

Projet TER par l'équipe ARS

TER ARTIST-RUN-SPACES

KARAMI Aya

JEBALI Anas

GUYON NOAH

VITOFFODJI Adjimon

EL OUALYDY Mohamed-Amine

Mars , 2025

Choix des visualisations

Dans le cadre de notre projet **TER ARTIST-RUN-SPACES**, nous avons pris soin de sélectionner et de créer des visualisations qui non seulement illustrent nos résultats, mais aussi permettent une exploration approfondie des données. Voici ce que nous avons réalisé et ce que nous comptons faire pour notre rapport final en termes de visualisations.

1. Objectif des visualisations

Les visualisations ont été conçues pour nous aider à explorer les données brutes des espaces artistiques, tout en facilitant la communication de notre message. Elles sont essentielles pour :

- Identifier des tendances importantes dans les données (par exemple, la répartition géographique des espaces, les services proposés, etc.).
- Rendre les résultats de notre modèle **BERTopic** plus accessibles, notamment pour visualiser la distribution des topics.
- Illustrer de manière claire et précise les différents aspects de notre analyse.

2. Visualisations réalisées et à venir

a. Carte de distance entre les topics

Nous avons utilisé le fichier [intertopic_distance_map.html](#), qui présente une carte interactive des relations entre les topics générés par notre modèle **BERTopic**. Cette carte nous permet de mieux comprendre les liens sémantiques entre les différents sujets abordés dans les textes. Nous prévoyons d'intégrer cette carte dans notre rapport final, car elle est particulièrement utile pour visualiser les clusters de topics et comprendre comment les différents sujets sont reliés.

b. Distribution des thèmes

Le graphique [distribution_themes.png](#) montre la répartition des différents thèmes extraits du corpus, ce qui nous aide à identifier les sujets dominants. Pour le rapport final, nous comptons utiliser ce graphique sous forme de **barplot** ou de **diagramme circulaire**, afin de présenter clairement la fréquence de chaque thème. Cela permet à notre auditoire de voir rapidement quels sont les sujets les plus abordés.

c. Distribution des topics sur une échelle interactive

Le fichier [topic_distribution.html](#) permet d'explorer de manière interactive la répartition des documents entre les topics. Ce graphique nous aide à visualiser comment les documents sont regroupés et si certains topics sont sous-représentés. Nous envisageons d'inclure cette visualisation dans notre rapport final, car elle fournit un aperçu complet de la couverture des topics dans le corpus.

d. Visualisation géographique des espaces d'art

Nous avons plusieurs fichiers de données ([spacesnew.csv](#), [spaces.csv](#), [fichier_a_jour.csv](#)) qui contiennent des informations géographiques sur les espaces d'art. Pour le rapport final, nous prévoyons de créer une **carte géographique** qui montre où se situent ces espaces, afin d'illustrer les concentrations géographiques et les tendances régionales. Cette carte sera un moyen puissant d'illustrer la répartition des espaces artistiques à travers différentes villes et pays.

e. Visualisation des services offerts par les espaces d'art

Nous avons également des données sur les services proposés par chaque espace d'art (par exemple, résidences, expositions, ateliers). Pour le rapport, nous comptons créer des graphiques à **barres** ou en **secteurs**, qui montreront la fréquence de chaque service dans les différents espaces. Cela permettra à notre auditoire de comprendre facilement les types de services les plus fréquents dans ces espaces artistiques.

3. Règles pour les visualisations du rapport final

Afin de rendre nos visualisations aussi claires et efficaces que possible, nous avons défini plusieurs règles à suivre pour notre rapport final :

- **Choisir le graphique adapté à la variable** : Pour chaque type de données, nous choisirons le graphique le plus adapté. Par exemple, pour les variables qualitatives (comme les services offerts), nous utiliserons des **diagrammes à barres** ou en **secteurs**. Pour les variables quantitatives, nous opterons pour des **histogrammes** ou des **boxplots**.
- **Prendre en compte l'impression en noir et blanc** : Nous veillerons à ce que nos graphiques restent lisibles même en version noir et blanc. Cela implique de choisir des **couleurs contrastées** et d'éviter les éléments graphiques trop complexes.
- **Clarté des étiquettes et des axes** : Chaque graphique aura des **titres clairs**, des **étiquettes d'axes** et, si nécessaire, des **légendes**. Cela permettra à notre auditoire de comprendre immédiatement ce que chaque graphique représente.
- **Précision des chiffres** : Nous veillerons à **aligner les points décimaux** dans une même colonne et à **éviter les chiffres non significatifs**. Cela contribuera à rendre les

visualisations plus professionnelles et faciles à lire.

- **Limiter les messages** : Nous nous assurerons de ne pas multiplier les messages sur une même figure. Chaque graphique doit avoir un objectif précis et ne doit pas être surchargé d'informations.
- **Référencer toutes les figures et tableaux** : Chaque figure et tableau dans le rapport final sera référencé dans le texte (par exemple, **Figure 2.3** et **Table 3.1**), pour assurer une bonne lisibilité et cohérence entre les analyses et les visualisations.

