

Analyse des indicateurs de l'égalité femme-homme avec Knime

Openclassrooms

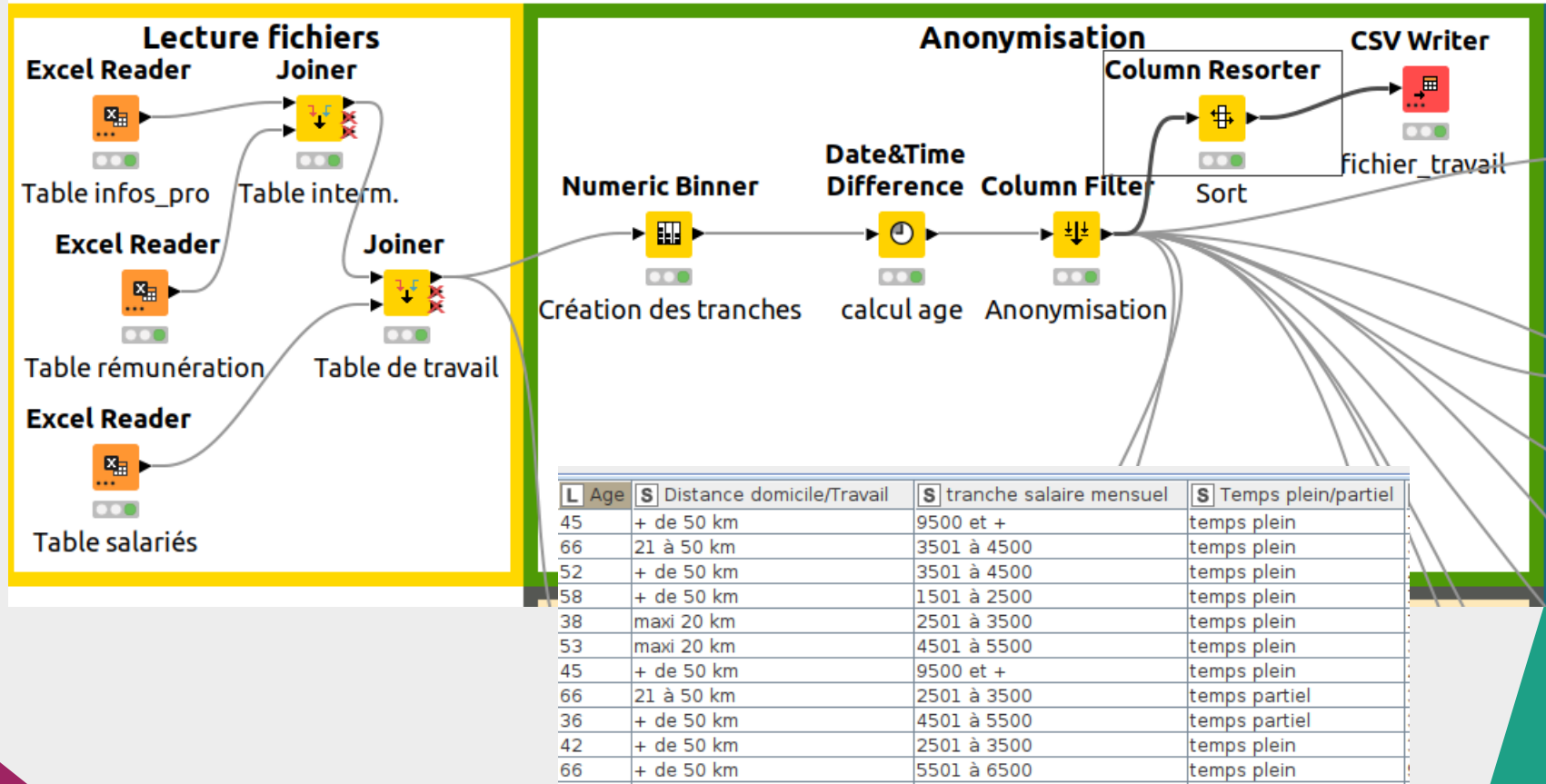
Formation Data Analyst

Projet 7

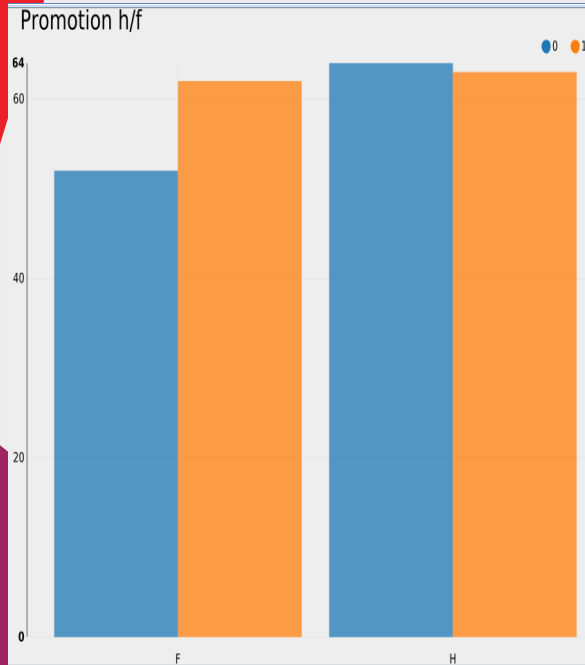
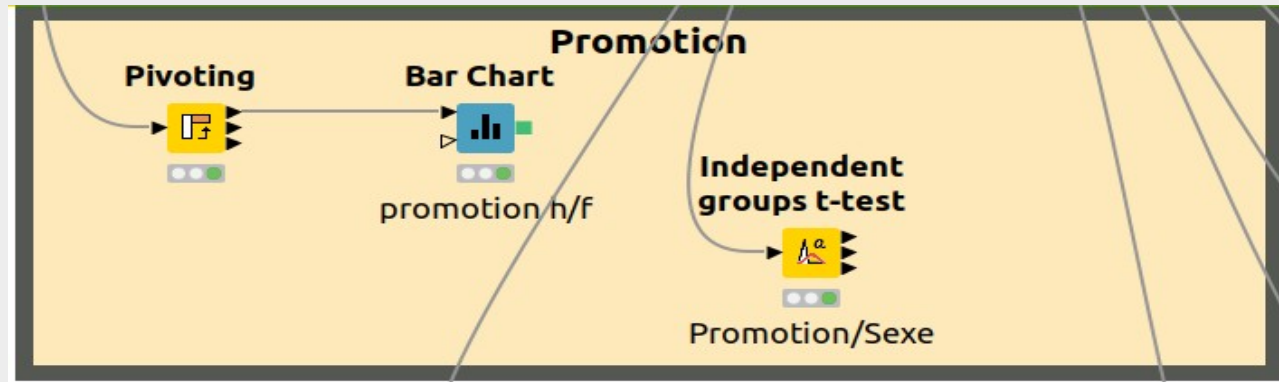
- Serge DAVISTER 01/2022

Lecture des fichiers csv et Anonymisation

- Lecture des fichiers csv et création d'une table de travail
- anonymisation du fichier de travail : suppression des champs id_salarié,nom/prénom et numéro de téléphone
- Ajout de champs tranches de salaire , temps partiel / temps plein et distance domicile/lieu de travail par tranche.
- création du fichier **anonymat.csv** avec les données préparées



Promotion



Group Statistics

	Group	N	Missing Count	Missing Count (Group Column)	Mean	Standard Deviation	Standard Error Mean
Promotion	F	114	11	0	0,5439	0,5003	0,0469
Promotion	H	127	4	0	0,4961	0,502	0,0445

Levene Test

The Levene Test is used to test for the equality of variances.

	F	df 1	df 2	p-Value
Promotion	0,9523	1	239	0,3301

Independent Groups Statistics

Confidence Interval (CI) Probability: 95.0%

Differences are reported of the groups: F - H

	Variance Assumption	t	df	p-value (2-tailed)	Mean Difference	Standard Error Difference	CI (Lower Bound)	CI (Upper Bound)
Promotion	Equal variances assumed	0,7392	239	0,4605	0,0478	0,0647	-0,0796	0,1752
Promotion	Equal variances not assumed	0,7393	236,3878	0,4604	0,0478	0,0646	-0,0796	0,1752

On vérifie si en moyenne, le nombre de promotions des femmes et des hommes sont identiques, à l'aide d'un t-test

