

cleanData

ELHAMDAOUI Abderrahim

Table des matières

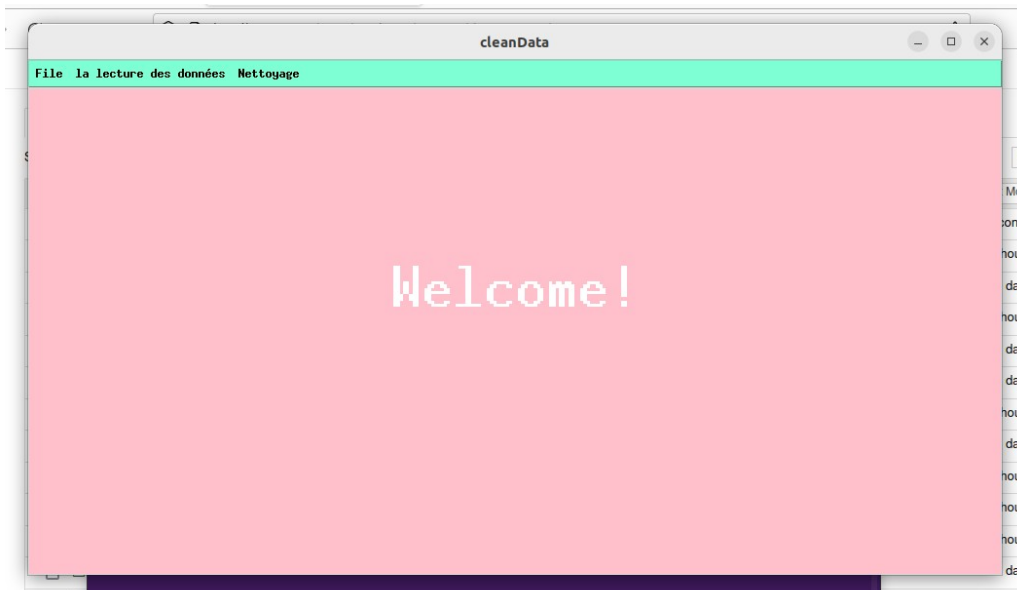
Définition.....	2
Les processus :.....	2
1. Lancer l'application :.....	2
2. choisir le dataset:.....	2
3. la lecture de dataset:.....	3
4. Nettoyage de donnée.....	6
5. Récupère le data nettoyer.....	8

Définition

cleanData une app desktop qui permet pour nettoyer une dataset par l'utilisation des algorithmes de classification de régression pour trouver une bon résultat.

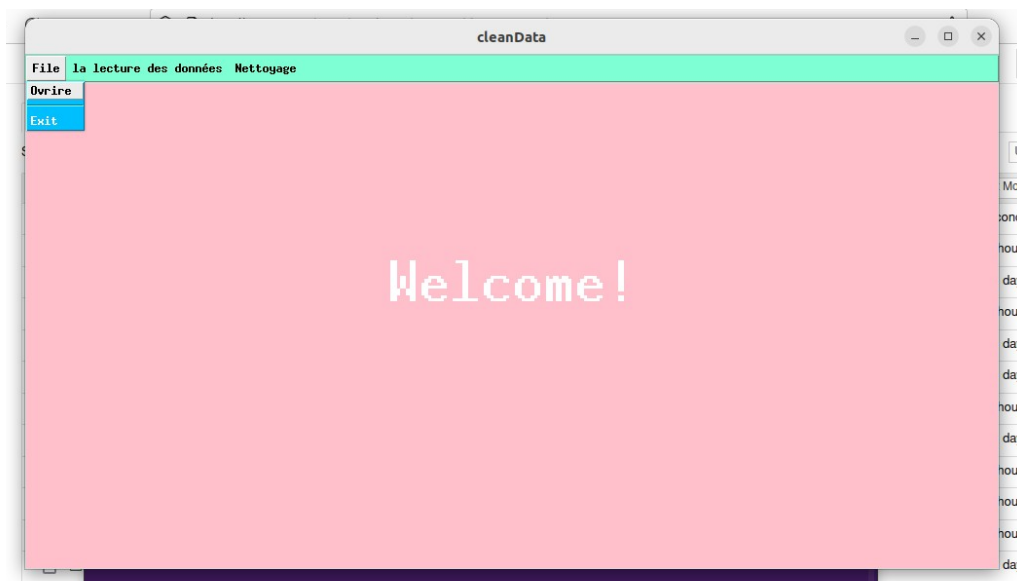
Les processus :

