

EXERCICE 1: Noms des animaux de compagnie de la ville de Seattle

Question: 1

```
dat <- read.table("Pet_Licenses.txt", sep=";", header=TRUE,
stringsAsFactors=FALSE, quote="\\"", fill=TRUE)
```

Question: 2 (nombre de variables et le nombre d'individus du jeu de données)

```
dim(dat)
```

```
## [1] 43683      7
```

```
head(dat)
```

```
##  License.Issue.Date License.Number Animal.s.Name Species
Primary.Breed
## 1  December 18 2015      S107948      Zen      Cat  Domestic
Longhair
## 2      June 14 2016      S116503      Misty     Cat
Siberian
## 3      August 04 2016     S119301      Lyra      Cat
Mix
## 4  February 13 2019      962273      Veronica  Cat  Domestic
Longhair
## 5      August 10 2019     S133113      Spider    Cat
LaPerm
## 6  November 21 2019     8002549      Maxx      Cat  American
Shorthair
##  Secondary.Breed ZIP.Code
## 1      Mix      98117
## 2      98117
## 3      98121
## 4      98107
## 5      98115
## 6      98125
```

Il y'a 7 variables et 43683 individus dans ce jeu de données

Question: 3 (nombre de modalités de la variable species et le nombre d'individus possédant chaque modalité)

Nous avons utilisé la fonction `table()` pour compter le nombre d'individus associés à chaque catégorie. Les résultats obtenus montrent la répartition suivante :

Cat : 13 935 individus

Dog : 29 729 individus

Goat : 16 individus

Pig : 3 individus

Cette distribution met en évidence une grande différence dans le nombre d'individus selon les espèces. Les catégories Cat et Dog sont largement dominantes, représentant la majorité des observations, tandis que Goat et Pig sont des catégories rares avec un nombre très limité d'individus.

```
dat$Species <- as.factor(dat$Species)
levels(dat$Species)
```

```
## [1] "Cat" "Dog" "Goat" "Pig"
```

```
table(dat$Species)
```

```
##
```

```
##   Cat   Dog  Goat   Pig
```

```
## 13935 29729    16     3
```

Question: 4(les 10 noms des chiens et les 10 noms des chats les plus populaires)

```
Chiens_popu<-names(sort(table(dat[dat$Species=="Dog",]
$Animal.s.Name),decreasing =TRUE))[1:10]
```

```
chats_popu<-names(sort(table(dat[dat$Species=="Cat",]
$Animal.s.Name),decreasing =TRUE))[1:10]
```

```
print("Les 10 chiens populaires :")
```

```
## [1] "Les 10 chiens populaires :"
```

```
print(Chiens_popu)
```

```
## [1] "Luna" "Charlie" "Lucy" "Daisy" "Bella" "Penny" "Ruby"
```

```
## [8] "Rosie" "Milo" "Cooper"
```

```
print("Les 10 chats populaires:")
```

```
## [1] "Les 10 chats populaires:"
```

```
print(chats_popu)
```

```
## [1] "Luna" "Lucy" "Charlie" "Bella" "Lily" "Loki" "Oliver"
```

```
## [8] "Leo" "Pepper" "Max"
```

******Pour avoir les 5 races primaires de chien on fait presque la meme chose

```
names(sort(table(dat[dat$Species=="Dog",]$Primary.Breed),decreasing =TRUE))
[1:5]
```

```
## [1] "Retriever, Labrador"    "Retriever, Golden"      "Chihuahua, Short
Coat"
## [4] "German Shepherd"        "Poodle, Miniature"
```

Question: 5 jeu de données dat_pop ne contenant que les licences de chiens de compagnie # dont le nom et la race primaire sont parmi les 30 plus populaires.

```
A<- names(sort(table(dat[dat$Species == "Dog", ]$Animal.s.Name), decreasing =
TRUE))[1:30]
B<- names(sort(table(dat[dat$Species == "Dog", ]$Primary.Breed), decreasing =
TRUE))[1:30]
dat_pop=dat[dat$Species == "Dog" & dat$Animal.s.Name %in% A &
dat$Primary.Breed %in% B, ]
dim(dat_pop)

## [1] 2708    7
```

