50	61.630
4	8.64542	34.8313	13.64430	76.900	292.04	81.875
5	18.65860	67.6031	62.16960	64.700	182.00	74.820
6	31.66980	286.5290	97.08830	580.350	115.00	814.000

	fonctionnel	Latitude	Longitude	Bassin_versant	Date
1	1	36.3144	8.70273	10300	2024-03-16
2	1	49.9167	26.58330	1850	2024-03-16

3	1	36.7471	8.74188	103	2024-03-16
4	1	36.7614	9.00230	101	2024-03-16
5	1	36.7341	8.53173	177	2024-03-16
6	1	36.5907	9.39706	7950	2024-03-16

```
donnees <- data.frame(
  "Utilisation.principale" = c("Hydroélectricité", "Irrigation", "Alimentation en eau",
  "Pourcentage" = c("5 %", "60 %", "30 %", "5 %")
)
```

### Outils utilisés pour Analyser les Données :

Dans cette étape, on a utilisé des techniques d'analyse statistique pour explorer les données de manière économiques. Cela inclut

- Des analyses de séries chronologiques : pour identifier les tendances et les variations dans les données au fil du temps , dans le cas du remplissage des barrages au cours de mois et du taux de précipitation au fil des années
- Des corrélations : entre différentes variables économiques pour comprendre les relations entre elles et identifier les facteurs qui influent sur l'économie.
- Des rapport mathématiques : pour évaluer les relations causales entre les variables .

## 3-Analyse et Visualisation des données & interpretation des Résultats :

### **1-Le volumes des précipitations et le remplissage des barrages :**

**\* code commenté correspondant aux volumes de précipitations et remplissage des barrages**  
**\***

a - volumes des precipitations :

```
#install.packages("sf")
#install.packages("ggplot2")
#install.packages("leaflet")
#install.packages("dplyr")
#pour determiner annee a partir de date
#install.packages("lubridate")
library(lubridate)
```

Attaching package: 'lubridate'



The following objects are masked from 'package:base':

date, intersect, setdiff, union

```
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(sf)
```

Linking to GEOS 3.11.2, GDAL 3.8.2, PROJ 9.3.1; sf\_use\_s2() is TRUE

```
library(ggplot2)
library(shiny)
library(leaflet)
###visualisation de barrages
# Charger les données des barrages à partir du data frame nettoyé 'barrage'
barrages_sf <- st_as_sf(barrage, coords = c("Longitude", "Latitude"), crs = 4326)

# Calculer le taux de remplissage
barrage$Taux <- (barrage$stock / barrage$Cap_tot_act) * 100

# Obtenir la gamme de valeurs de Taux
taux_range <- range(barrage$Taux, na.rm = TRUE)
my_palette <- rev(colorRampPalette(c("#FFFFFF", "#add8e6"))(100)) # Bleu clair
# Générer une palette de couleurs correspondant à la gamme de valeurs de Taux
color_palette <- colorNumeric(palette = my_palette, domain = taux_range)

# Créer une carte interactive avec leaflet
ma_carte_leaflet <- leaflet(data = barrages_sf) %>%
  addTiles() %>%
```

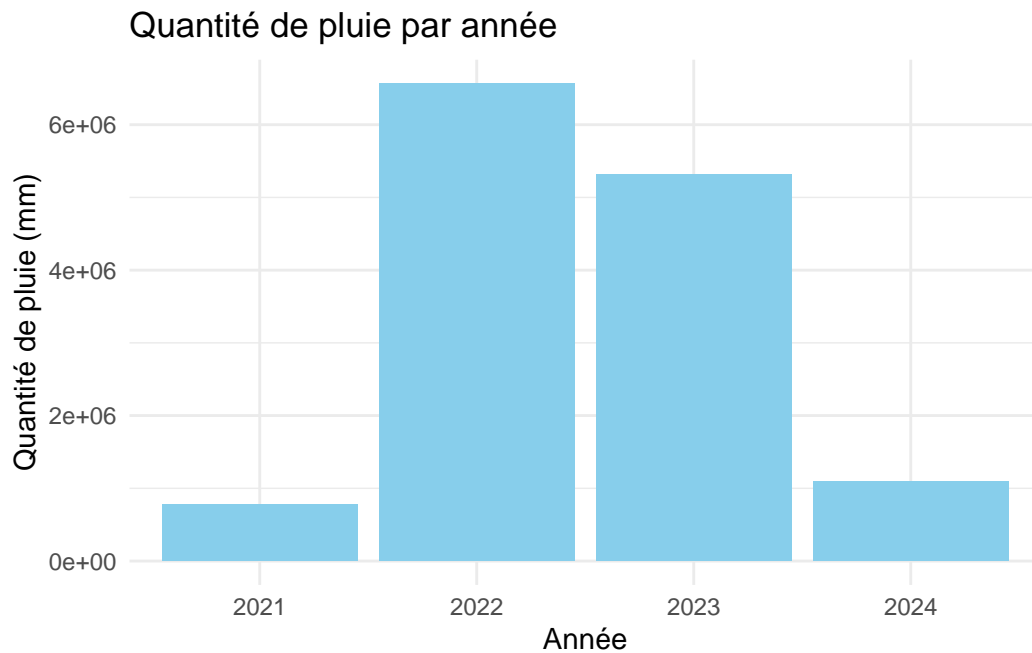
```

addCircleMarkers(
  fillColor = ~color_palette(barrage$Taux), # correction
  # Utiliser la palette de couleurs pour le taux de remplissage
  color = "black",
  weight = 1,
  radius = ~sqrt(Cap_tot_act) *2 , # Utiliser la racine carrée de la capacité pour le ray
  label = ~paste("Nom: ", Nom_Fr, "<br>",
                  "Stock: ", stock, "<br>",
                  "Capacité totale: ", Cap_tot_act, "<br>",
                  "Taux: ", barrage$Taux, "%<br>"),

)%>%
addLegend(position = "bottomright",
          pal = color_palette,
          opacity = 1,
          values = taux_range,
          title = "Taux de remplissage (%)")

# Afficher la carte interactive
#ma_carte_leaflet
###visualisation de pluviometrie
# Extraire l'année à partir de la date
pluviometrie$Annee <- format(pluviometrie$Date, "%Y")
# Agréger les données de pluviométrie par année
pluviometrie_an <- pluviometrie %>%
  group_by(Annee) %>%
  summarise(Cumul_periode_an = sum(Cumul_periode))
# Créer un graphique en barres pour représenter la quantité de pluie par année
ggplot(pluviometrie_an, aes(x = Année, y = Cumul_periode_an)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "Quantité de pluie par année",
       x = "Année",
       y = "Quantité de pluie (mm)") +
  theme_minimal()

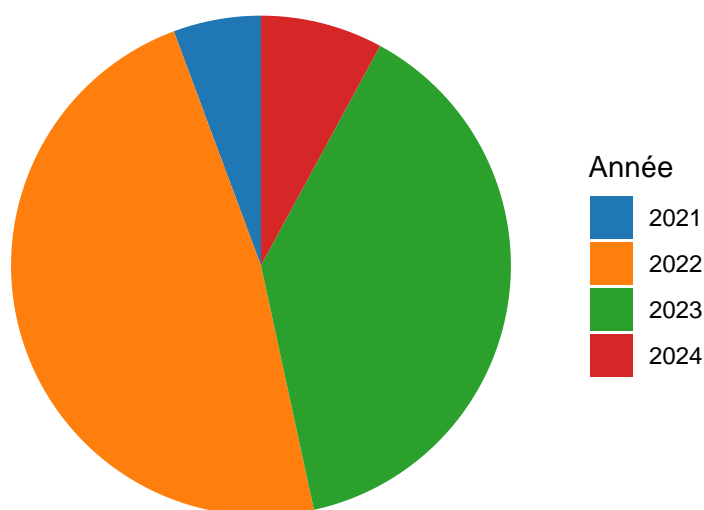
```



```
# Définir les couleurs pour chaque année
colors <- c("#1f77b4", "#ff7f0e", "#2ca02c", "#d62728")

# le pie chart
ggplot(pluviometrie_an, aes(x = factor(1), y = Cumul_periode_an, fill = Annee)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y") +
  scale_fill_manual(values = colors) +
  labs(title = "Répartition de la quantité de pluie par année",
       fill = "Année",
       y = "Quantité de pluie (cm)") +
  theme_void() # Supprimer les éléments de la grille et de l'axe
```

