

Analyse des résultat de sondage

Étude statistique avec la fonction effectifs

Le rôle de cette fonction est d'analyser le pourcentage des variables.

❖ En fonction de sexe

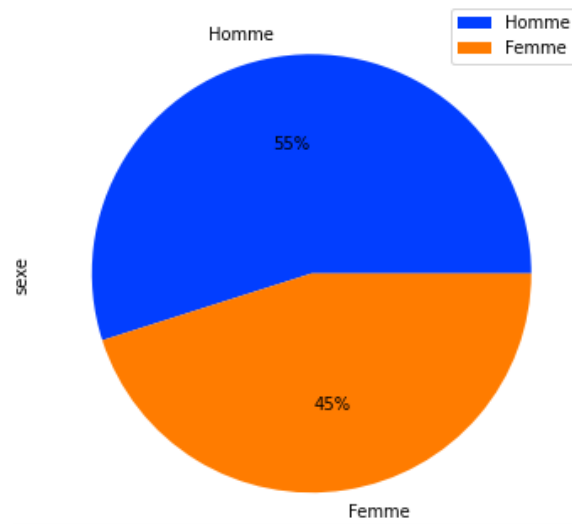


Figure 9 pie chart sur le sexe d'échantillonnage

On remarque que les hommes représentent 55% des réponses, Par contre les femmes avec 45, mais ce n'ai pas une grande différence chose qui rend les réponses plus au moins significatifs

❖ En fonction de l'âge

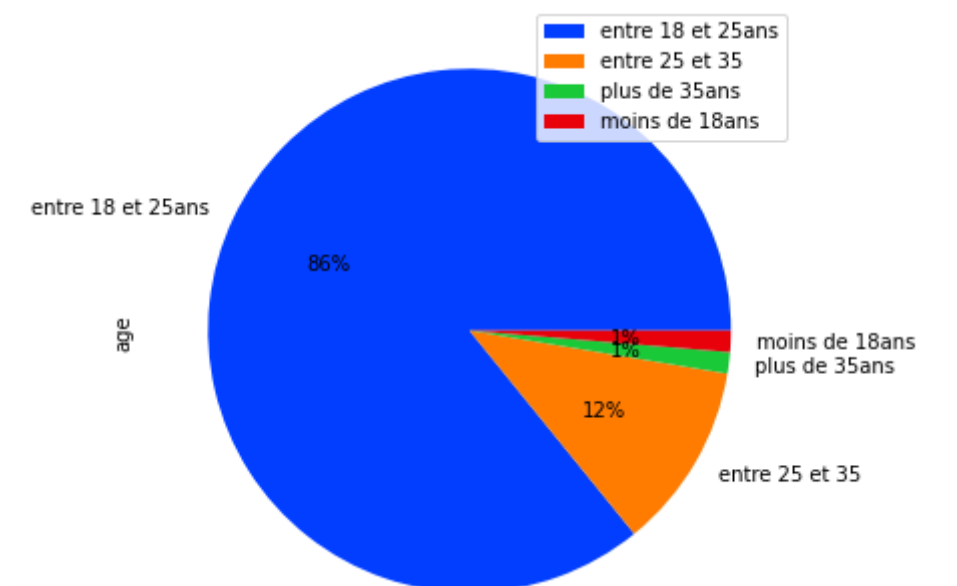


Figure 10 pie chart sur l'age

On remarque que l'âge entre 18 et 25 est la grand pourcentage dans notre

❖ En fonction de l'arrondissement

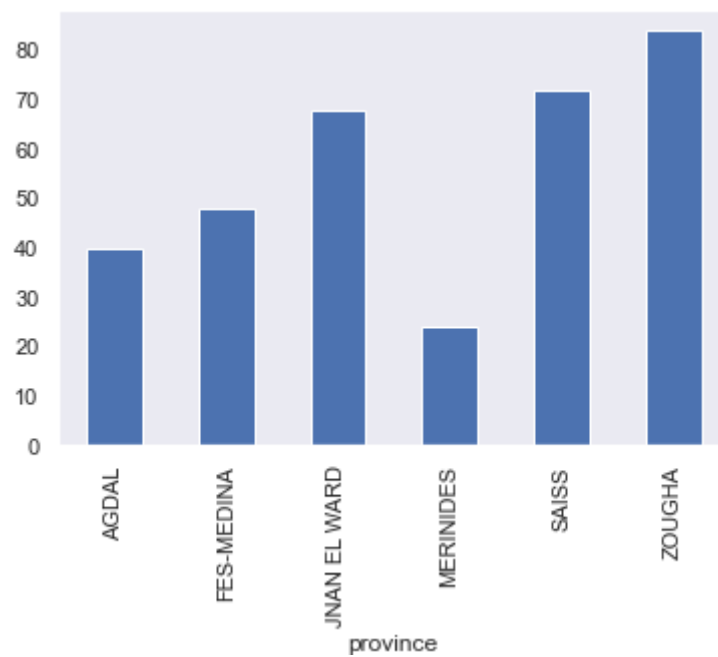


Figure 11 graphe de nombres des répons par arrondissement

On remarque zouagha et jnan el Ward est la grand percentage dans notre répons

❖ En fonction de les bureau de l'etat civil

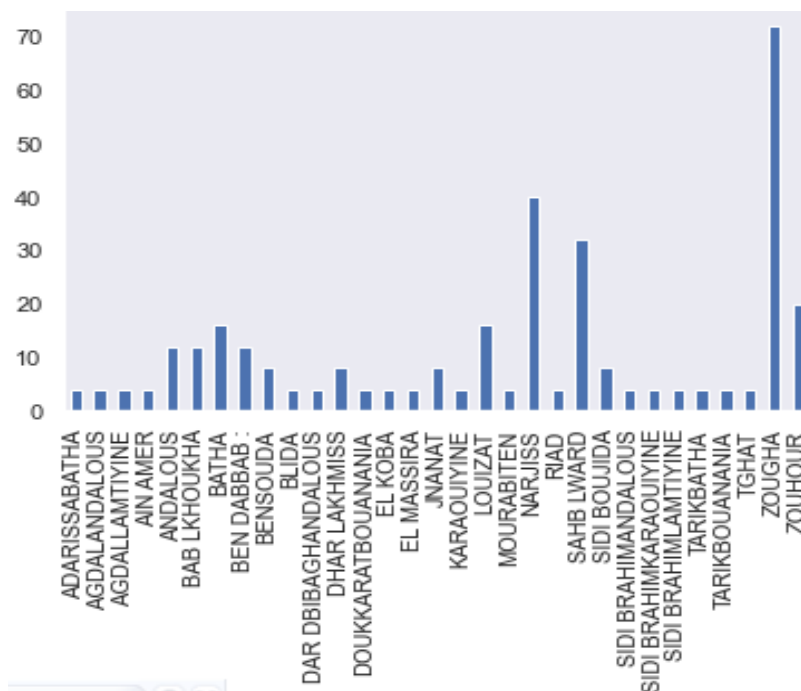


Figure 12 graphe nombre des répons par bureau

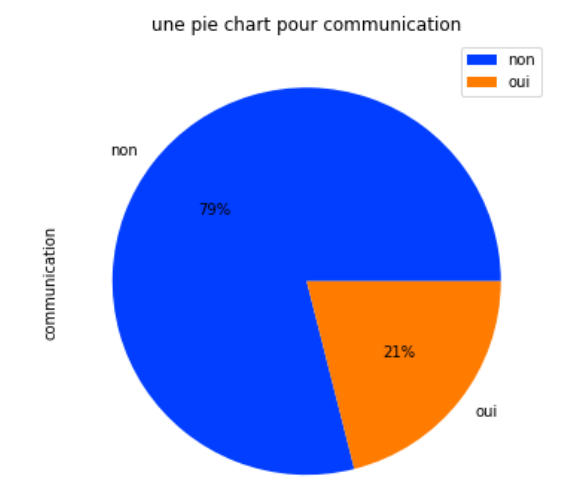
Ce graphe nous montre que les répons sont dans la même intervalle pour la plupart des bureaux sauf quelque cas exceptionnel comme **zouagha** et **narjis**

❖ En fonction de nombre de fois visite



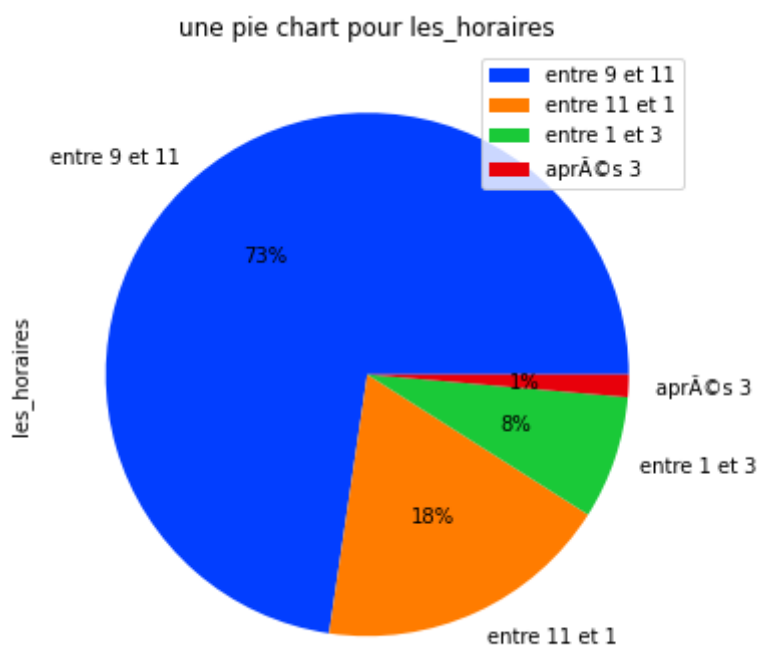
On a la plupart des personnes visite les bureaux de l'état civil rarement

