

MASTER 2 MIAGE

UX Données Massive

Documentation Projet Visualisation

Réalisé par:

EL BELGHITI Hamza, ESSALEHI Oussama

NAOUI Ilias, BARRY Bailo

Année Universitaire: 2024-2025

Utilisateurs Cibles:

- Chercheurs d'emplois (jeunes diplômés, personnes en reconversion etc.)
- Les RHs et les recruteurs
- Les institutions éducatives (universités, écoles d'ingénieurs, formations online, etc...)

Liste des tâches :

Chercheurs d'emploi	RH/Recruteurs	Inst. Éducatives	Présence des données
Identifier les technologies les plus utilisées dans le domaine de l'ingénierie data.		Identifier les technologies les plus utilisées dans le domaine de l'ingénierie data.	OUI
Comparer la différence des salaires selon la taille et la région de l'entreprise			OUI
Comparer les États afin de déterminer ceux qui ont plus de demande en ingénieurs data		Comparer les États afin de déterminer ceux qui ont plus de demande en ingénieurs data	OUI
Distribution des salaires par secteur, industrie, taille d'entreprise ou location			OUI
Cartographie des offres d'emploi par états			PARTIEL (manque coordonnées géographiques des états)
Classer les États par volume d'offres ou salaire moyen			OUI
		Identifier spécialités les plus demandées dans le métier de la data	OUI
Corréler le salaire avec X param de l'entreprise (corrélation entre la note de l'entreprise et le salaire)			OUI

Transformations:

Colonne	Définition	Transformation	Résultat
company	Nom de l'entreprise ayant poster l'offre d'emploi	Enlever le rating du nom de l'entreprise	
company_rating	Note de l'entreprise sur Glassdoor	Remplacer les valeurs nulles par la moyenne (3.9)	
location	Localisation de l'offre d'emploi Format: Ville, État ou Remote	Séparer la ville et l'état + ajouter les coordonnées (lat, long)	- colonne job_city - colonne job_state - colonne coord_city - colonne coord_state
job_title	Le nom du poste	1. Simplifier le titre (data engineer, data scientist, data analyst, mle, na) 2. Extraire le nom du métier sans la séniorité	- colonne job_title - colonne job_seniority
job_description	La description du poste	Extraire les technologies demandés, le niveau d'études et les années d'expériences et en déduire la séniorité	- colonne job_technologies - colonne job_education - colonne job_seniority - colonne job_experience
salary_estimate	Salaire estimé (format: \$X/hr ou \$X/yr)	1. Extraire le montant du salaire 2. Convertir les salaires horaires en annuel (×1800) 3. Remplacer les nuls par la moyenne 4. Arrondir à l'entier	
company_founded	Année de fondation de l'entreprise	1. Remplacer les nuls par -1 2. Calculer l'âge de l'entreprise (année actuelle - année de fondation)	

+ Fusion avec une table qui contient les coordonnées géographiques des états des USA

Types de visualisations :

1. [Choropleth Map: Salaires & Opportunités par État](#) (HAMZA)

Pour : Chercheurs d'emploi + RH/Recruteurs

Tâches :

- Comparer les États avec plus de demande en ingénieurs data
- Classer les États par volume d'offres ou salaire moyen
- Cartographie des offres d'emploi par états

Dimensions/Filtres :

- Couleur : Salaire moyen ou nombre d'offres
- Toggle : Basculer entre "salaire moyen" et "nombre d'offres"
- Filtres : Séniorité (senior/na), taille d'entreprise, secteur
- Tooltip : Détails au survol (état, salaire, nb offres, top 3 technologies)

Animation possibles :

- Transition de couleur smooth lors du changement de métrique (salaire ↔ offres)
- Zoom/pan sur les états
- Highlight au survol avec bordure épaisse
- Animation de remplissage progressif au chargement (effet "vague" d'ouest en est)

2. [Sunburst Diagram - Hiérarchie Technologies](#) (BAILO)

Pour : Chercheurs d'emploi + Institutions éducatives

Tâches répondues :

- Identifier les technologies les plus utilisées
- Identifier les spécialités les plus demandées

Structure hiérarchique :

Centre : "Technologies Data Engineering"

- ├─ Langages (python, sql, java, scala...)
- ├─ Cloud (aws, azure, gcp)
- ├─ Databases (postgresql, mongodb, snowflake...)
- ├─ Big Data (spark, hadoop, flink)
- ├─ Data Tools (airflow, kafka, dbt)
- ├─ DevOps (docker, kubernetes, terraform...)

Dimensions :

- Taille des segments : Fréquence d'apparition dans les offres
- Couleur : Catégorie technologique
- Filtres : État, séniorité, salaire min/max
- Click sur segment : Zoom hiérarchique + affiche détails (% d'offres, salaire moyen avec cette tech)

Animation possibles :

- Rotation fluide lors du zoom/dézoom sur segments
- Transition arc tween pour agrandir/réduire segments
- Fade in/out des labels
- Effet "explosion" au chargement (segments partent du centre)
- Mise à jour animée des tailles lors du filtrage

3. Parallel Coordinates - Analyse Multidimensionnelle Salaires (ILIASSE)

Pour : RH/Recruteurs + Chercheurs d'emploi

Tâches répondues :

- Distribution des salaires par secteur, industrie, taille d'entreprise
- Corréler le salaire avec X paramètres de l'entreprise
- Comparer différence des salaires selon taille et région

Axes parallèles :

- Salaire (0 - 200k\$)
- Company Rating (1.0 - 5.0)
- Company Age (-1 à 100+ ans)
- Company Size (catégoriel : 1-50, 51-200, 501-1000, 10000+)
- Experience (0-2, 2-5, 5-10, +10 ans)
- Job State (top 10 états seulement)

Interactions :

- Brush sur chaque axe pour filtrer
- Couleur des lignes : par état ou par séniorité
- Opacité réduite pour lignes non-sélectionnées
- Réordonnement des axes par drag & drop
- Toggle pour inverser un axe

Animation possibles :

- Transition fluide des lignes lors du filtrage (opacity + stroke-width)
- Animation "bundle" : lignes se regroupent progressivement
- Morphing des paths lors du réordonnement d'axes
- Effet "dessin progressif" au chargement (stroke-dashoffset)

4. Radial Stacked Bar Chart - Technologies par État (OUSSAMA)

Pour : Institutions éducatives + RH/Recruteurs

Tâches répondues :

- Identifier technologies les plus utilisées
- Comparer États avec plus de demande
- Distribution par location

Structure :

- Chaque rayon = un État (top 15 états)
- Hauteur du rayon = nombre total d'offres
- Segments empilés = catégories de technologies (langages, cloud, databases, big data, data tools, devops)
- Angle = disposition circulaire uniforme

Dimensions/Filtres :

- Couleur : Type de technologie
- Filtres : Séniorité, salaire min/max, taille entreprise
- Toggle : Vue absolue (count) vs relative (%)
- Hover : Tooltip détaillé + highlight du segment sur tous les rayons

Animation possibles :

- Rotation du cercle complet au chargement
- Transition radiale des hauteurs lors du filtrage
- Effet "onde" : barres s'animent séquentiellement (delay progressif)
- Morphing smooth entre vue absolue ↔ relative
- Pulse effect sur le rayon sélectionné