$\text{promotion_moy_f} = \text{promotion_moy_h}$

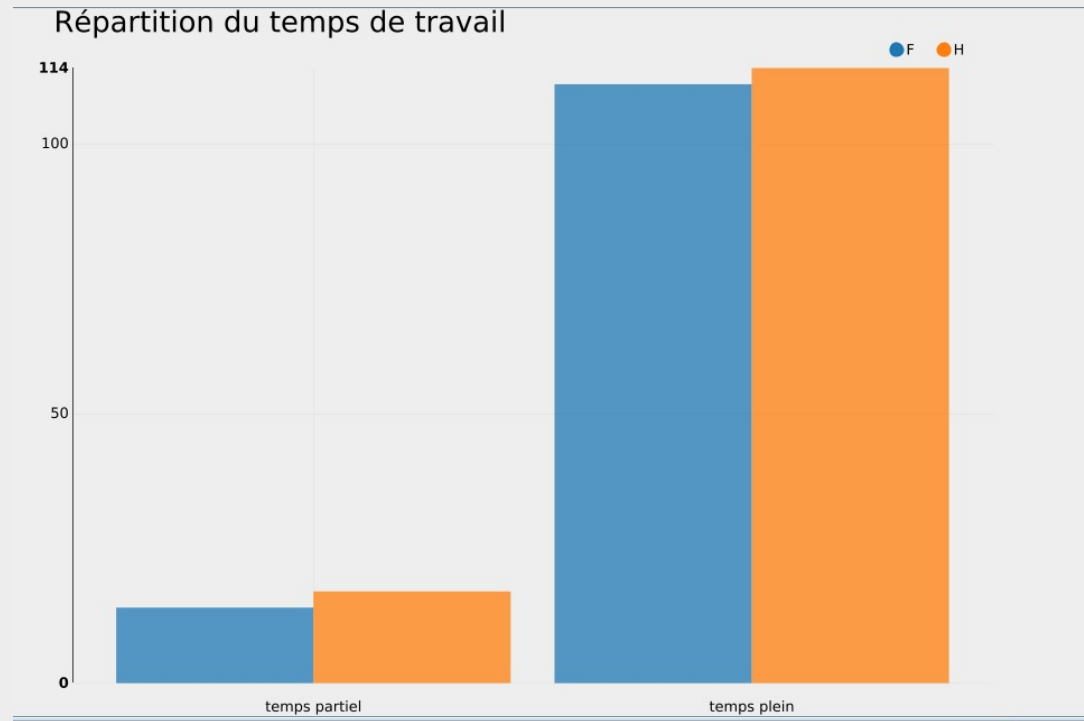
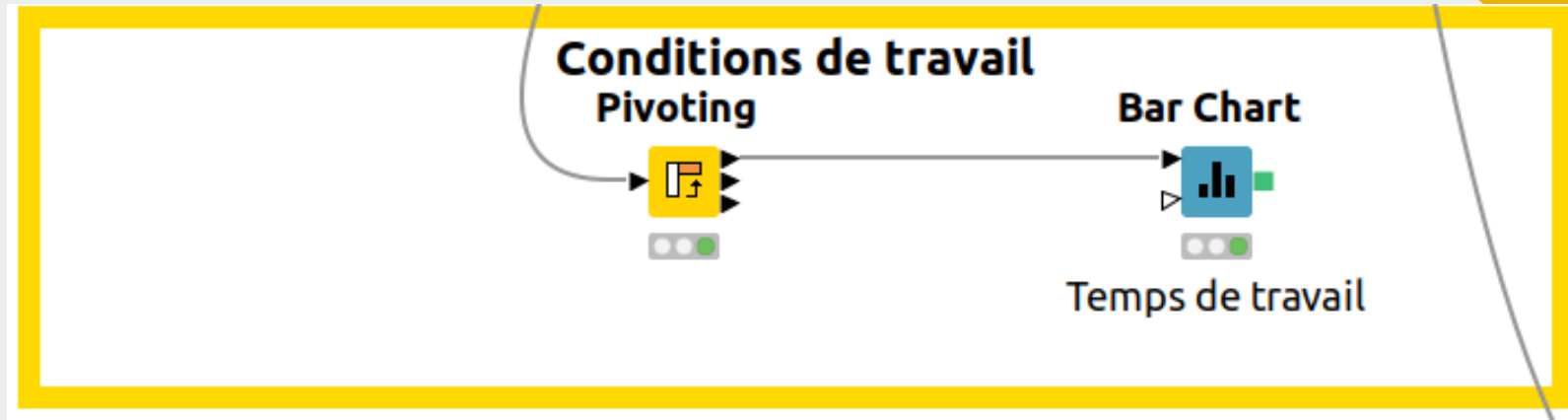
H0 : les variables sont indépendantes ; il n'existe aucune association entre les variables.

H1 : les variables ne sont pas indépendantes ; une association existe entre les variables et celles-ci sont dépendantes.

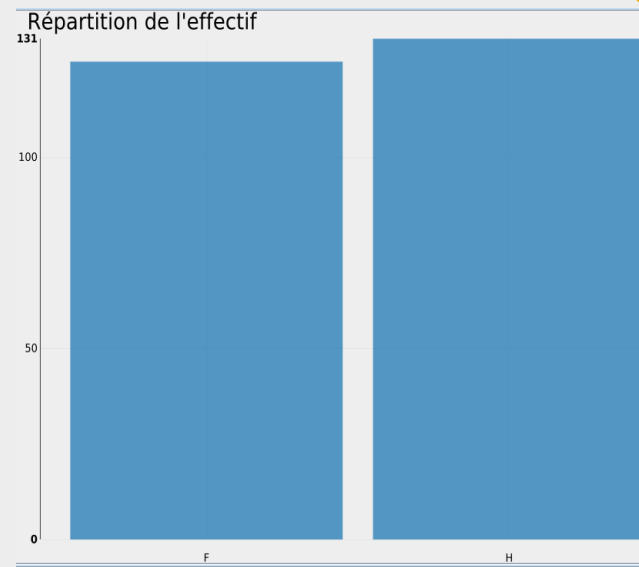
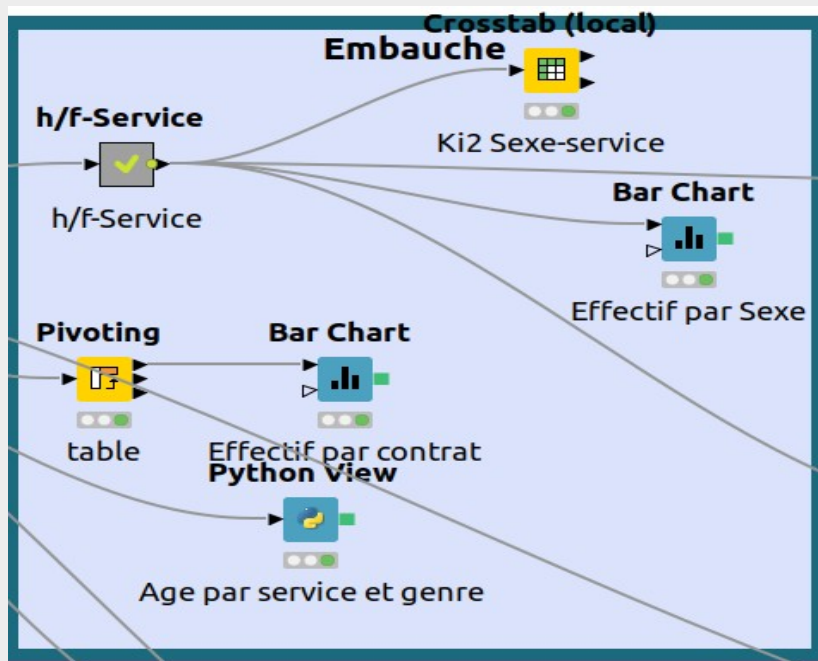
La p-value > 0.05 , on ne peut pas rejeter H0

