

**DIGITAL
SERVICE**

Diagnostic
égalité
femmes-
hommes

Descriptif Projet



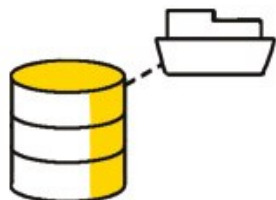
- Mettre en place un rapport diagnostique égalité Femme-Homme
 - Sources données : Système d'Informations des Ressources Humaines (SIRH)
 - Création fichier csv : Avec les données pour analyse avec autres outils (Tableau, Python.....)
 - ETL* & data viz* : Knime
- Contexte projet :
 - Conformité RGPD
 - S'inspirer des Outils mis à disposition par le législateur : [Guide_egalite](#) - [outil Excel](#)
 - Contrainte réglementaire : Code du travail : Egalité professionnelle entre les femmes et les hommes (Articles L1141-1 à L1146-3)

•Article L1142-8

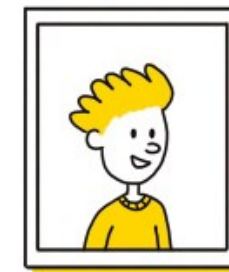
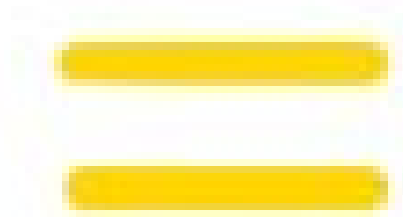
Modifié par LOI n°2021-1774 du 24 décembre 2021 - art. 13 (V)

Dans les entreprises d'au moins cinquante salariés, l'employeur publie chaque année l'ensemble des indicateurs relatifs aux écarts de rémunération entre les femmes et les hommes et aux actions mises en œuvre pour les supprimer, selon des modalités et une méthodologie définies par décret. Par dérogation aux articles L. 311-6 et L. 312-1-2 du code des relations entre le public et l'administration, l'ensemble de ces indicateurs est rendu public sur le site internet du ministère chargé du travail, dans des conditions déterminées par décret.

RGPD : Données personnelles



- ❑ Nom : ?
- ❑ Prénom : ?
- ☑ Sexe : masculin
- ☑ Âge : 19
- ☑ Adresse : 5 rue de la gare
79000 NIORT
- ☑ Lycée : Montaigne (Bordeaux)
- ☑ Passion : Le jazz



Marc PELLETIER



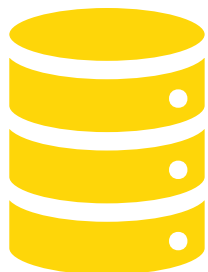
Je suis une base
de données personnelles



Je m'assure que
les données collectées
servent bien l'objectif prévu

Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données)

Process projet



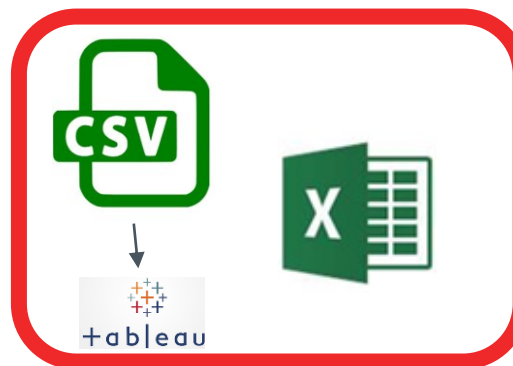
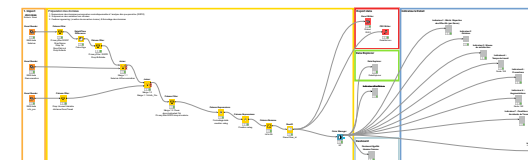
Source :



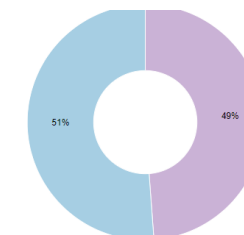
Exploration data set :



ETL :

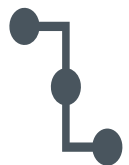


Data Viz:



Cliquer sur l'image ci-dessus pour accéder à l'ensemble des documents du projet.

