

Synthèse des résultats du TD1 R

TD1 – 2 [Données attributaires disponibles des IRIS parisiens]

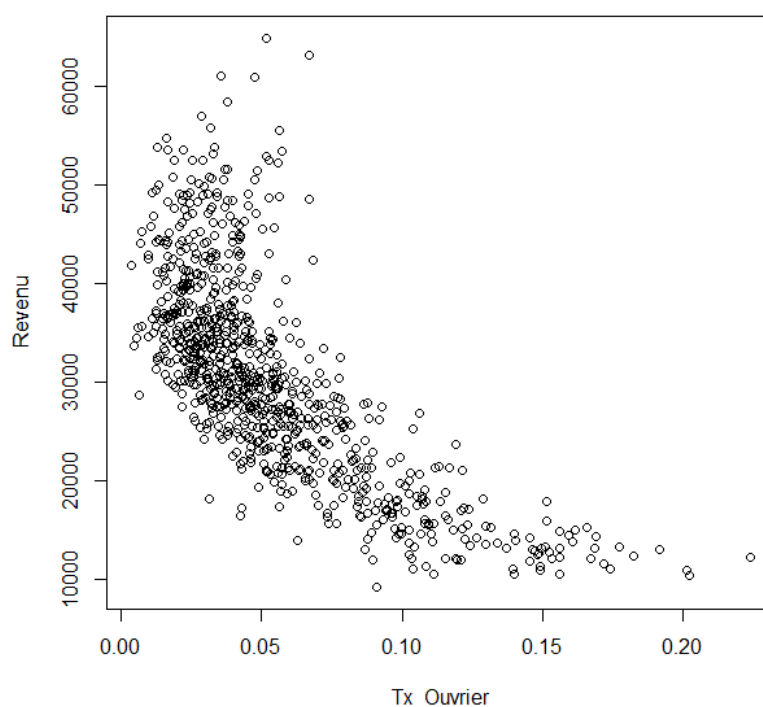
```

OBJECTID  N_SQ_IR  C_IR  L_IR  M2_POP  SHAPE_Area
0         20 750004033 751186922 Grandes Carrières 22 42891.78 80920.79
1         21 750004043 751186923 Grandes Carrières 23 51939.49 68661.34
2         22 750004027 751186924 Grandes Carrières 24 43233.73 73724.51
3         23 750004026 751186925 Grandes Carrières 25 37176.58 81608.21
4         24 750004862 751186926 Grandes Carrières 26 38308.33 176037.03
5         25 750004020 751186927 Grandes Carrières 27 21445.84 44308.01

Tx_Mono Tx_Ouvrier Revenu X Y
0 0.1699 0.0517 31376 651225.3 6866376
1 0.3265 0.0875 21229 650861.0 6866264
2 0.3491 0.0701 17436 650950.0 6866511
3 0.5014 0.1773 13180 651420.2 6866557
4 0.5323 0.1397 10465 651645.9 6866858
5 0.5561 0.1739 10986 651526.2 6866707

```

TD1 – 4 [Affichage des revenus des IRIS parisiens en fonction de leur taux d'ouvrier]



