

UNIVERSITE EPISCPALE D’HAITI

14,RUE LEGETIME,CHAMPS DE MARS(A COTE DU MUSEE D’ART HAITIEN

TEL :22277962-31298661

EMAIL :unephhaiti@hotmail.com

NOM  : Juillet

PRENOM : Jasnel

DATE  : 18 OCTOBRE 2021

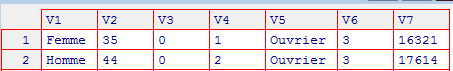
DEVOIR  : STATISTIQUE

PROF : GABRIEL DUVALSSAINT

1. **Importation et mise en forme**a) Importer les jeux de données

monFichier=read.table(file.choose(), header=FALSE,sep=",")

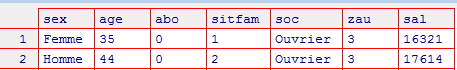
affichage : **View(**monFichier**)**



b) Renommer les variables du fichier importer

names(monFichier)[c(1,2,3,4,5,6,7)]=c("sex","age","abo","sitfam","soc","zau","sal")

affichage : **View(**monFichier**)**

****

c)Combien y-a-t-il de variables et d’observations.

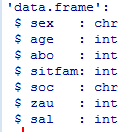
affichage : **str(**monFichier**)**

****

Cette table est composée de 7 variables qui sont les éléments présentes en lignes et 200 observations qui sont représentés en colonne.

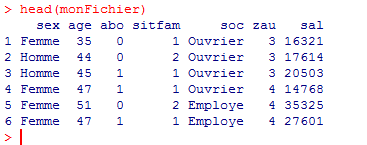
d) Dresser le dictionnaire des variables c.-à-d. présenter un tableau résumant le nom et le type de chaque variable.

affichage : **str(**monFichier**)**



e) Afficher les 6 premières lignes de chaque variable.

affichage : **head(monFichier)**



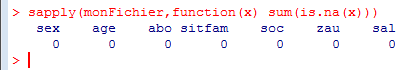
1. **Traitement des valeurs manquantes**

a) Vérifier s’il y a des valeurs manques dans le jeu de données. Si oui, pouvez-vous identifier à quelles variables correspondent ces valeurs manquantes.



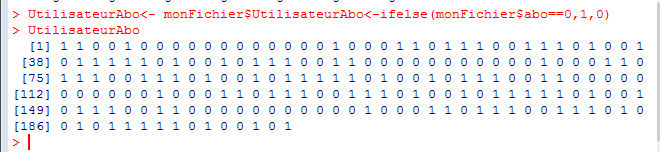
Il n ya pas de valeurs manques dans le jeu de données.

b) Y-a-t-il des valeurs aberrantes dans le jeu de données.

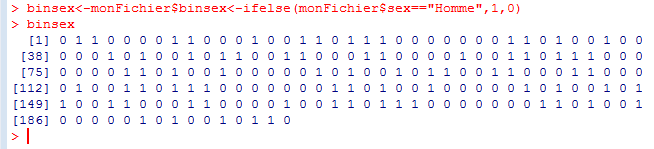


Il n ya pas des valeurs aberrantresdans le jeu de données.

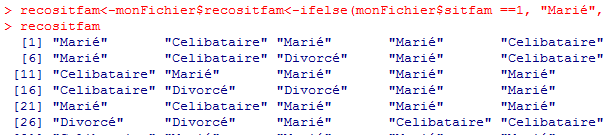
1. **Création de variables**
2. Une variable prenant la valeur 1 lorsque l’individu ne désire pas s’abonner



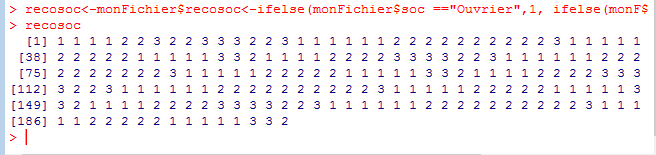
1. Des variables binaire associées à la variable caractérisant le sexe

****

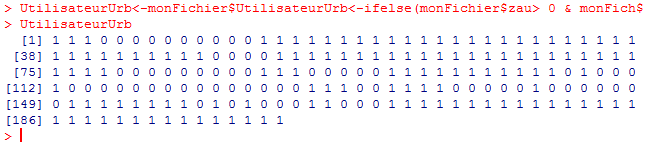
1. Recodifier la variable caractérisant la situation familiale : 1- Marié, 2-Celibataire 3-Divorcé



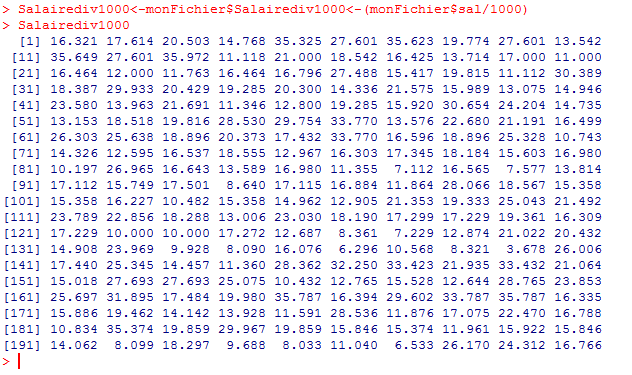
1. Recodifier à la variable caractérisant la CSP : Ouvrier-> 1, Employé-> 2, Cadre-> 3