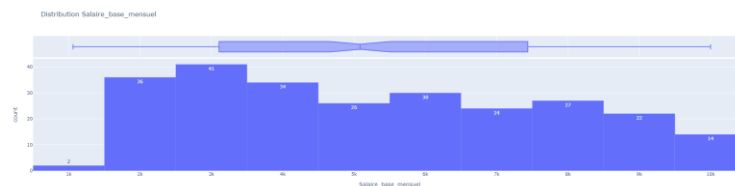
Exploration data Set (Notebook Google Collab)



Info DataFrame : df (256 Lignes, 5 Colonnes)

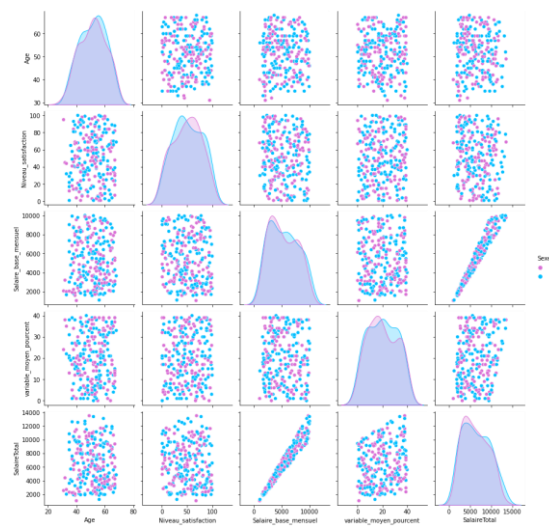
	Colonne	Nb_Unique	Nb_Doublons	Doublons_%	DataType	nbr_NaN	NaN_%	Commentaire	Min/first	Max/last
0	id_salarie	256	0	0	object	0	0	cléCandidate	16920105-1209	16451016-1336
1	Sexe	2	254	99	object	0	0	-	F	H
2	Age	37	219	86	int64	0	0	-	31	68
3	Age10	4	252	98	int64	0	0	-	30	60
4	Classe_Age	4	252	98	object	0	0	-	60ans_et_plus	40ans-49ans

1. Description forme, variable, recherche clefs candidate



3. Analyse Univariée et recherche outliers

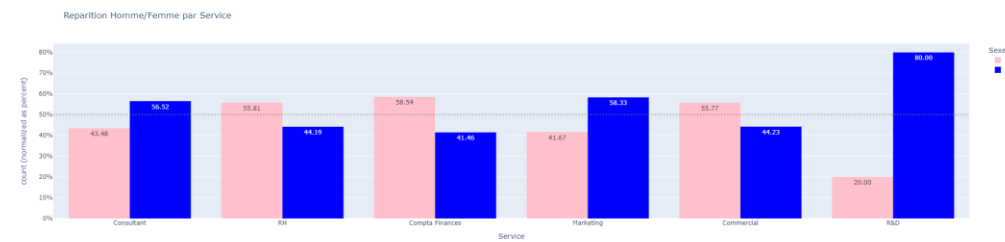
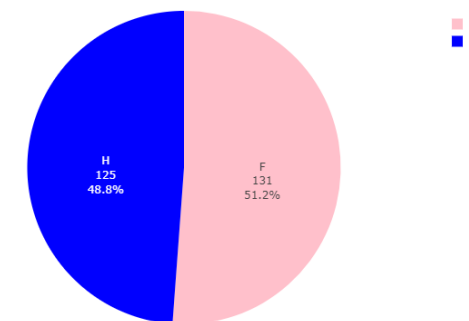
5. Analyse bivariée –recherche Corrélation (hue genre)