TD1 – 5 [Résultat de la régression linéaire simple : Y Revenus des IRIS parisiens ; X Taux d'ouvrier des IRIS parisiens]

```

Call:
lm(formula = Revenu ~ Tx_Ouvrier)

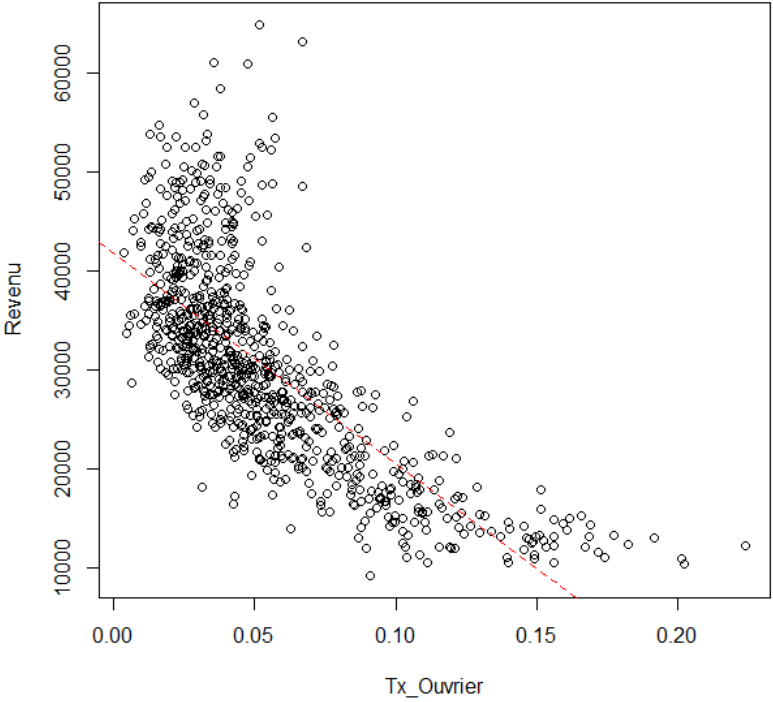
Residuals:
    Min       1Q   Median       3Q      Max
-17027  -4695  -1308    3146   35569

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  41869.8     440.6   95.02  <2e-16 ***
Tx_Ouvrier  -212907.8    6832.6  -31.16  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

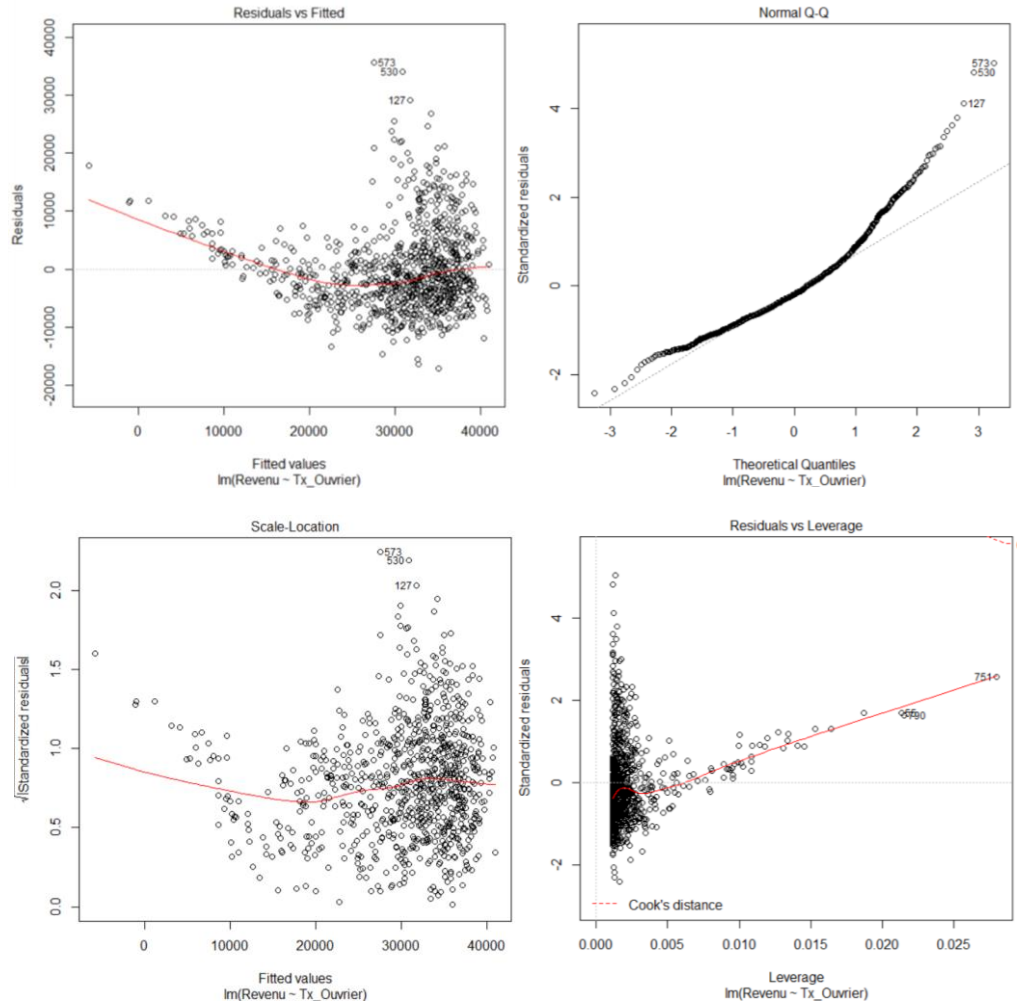
Residual standard error: 7083 on 862 degrees of freedom
Multiple R-squared:  0.5297,    Adjusted R-squared:  0.5292
F-statistic:  971 on 1 and 862 DF,  p-value: < 2.2e-16

```

TD1 – 7 [Affichage des revenus des IRIS parisiens en fonction de leur taux d'ouvrier et de la régression linéaire simple associée]



TD1 – 8 [Analyse des résidus de la régression précédente]



TD1 – 9 [Comparaison corrélation Spearman et Pearson]

```
> cor(Tx_Ouvrier, Revenu, method = "pearson")^2
[1] 0.5297259
> cor(Tx_Ouvrier, Revenu, method = "spearman")^2
[1] 0.5848331
```

TD1 – 11 [Résultat de la régression linéaire simple : Y Log des revenus des IRIS parisiens ; X Taux d'ouvrier des IRIS parisiens]