On peut considérer que les variables Promotion et Sexe sont indépendantes et qu'il n'y a pas de lien entre la Promotion et le Sexe

Conditions de travail



Embauche



Statistics for Table of Service by Sexe

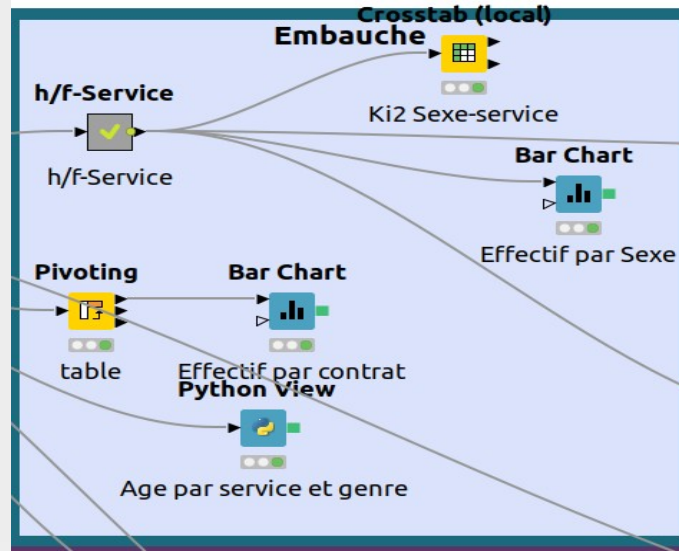
Statistic	DF	Value	Prob
Chi-Square	5	9,9076	0,0779

Total sample size: 256.0

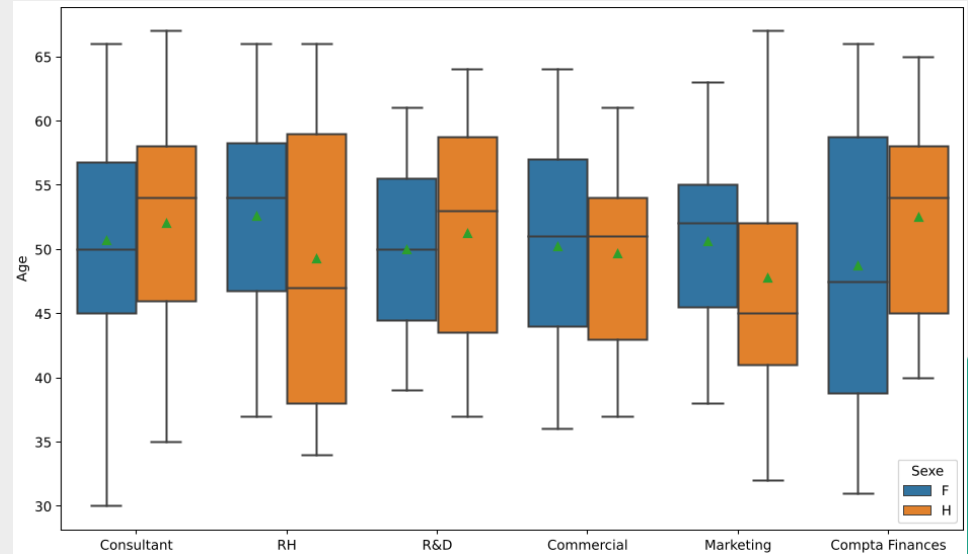
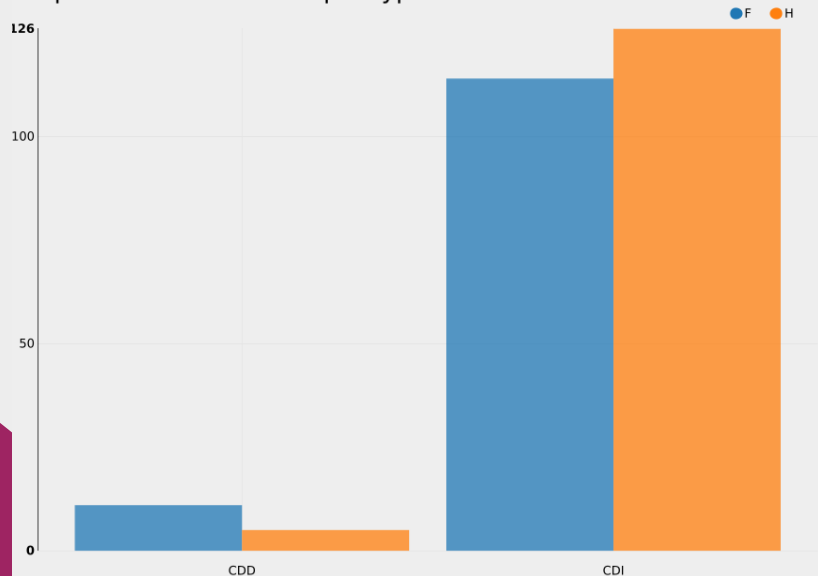
La p-value du chi square > 0.05

