



Sommaire

- ▶ Contexte et objectifs de la mission
- ▶ Présentation des données
- ▶ Présentation du RGPD
- ▶ Présentation du workflow Knime
 - ▶ Nettoyage
 - ▶ Traitement
 - ▶ Analyse d'indicateurs
 - ▶ Output .CSV

Contexte

- ▶ Data Analyste



- ▶ Cabinet de consultants spécialisé dans la transformation digital



- ▶ Plus de 150 salariés



Objectifs

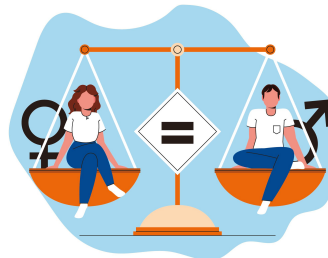
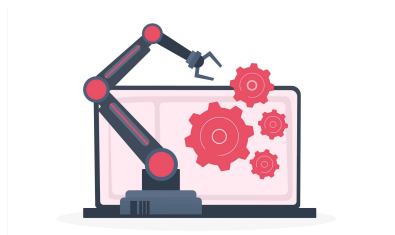
- ▶ Améliorer notre marque employeur



- ▶ Attirer des nouveaux talents



- ▶ Automatisation de la création d'un rapport diagnostique sur l'égalité professionnelle femmes/hommes



Ressources et outil

- ▶ Fichiers : 3 fichiers
 - ▶ Fichier info_pro
 - ▶ Fichier rémunération
 - ▶ Fichier salarié



- ▶ Logiciel : Knime

Présentation des données

Fichier info_pro

id_salarié	Ancienneté_an	Distance domicile/Travail	Service	Work_accident	Niveau de satisfaction
16000206-8278	12.6	59	Marketing	Non	67
16011205-0331	11.8	31	Commercial	Oui	66
16020327-1465	11.9	65	RH	Non	59
16020701-3376	10.2	59	Compta Finances	Non	60
16020725-8401	7.9	59	Consultant	Non	34
16021124-8356	8.6	34	RH	Oui	84
16021125-4271	14.9	36	Commercial	Non	62
16030220-7121	6.3	23	Consultant	Non	69
16030828-8604	18.2	23	Commercial	Oui	54
16031007-1667	23.1	49	Commercial	Oui	31
16050210-5695	23.8	40	Consultant	Non	9
16050407-7405	6.7	7	RH	Non	35
16050619-6252	23.9	70	Commercial	Oui	38
16050721-8014	18.6	16	RH	Non	30
16060812-9664	20.5	61	Compta Finances	Oui	70
16080130-8107	24.1	28	Marketing	Non	54
16080411-7877	13.3	17	Consultant	Non	9
16080824-2457	15.2	62	Compta Finances	Non	96
16081104-1011	1.1	31	Compta Finances	Non	18

- id_salarié (Clé primaire)
- Ancienneté (Année)
- Distance domicile/travail (en km)
- Service (6 services)
- Accident de travail (Oui/Non)
- Satisfaction (de 1 à 100)

Présentation des données

Fichier rémunération

id_salarié	Contrat	Durée hebdo	Salaire base mensuel	%variable_moyen	Augmentation	Promotion
16990924-2829	CDI	35	9 873.50 €	13	0	0
16990414-3527	CDI	35	4 296.72 €	39	0	1
16981206-3817	CDI	35	3 705.22 €	29	1	1
16970415-1050	CDI	35	1 697.15 €	17	1	0
16960906-7989	CDD	35	2 848.30 €	15	1	0
16960628-0338	CDI	35	5 464.55 €	34	0	0
16960619-9017	CDI	35	9 999.38 €	27	0	1
16951117-9468	CDI	24	3 116.88 €	37	0	1
16950828-2259	CDI	24	5 462.08 €	39	0	1
16950514-2464	CDI	35	2 769.43 €	33	0	0
16940521-7028	CDI	35	5 522.18 €	9	1	1
16940414-0312	CDI	35	7 308.82 €	12	0	1
16931211-0175	CDI	24	5 895.76 €	23	1	1
16931009-2151	CDI	35	7 081.32 €	26	0	1
16920917-6040	CDI	35	6 195.58 €	23	0	1
16920713-6616	CDI	35	5 517.63 €	23	1	1
16920528-2461	CDI	35	8 458.45 €	1	0	0
16920219-1855	CDI	35	2 132.43 €	18	1	1
16920105-1209	CDI	35	3 097.03 €	13	0	0

- id_salarié (Clé primaire)
- Contrat (CDI/CDD)
- Durée hebdo (en heure)
- Salaire mensuel (en euro)
- Variable (0 à 100)
- Augmentation (0 ou 1)
- Promotion (0 ou 1)

