

[Retour à mon espace](#)

Segmentez des clients d'un site e-commerce

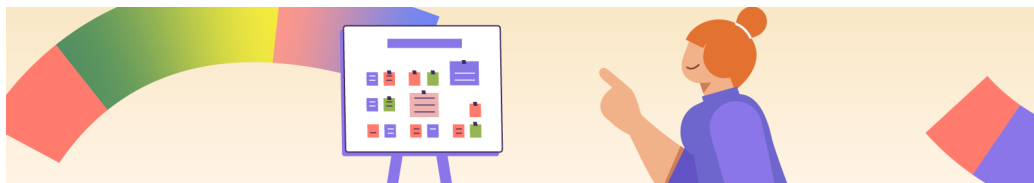
CONTENU

- 📄 Mission - Segmentez...
- 📄 Livrables et soutena...
- 📊 Évaluation

SUPPORTS PÉDAGOGIQUES

- 👤 Companion **Nouveau**
- 📖 Cours
- 📖 Ressources

Livrables et soutenance



Livrables

- Un **script** avec l'ensemble des requêtes SQL demandées par Fernanda
- Un **notebook de l'analyse exploratoire** (non cleané, pour comprendre votre démarche)
- Un **notebook** (ou code commenté au choix) **d'essais** des différentes approches de modélisation (non cleané, pour comprendre votre démarche)
- Un **notebook de simulation pour déterminer la fréquence nécessaire de mise à jour** du modèle de segmentation
- Un **support de présentation** pour présenter votre travail à un collègue

Pour faciliter votre passage devant le jury, déposez sur la plateforme, dans un dossier zip nommé **Titre_du_projet_nom_prénom**, vos livrables nommés comme suit : **Nom_Prénom_n° du livrable_nom du livrable_date de démarrage du projet**. Cela donnera :

- *Nom_Prénom_1_script_mmaaaa*
- *Nom_Prénom_2_notebook_exploration_mmaaaa*
- *Nom_Prénom_3_notebook_essais_mmaaa*
- *Nom_Prénom_4_notebook_simulation_mmaaaa*
- *Nom_Prénom_5_presentation_mmaaaa*

Par exemple, votre premier livrable peut être nommé comme suit :
Dupont_Jean_1_script_012022.

Soutenance

Pendant la soutenance, l'évaluateur jouera le rôle d'un collègue à qui vous présentez votre travail, avant de le présenter à Olist. Vous lui présenterez l'ensemble de votre travail.

- **Présentation (20 minutes)**
 - Présentation de la problématique, du cleaning effectué, du feature engineering et de l'exploration – 5 min
 - Présentation des différentes pistes de modélisation effectuées et du modèle final sélectionné – 10 min



[Retour à mon espace](#)

Segmentez des clients d'un site e-commerce

CONTENU

- Mission - Segmentez...
- Livrables et soutena...
- Évaluation

SUPPORTS PÉDAGOGIQUES

- Companion **Nouveau**
- Cours
- Ressources

- **Discussion (5 minutes)**

- L'évaluateur vous challengera sur vos choix.

- **Débriefing (5 minutes)**

- À la fin de la soutenance, vous pourrez débriefer ensemble.

Votre présentation devrait durer 20 minutes (+/- 5 minutes). Puisque le respect des durées des présentations est important en milieu professionnel, les présentations en dessous de 15 minutes ou au-dessus de 25 minutes peuvent être refusées.

Référentiel d'évaluation

Définir la stratégie d'élaboration d'un modèle d'apprentissage non supervisé, sélectionner et entraîner des modèles adaptés à une problématique métier afin de réaliser une segmentation ou une réduction de données.

- CE1 Les variables catégorielles identifiées ont été transformées en fonction du besoin (par exemple via OneHotEncoder ou TargetEncoder).
- CE2 Vous avez créé de nouvelles variables à partir de variables existantes.
- CE3 Vous avez réalisé des transformations mathématiques lorsque c'est requis pour transformer les distributions de variables.
- CE4 Vous avez normalisé les variables lorsque c'est requis.
- CE5 Vous avez défini la stratégie d'élaboration d'un modèle pour répondre à un besoin métier.
- CE6 Vous avez proposé un nombre de segments et une répartition adaptés à la problématique métier.
- CE7 Vous avez explicité la stratégie d'ajout de nouveaux clients.
- CE8 Vous avez pris en compte la nature des variables d'entrée dans le choix des algorithmes et de la distance.
- CE9 Vous avez testé et comparé plusieurs algorithmes.

Évaluer les performances des modèles d'apprentissage non supervisé selon différents critères (scores, temps d'entraînement, etc.) en adaptant les paramètres afin de choisir le modèle le plus performant pour la problématique métier.

- CE1 Vous avez choisi des métriques adaptées pour évaluer les valeurs candidates pour le nombre de segments (selon algorithme, par exemple : coefficient silhouette / elbow pour le KMeans, eps/ min samples pour le DBScan).
- CE2 Vous avez évalué la forme des clusters.
- CE3 Vous avez évalué la stabilité des clusters à l'initialisation.
- CE4 Vous avez optimisé les hyper-paramètres pertinents dans les différents algorithmes.
- CE5 Vous avez justifié le choix final d'algorithme et d'hyper-paramètres.
- CE6 Vous avez analysé la stabilité du modèle dans le temps afin d'en assurer sa maintenance
- CE7 La convention PEP8 est respectée. Le code est commenté (commentaires réguliers, docstrings dans les fonctions) .



[Retour à mon espace](#)

Segmentez des clients d'un site e-commerce

CONTENU

 Mission - Segmentez...

 Livrables et soutena...

 Évaluation

SUPPORTS PÉDAGOGIQUES

 Companion **Nouveau**

 Cours

 Ressources