on ne peut pas rejeter H0 et on considère que les variables Sexe et Service sont indépendantes

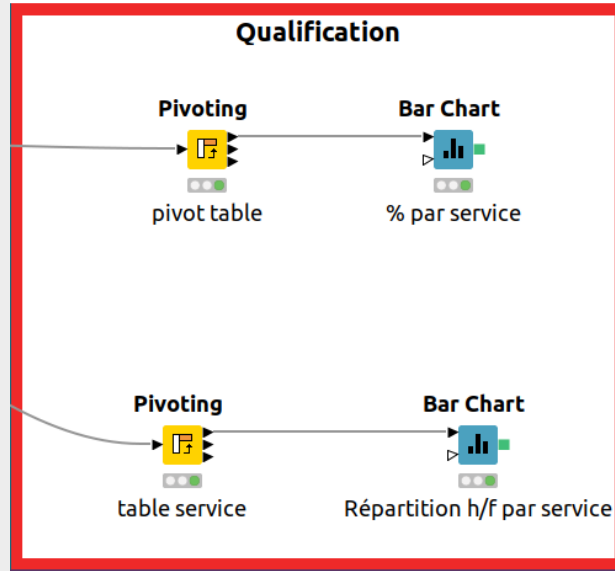
Embauche



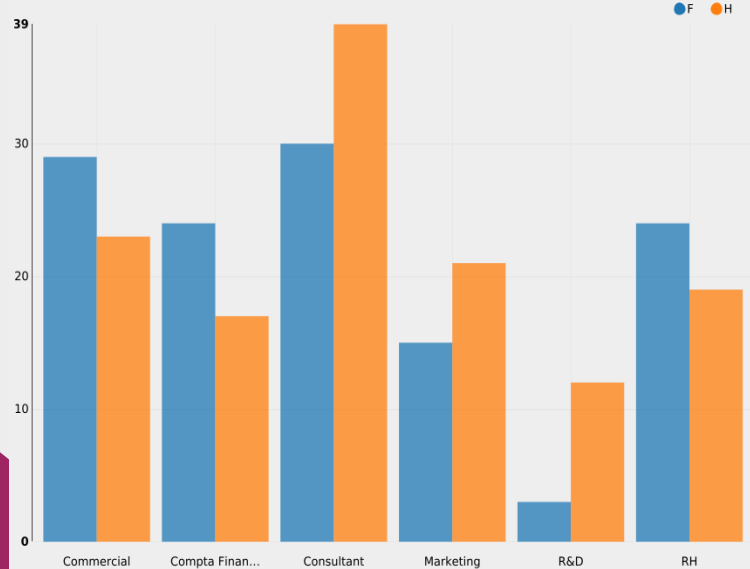
Répartition des effectifs par type de contrat