Présentation des données

Fichier salariés

id_salarié	Sexe	Prénom/Nom	Telephone	Date_naissance	Etat Civil	Enfants
16920105-1209	F	Solenne Delannoy	607958599	07/07/1955	Célibataire	1
16661123-8764	H	Nicolas Jaccoud	660029842	22/09/1970	Célibataire	3
16230718-4297	F	Alexandra Auguste	629910592	01/03/1965	Marié	1
16130305-6954	F	Marie-Claire Besnard	685888970	02/09/1963		1
16360125-5924	F	Marie Berengar	624840475	02/04/1959	Marié	3
16520612-7770	F	Marie-Madeleine Frér	697029629	18/12/1960	Célibataire	2
16860412-1965	H	Côme Martin	647351913	18/10/1975	Marié	1
16551001-4151	F	Agathe Berengar	608137743	03/11/1970	Marié	2
16220420-6227	H	Christophe Mallette	683794378	05/02/1975		0
16550904-2932	F	Julie Botrel	661890256	14/05/1969	Célibataire	3
16350127-5733	H	Jean-Marc Baumé	625304926	22/03/1966	Célibataire	3
16080130-8107	H	Thaddée Batteux	678070755	13/05/1974		0
16390301-2403	H	Gwenaël Brunet	666770026	25/11/1960	Marié	3
16861012-1389	H	Kévin Guilloux	681634010	26/06/1968	Célibataire	2
16800116-6613	H	Alceste Morel	692596498	30/11/1959	Célibataire	1
16050619-6252	F	Sonia Corne	621667514	19/11/1968	Célibataire	2
16750302-5590	F	Mélissa Lecerf	670801501	13/11/1977		1
16390127-6224	F	Lucie Gaudreau	641131944	25/10/1965		1
16050210-5695	F	Myriam Chapelle	671051572	18/12/1965	Marié	1

- id_salarié (Clé primaire)
- Sexe (F/H)
- Nom/Prénom
- Téléphone
- Date de naissance
- Etat civil (Célibataire/Marié)
- Enfants



Les 5 principes du RGPD

- ▶ Principe de finalité
- ▶ Principe de proportionnalité et de pertinence
- ▶ Principe d'une durée de conservation limitée
- ▶ Principe de sécurité et de confidentialité
- ▶ Droits des personnes

Workflow Knime

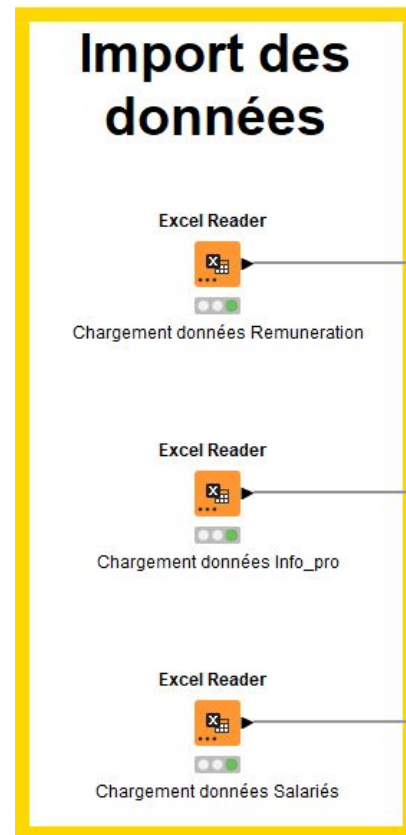


Indicateurs globaux

Indicateurs composé

Fichier .csv

Import des données



- ▶ 3 fichiers importer au forma Excel

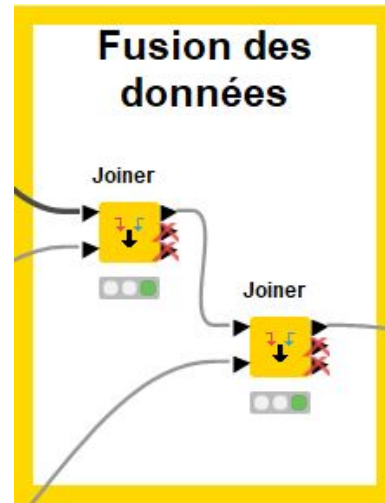
Nettoyage



Promotion	<div>Remove</div> <div>Fix Value</div> <div>Value 0</div>
Augmentation	<div>Remove</div> <div>Fix Value</div> <div>Value 0</div>

- ▶ Traitement des valeurs nulles ou manquantes
 - ▶ Remplacement (Valeur moyenne, Valeur par défaut, etc)
 - ▶ Suppression de ligne
- ▶ Suppression des doublons


Fusion des données



- ▶ Première fusion : inner (id_salarié)
Fichier remuneration + fichier info_pro
- ▶ Fusion finale : inner (id_salarié)
1^{er} fusion + Fichier salariés