Question: 6 On retire les modalités inutilisées des variables Animal.s.Name et Primary.Breed

```
dat_pop$Animal.s.Name <- as.factor(dat_pop$Animal.s.Name)
dat_pop$Primary.Breed<- as.factor(dat_pop$Primary.Breed)
dat_pop$Animal.s.Name <- droplevels(dat_pop$Animal.s.Name)
dat_pop$Primary.Breed <- droplevels(dat_pop$Primary.Breed)
length(levels(dat_pop$Animal.s.Name))

## [1] 30

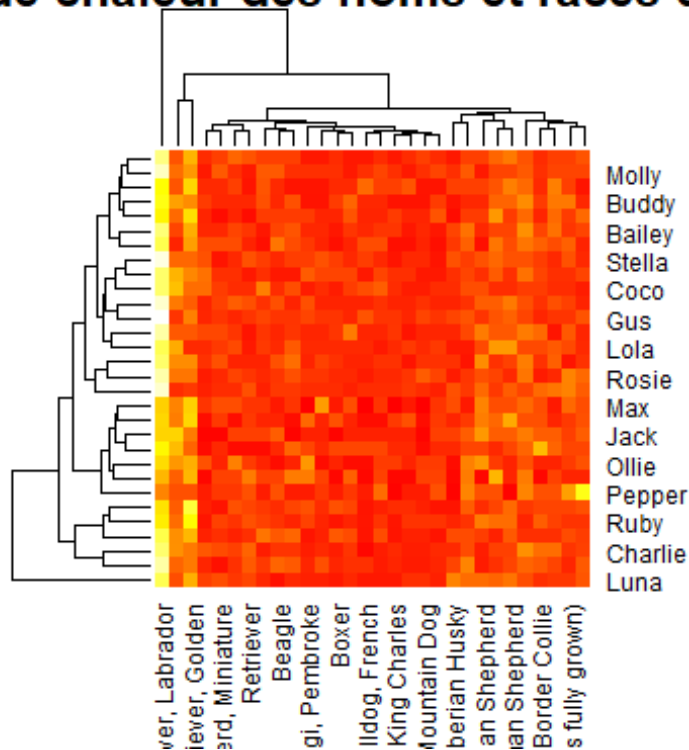
length(levels(dat_pop$Primary.Breed))

## [1] 30
```

Question 7: Table de contingence entre les variables Animal.s.Name et Primary.Breed

```
table_contingence <- table(dat_pop$Animal.s.Name, dat_pop$Primary.Breed)
heatmap(table_contingence, main="Carte de chaleur des noms et races de
chiens", col=heat.colors(256))
```

Carte de chaleur des noms et races de chien



```
table_contingence=table(dat_pop$Animal.s.Name, dat_pop$Primary.Breed)
print(table_contingence)
```

```
##
##      Australian Cattle Dog Australian Shepherd
##  Bailey                  3                   1
##  Bella                   3                   3
##  Buddy                   1                   1
##  Charlie                  5                   2
##  Coco                     0                   3
##  Cooper                   1                   0
##  Daisy                     5                   2
##  Gus                      3                   1
##  Jack                     3                   5
##  Leo                      0                   2
##  Lily                     2                   1
##  Lola                     1                   3
##  Lucy                     7                   1
##  Luna                    9                  12
##  Maggie                   3                   2
##  Max                      3                   5
##  Milo                     1                   6
##  Molly                    2                   0
##  Olive                    0                   4
##  Oliver                   2                   2
##  Ollie                    3                   6
##  Penny                    2                   8
```