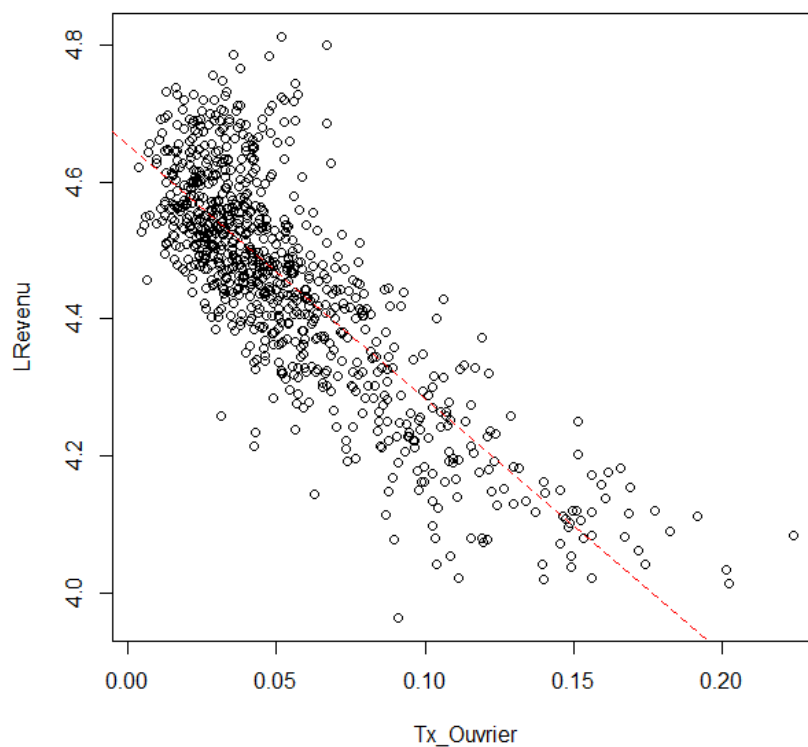
```
Call:
lm(formula = LRevenu ~ Tx_Ouvrier)

Residuals:
    Min       1Q   Median       3Q      Max
-0.35424 -0.05816 -0.00639  0.05106  0.39413

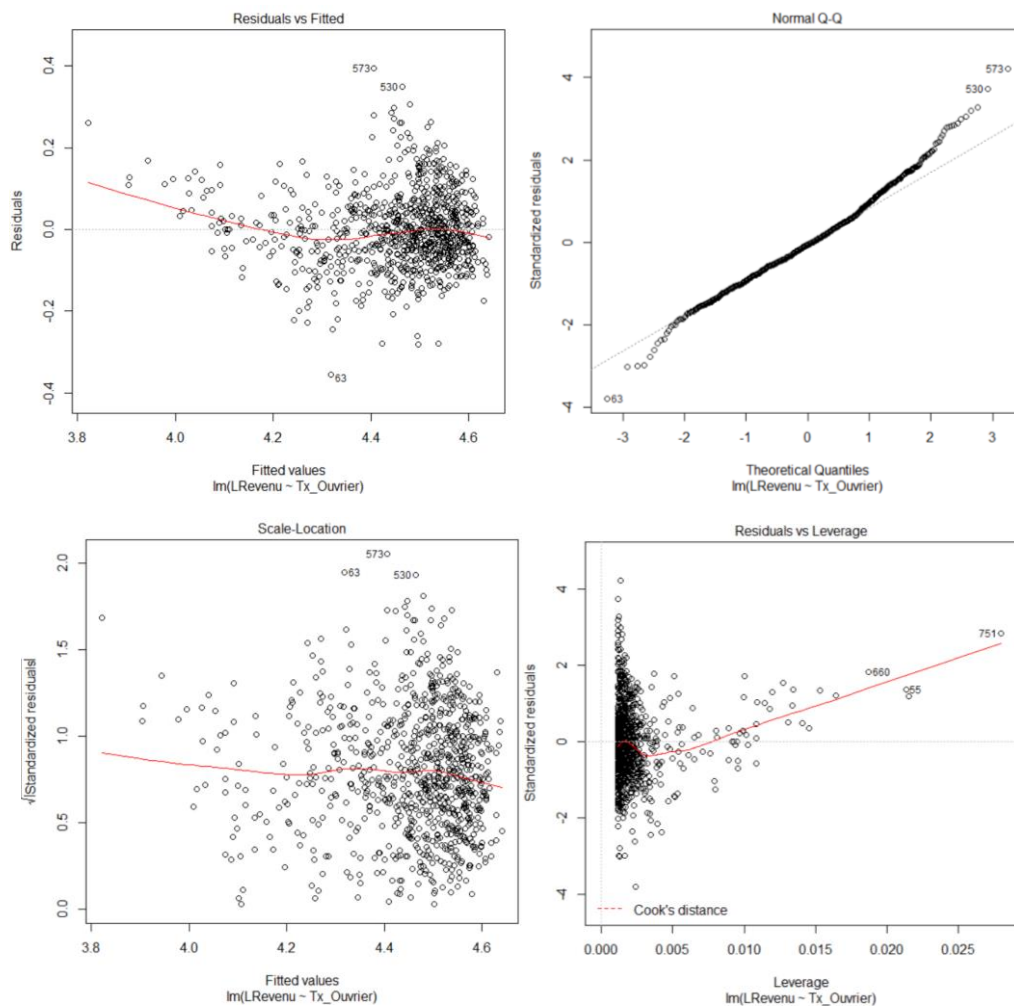
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.655753   0.005813   800.99  <2e-16 ***
Tx_Ouvrier   -3.721544   0.090131  -41.29  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09343 on 862 degrees of freedom
Multiple R-squared:  0.6642,    Adjusted R-squared:  0.6638
F-statistic: 1705 on 1 and 862 DF,  p-value: < 2.2e-16
```

TD1 – 11 [Affichage des logarithmes des revenus des IRIS parisiens en fonction de leur taux d'ouvrier et de la régression associée]



TD1 – 11 [Analyse des résidus de la régression précédente]



TD1 – 12 [Résultat régression linéaire multiple : Y Log des revenus des IRIS parisiens ; X1 Taux d'ouvrier des IRIS parisiens ; X2 Taux de familles monoparentales]

Call:

```
lm(formula = LRevenu ~ Tx_Ouvrier + Tx_Mono)
```

Residuals:

	Min	1Q	Median	3Q	Max
Residuals	-0.29297	-0.05558	-0.00878	0.04586	0.36247

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.687910	0.006157	761.37	<2e-16 ***
Tx_Ouvrier	-2.935410	0.109983	-26.69	<2e-16 ***
Tx_Mono	-0.348433	0.031298	-11.13	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08741 on 861 degrees of freedom