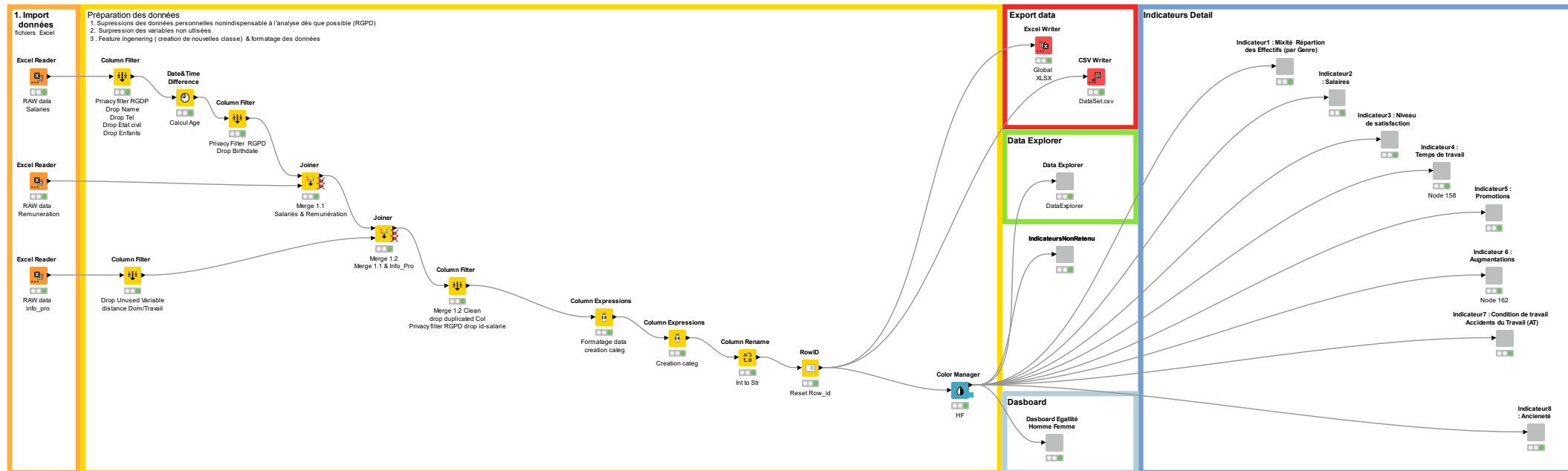
2. Heatmap des NaN => très peu de NaN sur le dataset
=> augmentation & Promotion pour 15 salariées

4. Analyse bivariée – graphiques Femme-Homme

Repartition par genre des salariés



ETL (Extract Transform Load)