##	Pepper	10	3
##	Poppy	4	2
##	Rosie	6	4
##	Ruby	2	7
##	Sadie	3	5
##	Scout	5	6
##	Stella	1	3
##	Teddy	1	0

##		Australian Shepherd,	Miniature Beagle	Bernese Mountain Dog
##	Bailey	1	3	1
##	Bella	2	5	1
##	Buddy	1	1	0
##	Charlie	6	4	2
##	Coco	2	1	0
##	Cooper	3	5	1
##	Daisy	0	2	1
##	Gus	1	1	2
##	Jack	0	3	0
##	Leo	1	1	0
##	Lily	1	0	0
##	Lola	2	1	1
##	Lucy	0	8	1
##	Luna	8	0	4
##	Maggie	2	2	2
##	Max	3	1	0
##	Milo	3	1	1
##	Molly	3	3	1
##	Olive	2	3	1
##	Oliver	1	1	0
##	Ollie	1	0	2
##	Penny	4	4	2
##	Pepper	2	3	1
##	Poppy	0	2	1
##	Rosie	1	2	5
##	Ruby	3	2	2
##	Sadie	2	1	0
##	Scout	1	2	1
##	Stella	0	1	1
##	Teddy	2	2	2

##		Border Collie	Boxer	Bulldog,	French Chihuahua,	Short Coat
##	Bailey	3	0		1	3
##	Bella	2	5		0	9
##	Buddy	2	2		1	7
##	Charlie	9	2		0	12
##	Coco	2	0		3	10
##	Cooper	4	0		3	1
##	Daisy	4	5		0	6
##	Gus	0	1		3	1

##	Jack	3	0	1	7
##	Leo	1	1	4	3
##	Lily	2	2	0	7
##	Lola	3	1	1	9
##	Lucy	7	2	3	8
##	Luna	4	1	1	8
##	Maggie	1	1	0	3
##	Max	3	1	0	5
##	Milo	5	2	1	4
##	Molly	1	1	1	4
##	Olive	1	0	0	6
##	Oliver	6	2	2	6
##	Ollie	1	1	0	6
##	Penny	2	2	1	8
##	Pepper	3	1	1	3
##	Poppy	1	1	1	4
##	Rosie	0	1	0	7
##	Ruby	5	0	0	6
##	Sadie	3	3	0	5
##	Scout	6	1	1	1
##	Stella	2	2	2	4
##	Teddy	0	0	0	3

##					
##		German Shepherd	Havanese	Maltese	Mix
##	Bailey	3	4	1	3
##	Bella	5	6	1	1
##	Buddy	4	3	2	1
##	Charlie	5	4	7	5
##	Coco	1	3	5	1
##	Cooper	3	3	3	3
##	Daisy	7	4	1	5
##	Gus	4	1	0	0
##	Jack	3	0	0	2
##	Leo	5	0	1	2
##	Lily	1	0	4	1
##	Lola	8	4	1	3
##	Lucy	5	7	2	3
##	Luna	13	3	6	5
##	Maggie	5	2	1	4
##	Max	3	2	2	3
##	Milo	4	2	3	2
##	Molly	2	1	0	1
##	Olive	6	0	0	4
##	Oliver	2	3	0	0
##	Ollie	1	1	3	5
##	Penny	3	6	2	1
##	Pepper	0	1	1	2
##	Poppy	2	2	1	1
##	Rosie	1	4	0	3
##	Ruby	6	1	0	2

##	Sadie	7	0	0	0
##	Scout	5	2	0	1
##	Stella	3	2	3	1
##	Teddy	1	4	0	2

Mixed Breed, Large (over 44 lbs fully grown)

##	Bailey	2
##	Bella	4
##	Buddy	5
##	Charlie	4
##	Coco	1
##	Cooper	4
##	Daisy	4
##	Gus	1
##	Jack	2
##	Leo	2
##	Lily	1
##	Lola	3
##	Lucy	2
##	Luna	5
##	Maggie	2
##	Max	4
##	Milo	4
##	Molly	2
##	Olive	2
##	Oliver	2
##	Ollie	2
##	Penny	0
##	Pepper	6
##	Poppy	2
##	Rosie	8
##	Ruby	2
##	Sadie	2
##	Scout	8
##	Stella	3
##	Teddy	1

Mixed Breed, Medium (up to 44 lbs fully grown)

##	Bailey	6
##	Bella	4
##	Buddy	1
##	Charlie	6
##	Coco	3
##	Cooper	4
##	Daisy	6
##	Gus	3
##	Jack	3
##	Leo	2
##	Lily	2
##	Lola	0