❖ En fonction de la qualité de communication



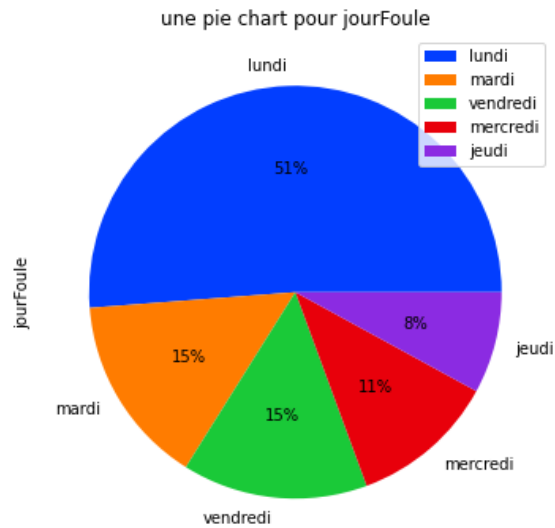
La plupart des personnes ont des problèmes de communication avec un pourcentage de 79%

❖ En fonction de temps de visite les bureau de l'état civil



On remarque que la plupart des personne visité les bureau de l'état civil entre 9 et 11h de matin

❖ On fonction du jour de visite le bureau de l'etat civil



On remarque que la plupart des personnes ont visité les bureaux de l'état civil lundi avec un grand pourcentage plus que 50%

❖ En fonction de taux de satisfait

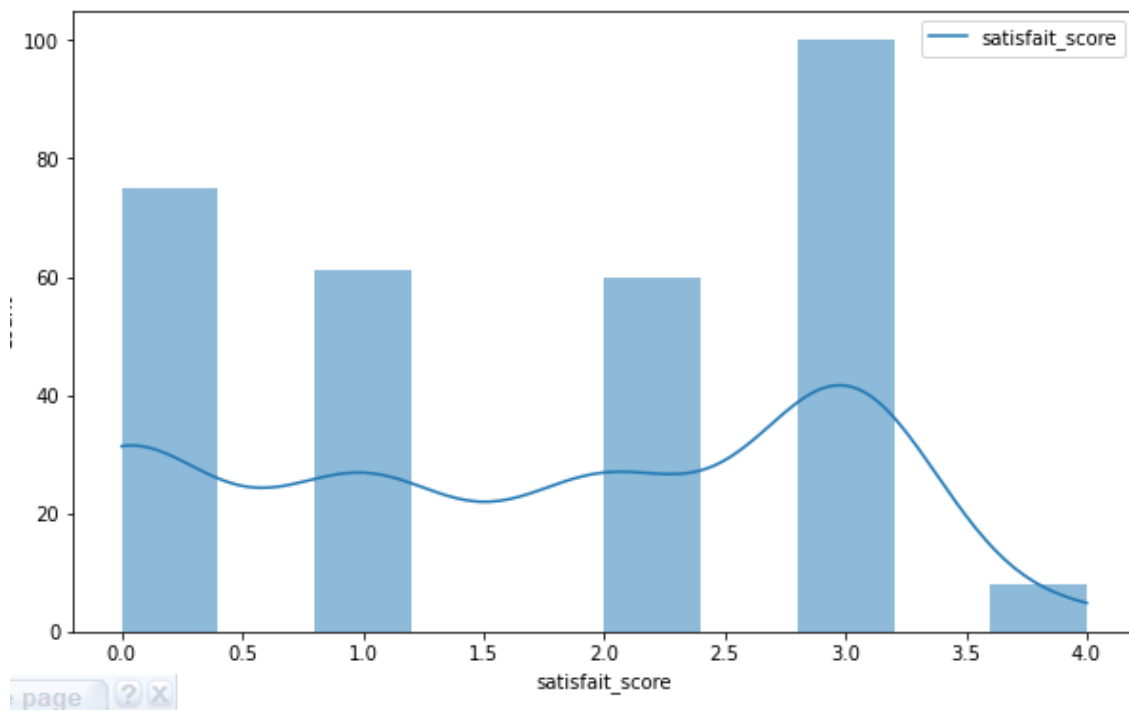


Figure 13 graphe de taux de satisfait selon les reponds

❖ Les problemes en fonction de group d'age



Figure 14 graphe de les problèmes les plus connus selon l'âge

On remarque que dans les personnes entre 18 et 5 le plus grand problème est de la mauvaise manière de traitement, mais pour les personnes entre 25 et 40 le plus grand problème est de la foule

Analyse des relations entre les variables

Etude de la relation entre les variables et le taux de satisfaction

Pour la réalisation de cette étude on doit diviser la population que nous avons en 2 catégories :

- ❖ Satisfait : qui ont un taux de satisfaction plus que 2
- ❖ Insatisfait : qui ont un taux de satisfaction moins que 2

Et on prend la même taille (100 individus) pour chaque catégorie, pour la réalisation de cette division on utilise python pour diviser la population et pour prendre un échantillon

```
satisfait=test[test['satisfait_score']>2]
insatisfait=test[test['satisfait_score']<2]

:
insatisfait_simple=insatisfait.sample(100)
satisfait_simple=satisfait.sample(100)
```

Figure 15 extrait de code pour la division de population

Des informations statistiques sur les échantillons

- ❖ Satisfait

	sexe	age	visite_score	communication	satisfait_score	la foule	Manque des employé	mauvaise maniere Traitement	les_horaires	jourFoule
count	100.000000	100.00	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	98.000000	98.000000
mean	0.440000	1.03	1.03000	0.390000	2.510000	0.490000	0.460000	0.660000	1.428571	2.326531
std	0.498888	0.30	0.70288	0.348735	0.810225	0.476095	0.500908	0.429235	0.642161	1.448608
min	0.000000	0.00	0.00000	0.000000	2.000000	0.000000	0.000000	0.000000	1.000000	1.000000
25%	0.000000	1.00	1.00000	0.000000	2.000000	0.000000	0.000000	1.000000	1.000000	1.000000
50%	0.000000	1.00	1.00000	0.000000	2.000000	0.000000	0.000000	1.000000	1.000000	2.000000
75%	1.000000	1.00	2.00000	0.000000	3.000000	1.000000	1.000000	1.000000	2.000000	3.000000
max	1.000000	2.00	2.00000	1.000000	4.000000	1.000000	1.000000	1.000000	3.000000	5.000000

Figure 16 information statistique de l'échantillonnage satisfait

❖ Insatisfait

	sexe	age	visite_score	communication	satisfait_score	la foule	Manque des employé	mauvaise maniere Traitement	les_horaires	jourFoule
count	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	96.000000	98.000000
mean	0.570000	1.130000	0.590000	0.210000	0.390000	0.480000	0.340000	0.780000	1.291667	1.653061
std	0.49757	0.366667	0.668105	0.40936	0.490207	0.502117	0.476095	0.416333	0.694591	1.252643
min	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	1.000000
25%	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
50%	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
75%	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.750000
max	1.000000	3.000000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000	4.000000	5.000000

Figure 17 information statistique de l'échantillonnage insatisfait

On remarque que on a des différents dans les moyens et écart type entre les deux échantillon, pour voir le différents plus détails on commence par les comparaison suivant

❖ La relation entre le sexe et le taux de satisfait

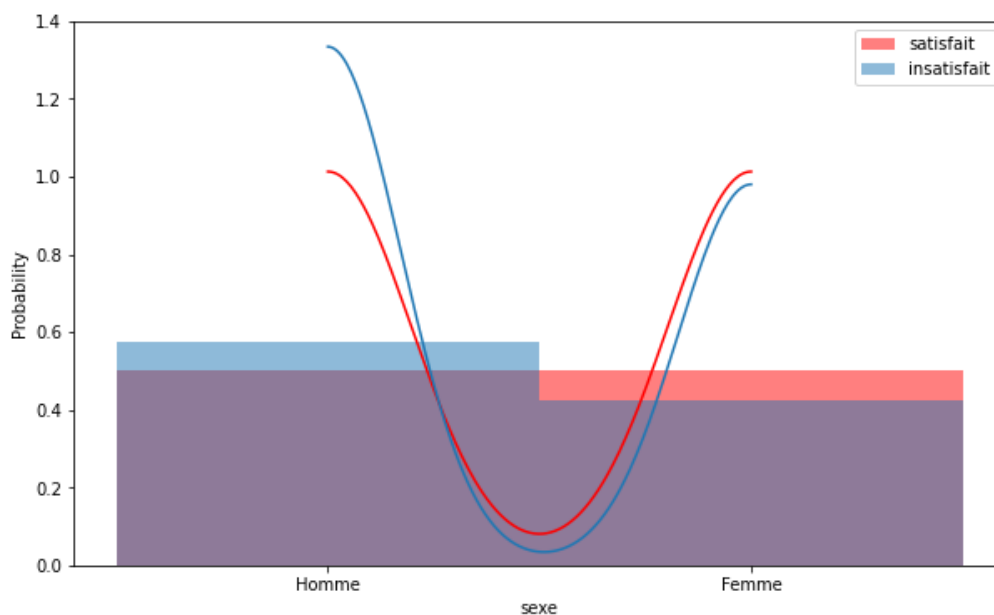


Figure 18 graphe de sexe selon le taux de satisfait

On remarque que les hommes sont plus insatisfaits, mais ce n'est pas une grande différence chose qui rend le taux de satisfait change seulement le sexe

