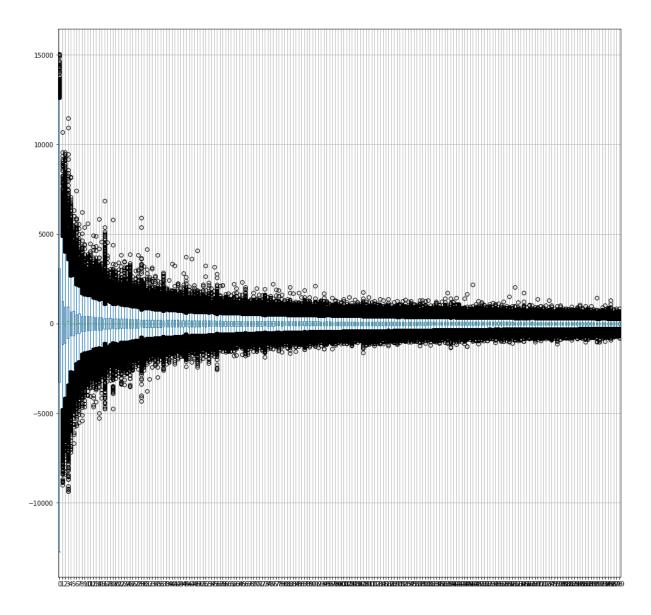
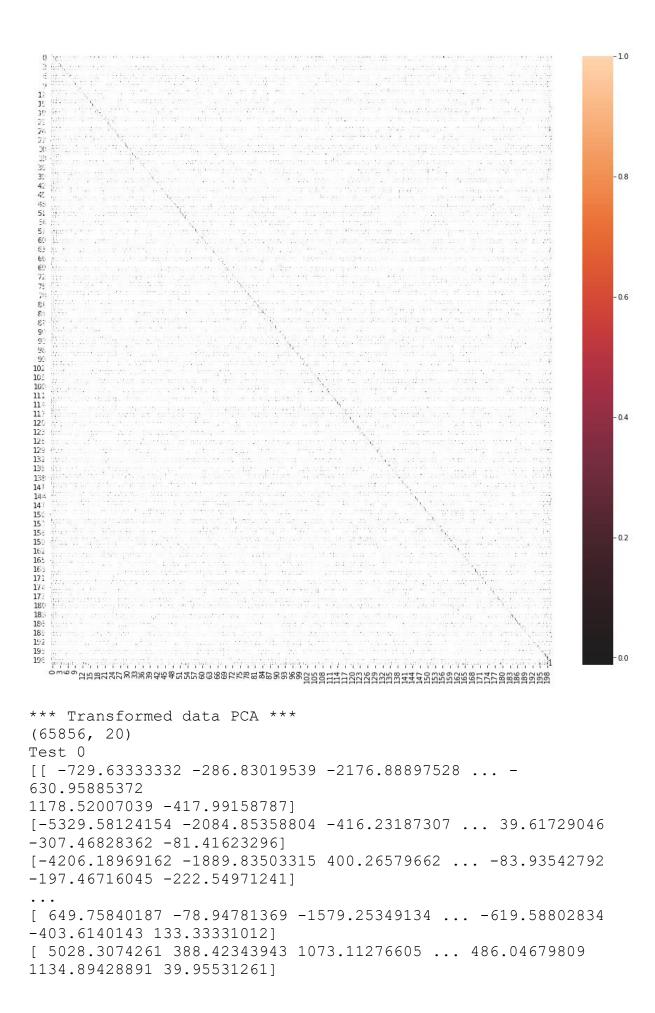
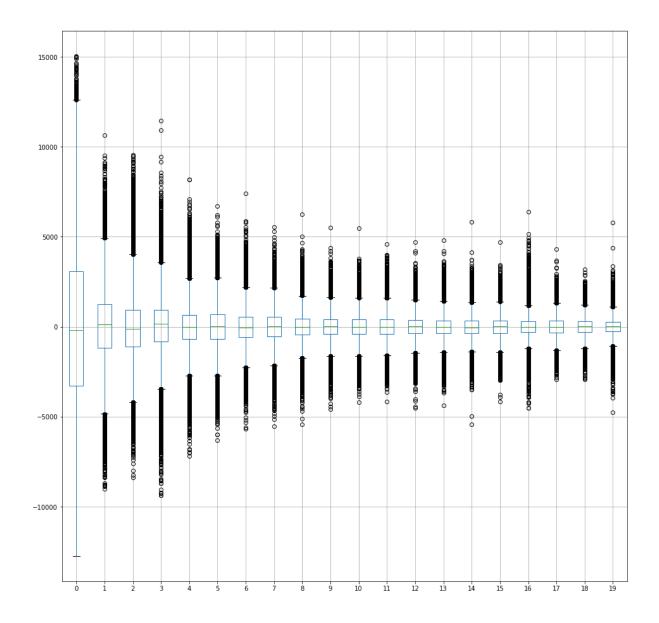
## SORTIE CONSOLE

```
In [16]:
runfile('C:/Users/isabe/Pictures/monetFINAL/starting kit/sampl
e code submission/preprocessingFinal.py',
wdir='C:/Users/isabe/Pictures/monetFINAL/starting kit/sample c
ode submission')
Reloaded modules: visualisation, data io, data converter,
data manager
Info file found : C:\Users\isabe\Downloads\monet-
master\starting kit\c1 input data\perso public.info
*** Original data ***
(65856, 200)
[[-720.579708 -330.571966 -2188.381016 ... -298.022503]
268.679125
-107.8633981
[-5322.93398 -2089.062676 -380.988992 ... 23.010128 26.57474
169.1943441
[-4200.092388 -1871.126468 447.135529 ... 67.240814 130.350498
11.2251831
[ 656.854711 -104.34349 -1564.462215 ... 111.361633 114.283328
158.923886]
5036.277568 413.900109 1060.588285 ... 87.56614 -157.06906
-174.1501621
[-1010.799311 3252.215798 -4499.276864 ... 6.165946 -70.050208
201.096046]]
```