Multiple R-squared: 0.7064, Adjusted R-squared: 0.7058

F-statistic: 1036 on 2 and 861 DF, p-value: < 2.2e-16

TD1 – 13 [Comparaison de nos deux modèles emboîtés concernant les logarithmes des revenus]

Analysis of Variance Table

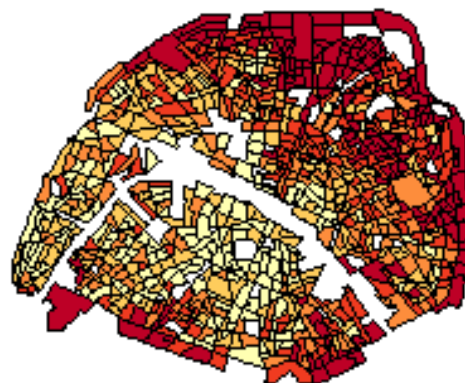
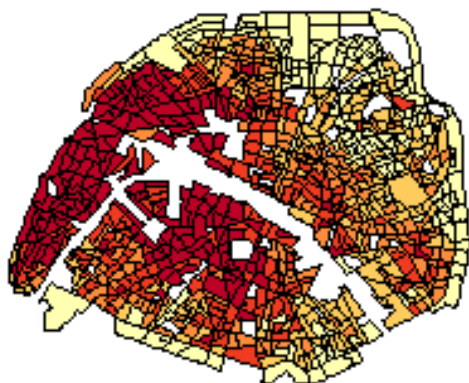
Model 1: LRevenu ~ Tx_Ouvrier + Tx_Mono

Model 2: LRevenu ~ Tx_Ouvrier

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	861	6.5781				
2	862	7.5250	-1	-0.94689	123.94	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TD1 – 14 [Affichage cartographique des logarithmes des revenus à gauche et des taux d'ouvrier à droite – 5 classes selon la méthode des quantiles]



Synthèse résultats TD2 R

TD2 – 2 [Analyse des variables Revenus et des coordonnées X des centroïdes des IRIS parisiens]

```
> summary(Revenu)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  9196  23240   30123   30375   36482   64854

> summary(X)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
645404  649702  652064  651840  654180  656987
```

TD2 – 3 et 4 [Ajout de champs qualitatifs « Riche » et « Ouest »]

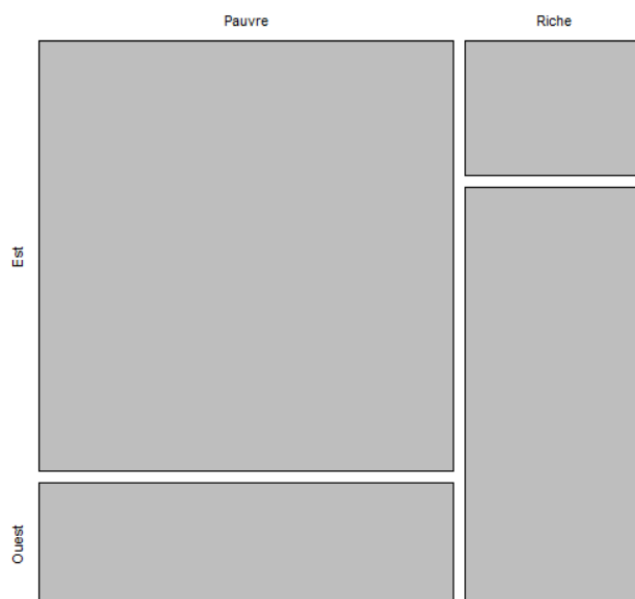
```
OBJECTID  N_SQ_IR  C_IR  L_IR  M2_POP SHAPE_Area Tx_Mono
0         20 750004033 751186922 Grandes CarriÃ`res 22 42891.78  80920.79  0.1699
1         21 750004043 751186923 Grandes CarriÃ`res 23 51939.49  68661.34  0.3265
2         22 750004027 751186924 Grandes CarriÃ`res 24 43233.73  73724.51  0.3491
3         23 750004026 751186925 Grandes CarriÃ`res 25 37176.58  81608.21  0.5014
4         24 750004862 751186926 Grandes CarriÃ`res 26 38308.33  176037.03  0.5323
5         25 750004020 751186927 Grandes CarriÃ`res 27 21445.84  44308.01  0.5561

Tx_Ouvrier Revenu      X      Y Riche Ouest
0         0.0517  31376 651225.3 6866376 Pauvre Est
1         0.0875  21229 650861.0 6866264 Pauvre Ouest
2         0.0701  17436 650950.0 6866511 Pauvre Ouest
3         0.1773  13180 651420.2 6866557 Pauvre Est
4         0.1397  10465 651645.9 6866858 Pauvre Est
5         0.1739  10986 651526.2 6866707 Pauvre Est
```

TD2 – 5 [Tableau de contingence entre la variable qualitative « Riche » et la variable qualitative « Ouest »]

```
      Est Ouest
Pauvre 468   133
Riche   64   199
```

TD2 – 6 [Représentation graphique du tableau de contingence entre la variable qualitative « Riche » et la variable qualitative « Ouest »]