❖ La relation entre l'arrondissement et le taux de satisfait

On utilise SPSS pour analyser la relation entre arrondissement et le taux de satisfait

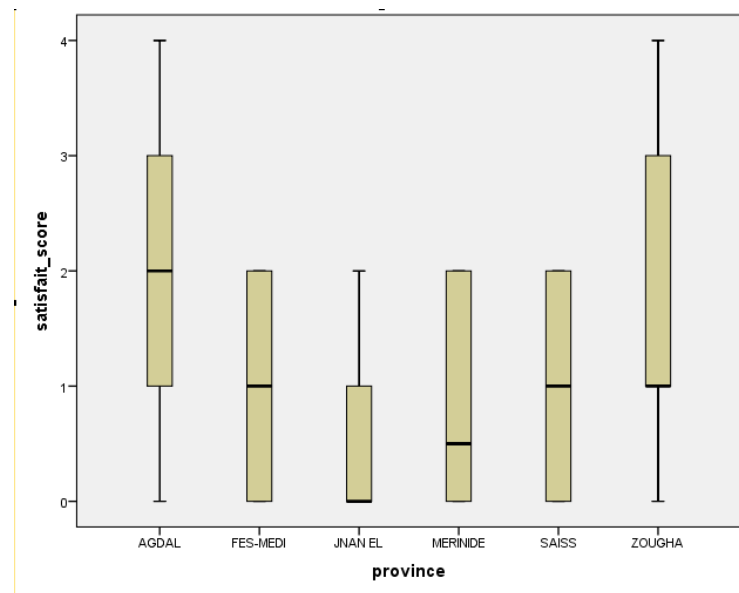


Figure 19 extrait de spss

on remarque que on a un grand différence entre le taux de satisfait dans zougba et Jnan el Ward et pour bien voir la différence on utilise python pour la création de la graphie suivant :

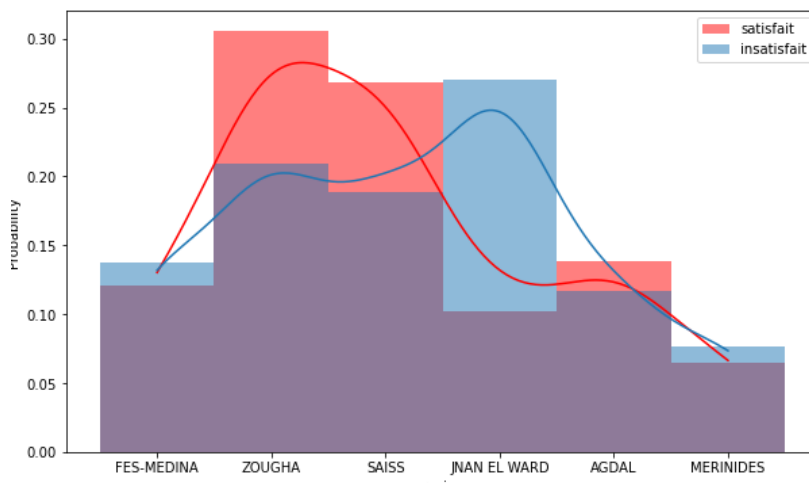
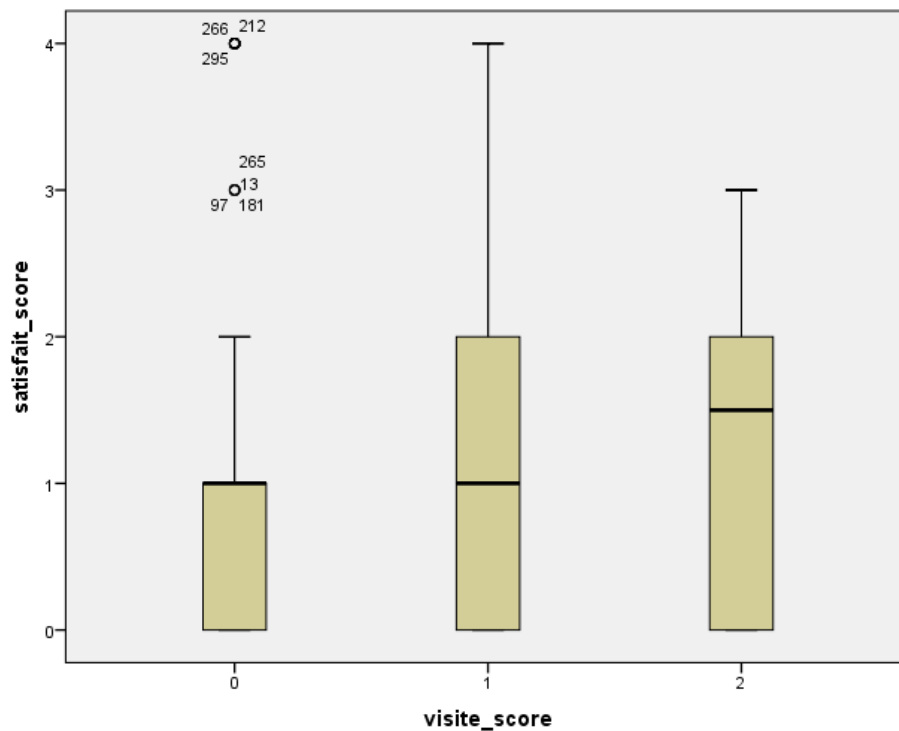


Figure 20 graphe de taux de satisfait en fonction de arrondissement

on remarque que on a un grand différence entre le taux de satisfait dans **zougba** et Jnan el Ward

❖ La relation entre le taux de visite et le taux de satisfait

On utilise aussi spss pour voir la relation entre ces deux variables



On peut remarqué que lorsque le visite score est élevé le moyen de taux de satisfait et aussi élevé avec la même façon , et la même chose pour le taux des visites , pou bien comprendre on a le graphe suivant :