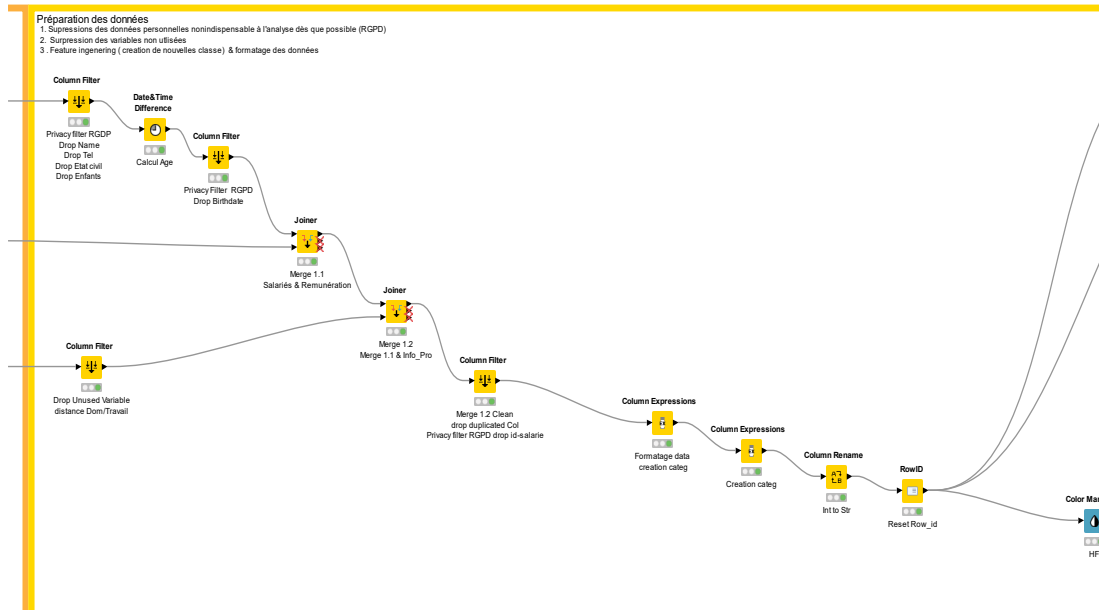
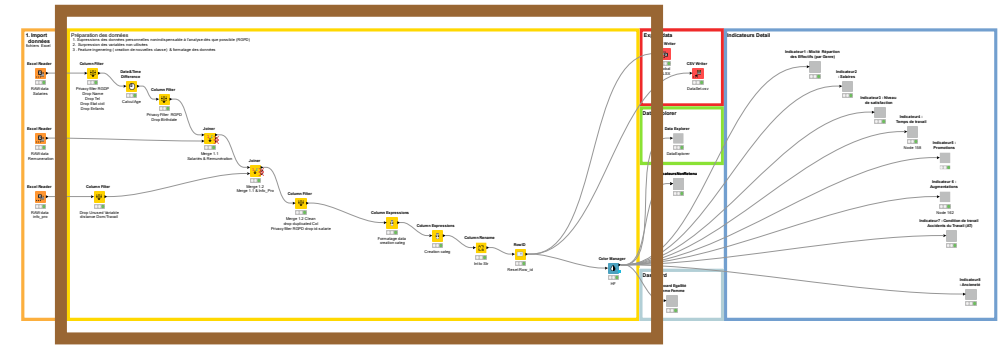
RAW data
info_pro

- **Extraction des données**
 - Importation des données issues des extraction Excel du :
Système d'Informations des Ressources Humaines (SIRH)