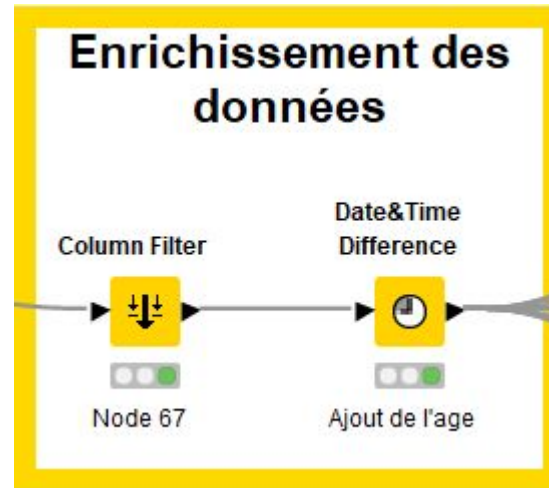
Enrichissement des données

Colonnes générées par les fusions



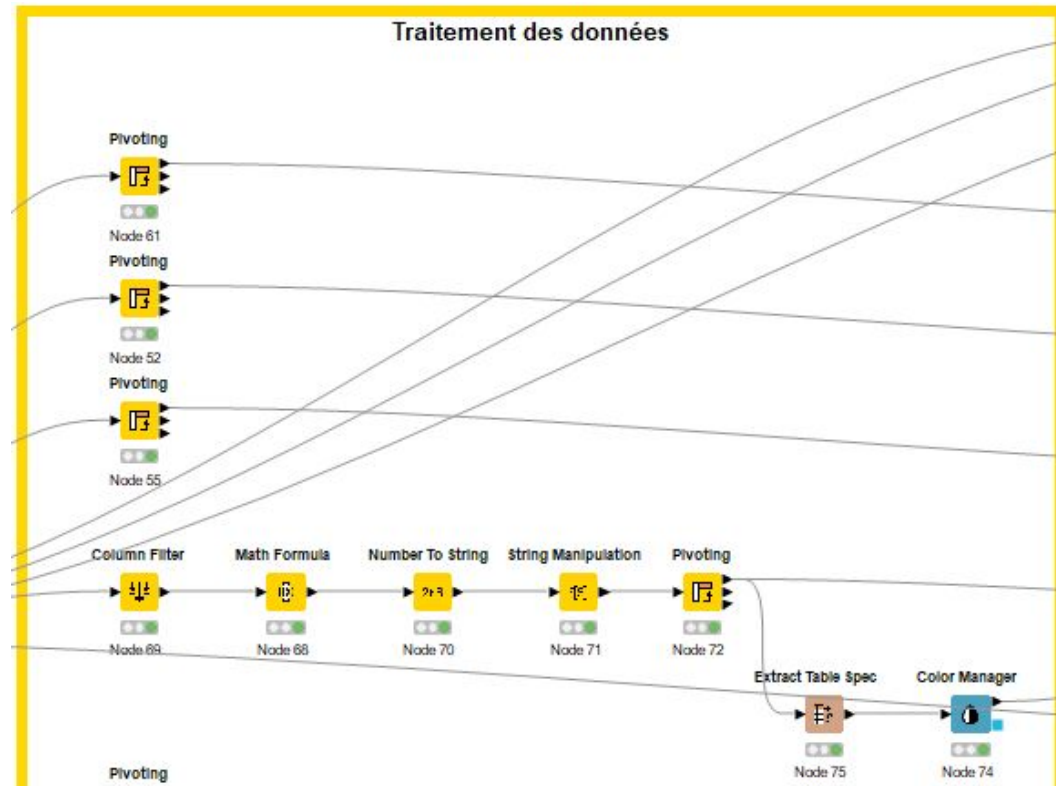
id_salarie (right)
16990924-2829
16990414-3527
16981206-3817

id_salarie (right) (right)
16990924-2829
16990414-3527
16981206-3817



- ▶ Suppression des colonnes générées par les fusions
- ▶ Calcul et ajout d'une colonne « Age »

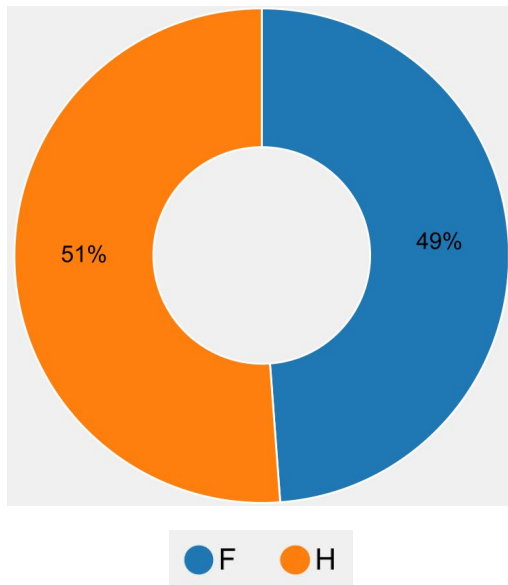
Traitement des données



- ▶ Calcul des indicateurs
- ▶ Classement des indicateurs
- ▶ Conversion des données