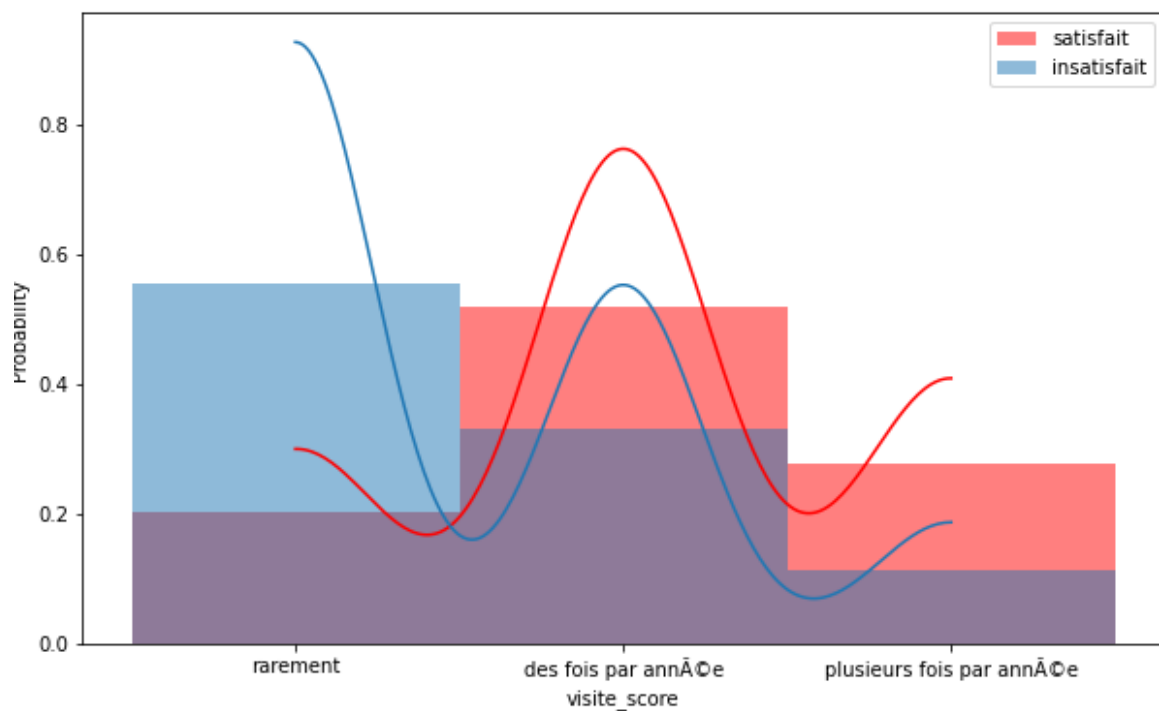
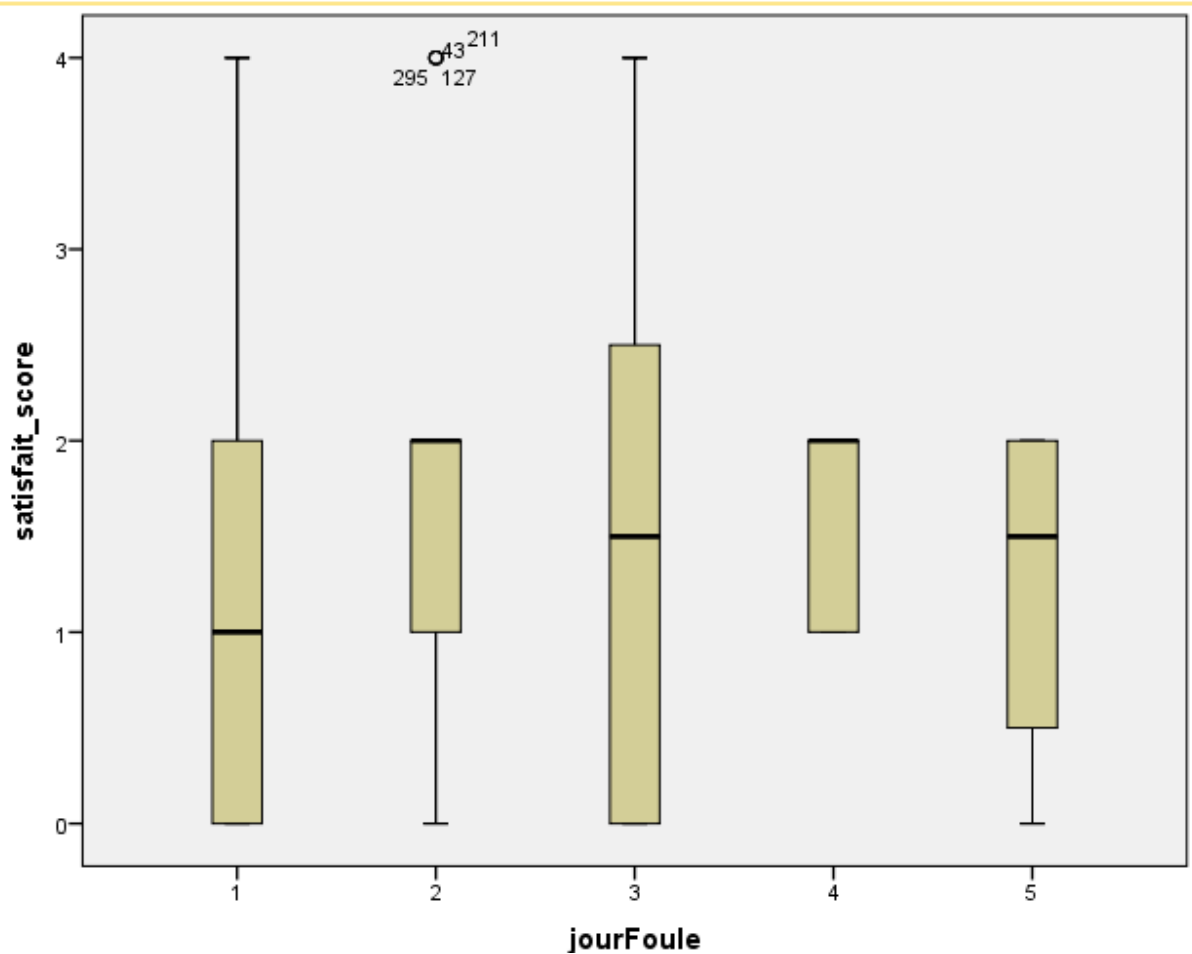


Figure 21 graphe de taux de satisfait en fonction de score de visites

En remarque que les personne qui ont visité les bureau de l'état civil plus est satisfait plus que les autres

❖ La relation entre le taux de satisfait et le jour foule

On utilise SPSS pour analysé la relation entre arrondissement et le taux de satisfait



On remarque que le pourcentage entre les jours est proche entre eux , ainsi qu'on remarque que le moyen de satisfait est moins dans le lundi , par contre dans les autres jours de la semaine , on voit ces remarque d'une façon plus claire dans le graphe suivant :

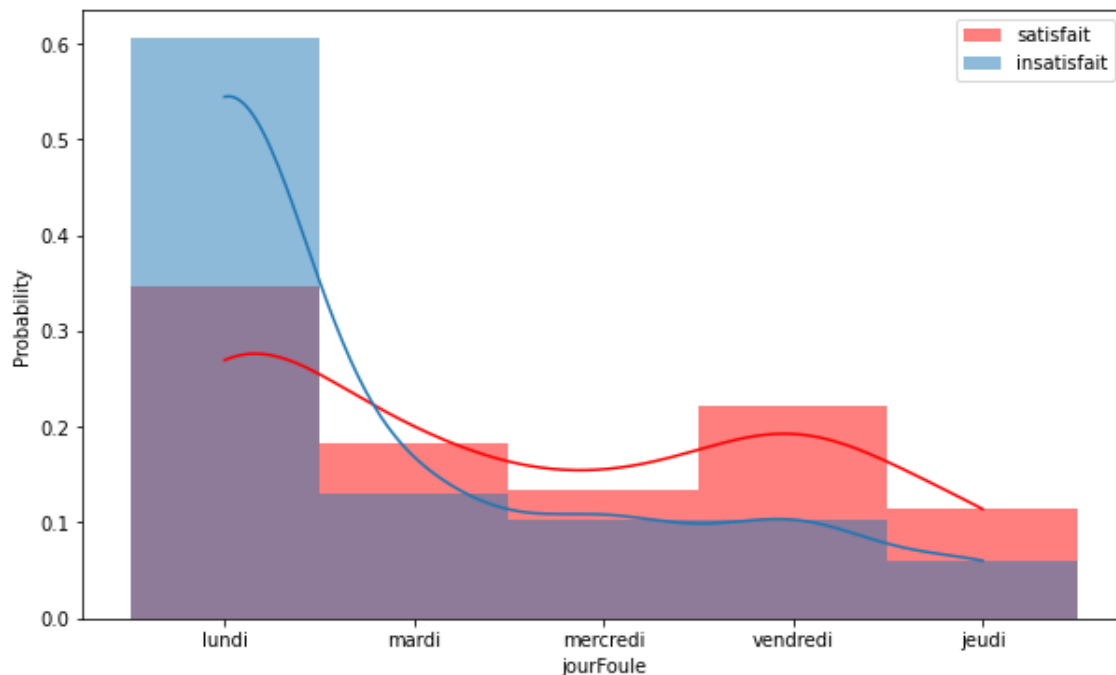


Figure 22 graphe de taux de satisfait en fonction de jour de visite

Comparaison entre deux arrondissement

On remarque déjà que on a un grand différent dans le taux de satisfait entre ZOUAGHA et JNAN EL WARD et pour ça on va prendre deux échantillonne le premier pour les habitants de l'arrondissement ZOUAGHA et l'autre pour les habitants de l'arrondissement JNAN EL WARD

❖ La relation entre la foule et l'arrondissement

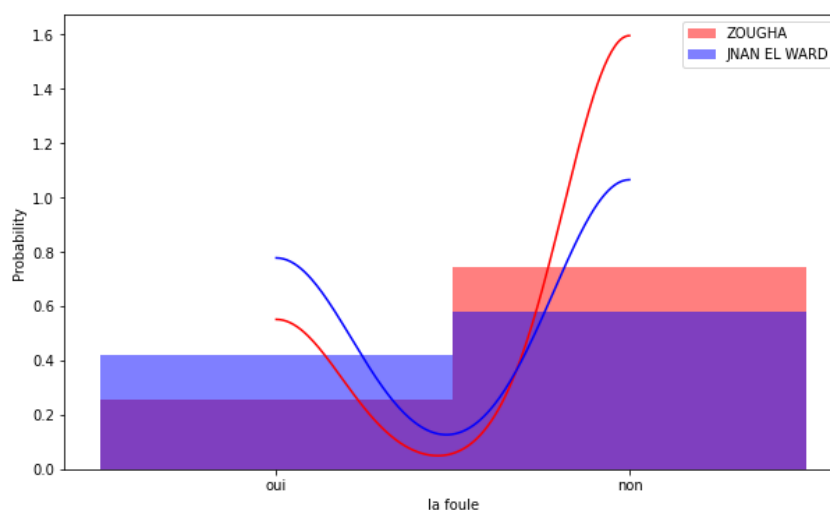


Figure 23 graphe de foule en fonction de arrondissement

On remarque que les habitants de JNAN EL WARD souffrent plus de la foule que les personne de ZOUAGHA

