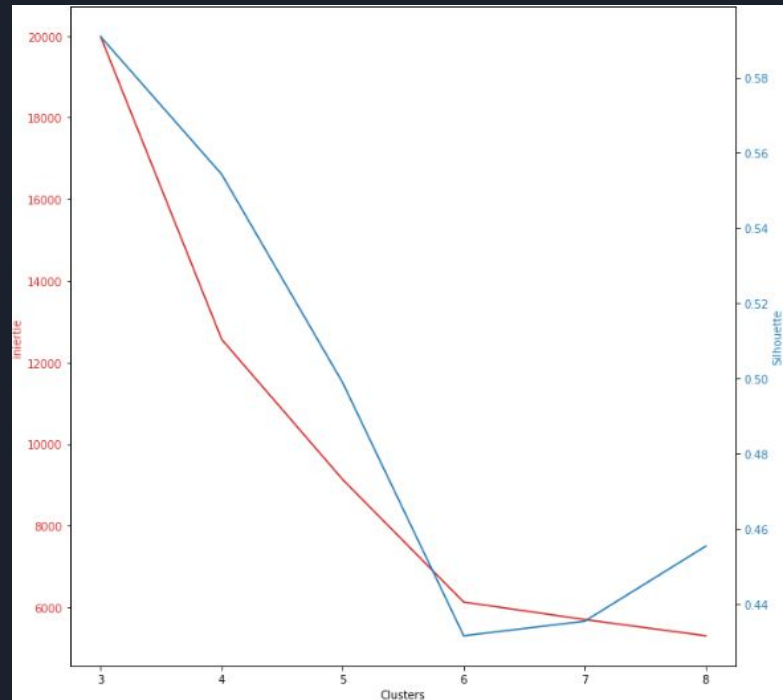
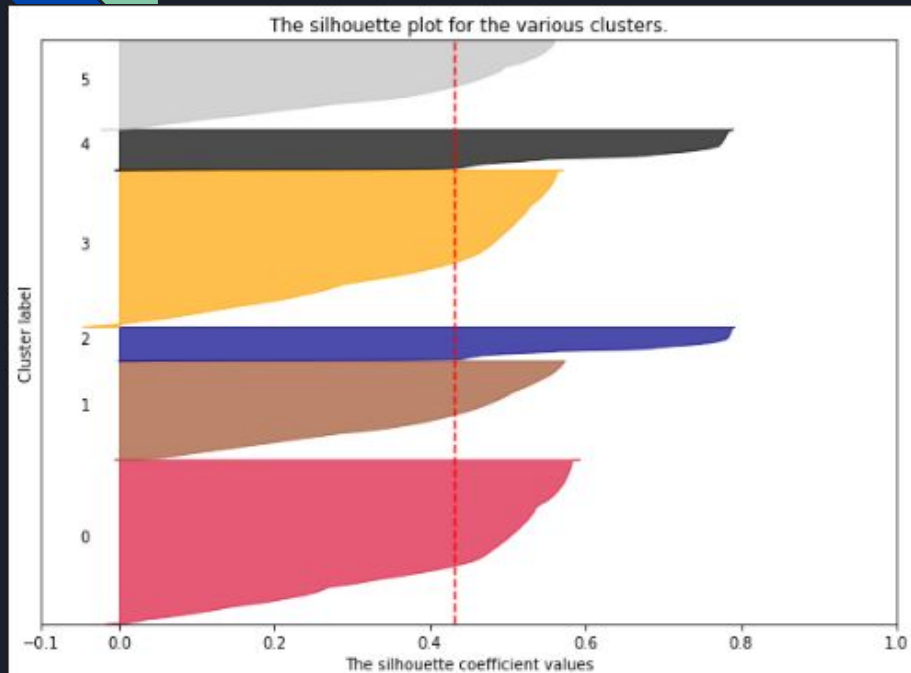