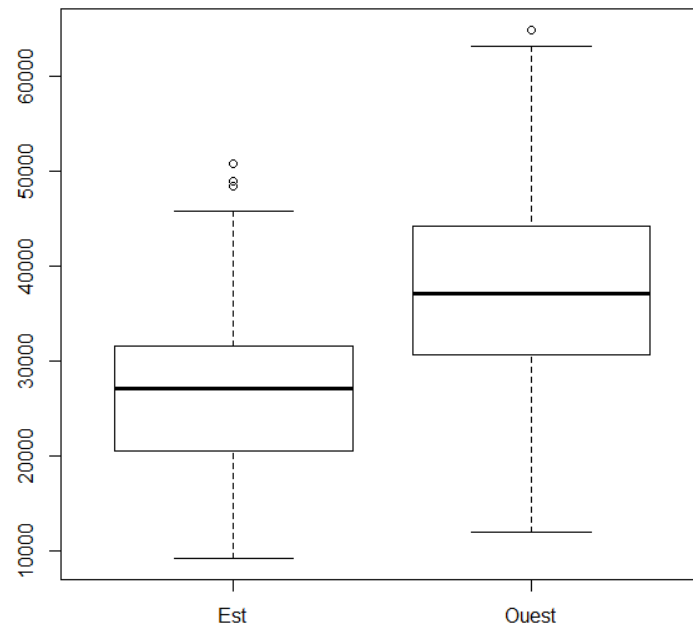
TD2 – 7 [Test du chi-2 entre la variable qualitative « Riche » et la variable qualitative « Ouest »]

```
Pearson's Chi-squared test with Yates' continuity correction

data:  contingence
X-squared = 219.35, df = 1, p-value < 2.2e-16
```

TD2 – 9 [Moyenne et distribution des revenus en fonction de la variable qualitative « Ouest »]

```
Est    Ouest
26112.33 37204.27
```



TD2 – 9 [Anova entre les revenus et la variable qualitative « Ouest »]

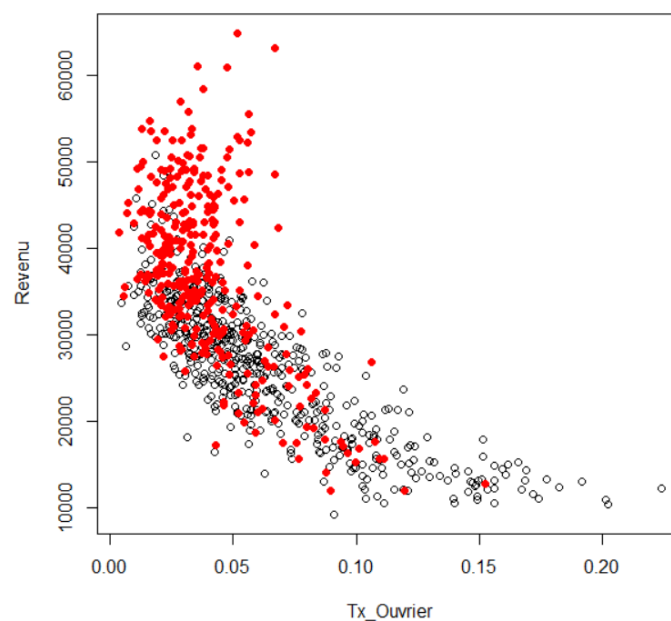
Analysis of Variance Table

Response: Revenu

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
data\$Ouest	1	2.5151e+10	2.5151e+10	324.53	< 2.2e-16 ***
Residuals	862	6.6805e+10	7.7500e+07		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TD2 – 10 [Représentation des revenus des IRIS de l'ouest parisien (en rouge) et de l'est parisien (en noir) en fonction du taux d'ouvrier]



TD2 – 11 [Ancova et modèle de régression tentant d'expliquer les revenus par la variable qualitative « Ouest » et la variable quantitative « Tx_Ouvrier »]

Analysis of Variance Table

Response: Revenu

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
data\$Ouest	1	2.5151e+10	2.5151e+10	620.42	< 2.2e-16 ***
Tx_Ouvrier	1	3.1901e+10	3.1901e+10	786.95	< 2.2e-16 ***
Residuals	861	3.4903e+10	4.0538e+07		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Call:

```
lm(formula = Revenu ~ data$Ouest + Tx_Ouvrier)
```

Residuals:

Min	1Q	Median	3Q	Max
-19337.4	-4090.8	-391.1	3273.4	31008.8

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	37622.7	494.5	76.08	<2e-16 ***
data\$OuestOuest	6760.7	471.3	14.34	<2e-16 ***
Tx_Ouvrier	-182361.2	6500.7	-28.05	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6367 on 861 degrees of freedom

Multiple R-squared: 0.6204, Adjusted R-squared: 0.6195

F-statistic: 703.7 on 2 and 861 DF, p-value: < 2.2e-16

TD2 – 11 [Anova comparant le modèle tentant d'expliquer les revenus par la variable qualitative « Ouest » et la variable quantitative « Tx_Ouvrier » un modèle se contentant de la variable qualitative]

Analysis of Variance Table

Model 1: Revenu ~ data\$Ouest + Tx_Ouvrier

Model 2: Revenu ~ data\$Ouest

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	861	3.4903e+10				
2	862	6.6805e+10	-1	-3.1901e+10	786.95	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TD2 – 12 [Ancova et modèle de régression tentant d'expliquer les revenus par la variable qualitative « Ouest », la variable quantitative « Tx_Ouvrier » et l'interaction entre ces deux variables]

Call:

```
lm(formula = Revenu ~ data$Ouest * Tx_Ouvrier)
```

Residuals:

Min	1Q	Median	3Q	Max
-19034	-4121	-542	3302	33194

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36662.7	517.8	70.807	< 2e-16 ***
data\$OuestOuest	10833.5	882.8	12.272	< 2e-16 ***
Tx_Ouvrier	-167151.8	6984.2	-23.933	< 2e-16 ***
data\$OuestOuest:Tx_Ouvrier	-94281.2	17388.9	-5.422	7.66e-08 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6265 on 860 degrees of freedom