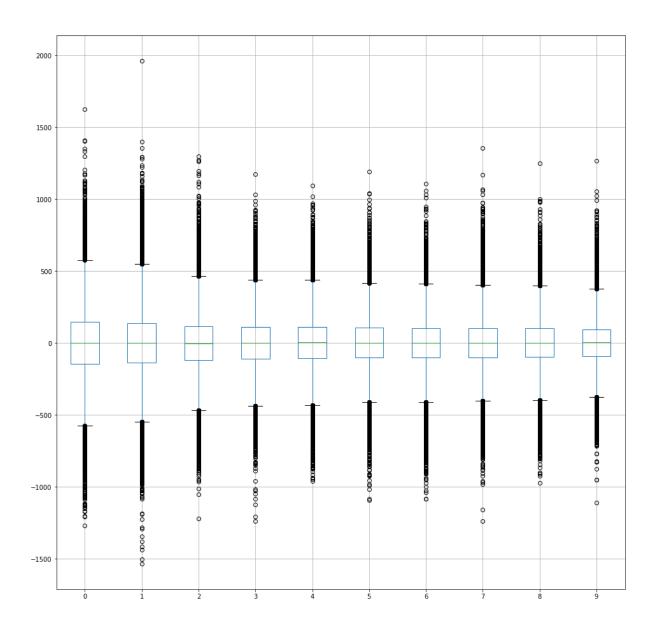
[-1015.38232173 3337.44691824 -4430.77078242 ... -129.44153824 151.47825469 153.98002354]]