## Répartition de la quantité de pluie par année



### Visualisation des Données :

Nous avons créer un diagramme en bâtons pour représenter les volumes de précipitations annuelles de 2021 à 2024. Chaque barre du diagramme représentera le volume de précipitations pour une année donnée.

Diagramme en Bâtons	Diagramme Circulaire
Chaque barre du diagramme représente le volume de précipitations pour une année donnée.	Chaque portion du cercle montre le taux des précipitations pour une année donnée.

### Interprétation des Résultats :

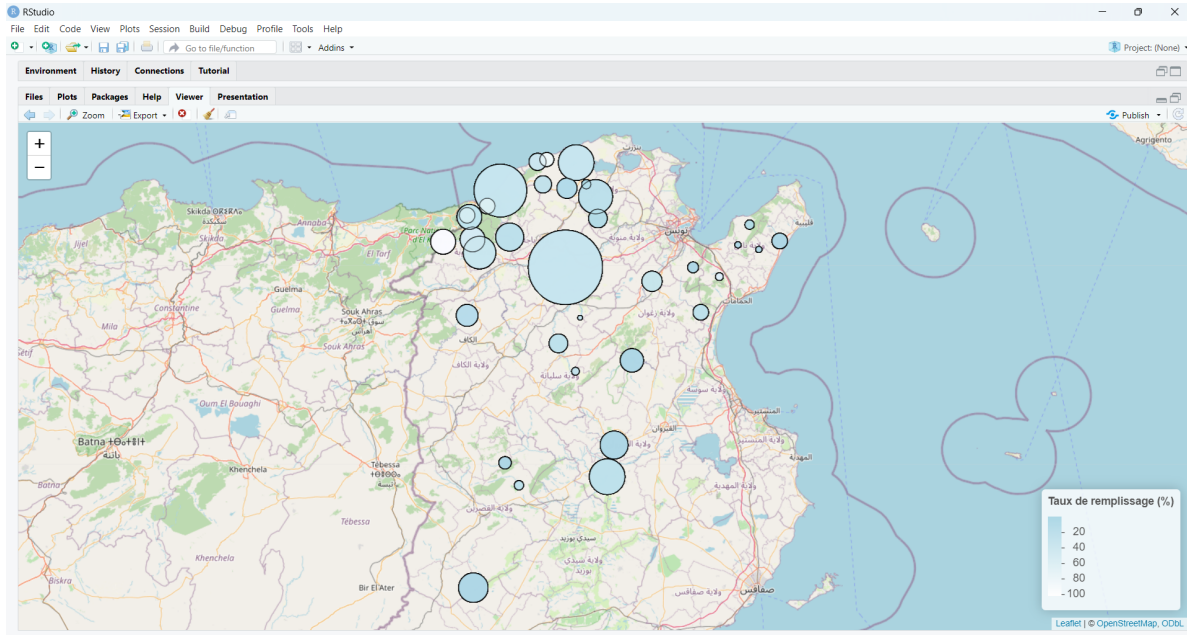
- **Année 2021** : Les volumes de précipitations ont été très faibles , ce qui peut avoir entraîné des défis en matière de disponibilité d'eau pour divers secteurs.
- **Année 2022** : Nous avons observé une augmentation significative des volumes de précipitations en 2022, dépassant même les 6 millions d'unités. Cette augmentation pourrait avoir contribué à une amélioration des ressources en eau.
- **Années 2023 et 2024** : Les volumes de précipitations ont diminué en 2023 et 2024 par rapport à 2022. Cette diminution peut avoir des implications sur la disponibilité des ressources en eau pour l'agriculture, l'approvisionnement en eau potable, etc.