1. Lancer l'application :

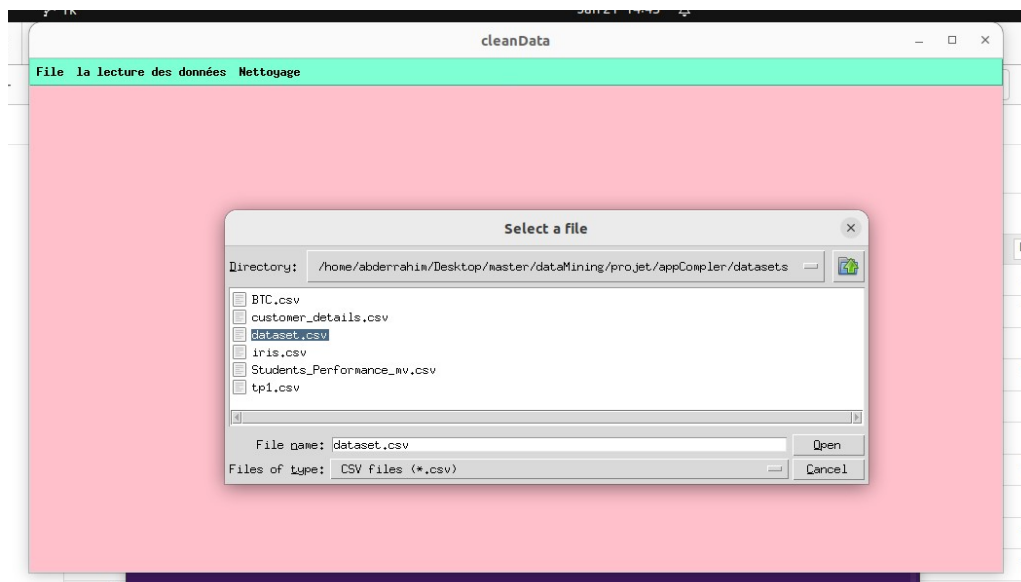


2. choisir le dataset:

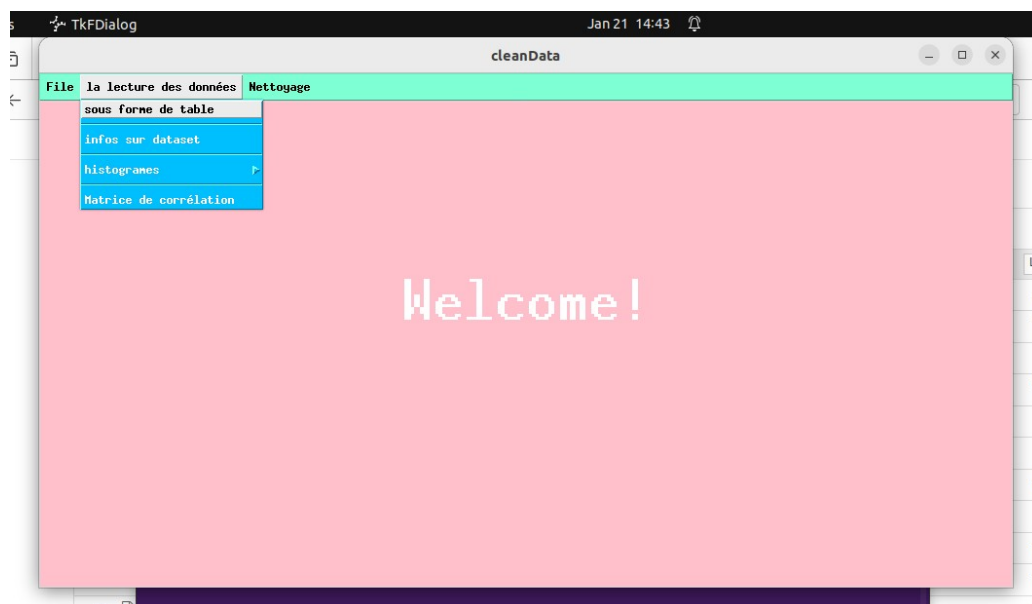
cliquer sur ouvrir :