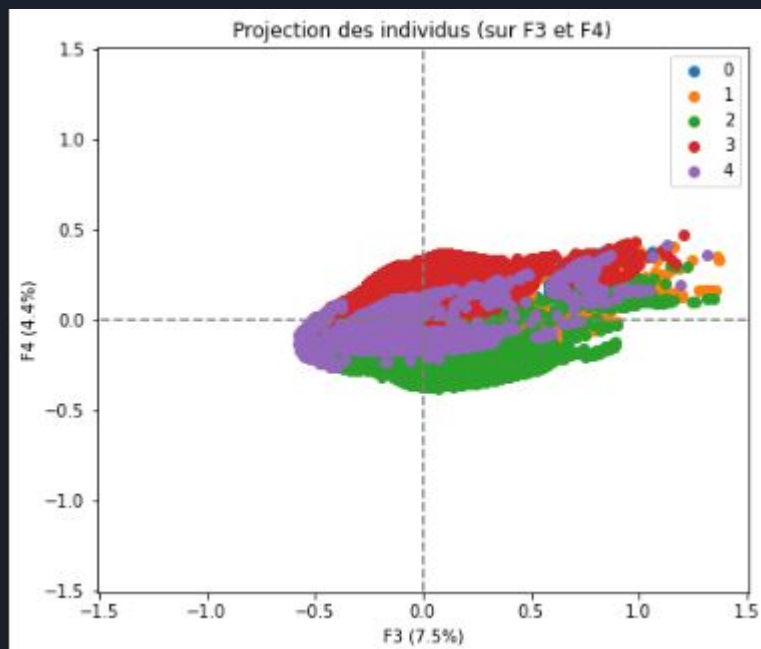
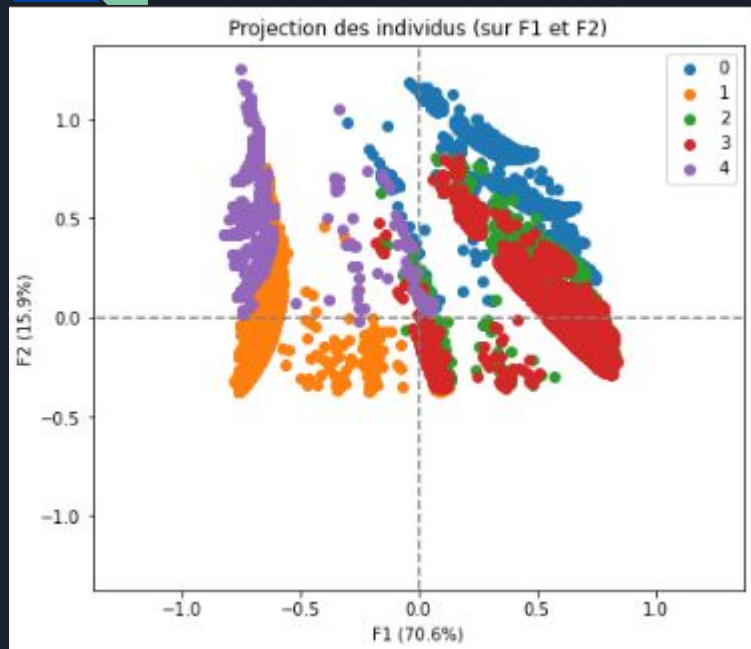
Segmentation clients pour Olist

Ayoub

6 Clusters



Visualisation des clusters





Analyses des 6 clusters

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
kmean_split								
1	402.041395	1.067289	179.084714	4.426240	0.996774	442.712490	-14.133165	4.973199
2	401.005823	1.123216	117.843573	4.440375	0.010451	441.788489	-13.516435	1.020724
3	133.631673	1.085382	112.549603	4.575246	0.006024	175.557574	-12.303336	1.010417
4	213.780461	1.108897	170.564849	1.351906	0.545297	247.809524	-0.381969	2.850127
5	132.700455	1.050920	181.591305	4.567968	0.996663	174.239845	-13.451059	4.452395
6	247.532298	2.799177	1224.904149	3.775068	0.658748	287.618539	-13.851685	5.218304