— Remarque : Étant donné que la saison d'hiver a pris fin pour 2024, les augmentations futures des précipitations ne sont pas attendues

### b-Remplissage des barrages :

pour l'analyse des données de remplissage des barrages on a effectué le Calcul du taux de remplissage en suivant cette formule :

$$=(\text{stockage du barrage en eau} / \text{capacité du barrage}) * 100$$



[\[ AFFICHER TOUTE LA CARTE \]](#)

### Visualisation des Données :

**Carte Interactive :** Nous avons représenté ces données sur une carte interactive pour permettre une visualisation dynamique et informative, chaque bulle dans la carte porte les informations telles que : Nom, Stockage, capacité totale et le taux de ce barrage.

### Interprétation des Résultats :

Dans l'année 2024, le barrage de Sidisalem se démarque en tant que le barrage le plus rempli parmi ceux étudiés. Cependant, malgré cette position, le pourcentage de remplissage de ce barrage n'a pas dépassé les 30 %, ce qui reste très faible et en deçà de la norme attendue. Cette observation souligne les défis persistants en matière de disponibilité en eau, même pour les barrages les mieux remplis.

### 3- Rapport entre les précipitations et le remplissage des barrages :

pour Analyser la relation entre les précipitations et le remplissage des barrages on a effectué le rapport entre les deux := remplissage barrage/ taux de precipitations

[ toute conversions necessaires des unités pour faire l'analyse était faite ]

\*\*\*code commenté correspondant au Rapport entre les précipitations et le remplissage des barrages\*\*

```
library(lubridate)
library(dplyr)
library(sf)
library(ggplot2)
library(shiny)
library(leaflet)

### Calculer le rapport entre le remplissage des barrages et les précipitations pour les ann
# Filtrer les données de pluviométrie pour les dates spécifiées
pluviometrie_2023 <- pluviometrie %>%
  filter(Date >= ymd("2023-12-20") & Date <= ymd("2024-03-16"))

# Extraire l'année à partir de la colonne Date pour les données de pluviométrie
pluviometrie_2023$Annee <- year(pluviometrie_2023$Date)

# Agréger les données de pluviométrie par année
pluviometrie_annee <- pluviometrie_2023 %>%
  group_by(Annee) %>%
  summarise(Total_precipitations = sum(Cumul_periode))

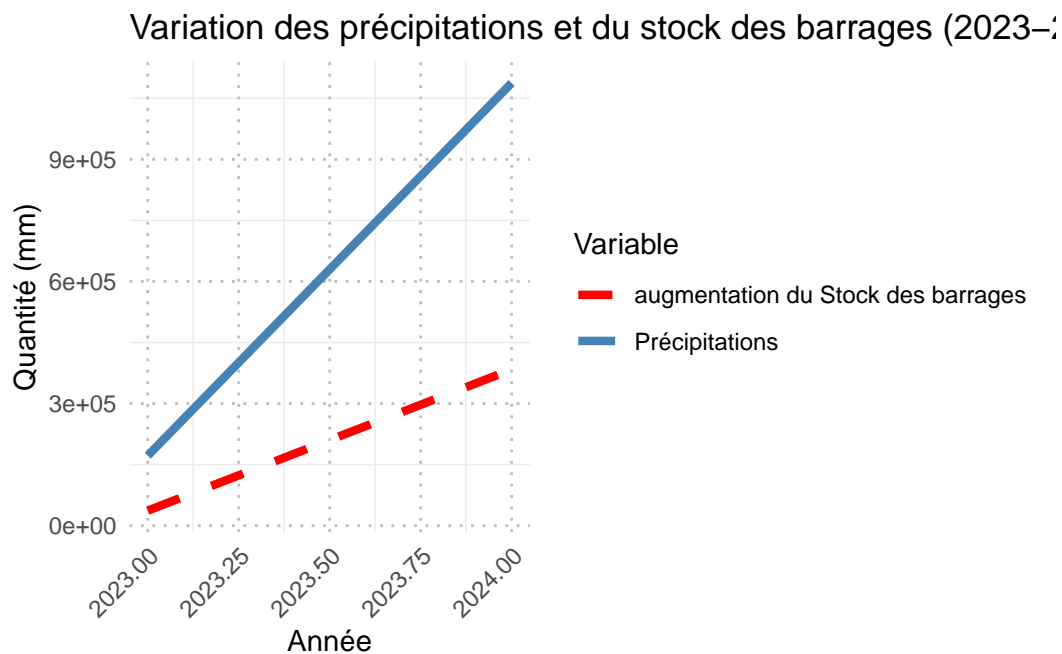
# Extraire l'année à partir de la colonne Date pour les données de barrage
barrage$Annee <- year(barrage$Date)

# Agréger les données des barrages par année et convertir les unités en millimètres cubes
barrages_par_annee <- barrage %>%
  group_by(Annee) %>%
  summarise(Total_apports_mm = sum(apports) * 1000) # Convertir en millimètres cubes

# Joindre les données de précipitations et de barrages par année
donnees_annee <- merge(pluviometrie_annee, barrages_par_annee, by = "Annee")

# Créer le graphique qui présente l'influence du pluie sur le stock
ggplot(donnees_annee, aes(x = Annee)) +
```

