

Projet Final : Les Français de l'étranger

Projet Final de Modèle Statistique Alexandre Lecuyer-Adlane Badr Lakhdari

Le jeu de donnée récupérer concerne les Français à l'étranger publié sur le site du Gouvernement Français "<https://www.data.gouv.fr/fr/datasets/francais-de-l-etranger-inscriptions-au-registre-des-francais-etablis-hors-de-france-2001-2013/>". Ces données ont été collectés par le Ministère des Affaires étrangères (Direction des Français de l'étranger et de l'administration consulaire). Les données contenues dans ces tableaux ont été récupéré via l'inscription au Registre, une démarche volontaire ouverte aux Français résidant à l'étranger, s'effectuant auprès du réseau des ambassades et consulats français. Toutefois, les Français ne s'étant pas inscrit n'apparaissent donc pas dans le jeu donnée. Il est intéressant de regarder comment évolue ces chiffres sur une période donnée (2001-2013). Ce fichier récupérer, il a été modifié, les pays sont regroupés par zones géographiques (zone définis par l'ONU).

Les zones sont :

-Afrique du Nord : Algérie, Libye, Maroc, Tunisie

-Afrique Francophone : Bénin, Burkina Faso, Burundi, Cameroun, République Centrafricaine, Comores, Congo, Côte d'Ivoire, Djibouti, Gabon, Guinée, Madagascar, Mali, Maurice, Mauritanie, Niger, République démocratique du Congo, Rwanda, Sénégal, Tchad, Togo

-Afrique Non Francophone : Afrique du Sud, Angola, Bostawana, Cap Vert, Éthiopie, Ghana, Guinée-Bissau, Guinée Équatoriale, Kenya, Mozambique, Namibie, Nigeria, Ouganda, Seychelles, Soudan, Soudan du Sud, Tanzanie, Zambie, Zimbabwe

-Proche et Moyen-Orient : Arabie saoudite, Bahreïn, Égypte, Émirats arabes unis, Iran, Irak, Israël, Jérusalem, Jordanie, Koweït, Liban, Oman, Qatar, Syrie, Yémen

-Asie-Océanie : Afghanistan, Australie, Bangladesh, Birmanie, Brunei, Cambodge, Chine, Corée du Sud, Fidji, Inde, Indonésie, Japon, Laos, Malaisie, Mongolie, Népal, Nouvelle-Zélande, Pakistan, Papouasie-Nouvelle-Guinée, Philippines, Singapour, Sri Lanka, Taïwan, Thaïlande, Vanuatu, Vietnam

-Europe hors UE : Albanie, Andorre, Arménie, Azerbaïdjan, Biélorussie, Bosnie-Herzégovine, Géorgie, Islande, Kazakhstan, Kosovo, Macédoine, Moldavie, Monaco, Monténégro, Norvège, Ouzbékistan, Russie, Saint-Siège, Serbie, Suisse, Tadjikistan, Turkménistan, Turquie, Ukraine

-Union Européenne : Allemagne, Autriche, Belgique, Bulgarie, Chypre, Croatie, Danemark, Espagne, Estonie, Finlande, Grèce, Hongrie, Irlande, Italie, Lettonie, Lituanie, Luxembourg, Malte, Pays-Bas, Pologne, Portugal, Roumanie, Royaume-Uni, Slovaquie, Slovénie, Suède, République tchèque

-Amérique Centrale-Sud : Argentine, Bolivie, Brésil, Chili, Colombie, Costa Rica, Cuba, République dominicaine, Équateur, Guatemala, Haïti, Honduras, Jamaïque, Mexique, Nicaragua, Panama, Paraguay, Pérou, Sainte-Lucie, Suriname, Trinité-et-Tobago, Uruguay, Venezuela

-Amérique du Nord : Canada, États-Unis d'Amérique

```
library(dplyr);
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(reshape2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")>%
mutate(Année = as.integer(gsub("X","",variable)))>%
select(-variable)>%
rename(ZoneGéographique = ZoneGéographique.Année)
```