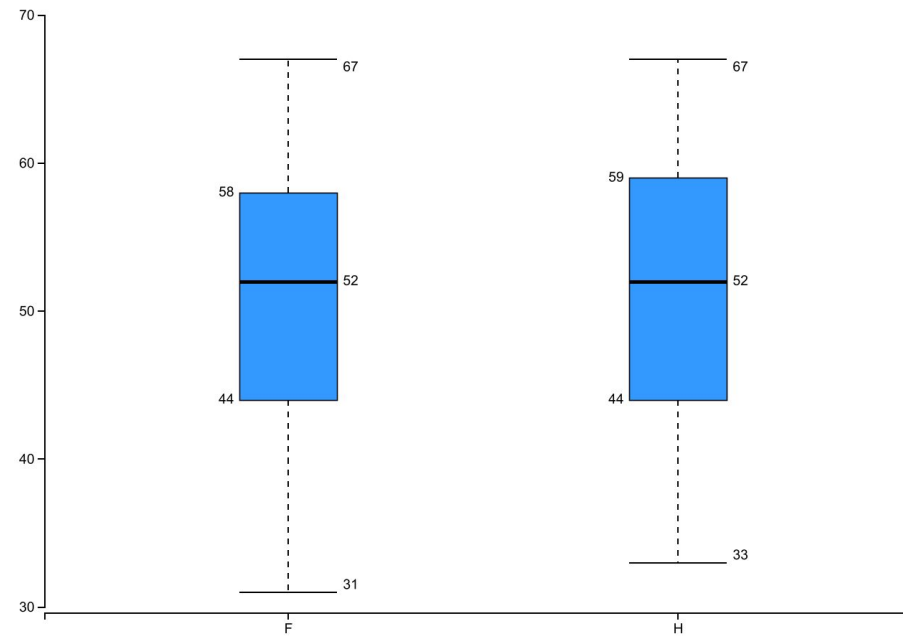
Indicateurs globaux

Répartition Homme/Femme au sein de l'entreprise



- ▶ Hommes légèrement plus nombreux
- ▶ Une quasi égalité Homme/Femme

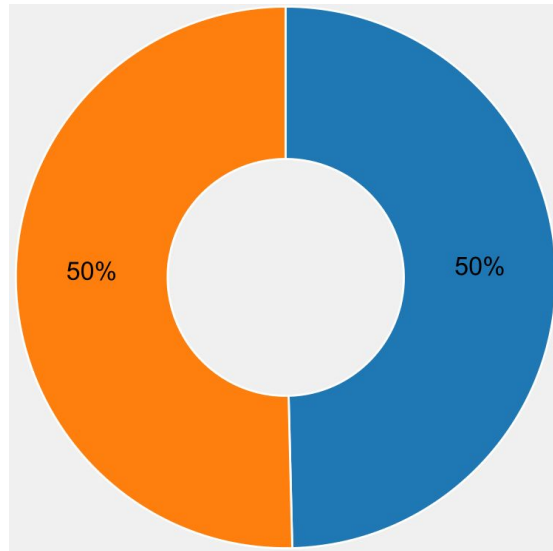
Répartition Femme/Homme par âge



- ▶ Femmes légèrement plus jeune
- ▶ Une quasi égalité des âges Homme/Femme

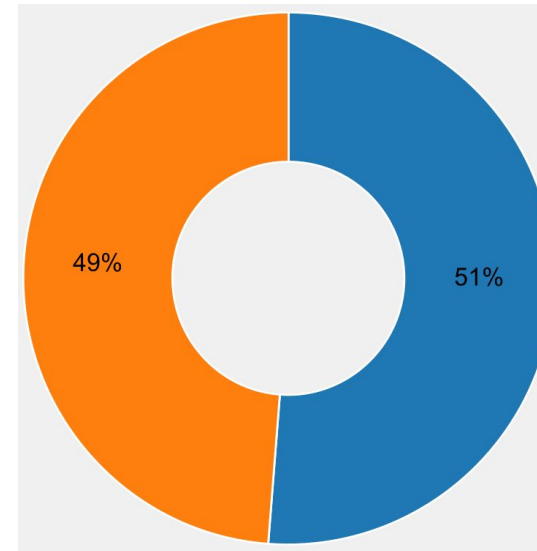
Indicateurs globaux

Répartition de promotions
Homme/Femme



Répartition des augmentations
Homme/Femme

● F ● H

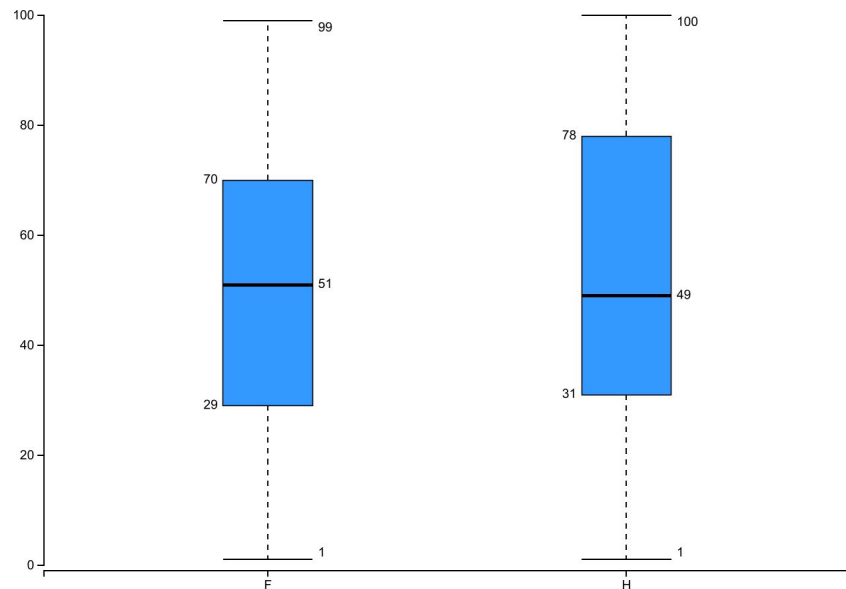


- ▶ Parfaite égalité Homme/Femme

- ▶ Femmes légèrement plus d'augmentations
- ▶ Une quasi égalité des âges Homme/Femme