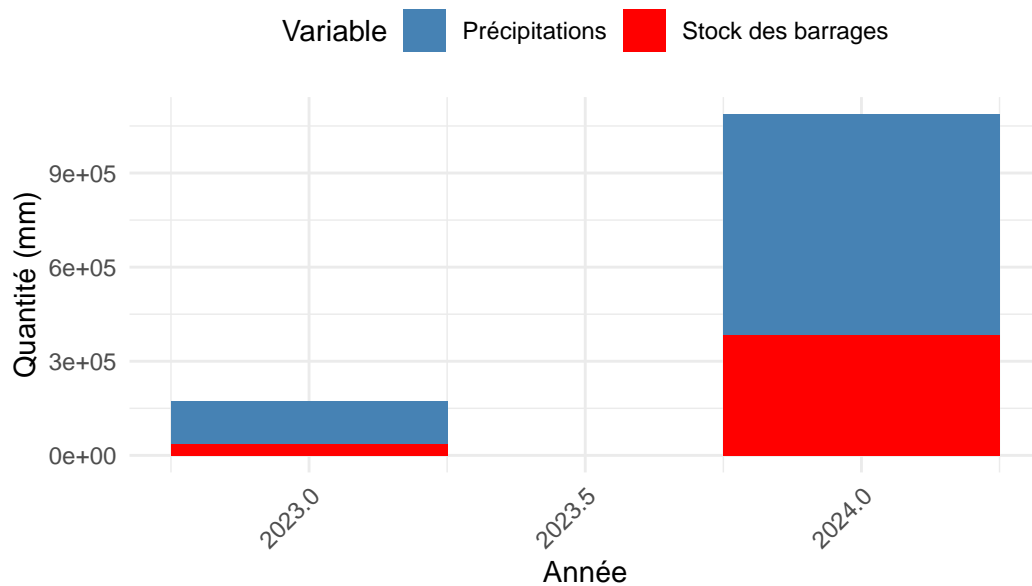
```
geom_line(aes(y = Total_precipitations, color = "Précipitations"), linewidth = 1.5) + # C
geom_line(aes(y = Total_apports_mm, color = "augmentation du Stock des barrages"), linewidth
labs(title = "Variation des précipitations et du stock des barrages (2023-2024)",
     x = "Année",
     y = "Quantité (mm)",
     color = "Variable") +
theme_minimal() +
theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Rotation des étiquettes de 1
scale_color_manual(values = c("Précipitations" = "steelblue", "augmentation du Stock des ba
theme(panel.grid.major = element_line(color = "gray", linetype = "dotted"))
```



```
## Créer un rapport
donnees_annee$Rapport <- (donnees_annee$Total_apports_mm / donnees_annee$Total_precipitations)
ggplot(donnees_annee, aes(x = Année)) +
  geom_bar(aes(y = Total_precipitations, fill = "Précipitations"), stat = "identity", width = 1) +
  geom_bar(aes(y = Total_apports_mm, fill = "Stock des barrages"), stat = "identity", width = 1) +
  labs(title = "Rapport entre stock des barrages et précipitations",
       x = "Année",
       y = "Quantité (mm)",
       fill = "Variable") +
  scale_fill_manual(values = c("Précipitations" = "steelblue", "Stock des barrages" = "red")) +
  theme_minimal() +
```

```
theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
theme(legend.position = "top") # Placer la légende en haut du graphique
```

## Rapport entre stock des barrages et précipitations



### Visualisation des Données :

- **Diagrammes en Bâtons et Courbe :** Nous avons utilisé deux diagrammes :

en bâtons pour représenter les volumes de précipitations sur les années étudiées, ainsi qu'une courbe pour illustrer l'évolution du taux de remplissage des barrages au fil du temps.

### Interprétation des Résultats :

- **Disparité entre Précipitations et Remplissage :** Cependant, malgré cette corrélation, il existe une grande disparité entre les volumes de précipitations et le taux de remplissage des barrages. Cette disparité est due à divers facteurs, tels que les pertes d'eau causées par l'évaporation, l'infiltration et les débits sortants des barrages.



#### 4-Répartition des utilisations principales de l'eau

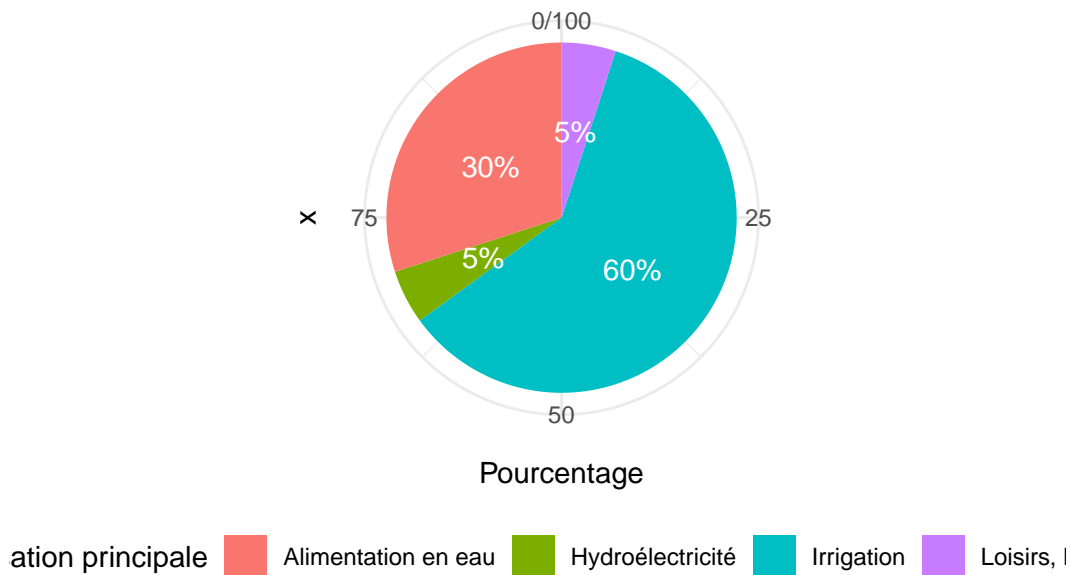
\*\*\*\*\* code commenté correspondant aux utilisations principales de l'eau \*\*\*\*\*

```
library(ggplot2)
# Changer le nom du dataset
donnees_eau <- data.frame(
  Utilisation.principale = c("Hydroélectricité", "Irrigation", "Alimentation en eau", "Loisirs"),
  Pourcentage = c(5, 60, 30, 5)
)

# Créer le pie chart avec le nouveau nom de dataset
pie_chart <- ggplot(donnees_eau, aes(x = "", y = Pourcentage, fill = Utilisation.principale)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  geom_text(aes(label = paste0(Pourcentage, "%")),
            position = position_stack(vjust = 0.5),
            color = "white",
            size = 4) +
  labs(title = "Répartition des utilisations principales de l'eau",
       fill = "Utilisation principale",
       y = "Pourcentage") +
  theme_minimal() +
  theme(legend.position = "bottom")

# Afficher le pie chart
print(pie_chart)
```

## Répartition des utilisations principales de l'eau



### Visualisation des Données :

Utilisation d'un diagramme circulaire pour illustrer la répartition des utilisations principales de l'eau entre différentes catégories.