Multiple R-squared: 0.633, Adjusted R-squared: 0.6317

F-statistic: 494.4 on 3 and 860 DF, p-value: < 2.2e-16

TD2 – 12 [Anova comparant le modèles avec prise en compte des interactions et sans leur prise en compte]

Analysis of Variance Table

```
Model 1: Revenu ~ data$Ouest * Tx_Ouvrier
Model 2: Revenu ~ data$Ouest + Tx_Ouvrier
```

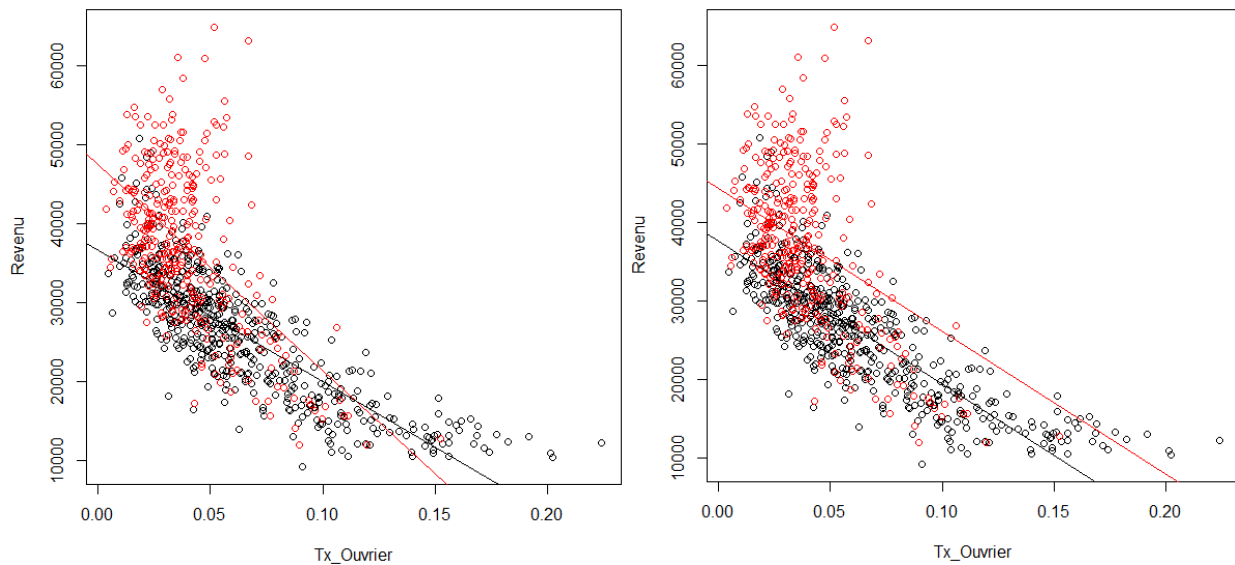
	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	860	3.3750e+10				
2	861	3.4903e+10	-1	-1153660172	29.397	7.664e-08 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

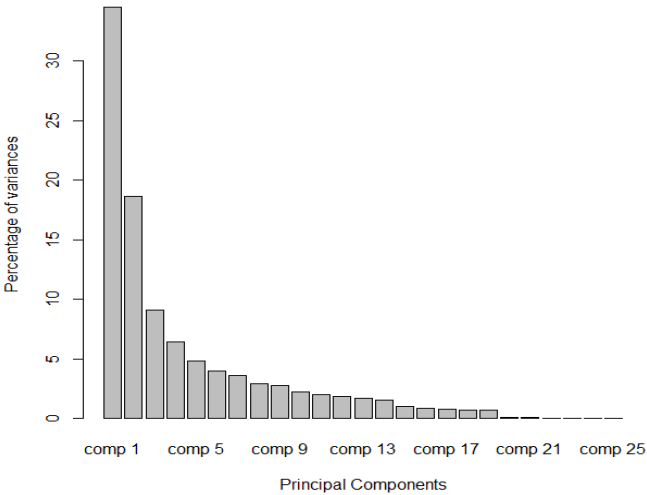
TD2 – 13 [Moyennes ajustées des revenus de l'ouest de parisien et de l'est parisien compte tenu des différences de taux d'ouvrier dans ces deux zones]

```
> c[1]+c[2] + (c[3]+c[4])* mean(Tx_Ouvrier)
(Intercept)
33380.93
> c[1] + c[3]* mean(Tx_Ouvrier)
(Intercept)
27637.85
```