	ZoneGéographique	value	Année
## 1	Afrique du Nord	47233	2001
## 2	Afrique Francophone	97568	2001
## 3	Afrique Non Francophone	11640	2001
## 4	Proche et Moyen-Orient	74063	2001
## 5	Asie-Océanie	52940	2001
## 6	Europe hors UE	131724	2001
## 7	Union Européenne	393288	2001
## 8	Amérique Centrale-Sud	63208	2001
## 9	Amérique du Nord	127119	2001
## 10	Total	998783	2001
## 11	Afrique du Nord	56014	2002
## 12	Afrique Francophone	100092	2002
## 13	Afrique Non Francophone	12092	2002
## 14	Proche et Moyen-Orient	82786	2002
## 15	Asie-Océanie	58423	2002
## 16	Europe hors UE	142528	2002
## 17	Union Européenne	438648	2002
## 18	Amérique Centrale-Sud	68972	2002
## 19	Amérique du Nord	140304	2002
## 20	Total	1099859	2002
## 21	Afrique du Nord	69031	2003
## 22	Afrique Francophone	104180	2003
## 23	Afrique Non Francophone	12934	2003
## 24	Proche et Moyen-Orient	90776	2003
## 25	Asie-Océanie	64785	2003
## 26	Europe hors UE	156120	2003
## 27	Union Européenne	487408	2003
## 28	Amérique Centrale-Sud	72255	2003
## 29	Amérique du Nord	162920	2003
## 30	Total	1220409	2003
## 31	Afrique du Nord	76716	2004
## 32	Afrique Francophone	101566	2004
## 33	Afrique Non Francophone	13621	2004
## 34	Proche et Moyen-Orient	92954	2004
## 35	Asie-Océanie	67564	2004
## 36	Europe hors UE	160140	2004
## 37	Union Européenne	495570	2004
## 38	Amérique Centrale-Sud	77756	2004
## 39	Amérique du Nord	167342	2004
## 40	Total	1253229	2004
## 41	Afrique du Nord	84049	2005
## 42	Afrique Francophone	102349	2005
## 43	Afrique Non Francophone	14033	2005
## 44	Proche et Moyen-Orient	96653	2005
## 45	Asie-Océanie	73531	2005
## 46	Europe hors UE	155951	2005

## 47	Union Européenne	495845	2005
## 48	Amérique Centrale-Sud	79255	2005
## 49	Amérique du Nord	166862	2005
## 50	Total	1268528	2005
## 51	Afrique du Nord	91092	2006
## 52	Afrique Francophone	110657	2006
## 53	Afrique Non Francophone	15916	2006
## 54	Proche et Moyen-Orient	104676	2006
## 55	Asie-Océanie	86160	2006
## 56	Europe hors UE	160046	2006
## 57	Union Européenne	537980	2006
## 58	Amérique Centrale-Sud	83823	2006
## 59	Amérique du Nord	183638	2006
## 60	Total	1373988	2006
## 61	Afrique du Nord	87280	2007
## 62	Afrique Francophone	100497	2007
## 63	Afrique Non Francophone	15605	2007
## 64	Proche et Moyen-Orient	107923	2007
## 65	Asie-Océanie	85238	2007
## 66	Europe hors UE	164059	2007
## 67	Union Européenne	509105	2007
## 68	Amérique Centrale-Sud	80773	2007
## 69	Amérique du Nord	175607	2007
## 70	Total	1326087	2007
## 71	Afrique du Nord	90118	2008
## 72	Afrique Francophone	107150	2008
## 73	Afrique Non Francophone	16602	2008
## 74	Proche et Moyen-Orient	121558	2008
## 75	Asie-Océanie	94035	2008
## 76	Europe hors UE	171093	2008
## 77	Union Européenne	553359	2008
## 78	Amérique Centrale-Sud	87980	2008
## 79	Amérique du Nord	185151	2008
## 80	Total	1427046	2008
## 81	Afrique du Nord	89789	2009
## 82	Afrique Francophone	108521	2009
## 83	Afrique Non Francophone	16875	2009
## 84	Proche et Moyen-Orient	127232	2009
## 85	Asie-Océanie	99080	2009
## 86	Europe hors UE	177301	2009
## 87	Union Européenne	572072	2009
## 88	Amérique Centrale-Sud	90694	2009
## 89	Amérique du Nord	188065	2009
## 90	Total	1469629	2009
## 91	Afrique du Nord	90153	2010
## 92	Afrique Francophone	112441	2010
## 93	Afrique Non Francophone	17475	2010
## 94	Proche et Moyen-Orient	133769	2010
## 95	Asie-Océanie	105365	2010
## 96	Europe hors UE	178876	2010
## 97	Union Européenne	586827	2010
## 98	Amérique Centrale-Sud	92633	2010
## 99	Amérique du Nord	186462	2010
## 100	Total	1504001	2010