----- cluster n° 1 -----

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
customer_unique_id								
0000f46a3911fa3c0805444483337064	542.0	1	86.22	3.0	1.0	570.0	-2.0	8.0
0000f6ccb0745a6a4b88665a16c9f078	326.0	1	43.62	4.0	1.0	360.0	-12.0	4.0
0004aac84e0df4da2b147fca70cf8255	293.0	1	196.89	5.0	1.0	331.0	-8.0	6.0
0005e1862207bf6cccc02e4228effd9a0	548.0	1	150.12	4.0	1.0	592.0	-28.0	3.0
0006fdc98a402fceb4eb0ee528f6a8d4	412.0	1	29.00	3.0	1.0	447.0	-12.0	2.0
...
ffedff0547d809c90c05c2691c51f9b7	522.0	1	32.42	5.0	1.0	559.0	-56.0	3.0
ffeefd086fc667aaf6595c8fe3d22d54	341.0	1	62.94	4.0	1.0	381.0	-8.0	3.0
fff1afc79f6b5db1e235a4a6c30ceda7	369.0	1	50.09	5.0	1.0	410.0	-16.0	4.0
fff3a9369e4b7102fab406a334a678c3	388.0	1	102.74	5.0	1.0	420.0	-7.0	2.0
ffffd2657e2aad2907e67c3e9daecbeb	489.0	1	71.56	5.0	1.0	534.0	-16.0	4.0

17462 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	17462.000000	17462.000000	17462.000000	17462.000000	17462.000000	17462.000000	17462.000000	17462.000000
mean	402.041395	1.067289	179.084714	4.426240	0.996774	442.712490	-14.133165	4.973199
std	92.587669	0.272221	157.898727	0.933743	0.039795	92.411021	7.561040	2.796428
min	238.000000	1.000000	20.080000	1.000000	0.333333	116.000000	-109.000000	2.000000
25%	322.000000	1.000000	77.570000	4.000000	1.000000	365.000000	-18.000000	3.000000
50%	394.000000	1.000000	130.905000	5.000000	1.000000	436.000000	-14.000000	4.000000
75%	471.000000	1.000000	217.765000	5.000000	1.000000	513.000000	-10.000000	7.000000
max	700.000000	4.000000	1246.420000	5.000000	1.000000	743.000000	42.000000	24.000000

----- cluster n° 2 -----

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
customer_unique_id								
00082cbe03e478190aadbea78542e933	288.0	1	126.26	5.0	0.0	330.0	-8.0	1.0
001147e649a7b1afd577e873841632dd	368.0	1	212.16	4.0	0.0	409.0	-14.0	1.0
00115fc7123b5310cf6d3a3aa932699e	590.0	1	76.11	4.0	0.0	633.0	-33.0	1.0
001926cef41060fae572e2e7b30bd2a4	382.0	2	182.42	4.0	0.0	426.0	-21.0	1.0
001a2bf0e46c684031af91fb2bce149d	431.0	1	36.73	4.0	0.0	480.0	-18.0	1.0
...
ffe9102bb78a76921ba0ff3c4659616a	451.0	1	416.36	4.0	0.0	494.0	-21.0	1.0
ffe0ffa736c7b3d9af741611089729b	462.0	1	139.07	5.0	0.0	505.0	-22.0	1.0
fff699c184bcc967d62fa2c6171765f7	367.0	1	55.00	4.0	0.0	409.0	-7.0	1.0
fffa431dd3fcdefea4b1777d114144f2	308.0	1	81.20	5.0	0.0	351.0	-7.0	1.0
ffff371b4d645b6ecea244b27531430a	573.0	1	112.46	5.0	0.0	611.0	-30.0	1.0

15972 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	15972.000000	15972.000000	15972.000000	15972.000000	15972.000000	15972.000000	15972.000000	15972.000000
mean	401.005823	1.123216	117.843573	4.440375	0.010451	441.788489	-13.516435	1.020724
std	96.490308	0.394971	120.059177	0.907791	0.071105	96.175086	7.489640	0.159519
min	252.000000	1.000000	10.070000	1.000000	0.000000	72.000000	-140.000000	1.000000
25%	316.000000	1.000000	50.847500	4.000000	0.000000	357.000000	-17.000000	1.000000
50%	387.000000	1.000000	79.845000	5.000000	0.000000	429.000000	-13.000000	1.000000
75%	476.000000	1.000000	138.142500	5.000000	0.000000	518.000000	-9.000000	1.000000
max	700.000000	5.000000	1323.790000	5.000000	0.666667	743.000000	38.000000	4.000000