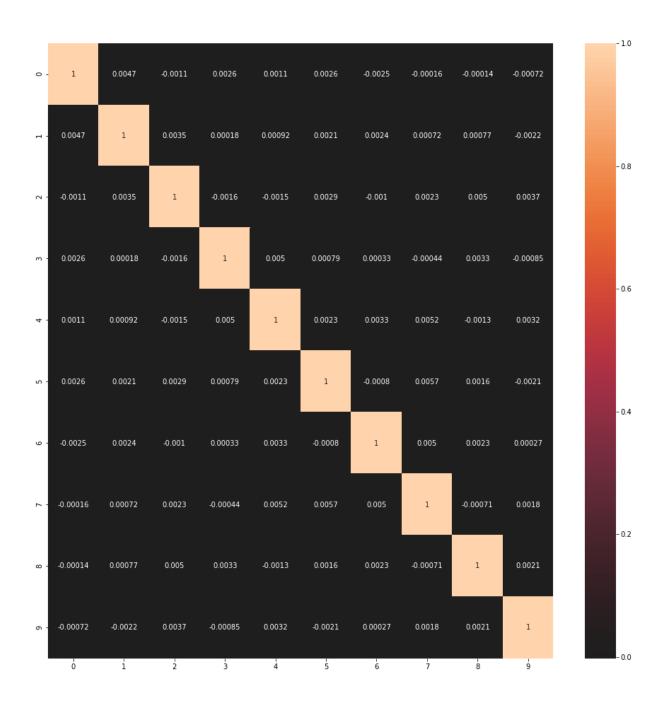


```
- 1.0
o - 1 -1.5e-171.9e-17-1e-17-1 3e-179.1e-182.7e-178.6e-186.1e-181.2e-172.7e-181.8e-171.3e-18-3e-171.3e-182.9e-173.5e-172.9e-201.9e-171.4e-17
... 1.5e-17 1 -1.7e-1덟.3e-1덟.9e-171.5e-188.6e-182.9e-171.4e-173.7e-1됩.7e-183.5e-185.1e-172.9e-171.2e-172.6e-171.4e-171.8e-18-2e-17-2.7e-1
🔼 -1.9e-171.7e-16 1 2.7e-16.7e-17 4e-17 4.6e-184.5e-182.1e-171.5e-17-6e-18-1.4e-171.8e-18 1e-17 -1.9e-186.7e-18 6e-19 1e-17 1.3e-173.2e-17
m --1e-17-1.3e-1@.7e-16 1 4.3e-1@.9e-172.1e-171.7e-172.1e-171.9e-172.4e-179.1e-181.5e-174.6e-171.2e-171.5e-181.1e-174.2e-181.5e-174.1e-17
                                                                                                                                                                                                                                                       - 0.8
9.1e-18.5e-18 4e-17 2.9e-17.5e-16 1 9.9e-173.8e-16.1e-165.1e-167.3e-172.1e-164.6e-191.4e-165.7e-171.5e-163.5e-171.8e-172.e-174.8e-18
vo -2.7e-178.6e-1¥4.6e-1¥2.1e-177.3e-189.9e-17 1 7e-18 8.9e-162.2e-168.9e-179.1e-177.7e-172.3e-161.2e-171.3e-1€2.1e-172.8e-171.7e-173.9e-18
~ -8.6e-182.9e-174.5e-181.7e-173.6e-173.8e-16.7e-18 1 1.5e-161.3e-161.6e-174.2e-165.3e-182.5e-173.7e-174.9e-182.8e-174.7e-172.7e-178.2e-19
                                                                                                                                                                                                                                                       - 0.6
∞ -6.1e-181.4e-172.1e-172.1e-172.1e-173.1e-173.1e-173.1e-168.9e-161.5e-16 1 -1.1e-161.9e-162.5e-192.9e-171.1e-161.6e-16 2e-16 4.3e-174.2e-183.2e-172.4e-17
o -1. 2e-173. 7e-181. 5e-171. 9e-175. 7e-175. 1e-162. 2e-161. 3e-161. 1e-16 1 -2. 6e-161. 5e-165. 5e-173. 8e-161. 5e-161. 5e-161. 7e-172. 5e-181. 5e-171. 4e-17
2.7e-18.7e-17-6e-182.4e-171.5e-177.3e-178.9e-171.6e-171.9e-162.6e-16 1 1.3e-161.2e-172.6e-165.9e-173.9e-169.4e-172.8e-178.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-181.3e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172.6e-172
18e-173.5e-181.4e-179.1e-181.1e-172.1e-169.1e-171.2e-162.5e-191.5e-161.3e-16 1 29e-16 6e-17-8.9e-178.1e-162.1e-17-3e-17 -1e-17-4.6e-17
                                                                                                                                                                                                                                                       - 0.4
1.3e-15.1e-17.8e-181.5e-17.4e-17.4.6e-197.7e-175.3e-18.9e-175.5e-171.2e-172.9e-16 1 -1.2e-165.7e-173.9e-163.5e-192.7e-174.3e-174.4e-17
g -3e-1729e-171e-17-6.6e-1722e-171.4e-1@.3e-1@.5e-171.1e-1@.8e-1@.6e-166e-17-1.2e-16 1 42e-164.5e-172.7e-1@.8e-174.7e-174.5e-17
크 1 3e-181.2e-171.9e-181.2e-171.7e-17.7e-171.2e-173.7e-171.6e-161.8e-165.9e-178.9e-175.7e-174.2e-16 1 8.3e-162.2e-178e-17-1.9e-161.6e-16
က္က -2.9e-172.6e-176.7e-18.5e-18.5e-18-1.5e-1வ.3e-1வ.9e-182e-16-1.5e-1ຜ.9e-1ຜ.9e-1ຜ.9e-164.5e-178.3e-16 1 3.4e-163.8e-167.1e-172.8e-17
                                                                                                                                                                                                                                                       - 0.2
9-3.5e-17.4e-17-6e-19-1.1e-17.3e-173.5e-172.1e-172.8e-173.3e-171.7e-179.4e-172.1e-173.5e-192.7e-162.2e-173.4e-16 1 2e-16-3.5e-168.9e-17
-2.9e-201.8e-181e-17 4.2e-183.5e-171.8e-171.8e-171.7e-174.2e-182.5e-182.8e-17-3e-17-2.7e-179.8e-17-3.8e-16 2e-16 1 4.4e-163.1e-1
2-13.5e-17.2e-17.3e-17.5e-17-2e-17.2e-17.17e-172.7e-173.2e-171.5e-178.6e-18-1e-174.3e-174.7e-171.9e-167.1e-173.5e-164.4e-16
g -1.4e-172.7e-173.2e-174.1e-175.7e-184.8e-183.9e-183.2e-192.4e-171.4e-171.3e-174.6e-174.4e-174.5e-171.6e-162.8e-178.9e-173.1e-171.4e-1
                                                           5 6 7 8 9 10 11 12 13 14 15 16 17 18
                                                                                                                                                                                                                    19
```

```
*** Transformed SELECTKBEST ***
(65856, 10)
Test 1
[[-422.205033 68.977914 -483.316293 ... -150.823587 -158.27646 -31.943726]
[-191.73689 259.108221 -35.791773 ... -64.802092 -35.126481 -89.511007]
[ 366.147176 -206.061556 -21.142321 ... 53.88305 18.238142 43.622182]
...
[ 214.769236 56.534752 -499.348094 ... 289.514517 340.60026
```