## 101	Afrique du Nord	94789	2011
## 102	Afrique Francophone	116862	2011
## 103	Afrique Non Francophone	18621	2011
## 104	Proche et Moyen-Orient	136679	2011
## 105	Asie-Océanie	116729	2011
## 106	Europe hors UE	190283	2011
## 107	Union Européenne	620209	2011
## 108	Amérique Centrale-Sud	97117	2011
## 109	Amérique du Nord	202014	2011
## 110	Total	1593303	2011
## 111	Afrique du Nord	98090	2012
## 112	Afrique Francophone	117378	2012
## 113	Afrique Non Francophone	18796	2012
## 114	Proche et Moyen-Orient	136394	2012
## 115	Asie-Océanie	120027	2012
## 116	Europe hors UE	193920	2012
## 117	Union Européenne	625206	2012
## 118	Amérique Centrale-Sud	97425	2012
## 119	Amérique du Nord	203818	2012
## 120	Total	1611054	2012
## 121	Afrique du Nord	102007	2013
## 122	Afrique Francophone	119723	2013
## 123	Afrique Non Francophone	18790	2013
## 124	Proche et Moyen-Orient	137441	2013
## 125	Asie-Océanie	125057	2013
## 126	Europe hors UE	199738	2013
## 127	Union Européenne	627911	2013
## 128	Amérique Centrale-Sud	99471	2013
## 129	Amérique du Nord	212815	2013
## 130	Total	1642953	2013