cliquer sur ouvrir et âpre choisir votre dataset



3. la lecture de dataset:



Afficher dataset sous forme de table :

The screenshot shows the same 'cleanData' window, but now the 'sous forme de table' option has been selected. The main area displays a table with 6 columns: 'gender', 'age', 'hypertension', and 'heart_disease'. The table contains 20 rows of data.

	gender	age	hypertension	heart_disease
	Male	3.0	0	0
	Male	58.0	1	0
	Female	8.0	0	0
	Female	70.0	0	0
	Male	14.0	0	0
	Female	47.0	0	0
	Female	52.0	0	0
	Female	75.0	0	1
	Female	32.0	0	0
	Female	74.0	1	0
	Female	79.0	0	0
	Male	79.0	0	1
	Female	37.0	0	0
	Female	37.0	0	0
	Female	40.0	0	0
	Male	35.0	0	0
	Female	20.0	0	0
	Female	42.0	0	0
	Female	44.0	0	0
	Female	79.0	0	1
	Female	65.0	1	0
	Female	57.0	1	0
	Female	49.0	0	0

afficher les informations de dataset :

- le nombre des lignes.
- le nombre des colonnes.
- les attribut numériques.
- les attribut catégoriales.
- les attributs qui contiennent des valeurs manquantes.

```
cleanData
File la lecture des données Nettoyage
Infos sur dataset
le nombre des lignes est : 43400
le nombre des colonnes est : 11
les attribut numérique : ['age', 'hypertension', 'heart_disease',
'avg_glucose_level', 'bmi', 'stroke']
les attribut categorical : ['gender', 'ever_married', 'work_type',
'Residence_type', 'smoking_status']
les attributs qui contient des valeurs manquante est :
    bmi : 1462
    smoking_status : 13292
```

Afficher les histogrammes :