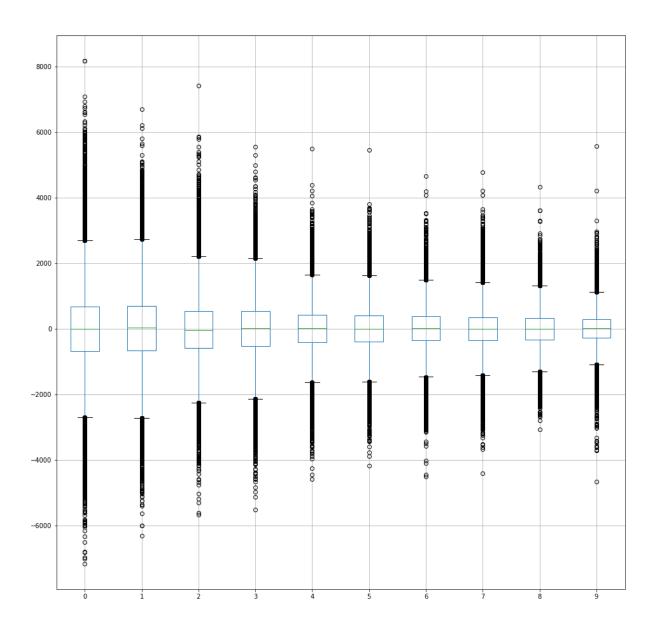
1.283359]
[ 285.649595 857.260039 -78.854101 ... -105.808715 -286.107867 22.838065]
[ -7.027664 428.999677 -197.878474 ... 126.456731 67.107772 -15.159047]]