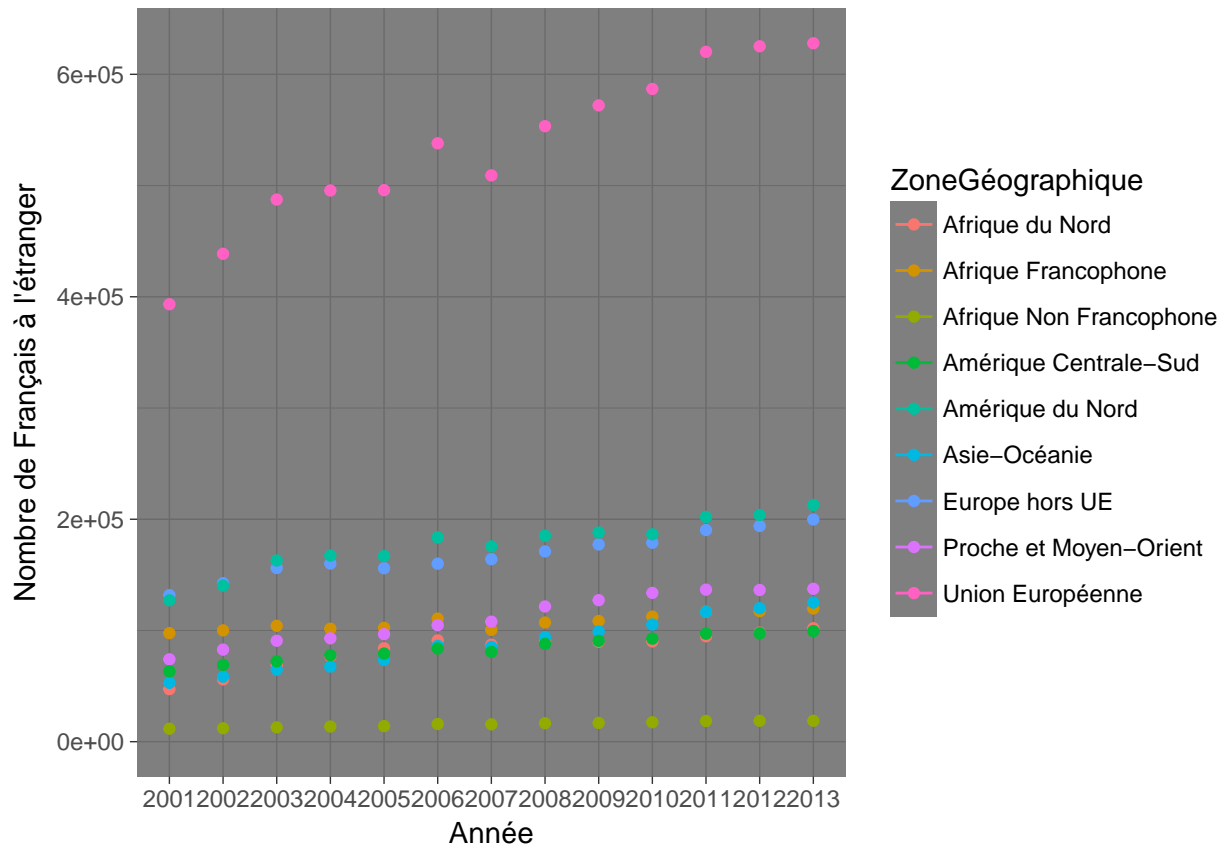
Le tableau ci-dessus affiche le nombre de français à l'étranger par zone et par année.

Le graphe ci-dessous, montre les différentes évolutions des Français sur la période 2001-2013 selon les différentes zones géographiques. On remarque immédiatement qu'il y a trois zones qui comportent le plus de Français, à savoir l'Union Européenne, l'Amérique du Nord et l'Europe hors UE. L'Afrique non Francophone, représente la zone comportant le moins de Français.