### Interprétation des Résultats :

- Le diagramme circulaire met en évidence la répartition de l'utilisation de l'eau entre différentes fonctions.
- Environ 60 % de l'eau est utilisée pour l'irrigation, ce qui représente la plus grande part de l'utilisation totale.
- Environ 30 % de l'eau est consacrée à l'alimentation humaine et animale.
- Environ 5 % de l'eau est utilisée pour la production d'hydroélectricité, et les activités de loisirs et de navigation utilisent également environ 5 % de l'eau.

» En résumé, la répartition des utilisations principales de l'eau met en évidence le rôle crucial de cette ressource pour l'agriculture, largement dépendante de l'eau des barrages. Cependant, les efforts en matière d'énergie renouvelable en Tunisie n'ont pas encore pleinement exploité le potentiel de collaboration avec le secteur de l'eau.

## 5-Tarification de la SONEDE :

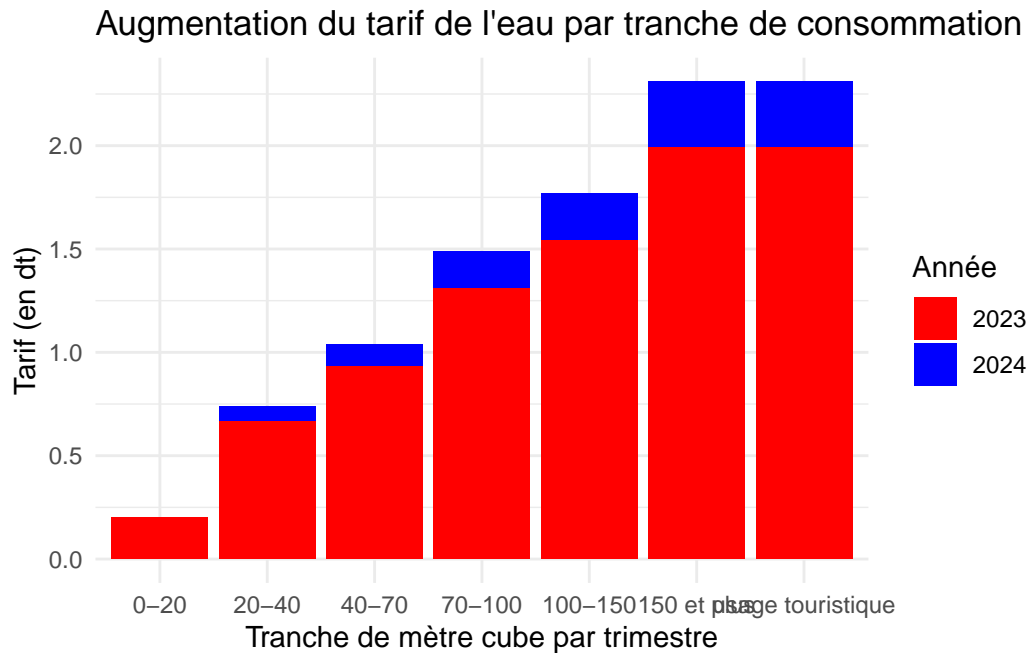
```
library(ggplot2)

# Convertir les virgules en points dans les valeurs numériques
sonede$`2023 (en dt)` <- as.numeric(gsub(",", ".", sonede$`2023 (en dt)`))
sonede$`2024 (en dt)` <- as.numeric(gsub(",", ".", sonede$`2024 (en dt)`))

# Convertir `tranche-metre-cube-par-trimestre` en facteur avec l'ordre approprié
sonede$`tranche-metre-cube-par-trimestre` <- factor(sonede$`tranche-metre-cube-par-trimestre`,
                                                    levels = c("0-20", "20-40", "40-70", "70-"))

# Créer le graphique avec les valeurs pivotées
tarif_water_plot <- ggplot(sonede, aes(x = `tranche-metre-cube-par-trimestre`)) +
  geom_bar(aes(y = `2024 (en dt)`, fill = "2024"), stat = "identity", position = "dodge") +
  geom_bar(aes(y = `2023 (en dt)`, fill = "2023"), stat = "identity", position = "dodge") +
  labs(title = "Augmentation du tarif de l'eau par tranche de consommation",
       x = "Tranche de mètre cube par trimestre",
       y = "Tarif (en dt)",
       fill = "Année") +
  scale_fill_manual(values = c("2023" = "red", "2024" = "blue")) +
  theme_minimal()

# Afficher le graphique
print(tarif_water_plot)
```



#### Visualisation des Données :

Utilisation d'un diagramme en bâtons pour illustrer l'augmentation dans chaque tranche de mètre cube par trimestre.

#### Interprétation des Résultats :

- Le diagramme en bâtons met en évidence l'évolution des tarifs de la SONEDE pour chaque tranche de mètre cube au fil des trimestres.
- Il permet de visualiser clairement les augmentations de tarifs, fournissant une représentation graphique de la progression des coûts pour les différents niveaux de consommation d'eau.