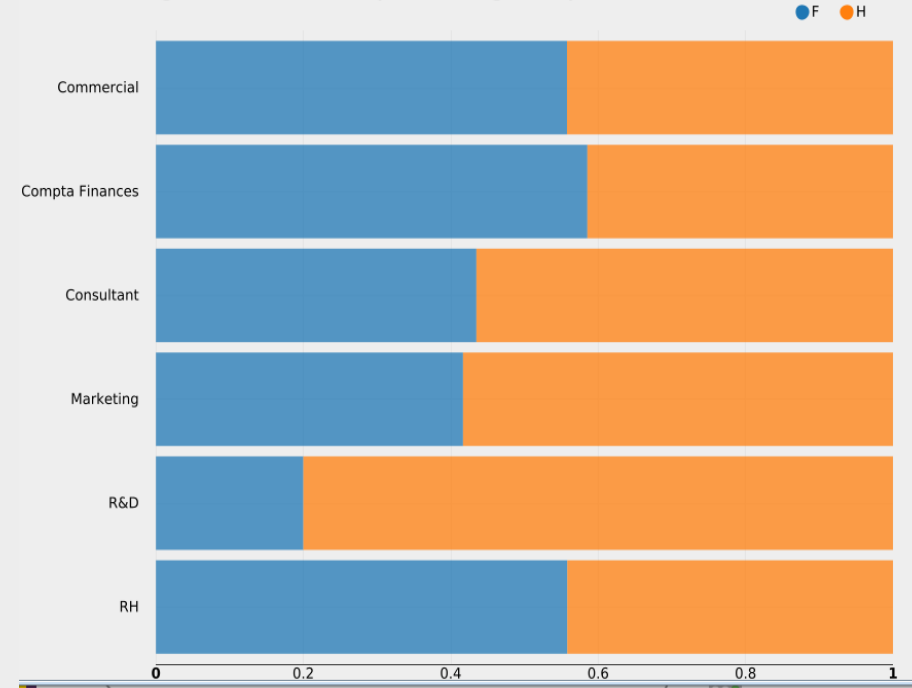
Qualification



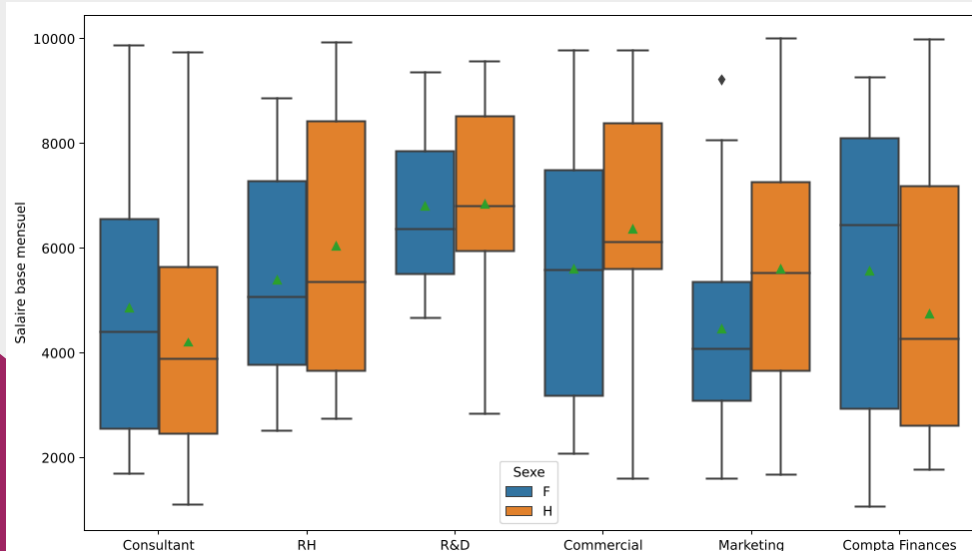
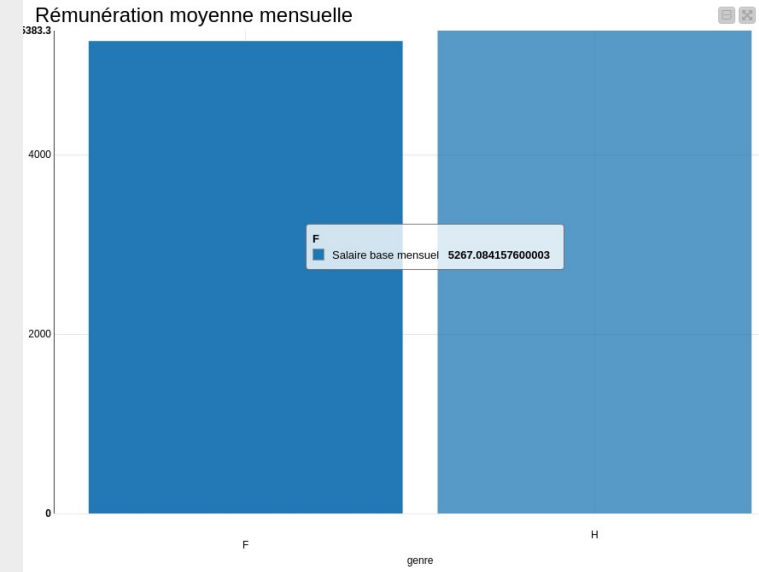
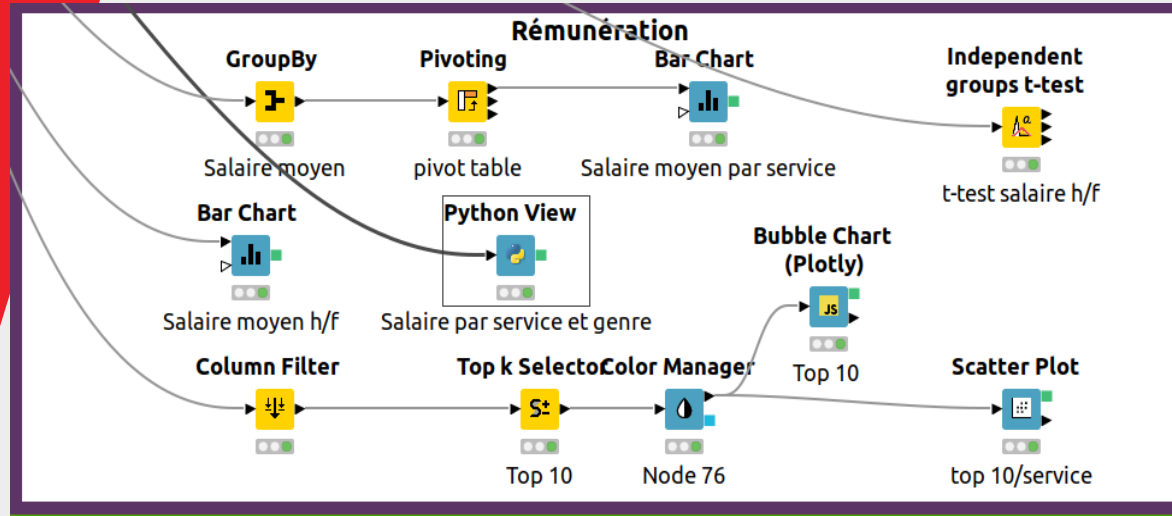
Répartition des effectifs par catégorie professionnelle