❖ La relation entre la foule et l'arrondissement

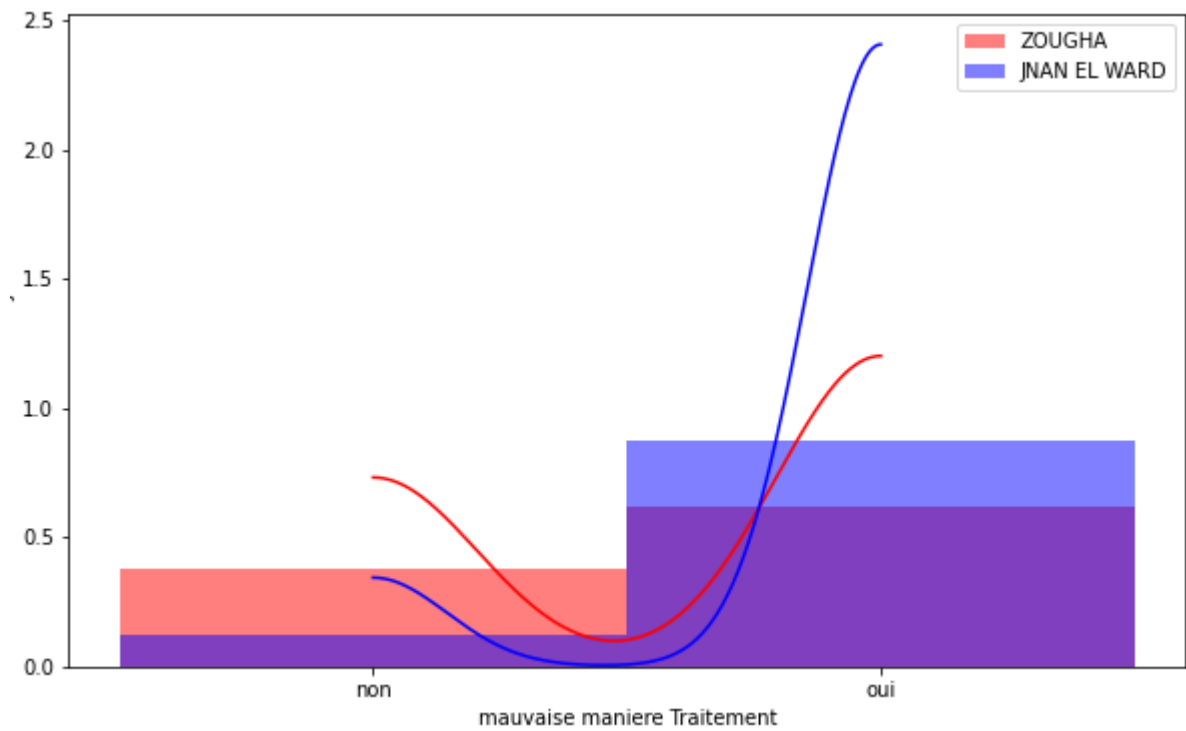


Figure 24 graphe de problème en fonction d'arrondissement

On remarque que les habitants de JNAN EL WARD souffrent plus de la mauvaise manière de traitement que les personnes de ZOUGHA

❖ La relation entre le jour foule et l'arrondissement

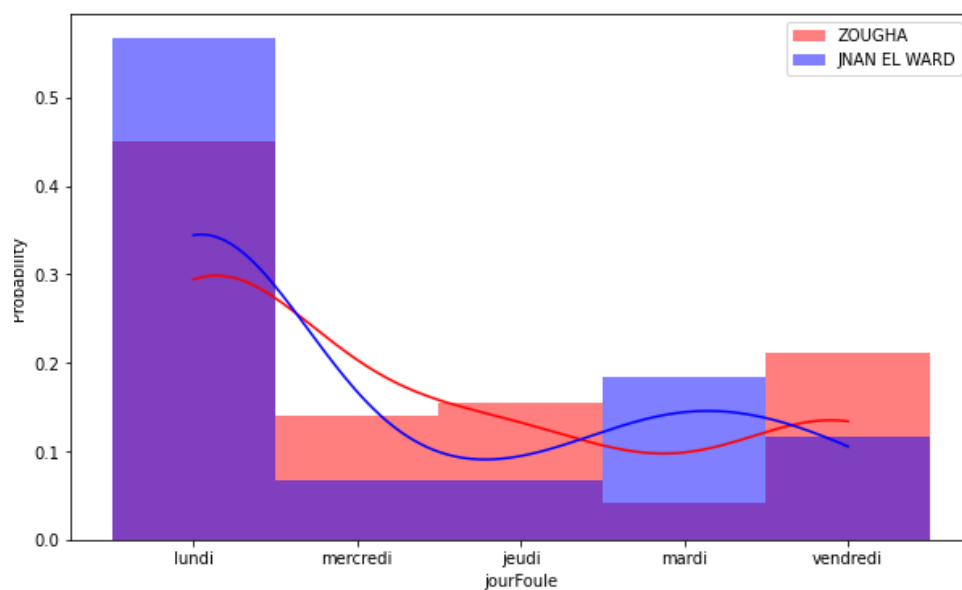


Figure 25 graphe des jours foule en fonction d'arrondissement

On remarque que les visites le lundi sont plus dans Jnan el Ward, ce qui est nous aide a connaitre la cause de la foule

Etude de la relation entre le taux de communication et l'age

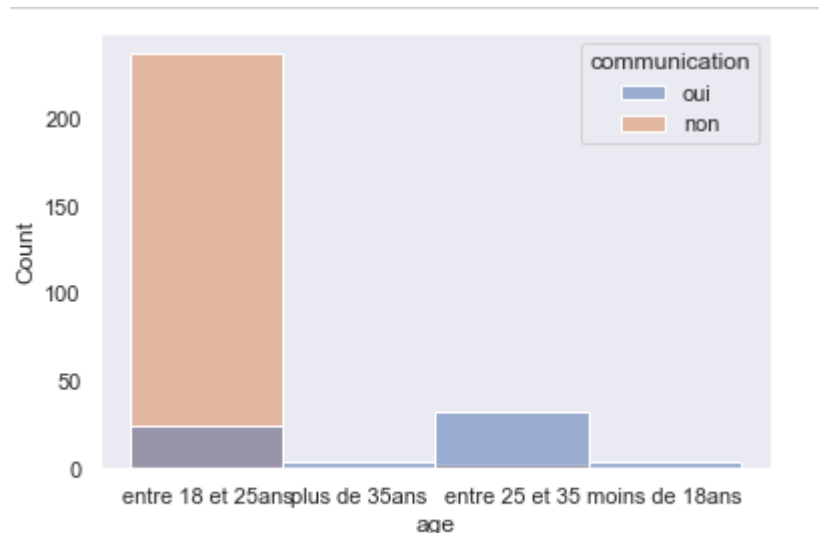


Figure 26 graphe de la relation entre le taux de communication et l'age

On remarque que la plupart des personnes qui ont des difficultés de communication sont dans la catégorie d'âge entre 18 et 25ans

Test statistique sur des hypothèse

Généralité sur les tests statistiques

Un test statistique est un mécanisme visant à trancher entre deux hypothèses a partir de résultats observés sur un ou plusieurs échantillon(s). On formule une hypothèse de départ, appelée hypothèse nulle et souvent notée (H_0) et il s'agit de décider si on rejette ou non cette hypothèse par opposition à une contre-hypothèse appelée hypothèse alternative et souvent notée (H_1).

On ne pourra jamais conclure avec certitude dans un test statistique. Il y aura toujours des erreurs de décision. Pour effectuer le test statistique, il faudra choisir un certain risque d'erreur qui est la probabilité de se tromper en prenant la décision retenue.