» l'analyse suggère que les tarifs de la SONEDE peuvent être influencés par le remplissage des barrages. **Une corrélation** entre ces deux facteurs pourrait être observée, où une augmentation du remplissage des barrages pourrait potentiellement conduire à des ajustements tarifaires de la part de la SONEDE

#### **Utilisation de la corrélation pour l'analyse :**

la corrélation entre les apports des barrages et les augmentations tarifaires est calculée pour chaque année à l'aide de la fonction `cor()`. La corrélation mesure la relation linéaire entre ces

deux ensembles de données, indiquant dans quelle mesure les variations dans les apports des barrages sont associées aux variations des tarifs de Sonned. » voici le code :

```
library(dplyr)

# Sélectionner les colonnes pertinentes
barrage_corr <- barrage[, c("Date", "apports")]
sonede_corr <- sonede[, c("id", "2023 (en dt)", "2024 (en dt)")]

# Filtrer les données pour chaque année
barrage_2023 <- filter(barrage_corr, Date >= "2023-01-01" & Date <= "2023-12-31")
barrage_2024 <- filter(barrage_corr, Date >= "2024-01-01" & Date <= "2024-12-31")
sonede_2023 <- filter(sonede_corr, !is.na(`2023 (en dt)`) )
sonede_2024 <- filter(sonede_corr, !is.na(`2024 (en dt)`) )

# Assurez-vous que les données ont la même longueur
min_length <- min(length(barrage_2023$apports), length(sonede_2023$`2023 (en dt)`) )
barrage_2023 <- head(barrage_2023, min_length)
sonede_2023 <- head(sonede_2023, min_length)

min_length <- min(length(barrage_2024$apports), length(sonede_2024$`2024 (en dt)`) )
barrage_2024 <- head(barrage_2024, min_length)
sonede_2024 <- head(sonede_2024, min_length)

# Calculer la corrélation entre la somme du stock des barrages et l'augmentation des tarifs
correlation_2023 <- cor(barrage_2023$apports, sonede_2023$`2023 (en dt)`)
correlation_2024 <- cor(barrage_2024$apports, sonede_2024$`2024 (en dt)`)

# Afficher les résultats
print(paste("Corrélation pour 2023 :", correlation_2023))
```

```
[1] "Corrélation pour 2023 : 0.669985522811643"
```

```
print(paste("Corrélation pour 2024 :", correlation_2024))
```

```
[1] "Corrélation pour 2024 : 0.611338829102315"
```

Ce code permet d'évaluer la corrélation entre les apports des barrages et les augmentations tarifaires pour les années 2023 et 2024, offrant ainsi des insights sur la relation entre ces deux variables dans le contexte donné où la corrélation entre les apports des barrages et les tarifs de Sonede diminue en 2024 et qui peut atteindre un jour une valeur négative, cela suggère

un changement dans la relation entre ces variables qui vont être opposées .(diminution du remplissage des barrages » rareté de l'eau » augmentation des tarifs )

## 6-Production des céréales :

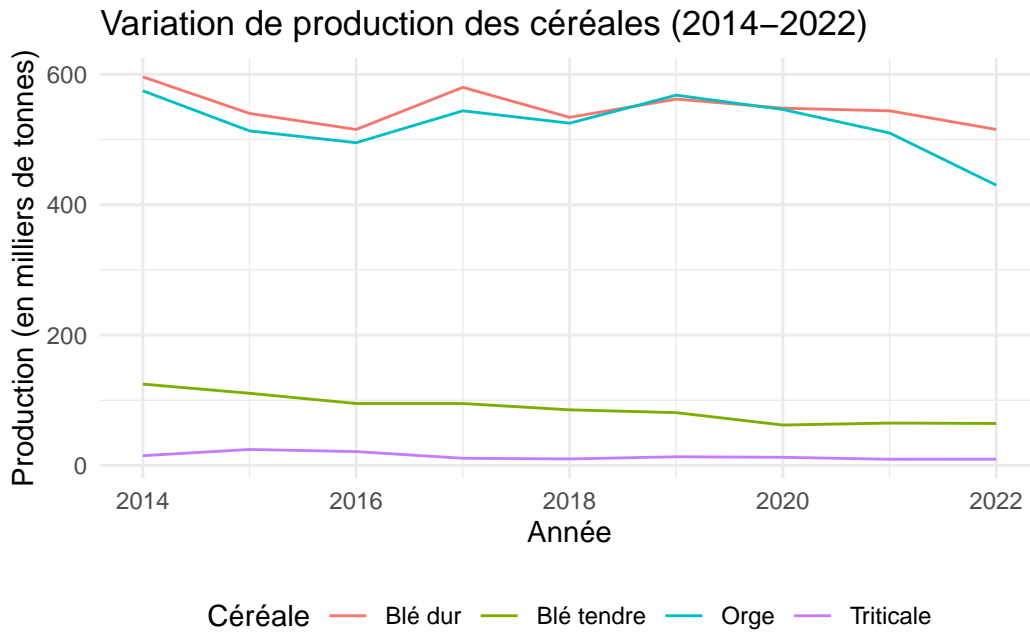
\*\*\*\*\* code commenté correspondant à la variation de la production des céréales\*\*\*\*\*

```
#install.packages("tidyr")
library(tidyr)
library(ggplot2)
# Renommer les colonnes
colnames(cereale) <- c("Céréale", "X2014", "X2015", "X2016", "X2017", "X2018", "X2019", "X2020", "X2021", "X2022")
head(cereale)
```

	Céréale	X2014	X2015	X2016	X2017	X2018	X2019	X2020	X2021	X2022
1	Blé dur	596.1	540.0	515.4	580	534.0	562.0	548.0	544.0	515.4
2	Blé tendre	124.8	110.7	95.0	95	85.3	81.0	62.0	65.0	64.4
3	Orge	574.7	513.2	495.1	544	525.0	568.0	546.0	510.0	429.9
4	Triticale	15.0	24.5	21.3	11	10.0	13.3	12.5	9.5	9.6

```
# Transformer le jeu de données en format long
cereale<- pivot_longer(cereale, cols = -Céréale, names_to = "Annee", values_to = "Production")
# Convertir l'année en numérique
cereale$Annee <- as.numeric(sub("X", "", cereale$Annee))

# Créer le graphique
ggplot(cereale, aes(x = Annee, y = as.numeric(Production), color = Céréale)) +
  geom_line() +
  labs(title = "Variation de production des céréales (2014-2022)",
        x = "Année",
        y = "Production (en milliers de tonnes)",
        color = "Céréale") +
  theme_minimal() +
  theme(legend.position = "bottom")
```



#### **Visualisation des Données :**

Utilisation d'un graphique en courbe pour représenter la production de céréales de 2014 à 2022.

#### **Interprétation des Résultats :**

- La courbe illustre une tendance générale à la baisse de la production de céréales au fil des années.
- Cette diminution peut être attribuée à divers facteurs tels que les conditions météorologiques défavorables, les changements dans les pratiques agricoles, les fluctuations des prix des intrants agricoles et surtout la diminution de l'eau dans les barrages.