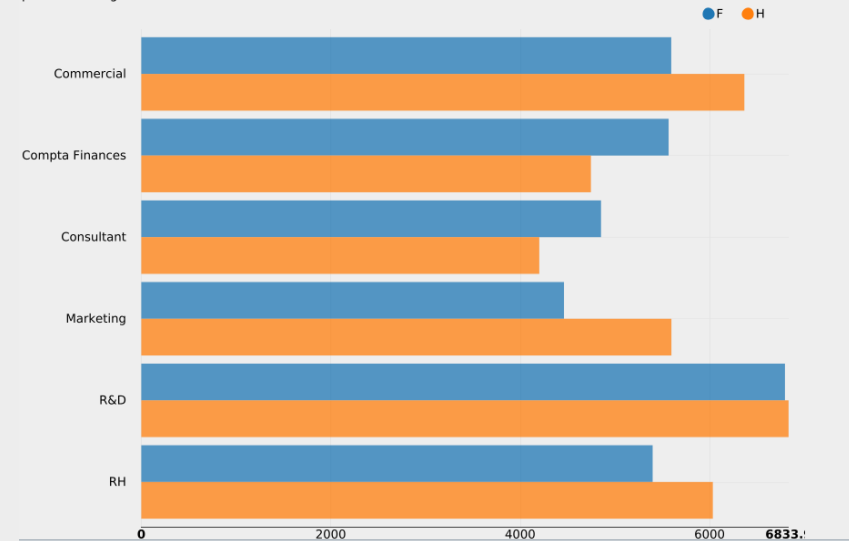
Pourcentage des effectifs par catégorie professionnelle