Groupe 1 + Groupe 2

Groupe 1(17 462 clients) + **Groupe 2**(15 972 clients) =

Groupe qui n'a pas été vu depuis + d'1an

Deux groupes qui sont bien **satisfait**, groupe 1 paye plutôt par crédit avec environ 5 mensualisations, et le groupe 2 paye plutôt en 1 fois. La moyenne des achats est situé entre 100 et 200\$

Recommandation: Ces deux clusters satisfait dans l'ensemble et c'est les deux groupes qui n'ont pas été vu depuis + d'1 an. Il serait pertinent de relancer ces deux groupes avec les produits du magasins.

----- cluster n° 3 -----

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
customer_unique_id								
0000b849f77a49e4a4ce2b2a4ca5be3f	119.0	1	27.19	4.0	0.0	171.0	-5.0	1.0
00050ab1314c0e55a6ca13cf7181fecf	136.0	1	35.38	4.0	0.0	181.0	-12.0	1.0
000949456b182f53c18b68d6babc79c1	133.0	1	82.05	4.0	0.0	173.0	-15.0	1.0
000d460961d6dbfa3ec6c9f5805769e1	239.0	1	36.68	5.0	0.0	287.0	-13.0	1.0
000e309254ab1fc5ba99dd469d36bdb4	70.0	2	78.42	3.0	0.0	108.0	-15.0	1.0
...
fff22793223fe80c97a8fd02ac5c6295	69.0	1	89.19	5.0	0.0	120.0	-18.0	1.0
fff3e1d7bc75f11dc7670619b2e61840	45.0	1	82.51	5.0	0.0	87.0	-9.0	1.0
fff7219c86179ca6441b8f37823ba3d3	250.0	1	265.80	4.0	0.0	281.0	-12.0	1.0
fffb09418989a0dbff854a28163e47c6	260.0	1	73.16	5.0	0.0	306.0	-23.0	1.0
fffea47cd6d3cc0a88bd621562a9d061	267.0	1	84.58	4.0	0.0	287.0	-3.0	1.0

23752 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	23752.000000	23752.000000	23752.000000	23752.000000	23752.000000	23752.000000	23752.000000	23752.000000
mean	133.631673	1.085382	112.549603	4.575246	0.006024	175.557574	-12.303336	1.010417
std	75.029223	0.324880	110.007247	0.699772	0.053891	72.868716	7.830274	0.107247
min	5.000000	1.000000	9.590000	1.000000	0.000000	0.000000	-71.000000	0.000000
25%	69.000000	1.000000	48.307500	4.000000	0.000000	113.000000	-17.000000	1.000000
50%	132.750000	1.000000	79.290000	5.000000	0.000000	175.000000	-12.000000	1.000000
75%	198.000000	1.000000	134.632500	5.000000	0.000000	238.000000	-7.000000	1.000000
max	459.000000	4.000000	1147.550000	5.000000	0.500000	316.000000	33.000000	4.000000

----- cluster n° 5 -----

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
customer_unique_id								
0000366f3b9a7992bf8c76cfd3221e2	116.0	1	141.90	5.0	1.0	161.0	-5.0	8.0
0004bd2a26a76fe21f786e4fbd80607f	151.0	1	166.98	4.0	1.0	202.0	-12.0	8.0
00090324bbad0e9342388303bb71ba0a	163.0	1	63.66	5.0	1.0	202.0	-8.0	3.0
000c8bdb58a29e7115cfc257230fb21b	265.0	1	29.00	5.0	1.0	304.0	-9.0	2.0
0010fb34b966d44409382af9e8fd5b77	182.0	1	61.80	4.0	1.0	212.0	2.0	6.0
...
ffeb904468642a1ce663a322629801cb	68.0	1	93.11	5.0	1.0	118.0	-16.0	3.0
ffec10ad4229ba46818560e1c8b40a68	151.0	1	135.72	5.0	1.0	189.0	-10.0	3.0
fff2ae16b99c6f3c785f0e052f2a9cfb	136.0	1	200.90	5.0	1.0	168.0	-21.0	8.0
fff96bc586f78b1f070da28c4977e810	19.0	1	63.42	4.0	1.0	68.0	-20.0	2.0
ffff5962728ec6157033ef9805bacc48	124.0	1	133.69	5.0	1.0	165.0	-14.0	5.0