» Malheureusement, Cette diminution soulève des préoccupations importantes quant à la sécurité alimentaire et à l'économie agricole. Il est crucial d'identifier les causes de cette diminution et de mettre en œuvre des mesures appropriées pour atténuer ses effets négatifs.

#### **4-Des prédictions :**

**Description de la prédiction :**

1. **Objectif** : L'objectif principal est de prédire les tarifs de la SONEDE pour l'année 2025 par tranche, en utilisant un modèle de régression linéaire.
2. **Méthode** : La méthode utilisée est la régression linéaire, qui est un modèle statistique couramment utilisé pour modéliser la relation entre une variable dépendante (dans ce cas, les tarifs de la SONEDE) et une ou plusieurs variables indépendantes (telles que les données historiques sur les tarifs, les coûts de production, etc.).
3. **Données** : Les données utilisées comprennent les tarifs historiques de la SONEDE, ainsi que d'autres variables pertinentes qui peuvent influencer les tarifs. Ces données sont divisées en ensembles d'entraînement et de test pour ajuster le modèle et évaluer sa performance.
4. **Ajustement du modèle** : Un modèle de régression linéaire est ajusté en utilisant les données d'entraînement, où les tarifs historiques sont la variable dépendante et les autres variables sont les variables indépendantes. Le modèle est ajusté pour minimiser l'erreur quadratique moyenne entre les valeurs observées et prédites.
5. **Évaluation du modèle** : La performance du modèle est évaluée en utilisant des mesures telles que le coefficient de détermination ( $R^2$ ) et l'erreur quadratique moyenne (RMSE) sur l'ensemble de test. Cela permet de vérifier si le modèle est capable de généraliser aux nouvelles données.
6. **Prédiction pour 2025** : Une fois que le modèle est validé, il est utilisé pour prédire les tarifs de la SONEDE pour l'année 2025 en fonction des valeurs des variables indépendantes pour cette année-là.
7. **Interprétation des résultats** : Les résultats de la prédiction sont interprétés pour comprendre les facteurs qui influencent les tarifs de la SONEDE en 2025 et formuler des recommandations en conséquence.

```
# Install and load the tidyverse package
#install.packages("tidyverse")
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v forcats 1.0.0      v stringr 1.5.1
v purrr  1.0.2      v tibble  3.2.1
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to
```

```
library(ggplot2)
```

```
# Create the dataframe with the provided data
```



```

sonede_prediction_data <- sonede

# Rename the columns for ease of use
colnames(sonede_prediction_data)[3:4] <- c("prix_2023", "prix_2024")

# Adjust a linear regression model to predict water prices per bracket for the years 2023 to 2024
model <- lm(prix_2024 ~ prix_2023, data = sonede_prediction_data)

# Predict water prices per bracket for the years 2025 to 2034
predictions <- predict(model, newdata = data.frame(prix_2023 = sonede_prediction_data$prix_2023))

# Add predictions to the dataframe
sonede_prediction_data$`2025` <- predictions

# Calculate the difference between the predicted prices for 2025 and the prices for 2024
difference <- sonede_prediction_data$`2025` - sonede_prediction_data$prix_2024

# Create a new dataframe containing the bracket and the calculated difference
difference_data <- data.frame(tranche = sonede_prediction_data$`tranche-metre-cube-par-an`, difference)

# Convert the bracket variable into a factor with a specific order
difference_data$tranche <- factor(difference_data$tranche, levels = c("0-20", "20-40", "40-60", "60-80", "80-100", "100-120", "120-140", "140-160", "160-180", "180-200", "200-220", "220-240", "240-260", "260-280", "280-300", "300-320", "320-340", "340-360", "360-380", "380-400", "400-420", "420-440", "440-460", "460-480", "480-500", "500-520", "520-540", "540-560", "560-580", "580-600", "600-620", "620-640", "640-660", "660-680", "680-700", "700-720", "720-740", "740-760", "760-780", "780-800", "800-820", "820-840", "840-860", "860-880", "880-900", "900-920", "920-940", "940-960", "960-980", "980-1000"))

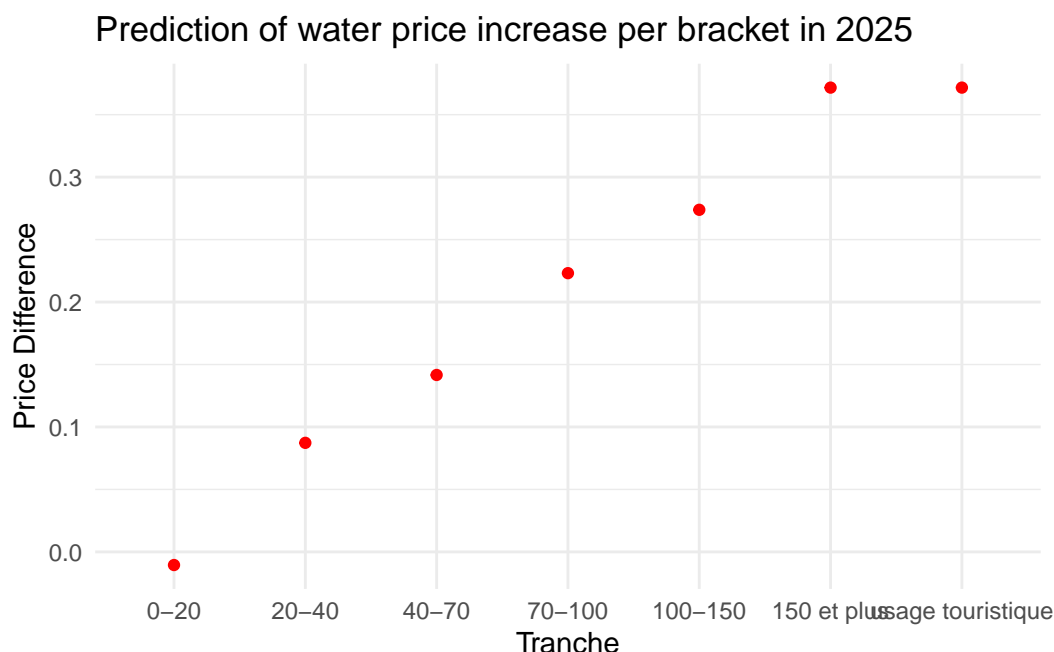
# Plot the curve with the bracket variable in the specified order
ggplot(difference_data, aes(x = tranche, y = difference)) +
  geom_line(color = "green") +
  geom_point(color = "red") +
  geom_smooth(method = "lm", se = FALSE, color = "blue") + # Add a linear regression line
  labs(x = "Tranche", y = "Price Difference", title = "Prediction of water price increase")
theme_minimal()