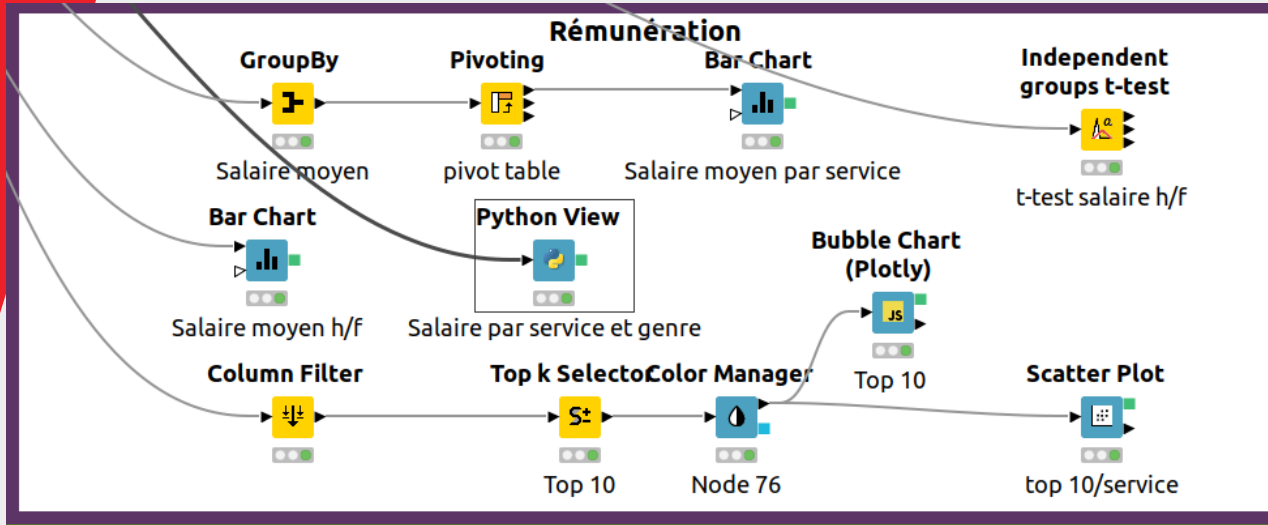
Rémunération



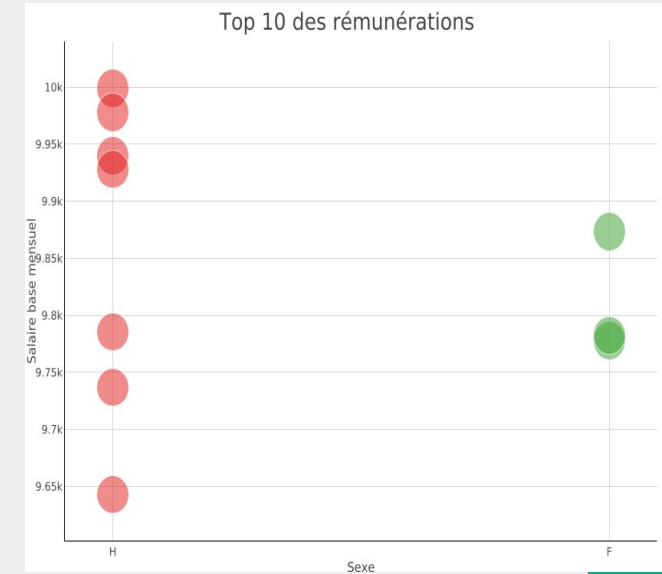
Salaire moyen
par Service et genre



Rémunération



S Service	D Salaire base mensuel	S Sexe
RH	9,642.967	H
Consultant	9,737	H
Commercial	9,777.95	F
Consultant	9,873.5	F
Consultant	9,782.5	F
Commercial	9,785.533	H
Marketing	9,999.383	H
Compta Finances	9,978.15	H
Marketing	9,940.233	H
RH	9,928.1	H



on vérifie si les salaires moyens des femmes et des hommes sont identiques.

à l'aide d'un t-test

$sal_moy_f = sal_moy_h$: les variables sont indépendantes ; il n'existe aucune association entre les variables.

H_1 : les variables ne sont pas indépendantes ; une association existe entre les variables et celles-ci sont dépendantes.

La p-value du test > 0.05 on ne peut pas rejeter H_0 et on considère que les variables Sexe et Salaire sont indépendantes.

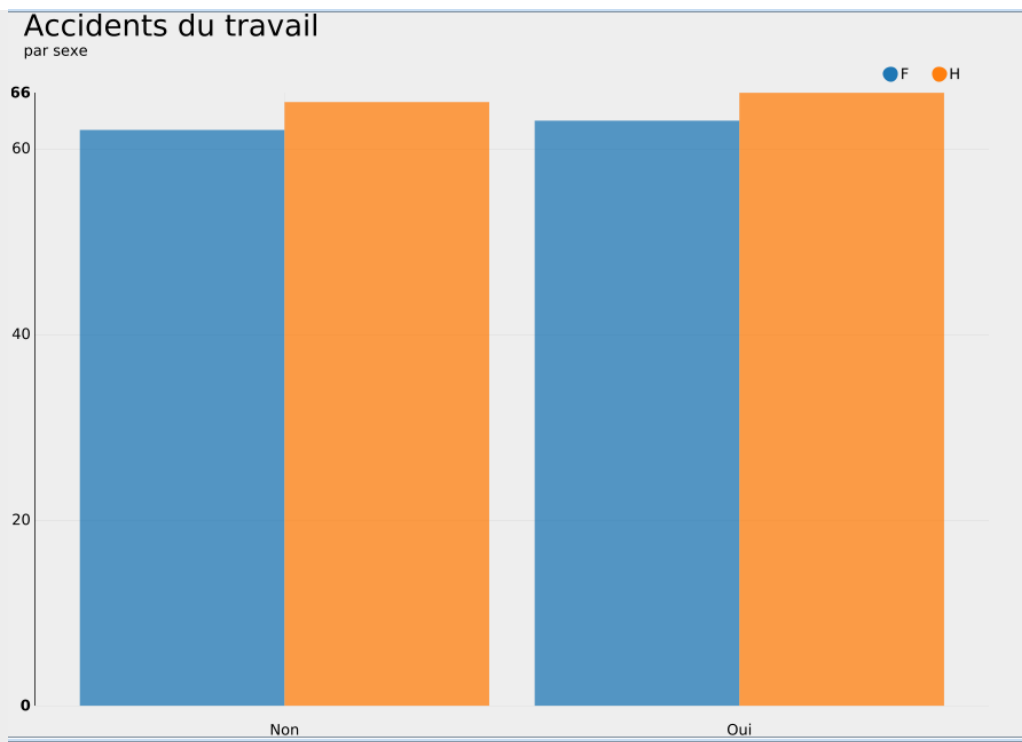
Independent Groups Statistics

Confidence Interval (CI) Probability: 95.0%

Differences are reported of the groups: F - H

	Variance Assumption	t	df	p-value (2-tailed)	Mean Difference	Standard Error Difference	CI (Lower Bound)	CI (Upper Bound)
Salaire base mensuel	Equal variances assumed	-0.3702	254	0.7115	-116.1702	313.8042	-734.1597	501.8193
Salaire base mensuel	Equal variances not assumed	-0.3706	253.9995	0.7112	-116.1702	313.447	-733.4563	501.1159

Sécurité au travail



Workflow KNIME égalité homme/femme