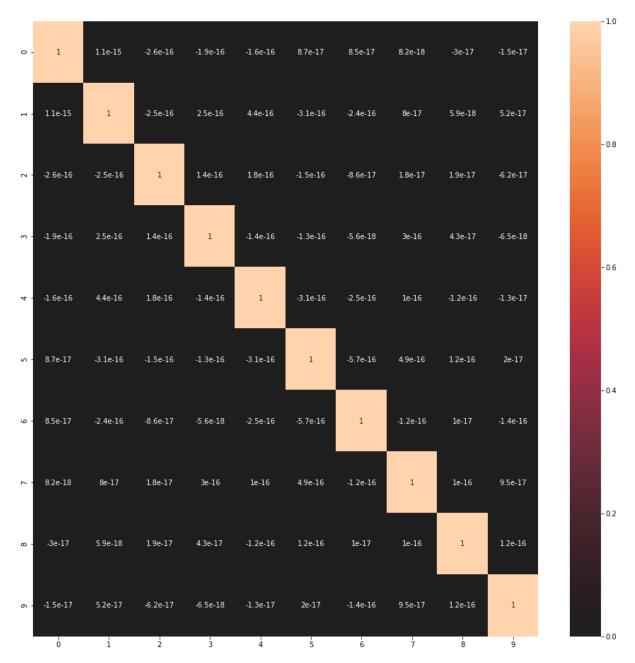




```
*** Transformed data PCA + SELECTKBEST ***
(65856, 10)
Test 2
[[ 408.23130234 -2894.87840623 970.52219793 ... 333.2207937 -638.24261047 -103.89087966]
[ -158.33853374 209.2166804 -512.2507911 ... 222.28799044 41.69091798 49.34923789]
[ -470.20853717 -56.73446203 -370.47596797 ... 293.25099416 -89.32577709 -382.782457 ]
...
[ -734.39941316 955.19105045 -819.48424445 ... 483.03990499 -627.53178658 114.84507754]
```

[-1746.31879862 -1566.40363456 -538.68385777 ... 462.47712546 472.63302157 91.96324298] [-1817.81182797 431.10840071 829.33293959 ... 426.35308882 -173.75324055 271.07035418]]