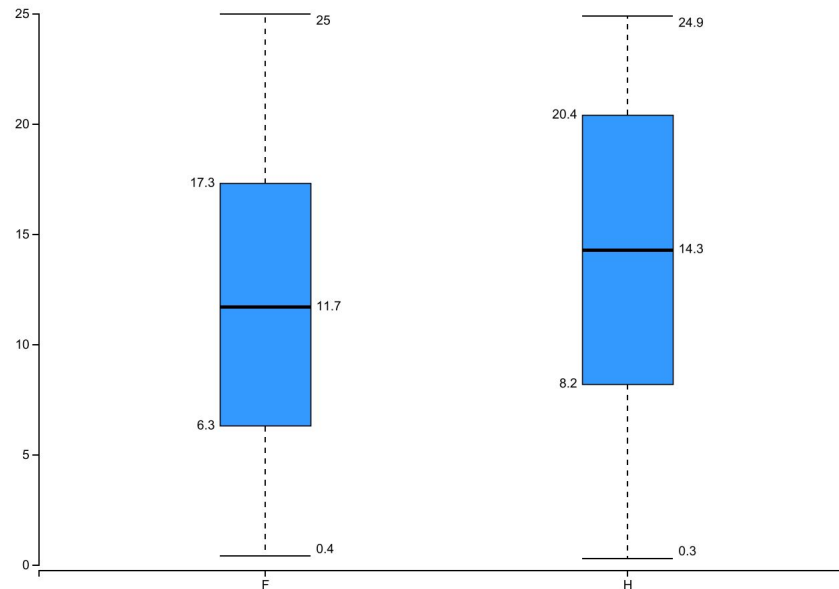
Indicateurs globaux

Répartition de la satisfaction
Femme/Homme



- ▶ 25% des femmes les plus satisfaites entre 70 et 99
- ▶ 25% des hommes les plus satisfaits entre 78 et 100

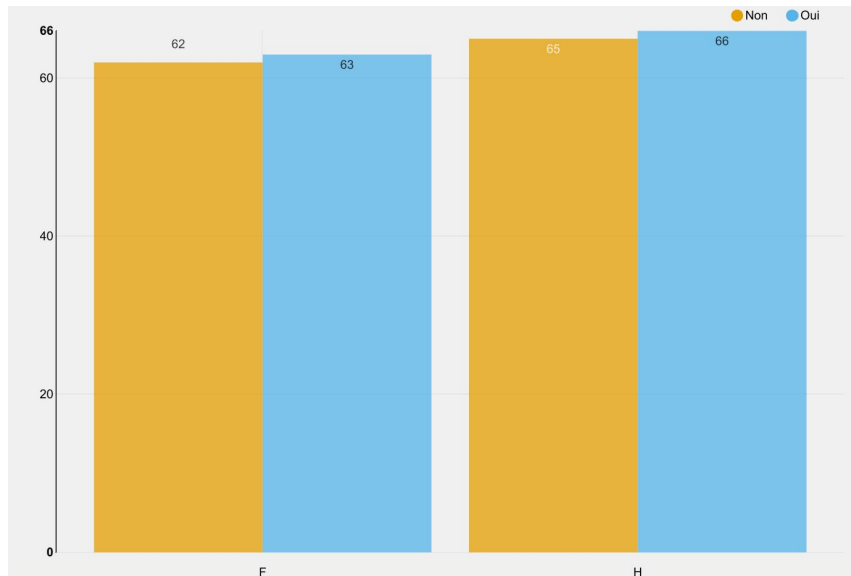
Répartition de l'ancienneté
Femme/Homme



- ▶ Pour 75% des femmes l'ancienneté la plus longue est de 17 ans
- ▶ Pour 75% des hommes l'ancienneté la plus longue est de 20 ans

