



Master 1 IDSM

Visualisation des données

TD5 – Bibliothèque matplotlib

Objectif : Dans ce TD, nous allons tracer des graphiques et améliorer des visualisations de données en utilisant la bibliothèque matplotlib (Python).

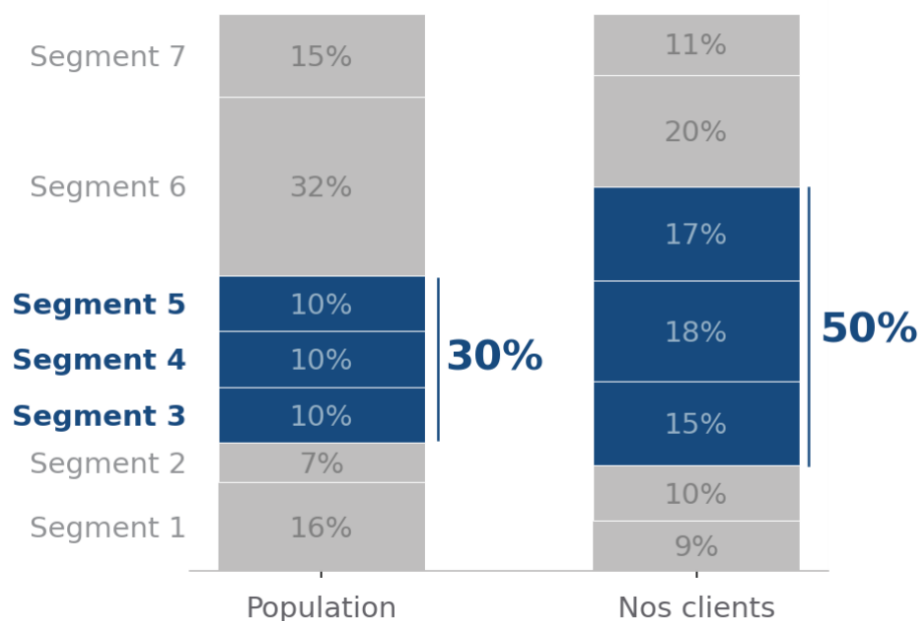
Exercice 1 : Diagramme en bâton

Vous travaillez pour un gros commerçant. Vous avez les données de la répartition de la population dans son ensemble et de « vos clients » pour sept segments de clientèle (par classe d'âge).

	Population	Nos clients
Segment 1	16%	9%
Segment 2	7%	10%
Segment 3	10%	15%
Segment 4	10%	18%
Segment 5	10%	17%
Segment 6	32%	20%
Segment 7	15%	11%

En utilisant la bibliothèque matplotlib, tracer le graphique ci-dessous :

Distribution par segment de clientèle



Exercice 2 : Nuage de points

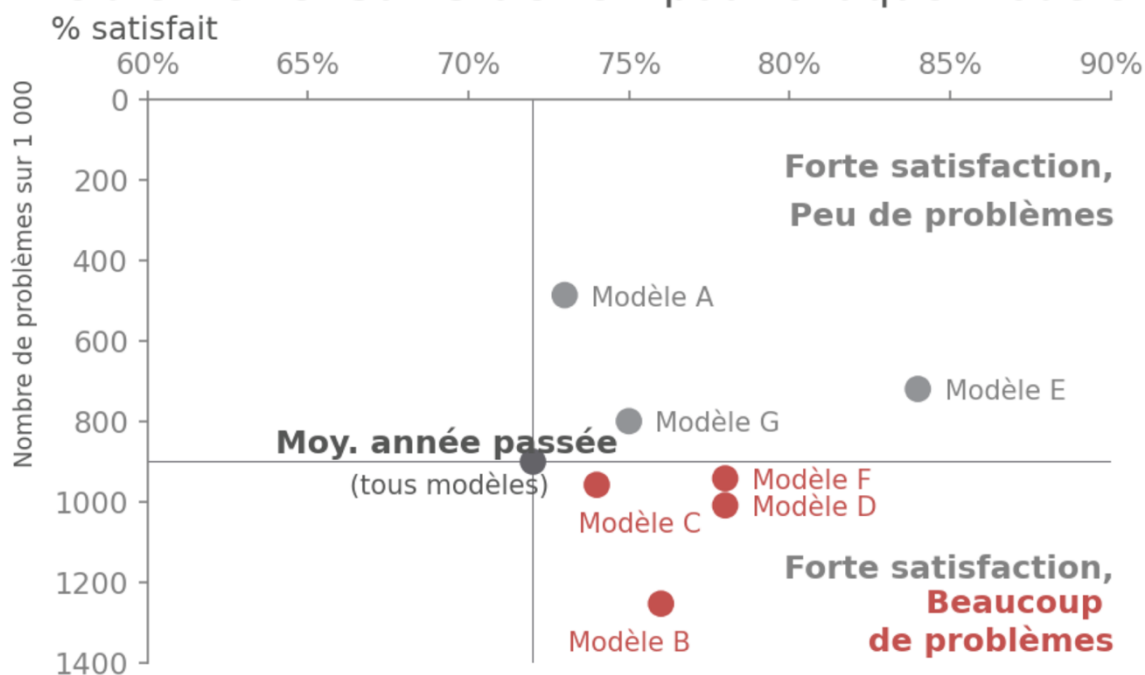
Scenario :

Vous êtes un entrepreneur dans le secteur de l'automobile. Pour estimer le succès d'un modèle de véhicule, vous utilisez principalement deux critères : 1) la satisfaction du client, 2) la fréquence des problèmes rencontrés.