## CODE :

# -\*- coding: utf-8 -\*-

\*\*\*\*\*\*\*\*

Created on Fri Mar 1 09:54:48 2019

@author: isabe

11 11 11

,, ,, ,,

Il s'agit d'une deux fonctions de préprocessings, suivi de quelques tests pour s'assurer que nos données ont bien étés modifiés. Dans le premier cas, un preprocessing avec PCA. Le second cas, une selection des features les plus pertinants.

```
Nous avons ensuite une fonction choixPrepro, qui nous sert a
selectionner le preprocessing voulu, chose que nous
utiliserons dans les tests, qui sont fait dans le main.
from sklearn.pipeline import Pipeline
import warnings
from sklearn.feature selection import SelectKBest
from sklearn.feature selection import f regression
from sklearn.decomposition import PCA
from visualisation import duTP
import pandas as pd
from sys import path;
with warnings.catch warnings():
    warnings.filterwarnings("ignore", category=DeprecationWarning)
    from sklearn.base import BaseEstimator
    from data manager import DataManager # The class provided by
binome 1
    # Note: if zDataManager is not ready, use the mother class
DataManager
"""nombre de component pour PCA, et nombre de features pour
selectKbest"""
nbcomponent = 20;
nbkfeatures = 10;
"""Il s'agit ici du preprocessing PCA"""
class Preprocessor(BaseEstimator):
    def init (self):
        self.transformer = PCA(n components=nbcomponent)
    def fit(self, X, y=None):
        return self.transformer.fit(X, y)
    def fit transform(self, X, y=None):
         return self.transformer.fit transform(X)
    def transform(self, X, y=None):
        return self.transformer.transform(X)
    """Notre deuxieme methode de preprocessing, le selectKbest"""
class Preprocessor2(BaseEstimator):
    def init (self):
```

```
self.transformer = SelectKBest(f regression, k=nbkfeatures)
    def fit(self, X, y=None):
        return self.transformer.fit(X, y)
    def fit transform(self, X, y=None):
         return self.transformer.fit_transform(X, y)
    def transform(self, X, y=None):
        return self.transformer.transform(X)
"""Une fonction pour appeler plus facilement les méthodes lors du
test, ou nous combinons les methodes"""
#Permet de tester les différents préprocesseurs
def choixPrepro(option):
    if option == 0:
        print("\n\n*** Transformed data PCA ***")
        Prepro = Preprocessor()
    elif option == 1:
        print("\n\n*** Transformed SELECTKBEST ***")
        Prepro = Preprocessor2()
    elif option == 2:
        print("\n\n*** Transformed data PCA + SELECTKBEST ***")
        Prepro = Pipeline([('PCASelectKBest',
Preprocessor()),('SelectKBest', Preprocessor2())])
    elif option == 3:
        print("\n\n*** Transformed data SELECTKBEST + PCA ***")
        Prepro = Pipeline([('PCASelectKBest',
Preprocessor2()), ('SelectKBest', Preprocessor())])
    return Prepro
"""C'est ici que nous testons le tout: nous checkons deja en sortie
sur console, si les données varient,
  et sur leur "shape" entre les données de base et les differentes
methode de preprocessing
  varient aussi comme on le souhaite. On double check en faisant
appel à une methode du fichier visualisation.py
  qui permet de checker de manière plus visuel.
  Enfin, on triple check en les mettant en format csv pour voir
s'ils sont bien formé en sortie."""
if name ==" main ":
# We can use this to run this file as a script and test the
Preprocessor
```

```
input_dir = "C:\\Users\\isabe\\Downloads\\monet-
master\\starting_kit\\c1_input_data"
    output dir = "./fichiers preprocesses"
    basename = "perso"
    #Pour le test unitaire, on doit réduire le parametre k)
    D = DataManager(basename, input dir) # Load data
    print("*** Original data ***")
    print(D.data['X train'].shape)
    print(D.data['X train'])
    Ddf= pd.DataFrame(D.data['X train'], D.data['Y train'])
    duTP(Ddf,True)
    for i in range(3):
        Prepro = choixPrepro(i)
        test = Prepro.fit transform(D.data['X train'],
D.data['Y train'])
        # Here show something that proves that the preprocessing
worked fine
        print(test.shape)
        print("Test ", i)
        print(test)
       df = pd.DataFrame(test)
       duTP(df,True)
       # nomfichier = 'test'+str(i)+'_train.data'
       # df.to csv(nomfichier, index=False, header=False)
```