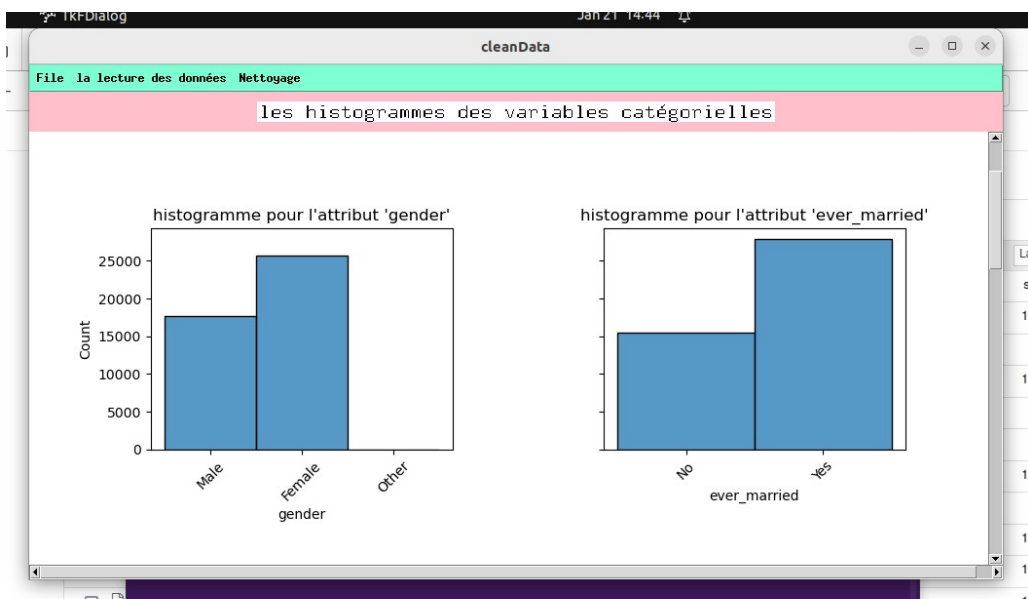


Figure : les attributs catégorielles



Figure : les attributs continues

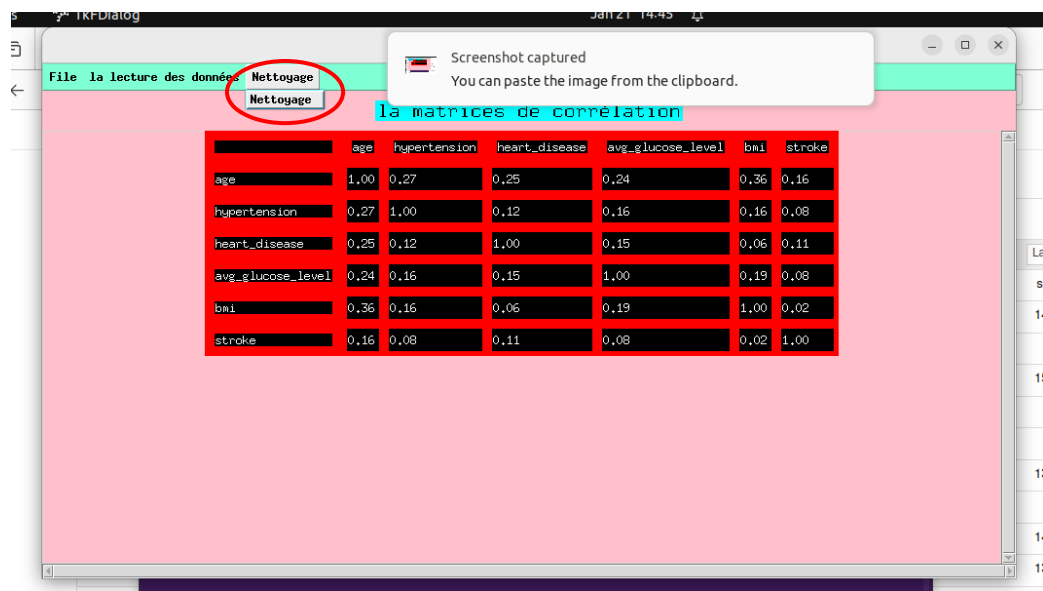
Afficher la matrice de corrélation :