****

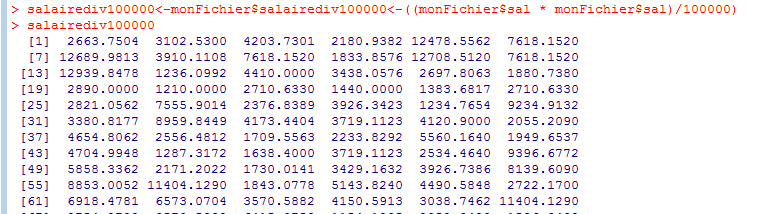
1. Une variable prenant la valeur 1 lorsque l’individu vie en zone urbaine

****

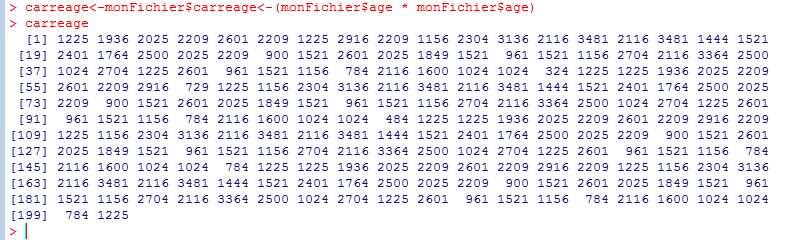
1. Une variable égale au salaire divisé par 1000



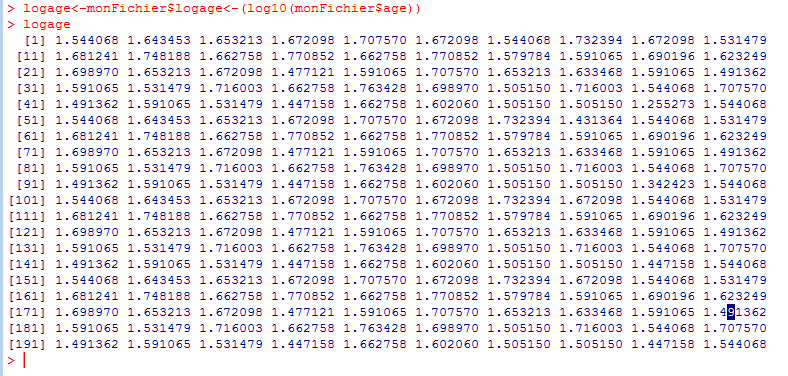
1. Une variable égale au salaire au carré divisé par 100000



1. Une variable égale au carré de l’âge



1. Une variable égale au logarithme de l’âge



1. Transformer la variable salaire en tranche de salaire en utilisant votre propre estimation de classes.
2. **Analyse descriptive**

**a)** Calculer la moyenne, médiane l’écart-type, minimum, maximum et les quartiles pour l’ensemble des variables continues du fichier de la base

Pour ce faire j’utilise les formules suivantes :

Moyenne : mean(monFichier$variable)

Médian : median(monFichier$ variable)

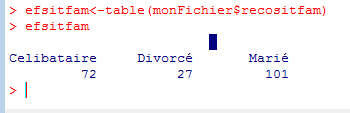
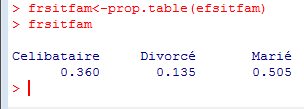
Max: max(monFichier$variable)

Min: min(monFichier$variable)

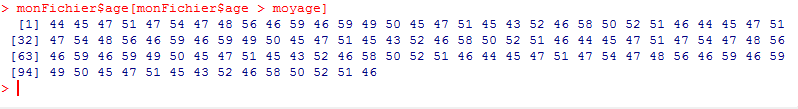
Ecart type : sd(monFichier$variable)

Quartiles : quantile(monFichier$variable)

**b)** Présenter un tableau d’effectif et de fréquence des variables « sitfam », « soc » qui sont recofidifié et « abo »

****

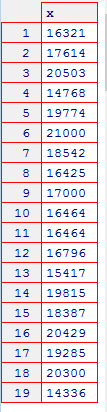
1. **Extraction**
2. Afficher tous les ages qui sont supérieur à la moyenne de cette même variable.



b) Afficher tous les salaires compris entre le premier et le troisième quartile

salaireComPreEtTroQua<-monFichier$sal[monFichier$sal > qsal[c(2)] & monFichier$sal < qsal[c(4)] ]