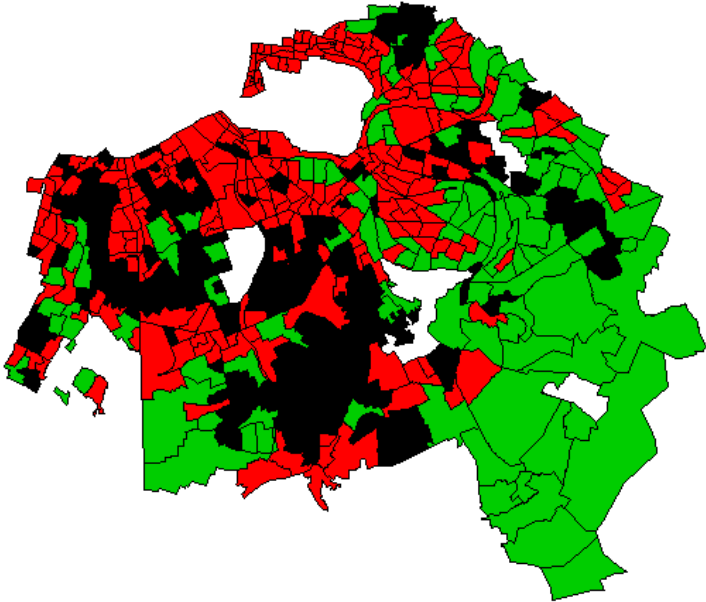
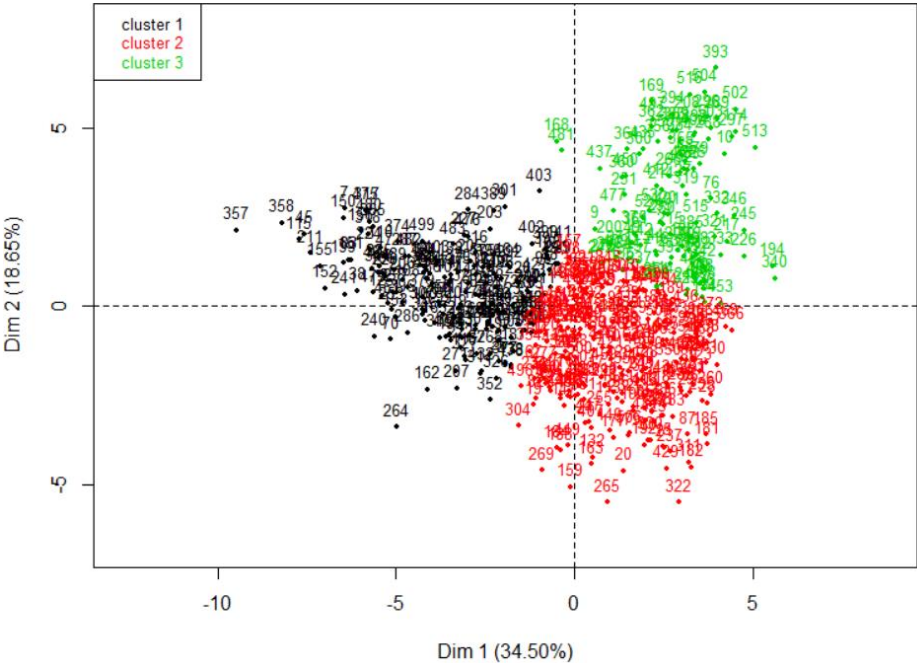
TD2 – 13 [Représentation graphique finale avec ou sans prise en compte des interactions entre variables explicatives]



TD3 – 3 [Barplot des pourcentages de variance (d’inertie) expliqués par les différentes composantes principalesACP]



TD3 – 5 [CAH en trois groupes sur l’ACP à deux dimensions]