##	Lucy	11	
##	Luna	12	
##	Maggie	2	
##	Max	1	
##	Milo	3	
##	Molly	3	
##	Olive	1	
##	Oliver	2	
##	Ollie	5	
##	Penny	0	
##	Pepper	5	
##	Poppy	2	
##	Rosie	2	
##	Ruby	2	
##	Sadie	2	
##	Scout	2	
##	Stella	2	
##	Teddy	4	
##			
##	Mixed Breed, Small (under 24 lbs fully grown)	Poodle, Miniature	
##	Bailey	0	8
##	Bella	0	2
##	Buddy	2	6
##	Charlie	3	12
##	Coco	1	2
##	Cooper	0	5
##	Daisy	0	3
##	Gus	1	2
##	Jack	2	4
##	Leo	0	4
##	Lily	0	4
##	Lola	2	3
##	Lucy	2	8
##	Luna	2	8
##	Maggie	1	3
##	Max	2	4
##	Milo	1	6
##	Molly	2	3
##	Olive	0	3
##	Oliver	1	3
##	Ollie	2	3
##	Penny	2	7
##	Pepper	3	5
##	Poppy	3	4
##	Rosie	3	7
##	Ruby	2	1
##	Sadie	2	3
##	Scout	0	3
##	Stella	0	3
##	Teddy	2	4

##		Poodle, Standard	Retriever	Retriever, Golden	Retriever, Labrador
##	Bailey	3	2	10	18
##	Bella	1	2	12	19
##	Buddy	4	1	6	12
##	Charlie	5	8	10	25
##	Coco	1	1	5	16
##	Cooper	5	1	7	13
##	Daisy	4	4	21	17
##	Gus	3	1	6	20
##	Jack	3	2	4	7
##	Leo	3	0	9	11
##	Lily	3	0	5	12
##	Lola	8	0	1	16
##	Lucy	3	7	10	28
##	Luna	11	7	21	35
##	Maggie	4	3	8	14
##	Max	3	2	8	8
##	Milo	4	0	4	17
##	Molly	2	0	6	14
##	Olive	4	3	4	20
##	Oliver	1	0	3	8
##	Ollie	2	2	7	9
##	Penny	2	1	8	18
##	Pepper	3	3	3	5
##	Poppy	6	0	9	10
##	Rosie	3	2	7	22
##	Ruby	6	4	16	15
##	Sadie	4	2	9	9
##	Scout	2	4	2	24
##	Stella	4	3	4	15
##	Teddy	6	4	6	6

##		Schnauzer, Miniature	Shepherd	Shih Tzu	Siberian Husky
##	Bailey	1	0	2	3
##	Bella	1	1	4	4
##	Buddy	0	0	4	0
##	Charlie	3	2	4	3
##	Coco	1	0	2	1
##	Cooper	1	0	1	2
##	Daisy	2	2	4	0
##	Gus	1	0	0	2
##	Jack	1	1	1	2
##	Leo	0	3	1	2
##	Lily	2	1	2	1
##	Lola	1	2	1	0
##	Lucy	6	1	3	5
##	Luna	2	4	3	15
##	Maggie	0	1	0	0
##	Max	6	1	3	3