22074 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	22074.000000	22074.000000	22074.000000	22074.000000	22074.000000	22074.000000	22074.000000	22074.000000
mean	132.700455	1.050920	181.591305	4.567968	0.996663	174.239845	-13.451059	4.452395
std	75.314577	0.232069	153.687564	0.726474	0.040465	73.237411	8.219207	2.531363
min	5.000000	1.000000	20.190000	1.000000	0.500000	5.000000	-147.000000	1.500000
25%	68.000000	1.000000	84.650000	4.000000	1.000000	112.000000	-18.000000	2.000000
50%	130.000000	1.000000	138.115000	5.000000	1.000000	171.000000	-13.000000	4.000000
75%	197.000000	1.000000	215.580000	5.000000	1.000000	236.000000	-8.000000	6.000000
max	485.000000	3.000000	1159.650000	5.000000	1.000000	323.000000	28.000000	24.000000



Groupe 3 + Groupe 5

Groupe 3 : (23 752 clients) => Un groupe qui paye plutôt en 1 fois à la différence du groupe suivant, sinon sont plutôt satisfait. Vu récemment en moyenne (**environ 3-4 mois**)

Groupe 5 : (22 074 clients) => Idem que le groupe précédent mais paye en plusieurs fois, (environ 2.6 mensualité par achat par client)

cluster n° 4

customer_unique_id	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
0005ef4cd20d2893f0d9fbd94d3c0d97	175.0	1	129.76	1.0	1.0	206.0	31.0	4.0
0012929d977a8d7280bb277c1e5f589d	321.0	1	155.65	1.0	0.0	369.0	-15.0	1.0
0019da6aa6bcb27cc32f1249bd12da05	204.0	1	87.71	1.0	0.0	254.0	-10.0	1.0
0025795df7a7d077c4c90162fa820085	41.0	2	428.16	1.0	1.0	91.0	-18.0	3.0
002aba8c1af80acacef6e011f9f23262	188.0	1	217.74	1.0	1.0	212.0	29.0	3.0
...
ffb3bb45b560b9ff08bdde6786987c796	147.0	1	188.12	1.0	1.0	184.0	25.0	8.0
ffc3b3eb07f2a033d45949fcedd160d98	114.0	1	206.08	2.0	1.0	159.0	-10.0	3.0
ffec490ab531184a483efe2eedd68908	31.0	1	57.98	1.0	0.0	73.0	9.0	1.0
ffeddf8aa7cdecf403e77b2e9a99e2ea	113.0	2	204.20	2.0	0.5	136.0	6.0	3.0
fff1bdd5c5e37ca79dd74deeb91aa5b6	191.0	1	172.98	1.0	0.0	216.0	32.0	1.0

9835 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	9835.000000	9835.000000	9835.000000	9835.000000	9835.000000	9835.000000	9835.000000	9835.000000
mean	213.780461	1.108897	170.564849	1.351906	0.545297	247.809524	-0.381969	2.850127
std	108.221999	0.350538	161.253036	0.686559	0.494954	106.316111	16.664255	2.504239
min	5.000000	1.000000	13.890000	1.000000	0.000000	27.000000	-34.000000	1.000000
25%	149.000000	1.000000	68.130000	1.000000	0.000000	183.000000	-11.000000	1.000000
50%	200.000000	1.000000	119.020000	1.000000	1.000000	231.000000	-3.000000	2.000000
75%	281.000000	1.000000	205.110000	1.000000	1.000000	312.000000	7.000000	4.000000
max	604.000000	5.000000	1354.090000	5.000000	1.000000	613.000000	188.000000	22.000000



Groupe 4

Groupe 4: (9 835 clients) => Un groupe peu satisfait, c'est le groupe qui a le plus de retard sur les délai de livraison.