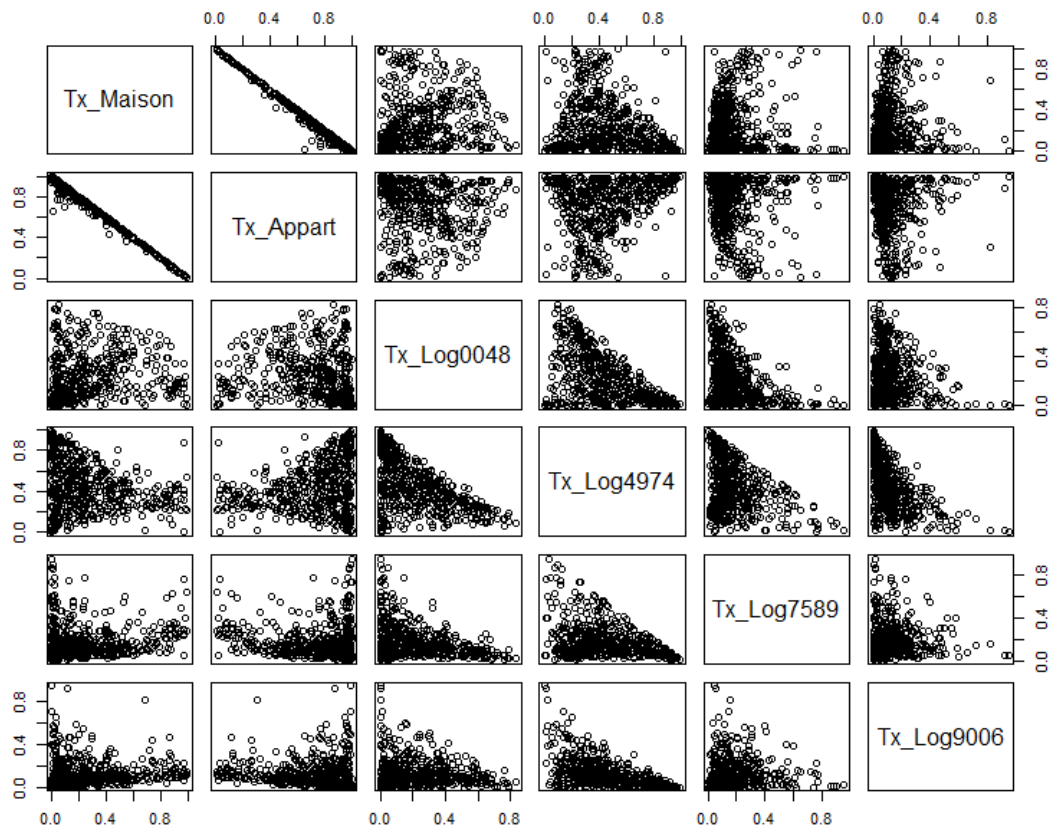
0.6494907

Description of each cluster by quantitative variables

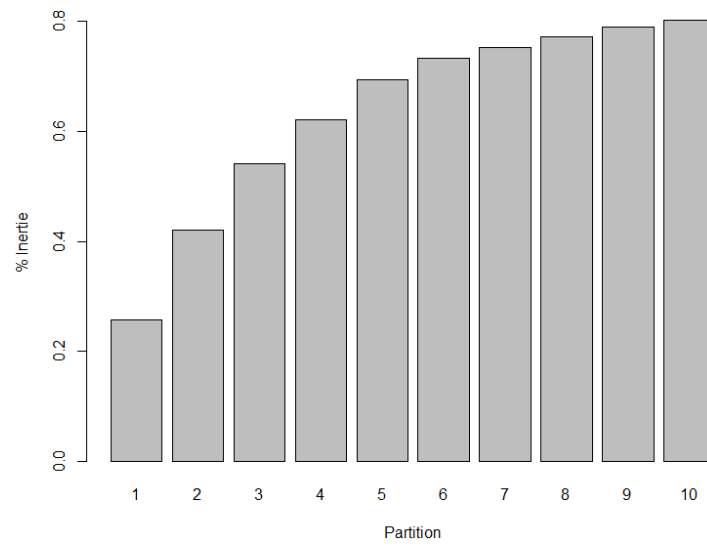
\$`1`

	v.test	Mean in category	Overall mean	sd in category	Overall sd
Tx_HLM	17.892648	0.59307742	0.28190842	0.23478361	0.28611294
Tx_Employ	17.398087	0.40115112	0.30033050	0.05863652	0.09533749
Tx_Monopa	17.162633	0.17695994	0.12430170	0.04364771	0.05047755
Tx_Ouvrie	14.529671	0.22172640	0.15392355	0.06461937	0.07677291
Tx_Cou4.E	14.157102	0.03811225	0.02080349	0.02410531	0.02011441
Tx_0014	13.484352	0.22499742	0.19553226	0.03270604	0.03594963
Tx_Log4974	10.978845	0.61399663	0.45075689	0.26489872	0.24461611
Tx_Appart	9.530542	0.88013764	0.72941127	0.13354420	0.26018811
Tx_CouAEn	4.400326	0.31862433	0.29961832	0.06204897	0.07105944
Tx_1529	3.896346	0.20703843	0.19773855	0.02983622	0.03926769
Tx_3044	-2.281746	0.21104084	0.21569112	0.03123880	0.03352952
Tx_Log9006	-3.770420	0.09411129	0.12535299	0.14613624	0.13632037
Tx_Pseul	-5.320921	0.31234680	0.34079637	0.08092218	0.08796396
Tx_6074	-6.869760	0.10947904	0.12255803	0.02789508	0.03132191
Tx_Artisa	-9.154864	0.03709006	0.04872299	0.01536195	0.02090515
Tx_75P	-9.236880	0.05431157	0.07250170	0.02661050	0.03239860
Tx_Maison	-9.403156	0.10681270	0.25677344	0.13112464	0.26237338
Tx_Loc	-9.951044	0.14919702	0.22839971	0.09708606	0.13094444
Tx_Log0048	-10.970072	0.09215281	0.22685313	0.12315467	0.20201097
Tx_ProInt	-11.381185	0.23079382	0.26363629	0.04767284	0.04747492
Tx_CouSEn	-13.987478	0.16323826	0.20778981	0.03654895	0.05240095
Tx_Cadre	-15.708400	0.10924270	0.23339035	0.05893959	0.13002366
Tx_Propri	-15.759299	0.22743809	0.45996006	0.16671611	0.24274087

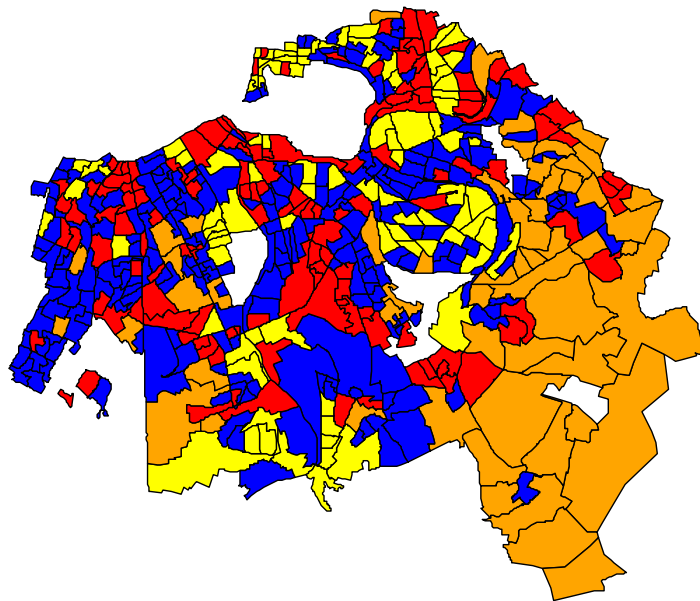
TD3 – 8 [Analyse de corrélation par paire de 6 variables]



TD3 – 8 [Analyse de la qualité des partitions obtenus par CAH sur 5 variables]



TD3 – 8 [CAH sur 5 variables en 4 classes]



0.5404948