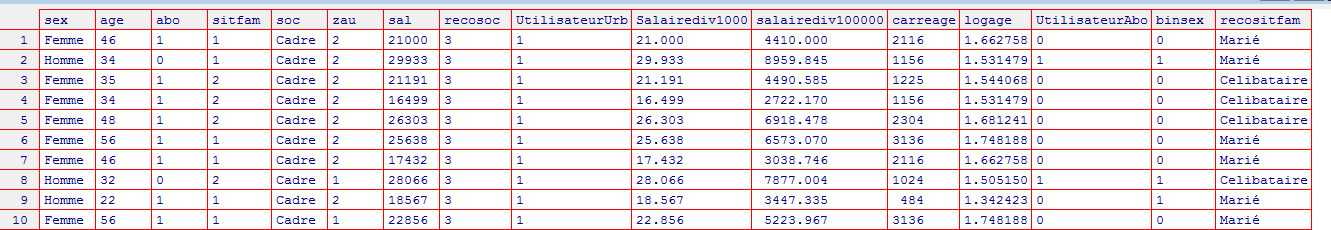
View(salaireComPreEtTroQua)



c) Afiicher la liste des cadres qui habite en zone urbaine

listCadre<-monFichier[monFichier$soc=="Cadre" & monFichier$UtilisateurUrb ==1,]

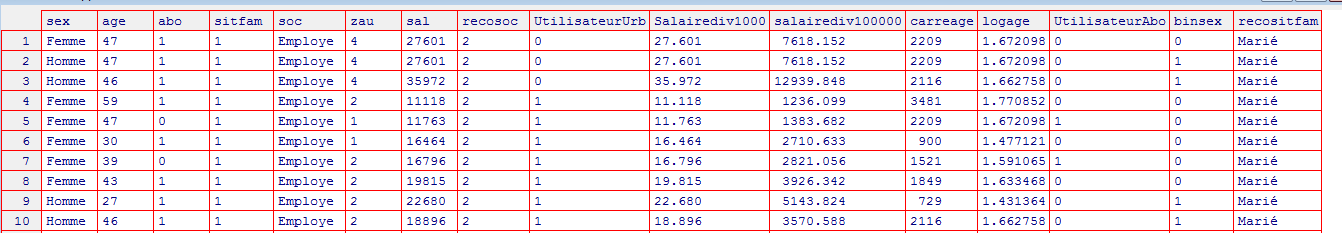
View(listCadre)



d) Afiicher la liste des employés qui sont mariés

lisemppasmarie<-monFichier[monFichier$soc=="Employe" & monFichier$recositfam == "Marié",]

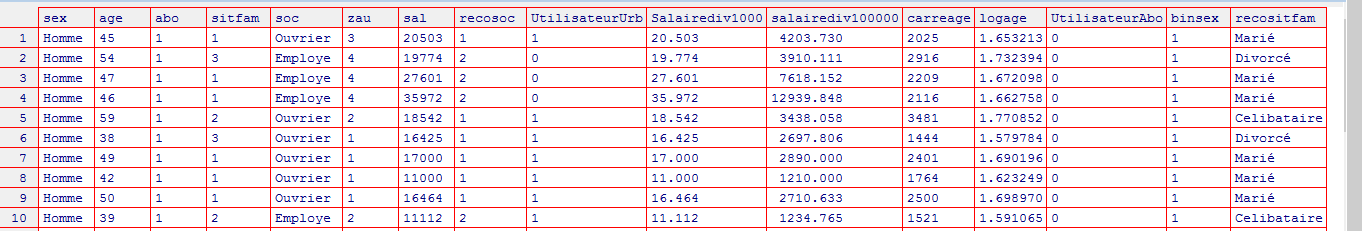
View(lisemppasmarie)



e) Extraire la liste des abonnées qui sont des hommes

listAboHom<-monFichier[monFichier$abo==1 & monFichier$sex=="Homme",]

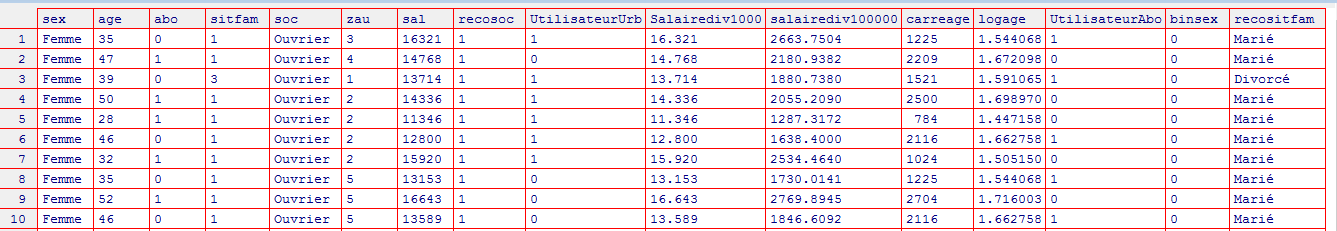
View(listAboHom)



f) Extraire la liste des ouvriers qui sont des femmes et qui touche un salaire inferieure à la médiane de cette dernière variable

listOuvFemmSalInfMed<-monFichier[monFichier$soc=="Ouvrier" & monFichier$sex=="Femme" & monFichier$sal < medsal,]

View(listOuvFemmSalInfMed)



g) la liste des abonnés célibataires dont l’age ne dépasse pas la moyenne d’age.

listAboCelAgeInfMoyAge<-monFichier[monFichier$abo==1 & monFichier$recositfam == "Celibataire" & monFichier$age < moyage,]

View(listAboCelAgeInfMoyAge)