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique!="Total")%>%
ggplot(aes(x=as.factor(Année),y=value, colour=ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark()+
  xlab("Année") +
  ylab("Nombre de Français à l'étranger")
```

geom_path: Each group consists of only one observation. Do you need to

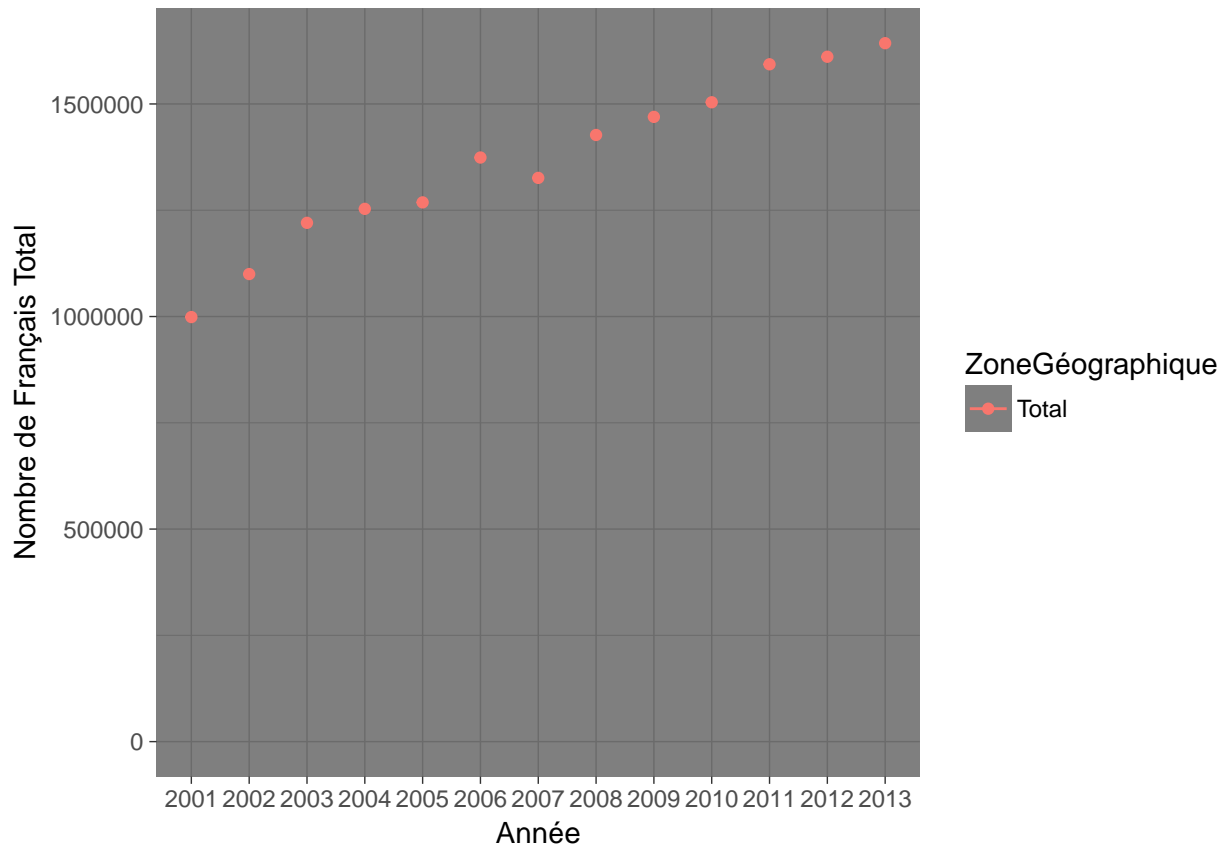
```
## adjust the group aesthetic?
```



Ce graphe nous montre l'évolution total du nombre de Français à l'étranger au cours de la période 2001-2013. On remarque que l'évolution du nombre de français à l'étranger est globalement assimilable à une courbe strictement croissante. Toutefois, l'année 2007 fait office de baisse.

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")>%
mutate(Année = as.integer(gsub("X","",variable)))>%
select(-variable)>%
rename(ZoneGéographique = ZoneGéographique.Année)>%
filter(ZoneGéographique == "Total")>%
ggplot(aes(x=as.factor(Année),y=value,colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français Total")
```

```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```



Analyse :

Grâce à nos deux graphes et aux annexes, on observe que le nombre de français à l'étranger augmente globalement, hormis une baisse en 2007 comme on a pu le constater avec le graphe précédent. Cette diminution est également remarqué dans la plupart des annexes (graphe ciblant une zone précise).