Recommandation: Gagner en précision sur l'estimation des délais de livraison pour assurer un meilleur score et une meilleur fidélité

----- cluster n° 6 -----

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
customer_unique_id								
00053a61a98854899e70ed204dd4bafef	187.0	2	838.36	1.000000	1.0	226.000000	-10.0	3.000000
005226be8d5b972d0f72b4e8a79f322b	155.0	1	1600.51	5.000000	1.0	203.000000	-19.0	10.000000
0085139b503668b3addcfe826d91be8d	216.0	2	1008.94	2.000000	1.0	255.000000	-6.0	5.000000
00a9fd000ff87bb48b989df819c418f5	167.0	4	259.40	5.000000	0.0	216.000000	-35.0	1.000000
00aded9b742746c0c66e10d00ea1b74	300.0	2	4296.80	1.000000	1.0	347.000000	-21.0	10.000000
...
ff8892f7c26aa0446da53d01b18df463	343.0	3	396.88	2.666667	1.0	369.666667	11.0	7.333333
ff8eddc3ff272a5724e2ac6635cb7df2	64.0	2	1983.48	3.000000	1.0	92.000000	-35.0	8.000000
ffba9f9dff87b05e310ecc46c8591044	553.0	1	1626.83	5.000000	1.0	586.000000	-23.0	5.000000
fff5eb4918b2bf4b2da476788d42051c	63.0	1	2844.96	5.000000	0.0	102.000000	-16.0	1.000000
fffcf5a5ff07b0908bd4e2dbc735a684	452.0	2	4134.84	5.000000	1.0	475.000000	-27.0	10.000000

2186 rows × 8 columns

	recence	frequence	montant	review_score	by_credit	last_review	delay	mean_installment
count	2186.000000	2186.000000	2186.000000	2186.000000	2186.000000	2186.000000	2186.000000	2186.000000
mean	247.532298	2.799177	1224.904149	3.775068	0.658748	287.618539	-13.851685	5.218304
std	144.143773	2.673614	961.376253	1.494895	0.451705	142.982197	9.316334	3.885374
min	8.000000	1.000000	19.000000	1.000000	0.000000	52.000000	-55.333333	1.000000
25%	132.000000	1.000000	659.947500	3.000000	0.000000	172.000000	-19.000000	1.000000
50%	236.000000	2.000000	1104.675000	4.333333	1.000000	276.000000	-14.000000	5.000000
75%	351.000000	4.000000	1521.540000	5.000000	1.000000	389.000000	-9.000000	9.000000
max	699.000000	54.000000	13664.080000	5.000000	1.000000	737.000000	96.000000	24.000000



Groupe 6

Groupe 6: (2 186 clients) => Groupe qui achète le plus avec une fréquence moyenne de 2.8 commandes par client, c'est aussi le groupe avec la plus haute moyenne d'achat à environ 1200\$, payer environ en 5 fois.

Note moyenne un peu plus basse, que dans les autres clusters, noté à 3.77/5. **Groupe non vu depuis environ 8 mois**

Recommandation: Groupe le plus fidèle pour Olist à ces périodes d'achats. Comprendre pourquoi la note est un peu plus basse afin de maintenir la fidélité ? Questionnaires par mail ?

Résumé des groupes

	Groupe 1	Groupe 2	Groupe 3	Groupe 5	Groupe 4	Groupe 6
Clients	17 462	15 972	23 752	22 074	9 835	2 186
Recense	> 1 an	> 1 an	4 mois	4 mois	7 mois	8 mois
Budget M	180	110	110	180	170	1 224
shipment	Ok	Ok	Ok	Ok	Delay	Ok
Satisfait	OK	OK	OK	OK	Non	3.77 / 5
Credit	Yes	No	No	Yes	0.5	Yes
Result	A relancer	A relancer	Client satisfait qui achète en 1 fois	Client satisfait qui achète par crédit	Promo suite aux retards de livraison ?	Les + fidèles du passé, à relancer



SCORE ARI

```
from sklearn.metrics import adjusted_rand_score
```

```
# Score ARI qu'on peut plot sur le temps et comparer l'évolution et voir quand il faut remettre les clusters a jours
```

```
predict = k_pipe.predict(df_model.iloc[:,0:14])  
adjusted_rand_score(df_model['kmean_split'], predict)
```

021 ✓ 0.4s

• 0.7977385085087856