On donne ces hypothèses avec risque d'erreur de 5%

Test sur la relation entre le taux de satisfait et le moyen de visites par années

On pose :

- H0 : le moyen de visites des personnes a les bureau est le meme dans les deux échantillonne
- H1 : le moyen de visites des personnes a les bureau n'a pas le même dans les deux échantillonne

Statistiques sur échantillon unique

	N	Moyenne	Ecart-type	Erreur standard moyenne
visite_score	100	,59	,668	,067

Test sur échantillon unique

	Valeur du test = 1.03					
	t	ddl	Sig. (bilatérale)	Différence moyenne	Intervalle de confiance 95% de la différence	
					Inférieure	Supérieure
visite_score	-6,586	99	,000	-,440	-,57	-,31

Figure 27 resultat de test sur visite score

On a 0 est n'a pas dans l'intervalle de confiance 95% de la différence et sig = 0,0 < 0,05 , Donc on va accepter H1 et rejeté H0

Alors on conclut que les personnes qui visites plus les bureau de l'état civil sont plus satisfait des services

Test sur la relation entre le taux de satisfait et le moyen de score communication

On pose :

- H0 : le moyen de communication des personnes avec les bureau est le même dans les deux échantillonne
- H1 : le moyen de communication des personnes avec les bureau n'a pas le même dans les deux échantillonne

Statistiques sur échantillon unique

	N	Moyenne	Ecart-type	Erreur standard moyenne
communication	100	,21	,409	,041

Test sur échantillon unique

	Valeur du test = 0.39					
	t	ddl	Sig. (bilatérale)	Différence moyenne	Intervalle de confiance 95% de la différence	
					Inférieure	Supérieure
communication	-4,397	99	,000	-,180	-,26	-,10

Figure 28 résultat de test sur communication

On a 0 est n'a pas dans l'intervalle de confiance 95% de la différence et $\text{sig} = 0,0 < 0,05$, Donc on va accepter H1 et rejeté H0

Alors on conclut que les personnes qui ont un score de communication plus avec les bureau de l'état civil sont plus satisfait des services

Test sur la relation entre le taux de satisfait et la foule

On pose :

- H0 : le moyen de foule est le même dans les deux échantillonne
- H1 : le moyen de foule est n'a pas le même dans les deux échantillonne

Statistiques sur échantillon unique

	N	Moyenne	Ecart-type	Erreur standard moyenne
lafoule	100	,48	,502	,050

Test sur échantillon unique

	Valeur du test = 0.49					
	t	ddl	Sig. (bilatérale)	Différence moyenne	Intervalle de confiance 95% de la différence	
					Inférieure	Supérieure
lafoule	-,199	99	,843	-,010	-,11	,09

Figure 29 resultat de test sur la foule

On a 0 est dans l'intervalle de confiance 95% de la différence et $\text{sig} = 0,843 > 0,05$, Donc on va accepter H0 et rejeté H1

Alors on conclut que la foule n'a aucune effet sur le taux de satisfaction

Test sur la relation entre le taux de satisfait et la manière de traitement des problèmes

On pose :

- H0 : le moyen de la manière de traitement des problèmes est le même dans les deux échantillonne
- H1 : le moyen de la manière de traitement des problèmes est n'a pas le même dans les deux échantillonne

Statistiques sur échantillon unique

	N	Moyenne	Ecart-type	Erreur standard moyenne
mauvaisemaniereTraitement	100	,78	,416	,042

Test sur échantillon unique

	Valeur du test = 0.66					
	t	ddl	Sig. (bilatérale)	Différence moyenne	Intervalle de confiance 95% de la différence	
					Inférieure	Supérieure
mauvaisemaniereTraitement	2,882	99	,004	,120	,04	,20

Figure 30 resultat de test sur traitement des problemes

On a 0 est n'a pas dans l'intervalle de confiance 95% de la différence et $\text{sig} = 0,004 < 0,05$, Donc on va accepter H1 et rejeté H0

Alors on conclut que la manière de traitement fait un grand effet sur le taux de satisfaction

.

Analyse de données sur TANAGRA :

la méthode de Classification ascendante hiérarchique HAC

HAC (Hierarchical Agglomerative Clustering) est une méthode de clustering qui produit des groupes « naturels » d'exemples caractérisés par des attributs. Un arbre, appelé dendrogramme, où les agglomérations successives sont affichées, à partir d'un exemple par cluster, jusqu'à ce que l'ensemble des données appartienne à un cluster, décrit le processus de clustering. L'objectif : Montrer la complémentarité des méthodes de fouille de données (clustering) et de visualisation (analyse factorielle).

Les données a traité :

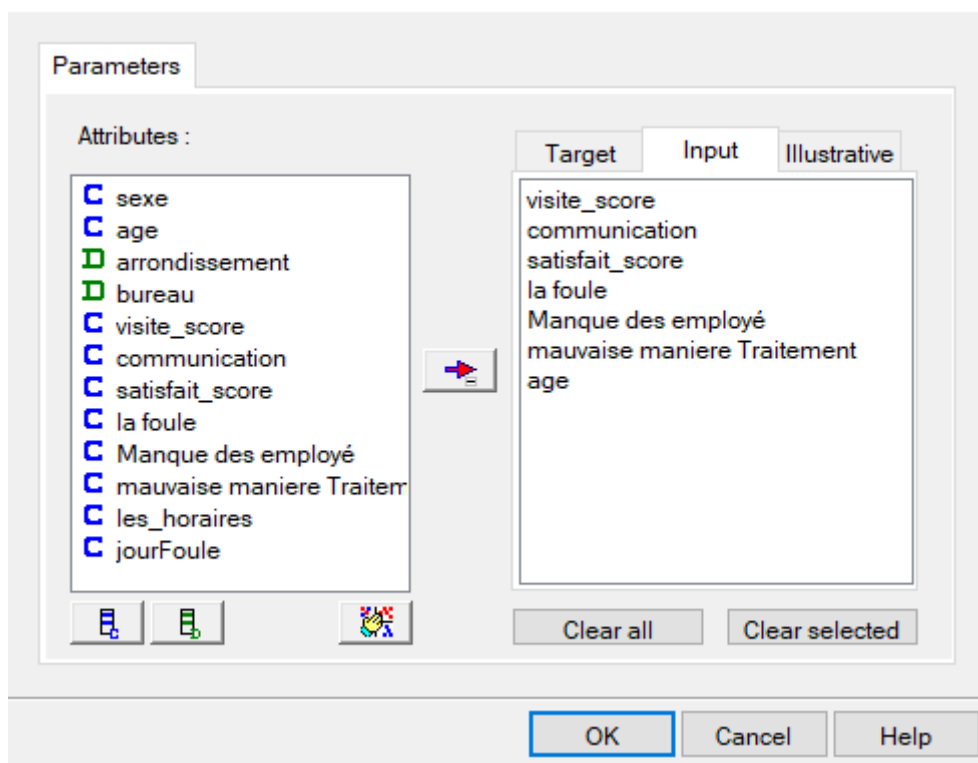
Nous avons travaillé sur 300 personnes qui sont déjà répond dans notre formulaire, et nous avons comme un but d'analyser les caractéristiques suivantes : Age, visite score, satisfait

score, communication , la foule, manière de traitement des problemes,et le manque des employés

Application de la méthode HAC

Nous choisissons comme input les variables quantitatives

Define attribute statuses



Après exécution (menu VIEW), les résultats s'affichent avec le dendrogramme est maintenant dessiné. La partition en trois classes semble être effectivement la plus évidente.

-Donc on déduit qu'il y a 3 classes pour cette classification, la première classe contient 36 individus, La deuxième 80 individus et la troisième 204 .

Clustering results

Clusters	From the dendrogram	After one-pass relocation
cluster n°1	54	54
cluster n°2	39	39
cluster n°3	54	54

Best cluster selection

Clusters	BSS ratio	Gap
1	0.0000	0.0000
2	0.2607	0.2740
3	0.4528	0.1390
4	0.6103	0.1179
5	0.7382	0.1406
6	0.8310	0.0894
7	0.9014	0.1346
8	0.9382	0.0704
9	0.9574	0.0400
10	0.9666	0.0088

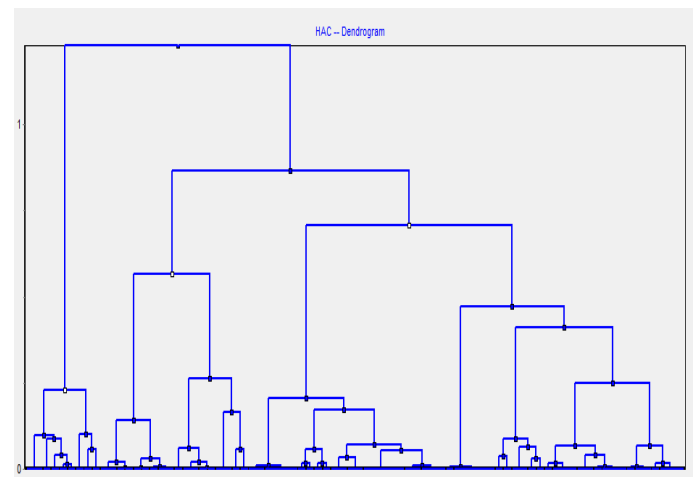


Figure 32 result de hac

Figure 31le dendrogramme HAC

Caractérisation des groupes :

L'étape suivante consiste à caractériser ces groupes. L'outil GROUP CHARACTERIZATION est le plus approprié pour cela, il permet de comparer les indicateurs (moyenne ou proportion) marginaux et conditionnellement aux groupes. Pour ce faire, nous insérons dans un premier temps un nouveau composant DEFINE STATUS. Nous plaçons en TARGET, la classification produite par la CAH ; en INPUT, les autres variables à l'exception de la variable CAR qui caractérise les touristes