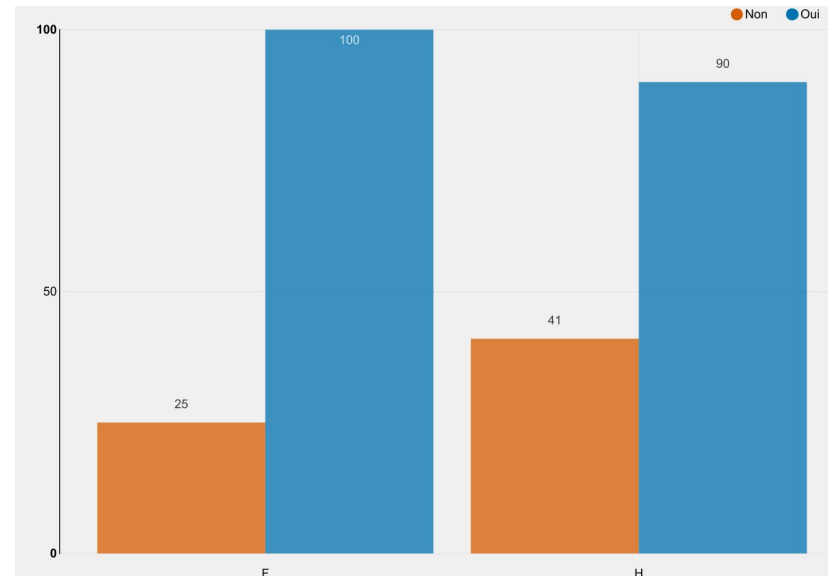
Indicateurs globaux

Répartition des accidents de travail
Femme/Homme



- ▶ Beaucoup d'accidents de travail (+ de 120 accidents)