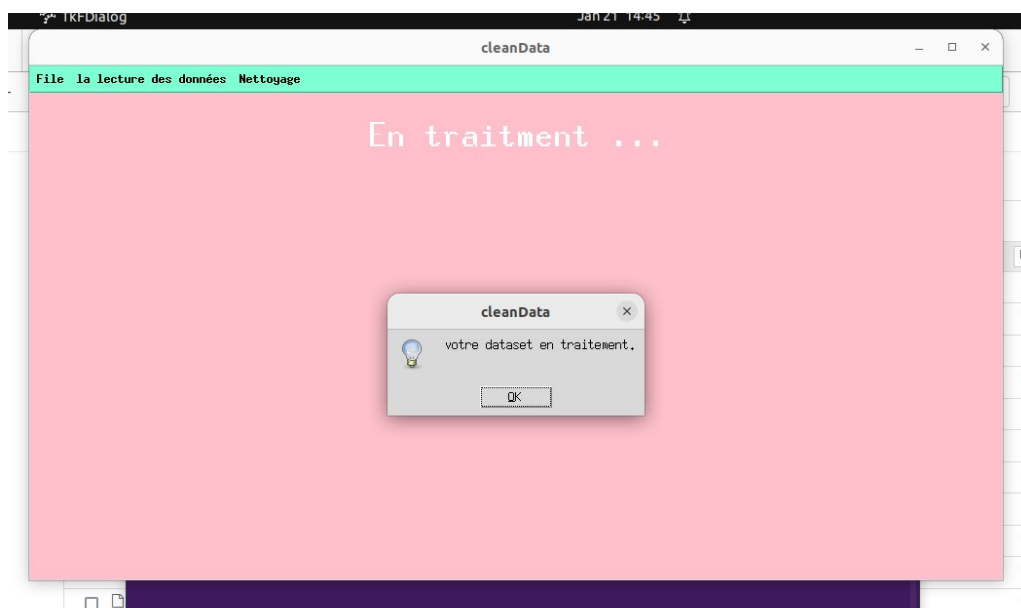
Figure : Afficher la matrice de corrélation :

4. Nettoyage de donnée.

cliquer sur netouayer



et après cliquer sur ok

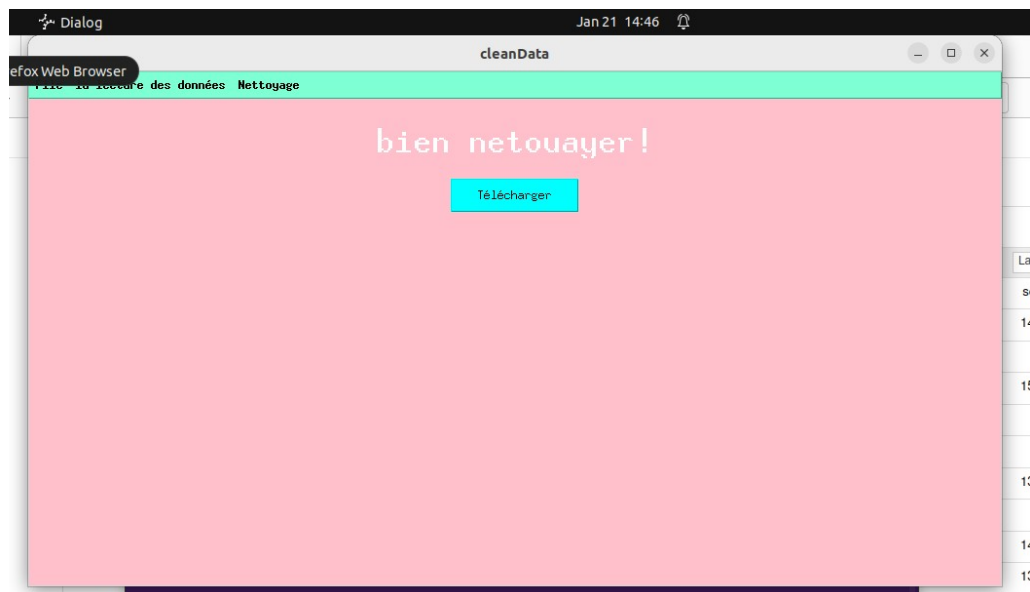


et attendre quelque minute jusqu'à le traitement est fini.



une fois le traitement est fini, on va afficher une message qui vous dit que le traitement est fini.

avec une bouton pour télécharger le dataset nettoyer.



et après cliquer sur ok pour le charger votre dataset nettoyer.



5. Récupère le data nettoyer.

Enfinement aller le dossier téléchargement à cote de l'application pour récupérer le data nettoyer.