Results											
Description of "Cluster_HAC_1"											
Cluster_HAC_1=c_hac_1				Cluster_HAC_1=c_hac_2				Cluster_HAC_1=c_hac_3			
Examples		[36.7 %] 54		Examples		[26.5 %] 39		Examples		[36.7 %] 54	
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy			
visite_score=rarement	3.64	[52.2 %] 66.7 %	46.9 %	age=moins de 18ans	2.90	[100.0 %] 7.7 %	2.0 %	visite_score=des fois par année	2.12	[47.4 %] 50.0 %	38.8 %
age=entre 25 et 35	2.85	[62.5 %] 27.8 %	16.3 %	visite_score=plusieurs fois par année	1.82	[42.9 %] 23.1 %	14.3 %	age=entre 18 et 25ans	1.73	[40.0 %] 88.9 %	81.6 %
sexe=Homme	0.65	[38.7 %] 66.7 %	63.3 %	sexe=Homme	0.90	[29.0 %] 69.2 %	63.3 %	sexe=Femme	1.47	[44.4 %] 44.4 %	36.7 %
age=plus de 35ans	0.00	[0.0 %] 0.0 %	0.0 %	age=entre 18 et 25ans	0.56	[27.5 %] 84.6 %	81.6 %	visite_score=plusieurs fois par année	0.63	[42.9 %] 16.7 %	14.3 %
sexe=Femme	-0.65	[33.3 %] 33.3 %	36.7 %	age=plus de 35ans	0.00	[0.0 %] 0.0 %	0.0 %	age=plus de 35ans	0.00	[0.0 %] 0.0 %	0.0 %
age=moins de 18ans	-1.33	[0.0 %] 0.0 %	2.0 %	visite_score=des fois par année	-0.05	[26.3 %] 38.5 %	38.8 %	age=entre 25 et 35	-1.30	[25.0 %] 11.1 %	16.3 %
visite_score=des fois par année	-2.08	[26.3 %] 27.8 %	38.8 %	sexe=Femme	-0.90	[22.2 %] 30.8 %	36.7 %	age=moins de 18ans	-1.33	[0.0 %] 0.0 %	2.0 %
age=entre 18 et 25ans	-2.24	[32.5 %] 72.2 %	81.6 %	visite_score=rarement	-1.23	[21.7 %] 38.5 %	46.9 %	sexe=Homme	-1.47	[32.3 %] 55.6 %	63.3 %
visite_score=plusieurs fois par année	-2.30	[14.3 %] 5.6 %	14.3 %	age=entre 25 et 35	-1.70	[12.5 %] 7.7 %	16.3 %	visite_score=rarement	-2.51	[26.1 %] 33.3 %	46.9 %

Figure 33 : les groupes de caractérisation

Le premier groupe (C_HAC_1) est composé de 52% des individus visite les bureaux de la commune rarement de niveau d'existence de 66%, et avec 62% des individus qui ont un âge entre 25 et 35ans ainsi que la majorité de sexe homme au niveau de 38% dans le groupe.

Le deuxième groupe (C_HAC_2) est composé de 42% des individus visite les bureaux de la commune plusieurs fois par années de niveau d'existence de 23%, et avec 100% des individus qui ont un âge moins de 18ans ainsi que la majorité de sexe homme au niveau de 29% dans le groupe.

Le troisième groupe (C_HAC_3) est composé de 7% des individus visite les bureaux de la commune des fois par années de niveau d'existence de 50%, et avec 40% des individus qui ont un âge entre 18 et 25ans ainsi que la majorité de sexe femme au niveau de 29% dans le groupe.

Remarques et conclusions observées

Selon les analyses les graphes les comparaison et les test statistique que on a fait on a des conclusion et des remarque sur plusieurs chose :

Remarque sur les visiteurs des bureaux communautaires :

- ❖ Selon les répons on a l'arrondissement qui est avec plus visites est :ZOUAGHA
- ❖ La plupart des personnes visites les bureaux communautaires rarement avec percentage de 43%
- ❖ La plupart des personnes visites les bureaux communautaires dit qu'il Ya un problème de communication avec percentage de 79%
- ❖ Plus que 70% des visites des bureaux communautaires sont entre 9 et 11 de matin , et 51% des visites sont au lundi
- ❖ Pour les problèmes on a « mauvaise manière de traitement des problèmes » est le problèmes le plus courants, suivi par le manque des employés

Remarque sur les élément qui influence le taux de satisfait :

- ❖ On remarque que le taux de satisfait est déférente entre les arrondissement et les bureaux
- ❖ On remarque que les personnes qui visites rarement les bureaux communautaires sont les plus insatisfait, et les personnes qui visites plusieurs fois les bureaux communautaires sont les plus satisfait
- ❖ On a un relation avec les visiteurs au lundi et l'insatisfait, et aussi avec la foule
- ❖ D'après les test statistique que on a fait, on peut dire que le moyen de visites par années , score communication , la manière de traitement des problèmes sont impact le taux de satisfait

Autres remarques :

- ❖ La plupart des personnes qui ont des difficultés de communication sont dans la catégorie d'âge entre 18 et 25ans
- ❖ La foule et la visites des bureaux communautaire au lundi sont

CHAPITRE 4 : La réalisation « Approche Développement Web »

Introduction

Pour présenter les résultats obtenues, nous avons utilisé **python** avec **streamlit** html et un peu de CSS , qui nous permet de créer un dashboard qui facilite la lecture de données et représenter les resultats d'une manière simple et compréhensible.

Le Dashboard que nous avons construit contient un navbar et un Dashboard, le navBar pour choisé la page, et les critique de filtrage

Dans ce qui suit, des captures d'écran présentant notre application web Dans la page d'accueil on a un Dashboard responsive sur la population

Présentation de l'application

Premier page

La première page affiche le résumé de tout ce qui concerne l'étude :

- ❖ moyenne des notes
- ❖ nombre des personnes répons au sondage
- ❖ les visiteurs de les bureau de la commune selon: le sexe, l'Age, arrondissement, bureau de l'état civil, les problemes...

❖ satisfait score selon le sexe, l'Age, arrondissement, bureau de l'état civil...