```

`geom\_smooth()` using formula = 'y ~ x'

`geom\_line()`: Each group consists of only one observation.

i Do you need to adjust the group aesthetic?



» **Conclusions tirées a partir de la courbe de la prédiction :**

1. **Tendance à la hausse des prix :** La régression linéaire indique une tendance à la hausse des prix des tranches du Sonede dans le temps. Cela suggère que les tarifs de l'eau devraient augmenter progressivement pour refléter les coûts croissants de gestion et de distribution de l'eau.
2. **Impact de la rareté de l'eau :** L'augmentation des prix peut être liée à la rareté croissante de l'eau, qui peut être due à des facteurs tels que la diminution des ressources en eau disponibles ou l'augmentation de la demande en raison de la croissance démographique.
3. **Nécessité d'investissements :** L'augmentation des prix peut également refléter la nécessité pour Sonede d'investir dans des infrastructures de gestion de l'eau plus efficaces ou dans des technologies de dessalement pour répondre à la demande croissante en eau potable.
4. **Conséquences pour les consommateurs :** Une augmentation des prix de l'eau peut avoir des conséquences financières pour les consommateurs, en particulier pour ceux à revenu fixe ou à faible revenu. Cela souligne l'importance de politiques sociales pour atténuer l'impact sur les ménages vulnérables.
5. **Incitation à la conservation de l'eau :** Des prix plus élevés peuvent inciter les consommateurs à adopter des pratiques de conservation de l'eau et à utiliser l'eau de

manière plus efficace, contribuant ainsi à la durabilité à long terme de la gestion des ressources en eau.

## **5-Solutions :**

Propositions de quelques solutions qu'on peut les adapter au futur pour diminuer le danger de ce manque remarquable des ressources en eau:

**1- la machine Kumulus :** qui révolutionne l'accès à l'eau potable en produisant jusqu'à 30 litres par jour, même dans les régions désertiques. Cette machine exploite l'humidité de l'air pour fournir une source d'eau sûre et fiable. Cette solution autonome offre une réponse cruciale au défi mondial de la pénurie d'eau, en offrant une solution durable et accessible aux communautés isolées et aux zones arides.

[point négatif] :le coût élevé de la machine peut limiter son accessibilité , on peut par exemple opter pour une stratégie d'encouragement des entreprises

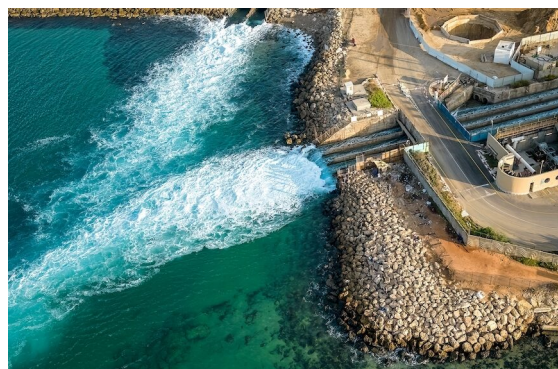


**2- L'installation de millions de boules en plastique dans les réservoirs d'eau:** offre une solution innovante pour lutter contre l'évaporation, préservant ainsi les ressources en eau. Cette approche ingénieuse réduit l'exposition de l'eau à l'air et au soleil, limitant ainsi la perte d'eau par évaporation. [solution utilisée en USA]



### 3-La dessalination de l'eau de mer :

représente une solution efficace pour répondre à la demande croissante en eau potable dans les régions souffrant de pénurie d'eau. En utilisant des technologies telles que l'osmose inverse, cette méthode élimine les sels et autres impuretés de l'eau de mer, la rendant ainsi sûre à boire.



**4- La création d'une serre** offre une solution pratique et durable pour cultiver des fruits, des légumes et des plantes tout au long de l'année, même dans des climats difficiles. elle protège les cultures des ravageurs et des variations de température, favorisant ainsi des rendements

plus élevés des plantes. En plus elle contribue à la conservation de l'eau en permettant un contrôle plus précis de l'irrigation et en réduisant la perte d'eau par évaporation.



---

### **Conclusion generale du projet :**

L'analyse des données météorologiques et hydrologiques met en évidence l'importance cruciale des précipitations pour les ressources en eau en Tunisie. Nous avons observé l'impact significatif des précipitations sur le remplissage des barrages, ainsi que ses répercussions sur la consommation d'eau quotidienne, l'irrigation agricole et l'économie nationale. Cette étude souligne l'urgence d'adopter des mesures efficaces pour une gestion durable de l'eau, notamment en renforçant les infrastructures hydrauliques, en promouvant des pratiques agricoles efficaces et en sensibilisant le public à l'importance de la préservation des ressources en eau. En outre, il est impératif de mettre en œuvre des politiques et des stratégies intégrées pour faire face aux défis croissants liés au changement climatique et à la pression démographique, afin de garantir un approvisionnement en eau adéquat pour les générations futures.

---

***“Mieux vaut prendre le changement par la main avant qu’il ne nous prenne par la gorge”***

Winston Churchill

---

## **BIBLIOGRAPHIE :**

- AgriData : <http://www.agridata.tn/>
- Data.Gov.tn :<https://Data.Gov.tn/fr/tn>
- Ministère de l’agriculture, des ressources hydrauliques et de la pêche :<http://www.agriculture.tn>