	Satisfaction	Nombre de problèmes sur 1000
Moyenne de l'année passée	72%	895
Modèle A	73%	487
Modèle B	76%	1253
Modèle C	74%	958
Modèle D	78%	1009
Modèle E	84%	720
Modèle F	78%	942
Modèle G	75%	800

Utilisez les données présentées dans le tableau pour faire le graphique ci-dessous :

Problème vs. Satisfaction pour chaque modèle

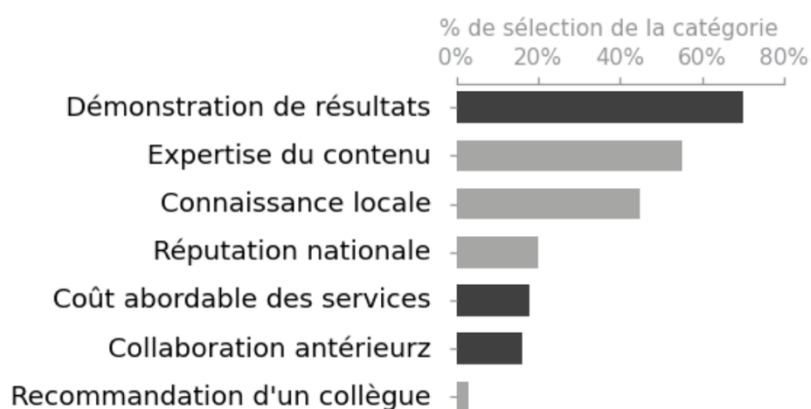


Exercice 3 : Diagramme horizontal

Vous devez analyser les résultats d'une enquête sur les critères utilisés par les associations pour choisir un prestataire de services. Ces données se trouvent dans un tableau ci-dessous.

Critère	% de sélection de la catégorie
Recommandation d'un collègue	3
Expertise du contenu	16
Connaissance locale	18
Réputation nationale	20
Coût abordable des services	45
Collaboration antérieure	55
Recommandation d'un collègue	70

En utilisant la bibliothèque matplotlib, tracer le graphique suivant :



Exercice 4 : Diagramme horizontal empilé

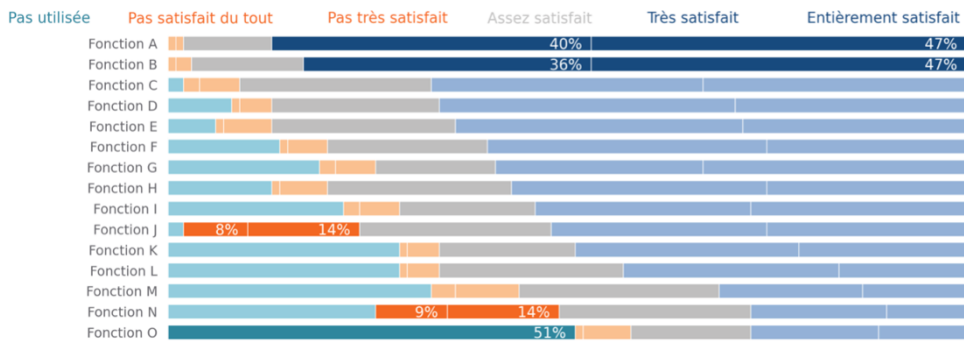
Vous travaillez dans une entreprise qui vend un produit ayant plusieurs fonctions. Vous avez réalisé un sondage auprès des utilisateurs pour savoir s'ils s'en servent et s'ils en sont satisfaits. Vous voulez maintenant explorer les résultats.

Fonction	Pas utilisé	Pas satisfait du tout	Pas très satisfait	Assez satisfait	Très satisfait	Entièrement satisfait
A	0	1	1	11	40	47
B	0	1	2	14	36	47
C	2	2	5	24	34	33
D	8	1	4	21	37	29
E	6	1	6	23	36	28
F	14	1	5	20	35	25
G	19	2	5	15	26	33
H	13	1	6	23	32	25
I	22	2	5	17	27	27
J	2	8	14	24	27	25
K	29	1	4	17	28	21

L	29	1	4	23	27	16
M	33	3	8	25	18	13
N	26	9	14	24	17	10
O	51	1	6	15	16	11

Proposez une visualisation de données qui permet ressortir des chiffres les plus intéressants de ces données en utilisant le diagramme horizontal empilé.
 Par exemple, il est possible de faire ressortir les trois points les plus important en utilisant ce graphique (si vous devez présenter votre analyse dans un document écrit).

Satisfaction de l'utilisateur pour le produit X : **Fonctions**



Les fonctions A et B continuent à obtenir la meilleure satisfaction.

Les fonctions J et N sont les moins satisfaisantes pour les utilisateurs. Quelle amélioration réaliser pour y répondre ?

La fonction O est la moins utilisée. Quelle initiative prendre pour augmenter son utilisation par les utilisateurs actuels ?