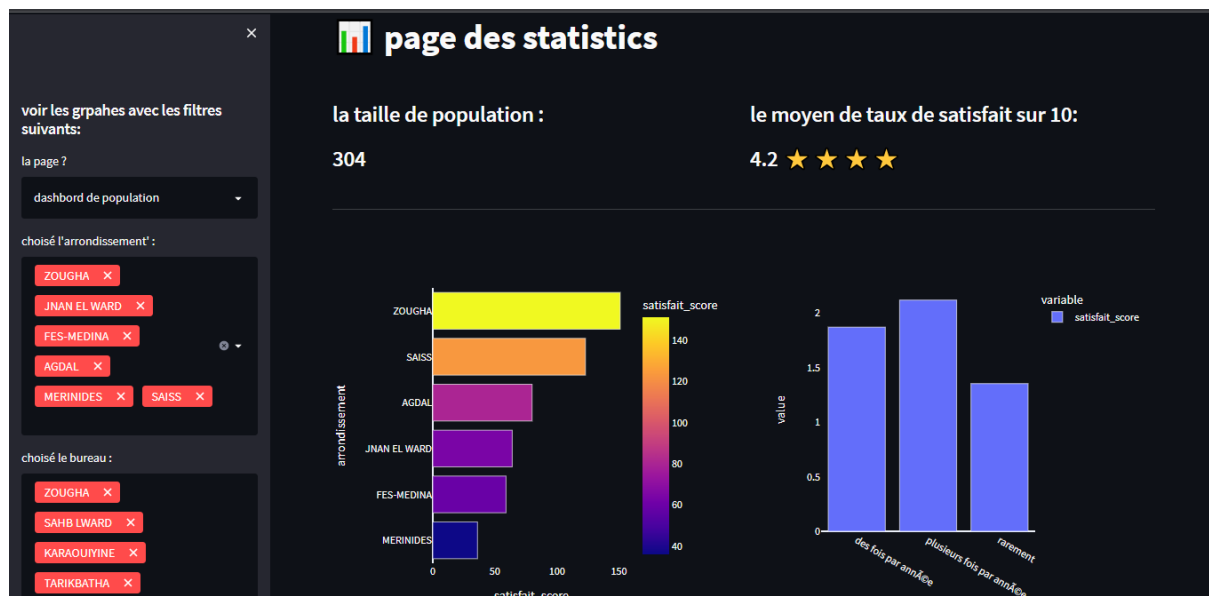


Figure 34 page d'accueil

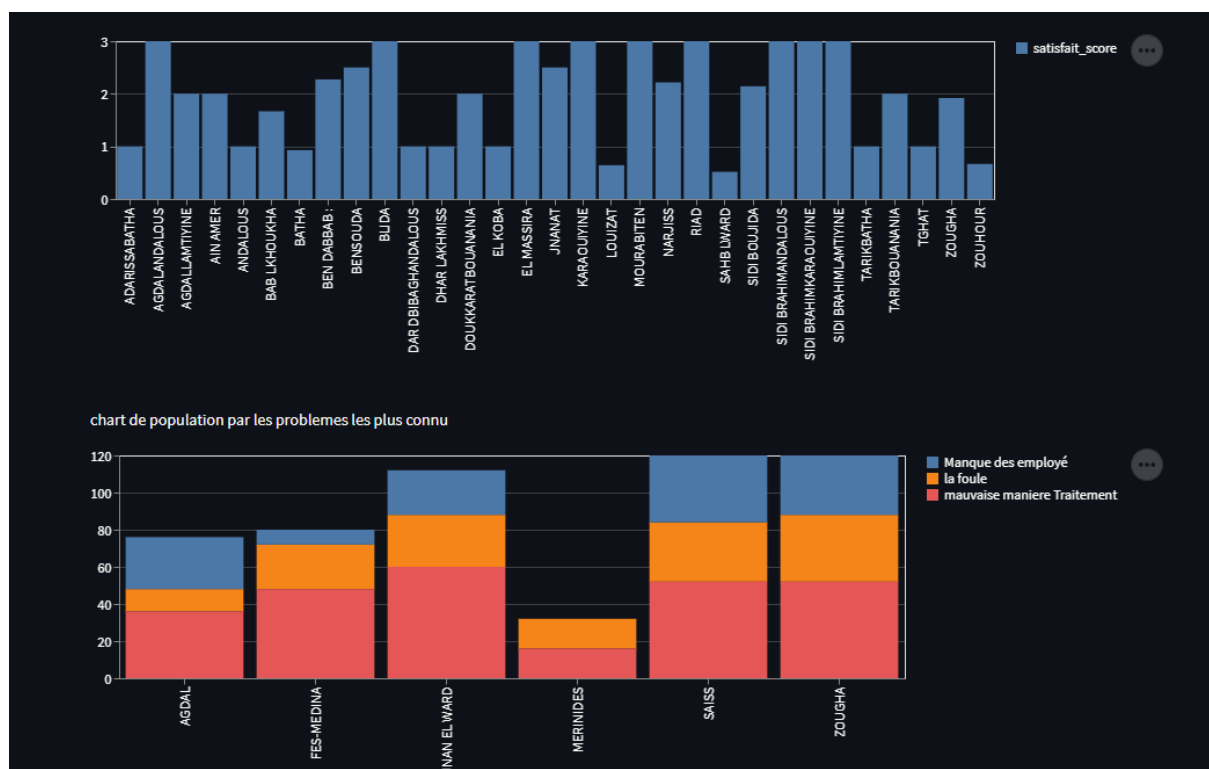


Figure 35 extrait des graphes de premier page

La deuxième page

La deuxième page affiche les différents graphes pour le taux de satisfaction en fonction de les autres variables age,taux visites, arrondissement...



Figure 36 extrait des graphes de deuxième page

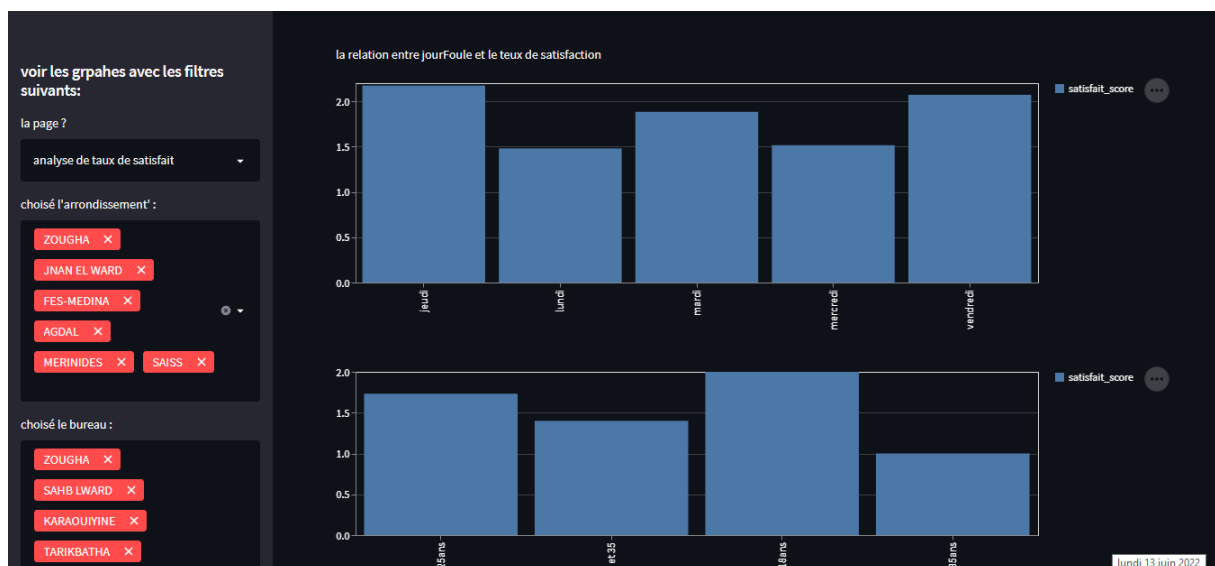
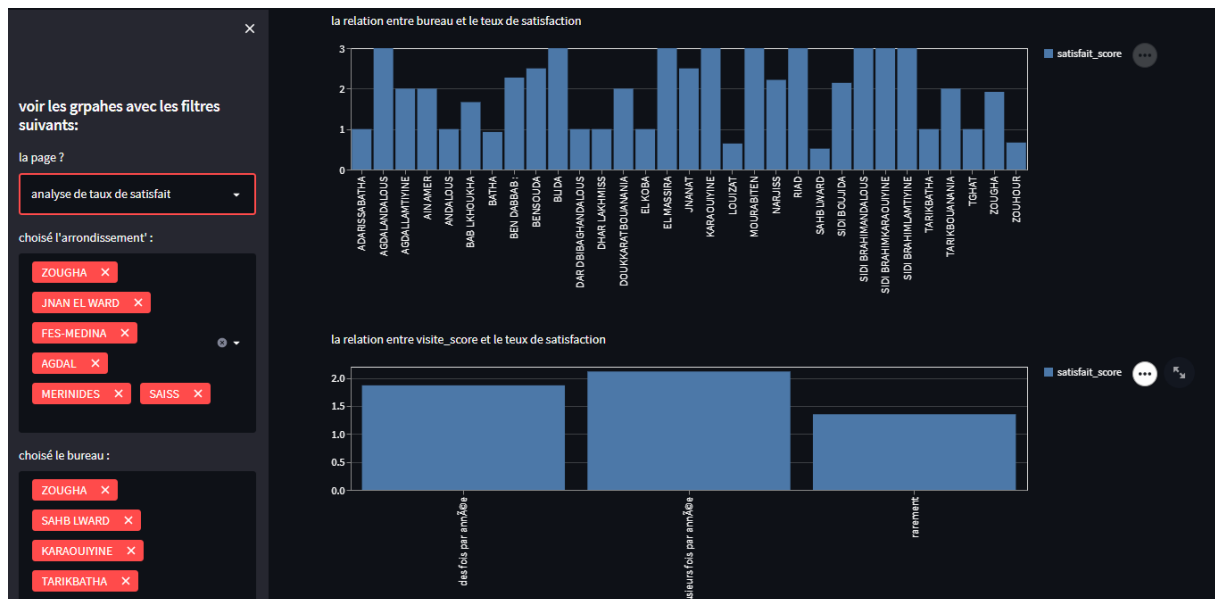


Figure 37 des autres graphe de deuxième page

Troisième page

Dans cette page on a des graphes sous forme des images, mais aussi relative en fonction de **navbar**, cet page pour la comparaison entre les arrondissement on fonction des arrondissement que on a choisi dans le **navbar**

Conclusion

Pour conclure Cette expérience m'a permis de répondre aux questionnements que j'avais en ce qui concerne comment transforme des données à un Dashboard active avec des nouvelles techniques par les entreprises pour s'adapter au changement de leur environnement.

période que j'ai passé au sein de la Commune de Fès m'était très bénéfique et m'a permis d'approfondir les notions théoriques que j'ai apprises durant ma période de formation à l'école supérieure du technologie de Fès, et d'enrichir mes connaissances aussi bien au niveau pratique qu'au niveau théorique. J'ai appris que le domaine théorique et le domaine pratique sont complémentaires, et que dans le monde de travail il n'est pas suffisant d'avoir des bases théoriques, mais il faut surtout être sociable et coopératif et avoir un esprit d'équipe pour mener à bien toutes les tâches. On peut conclure que le stage est une occasion très avantageuse pour le stagiaire, malgré l'insuffisance de sa durée.

Bibliographie

- [1] : <https://www.python.org/docs>
- [2] : <https://pandas.pydata.org/docs/>
- [3] : <https://seaborn.pydata.org/>
- [4] : <https://docs.streamlit.io/>
- [5] : <https://www.google.com/intl/fr/forms/about/>
- [6] : <https://console.cloud.google.com/apis>
- [7] : <https://sites.google.com/site/redouanequesmi/teaching>
- [8] : <https://developer.mozilla.org/en-US/docs/Web/HTML/Element>
- [9] : <https://help.talend.com/r/fr-FR/7.3/release-notes-data-integration-products/documentation>
- [10] : <https://sites.google.com/site/redouanequesmi/teaching>
- [11] : <https://developer.mozilla.org/en-US/docs/Web/HTML/Element>
- [12] : <https://help.talend.com/r/fr-FR/7.3/release-notes-data-integration-products/documentation>
- [13] : <https://www.ibm.com/support/pages/ibm-spss-statistics-27-documentation>
- [14] : <https://numpy.org/doc/>
- [15] : <https://www.anaconda.com/open-source>