Répartition Femme/Homme ayant des enfants

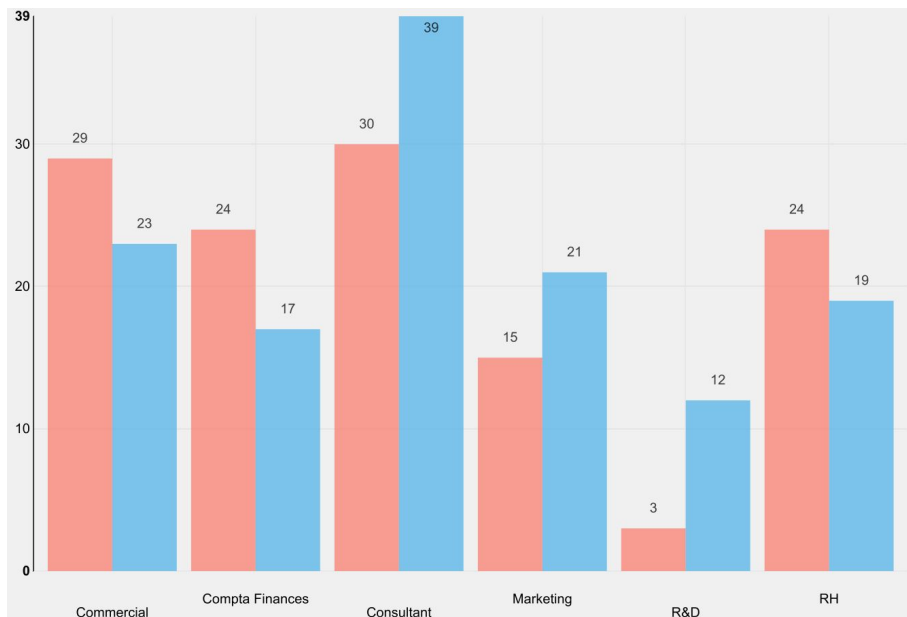


- ▶ Pas de discrimination envers les femmes ayant des enfants

Indicateurs composés

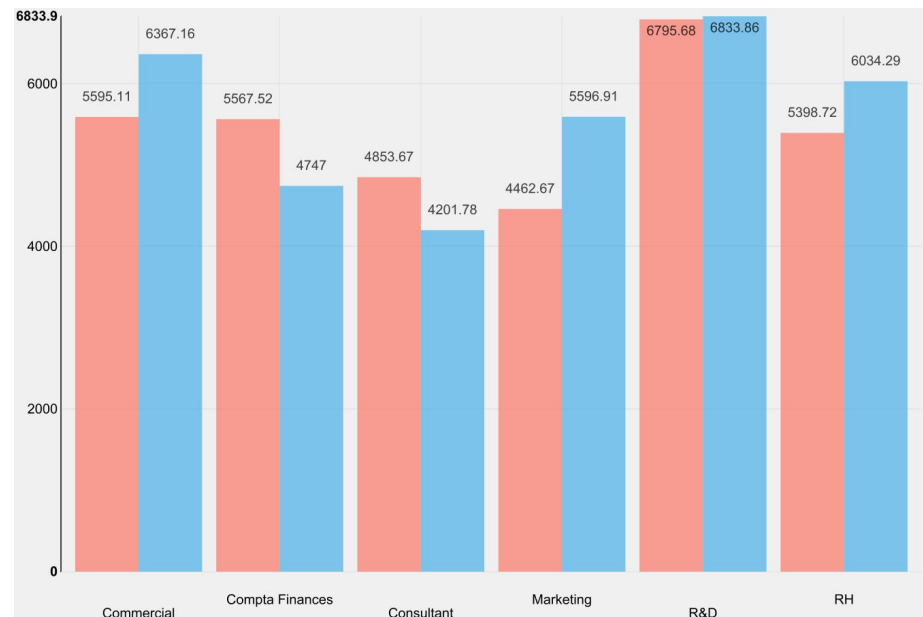
Analyse par service

Répartition des effectifs Femme/Homme par services



- Pas d'égalité parfaite au sein des services

Répartition du salaire moyen Femme/Homme par services



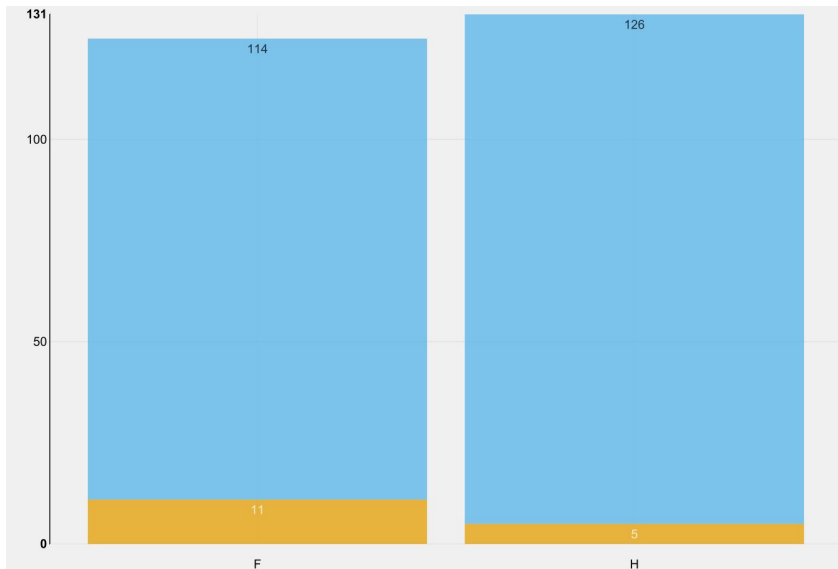
- Les femmes ont un salaire moyen plus élevé dans certains services