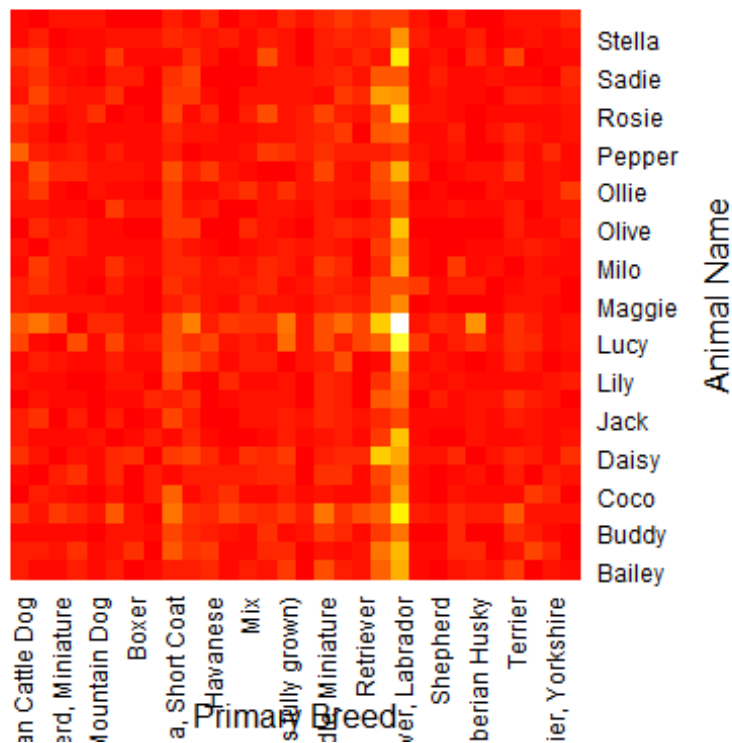
##	Milo	1	0	6	1
##	Molly	1	0	0	1
##	Olive	0	0	0	0
##	Oliver	1	1	2	1
##	Ollie	0	1	0	0
##	Penny	3	0	1	2
##	Pepper	2	1	2	0
##	Poppy	1	1	3	0
##	Rosie	2	2	1	0
##	Ruby	1	2	1	1
##	Sadie	1	3	1	1
##	Scout	1	2	0	4
##	Stella	3	1	1	3
##	Teddy	2	0	3	0
##					
##		Spaniel, Cavalier King Charles Terrier Terrier, American Pit Bull			
##	Bailey		2	6	
0					
##	Bella		0	2	
7					
##	Buddy		0	5	
3					
##	Charlie		3	9	
2					
##	Coco		1	1	
6					
##	Cooper		0	3	
0					
##	Daisy		2	4	
3					
##	Gus		1	1	
2					
##	Jack		1	3	
2					
##	Leo		2	5	
3					
##	Lily		1	1	
1					
##	Lola		1	4	
2					
##	Lucy		2	5	
4					
##	Luna		1	5	
3					
##	Maggie		0	2	
2					
##	Max		0	1	
2					
##	Milo		2	0	

2			
##	Molly	1	2
3			
##	Olive	0	3
1			
##	Oliver	1	3
0			
##	Ollie	2	1
1			
##	Penny	2	5
1			
##	Pepper	0	3
1			
##	Poppy	2	4
2			
##	Rosie	1	1
1			
##	Ruby	0	3
3			
##	Sadie	2	1
1			
##	Scout	1	7
0			
##	Stella	0	1
2			
##	Teddy	0	2
2			
##			
##	Terrier, Yorkshire Welsh Corgi, Pembroke		
##	Bailey	2	1
##	Bella	4	0
##	Buddy	1	0
##	Charlie	2	2
##	Coco	4	0
##	Cooper	3	2
##	Daisy	1	5
##	Gus	1	2
##	Jack	1	1
##	Leo	2	0
##	Lily	2	3
##	Lola	0	1
##	Lucy	1	2
##	Luna	1	2
##	Maggie	1	0
##	Max	2	0
##	Milo	1	1
##	Molly	2	1
##	Olive	1	3
##	Oliver	2	2
##	Ollie	2	6

```
## Penny 2 1
## Pepper 4 1
## Poppy 1 2
## Rosie 1 2
## Ruby 2 3
## Sadie 0 4
## Scout 1 1
## Stella 1 2
## Teddy 2 4
```

```
heatmap(as.matrix(table_contingence),
        Rowv = NA,
        Colv = NA,
        scale = "none",
        col = heat.colors(256),
        xlab = "Primary Breed",
        ylab = "Animal Name",
        main = "Carte de chaleur de Animal.s.Name vs Primary.Breed")
```