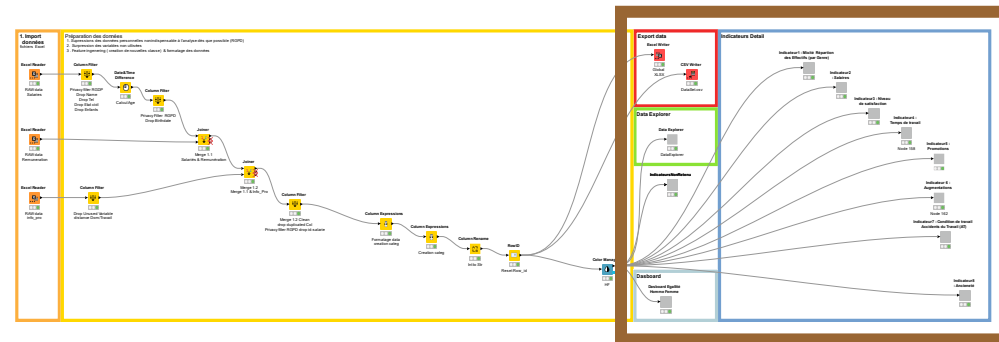
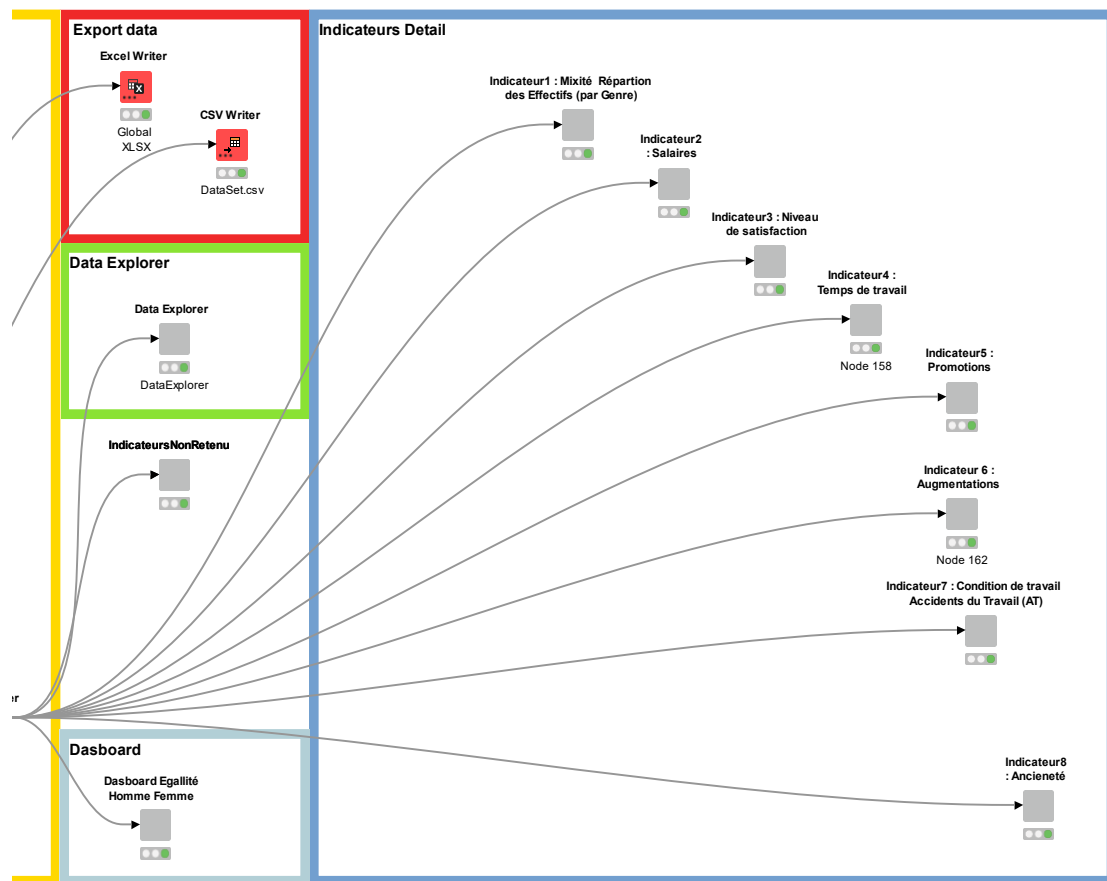
DIGITAL
SERVICE

ETL (Extract Transform Load)



- **Transformation des données**
 - **Jointures** des différentes tables pour croiser l'ensemble des informations
 - **Suppression des données personnelles** le plus tôt possible dans le process d'acquisition/ transformation des données en ne conservant que les données impérativement nécessaire à la construction des indicateurs.
 - **Feature engineering** : creation de nouvelles variables (Salaire total, Temps Complet /Temps Partiel, catégorisation de variables , ...)

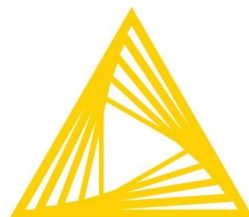
ETL (Extract Transform Load)



- Chargement des données & Data Viz

- **Export des données** au format CSV et au format Excel qui est le format choisi par le service RH pour l'import/ l'ensemble des informations
- **Explorateur de données** visuels dynamique pour exploration univarié et bivariée des données
- **Tableau de bord**: principaux visuels pour réaliser un diagnostique de l'égalité Femme-Homme dans l'entreprise
- **Indicateur détaillés**: visuels détaillé celons 8 thématiques, graphiques complémentaires, tests statistiques,...

Workflow KNIME



Open for Innovation

KNIME