Indicateurs composés

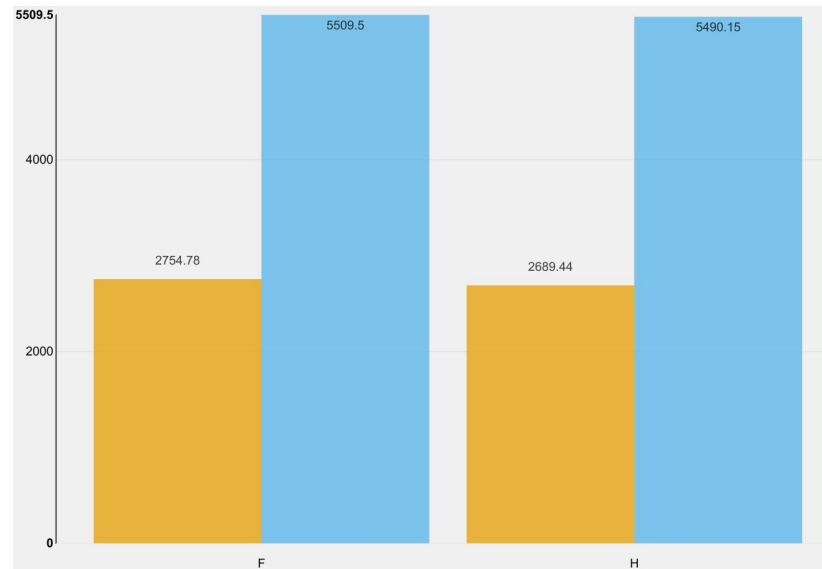
Analyse des types de contrat

Répartition des types de contrats
Femme/Homme



- Plus de CDI chez les hommes
- Plus de CDD chez les femmes


Répartition du salaire moyen
Femme/Homme par types de contrats



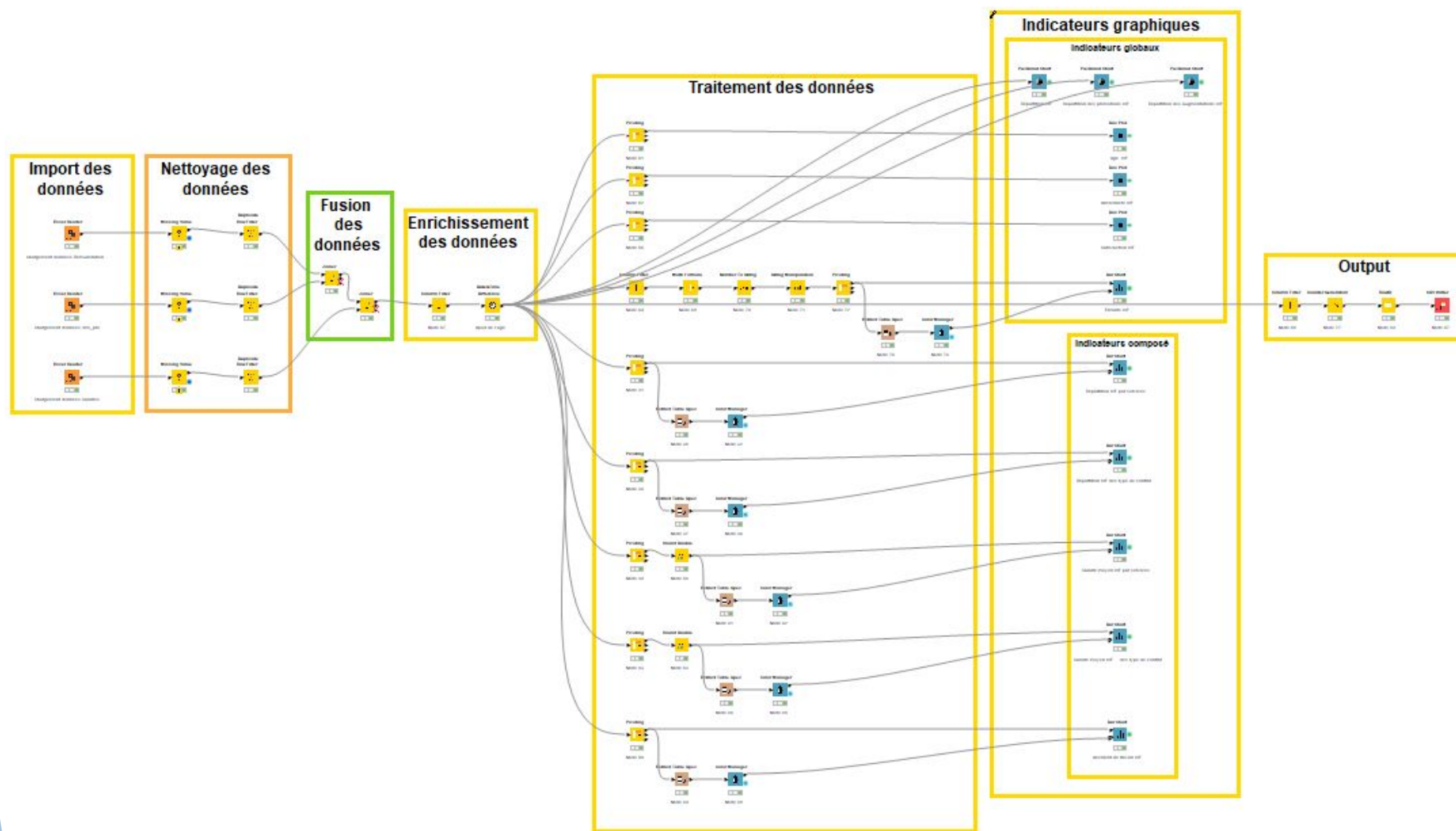
- Salaires moyens légèrement plus élever pour les femmes

Fichier .CSV

row ID	Contrat	Durée hebdo	Salaire base mensuel	Augmentation	Promotion	Ancienneté_an	Service	Work_accident	Niveau de satisfaction	Sexe	Enfants	Age
0	CDI	35	9873.5	0	0	22.5	Consultant	Non	60	F	3	46
1	CDI	35	4296.7167	0	1	10.8	Consultant	Non	39	H	0	67
2	CDI	35	3705.2167	1	1	13	RH	Oui	62	F	1	52
3	CDI	35	1697.15	1	0	21.7	Consultant	Non	98	H	3	59
4	CDD	35	2848.3	1	0	3.3	R&D	Oui	81	H	0	39
5	CDI	35	5464.55	0	0	10.5	Commercial	Non	47	H	0	54
6	CDI	35	9999.3833	0	1	17.8	Marketing	Oui	34	H	0	46
7	CDI	24	3116.88	0	1	8.2	RH	Oui	48	F	3	67
8	CDI	24	5462.08	0	1	7.5	Marketing	Non	25	H	2	37
9	CDI	35	2769.4333	0	0	18.9	Compta Fina	Oui	89	H	2	43
10	CDI	35	5522.1833	1	1	5.9	RH	Oui	7	F	2	67
11	CDI	35	7308.8167	0	1	20.2	Consultant	Non	24	F	3	43
12	CDI	24	5895.76	1	1	22.4	Commercial	Non	15	F	3	57
13	CDI	35	7081.3167	0	1	20	R&D	Non	31	H	0	59
14	CDI	35	6195.5833	0	1	13.1	Commercial	Oui	56	H	1	52
15	CDI	35	5517.6333	1	1	23.3	Marketing	Non	81	H	1	51
16	CDI	35	8458.45	0	0	15.2	R&D	Non	35	H	2	43
17	CDI	35	2132.4333	1	1	3.2	Consultant	Oui	60	F	3	52
18	CDI	35	3097.0333	0	0	25	Consultant	Non	19	F	1	67

- Respect des 5 principes du 
- Données supprimé toutes les 5 années
- Eventuel ajout d'un mot de passe sur le fichier
- Limiter les accès aux personnes concernées

Screen workflow



Synthèse

- ▶ Lecture des données:
 - ▶ Données manquantes
- ▶ Indicateurs globaux:
 - ▶ Respect de l'égalité homme/femme
 - ▶ Aucune discrimination visible
- ▶ Indicateurs composés:
 - ▶ Aucune égalité parfaite au sein des services
 - ▶ Plus de stabilité chez les hommes que chez le femme (femme + CDD)
 - ▶ Quasi parfaite égalité du salaire moyen par type de contrat
- ▶ Fichier .CSV
 - ▶ Respect du RGPD