chaleur de Animal.s.Name vs Primary.Bre

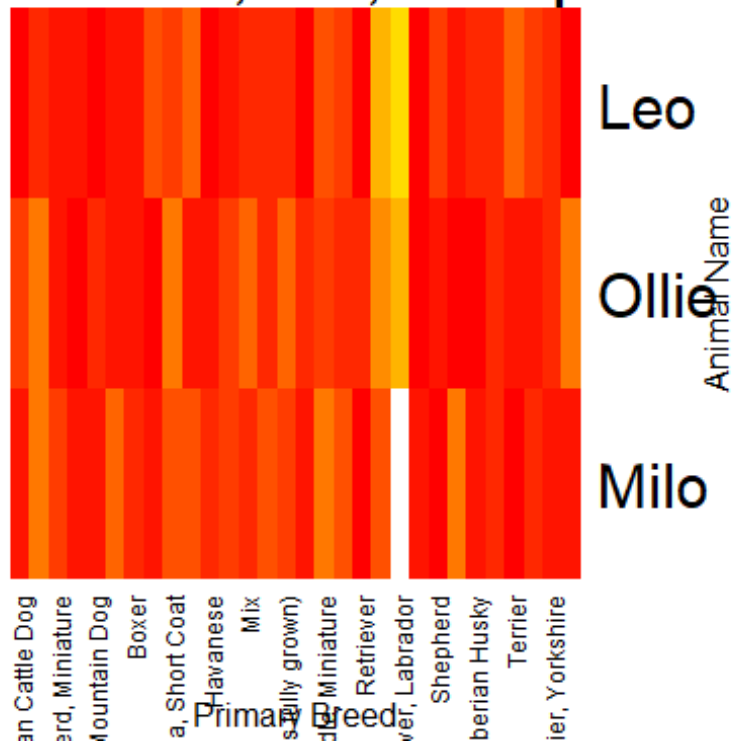


Question 8 : Ici on a filtré la table pour mieux visualiser et on trouve que (La couleur jaune représente les valeurs les plus élevées et la couleur plus foncée (vers le rouge) représente des valeurs plus faibles. Et on remarque que pour les 3 noms Milo, Ollie, Leo la race de chien la plus populaire est le Labrador

```
table_filtrée <- table_contingence[c("Milo", "Ollie", "Leo"), ]
heatmap(as.matrix(table_filtrée),
        Rowv = NA,
```

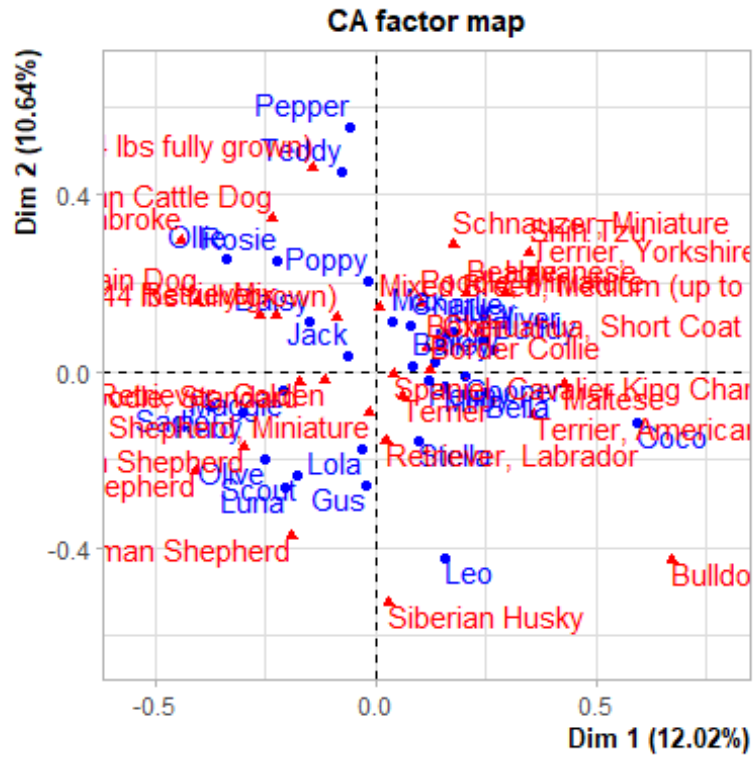
```
Colv = NA,
scale = "none",
col = heat.colors(256),
xlab = "Primary Breed",
ylab = "Animal Name",
main = "Popularité des noms Milo, Ollie, et Leo par race de chien")
```

des noms Milo, Ollie, et Leo par race de cl



#Question 9

```
library(FactoMineR)
afc_result <- CA(table_contingence, graph=TRUE)
```



Le pourcentage d'inertie est= 12,02+10,64=22,66 qui est tres faible et donc la qualité de cette AFC n'est pas bonne