Cependant, on constate les limites de notre jeu de donnée : - Les données sont incomplètes, les français ne s'étant pas inscrit au registre ne figurent pas dans le tableau - La baisse observé en 2007 n'a aucune explication. On peut émettre l'hypothèse que celle-ci est due à la crise financière mais ceci n'est confirmé ou confirmable, donc on ne peut avoir une réponse précise et juste. - L'augmentation constante du nombre de français à l'étranger ne peut être clairement et explicitement expliqué.

Le jeu de données étant relativement simpliste qui montre seulement un nombre de personnes sans aucune information sur ces dites personne (sexe,age,catégorie socio-professionnelle, raison du départ), nous ne pouvons donc apporter une réponse satisfaisante, claire et précise sur comment à évoluer le nombre de français à l'étranger sur la période 2001-2013.

En conclusion, comme il a été dit dans notre analyse, la réponse à la question de l'évolution des Français à l'étranger est trop simple, non complète. Le jeu de donnée ne permet absolument une bonne analyse de la question, une connaissance approfondie des personnes inscrites et du nombre exact de français résidant à l'étranger.

Les graphes suivant (les annexes 1 à 9), quand à eux, affiche le nombre de français à l'étranger par année pour une zone ciblée (soit 9 zones).

Annexe 1 (Évolution du nombre de Français dans la Zone Afrique du Nord) :

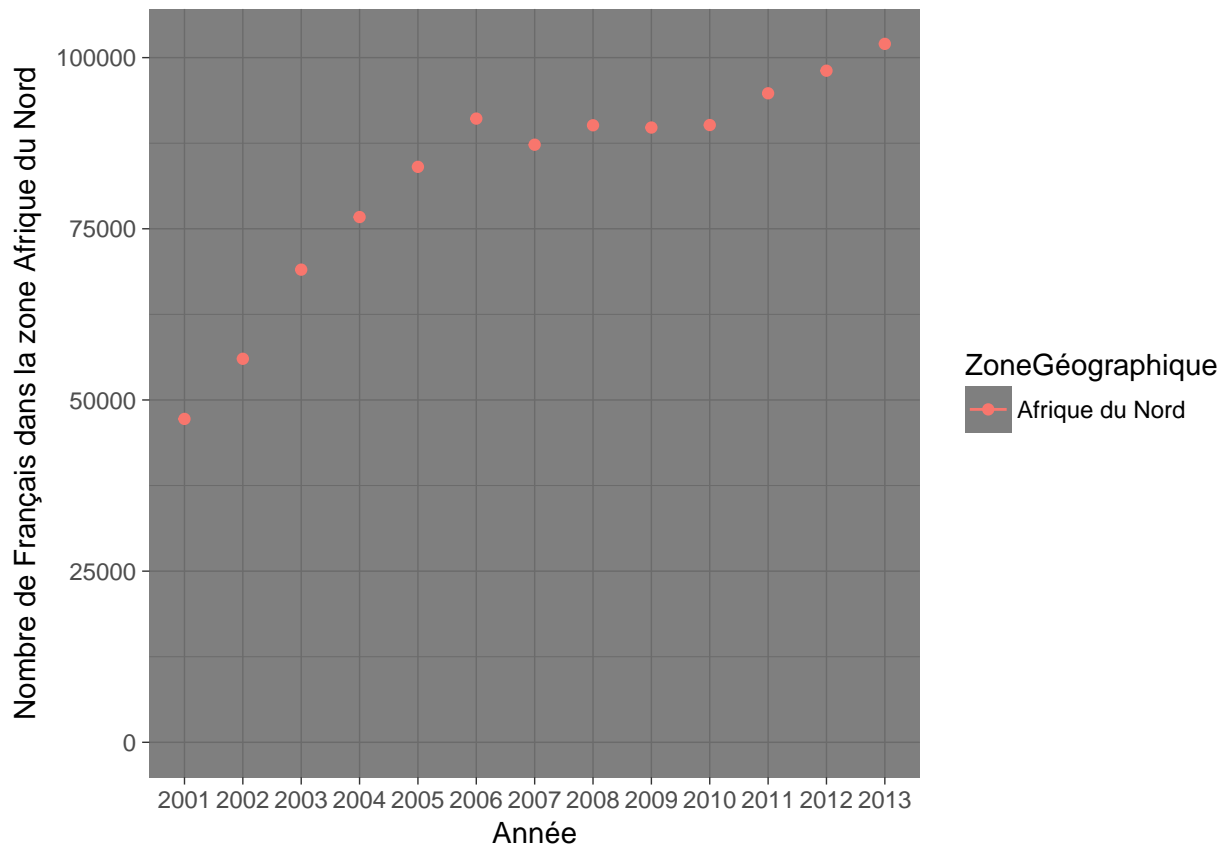
```
library(dplyr);
library(reshape2);
library(ggplot2);
```

```

dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Afrique du Nord")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Afrique du Nord")

```

geom_path: Each group consists of only one observation. Do you need to
adjust the group aesthetic?



Annexe 2 (Évolution du nombre de Français dans la Zone Afrique francophone) :

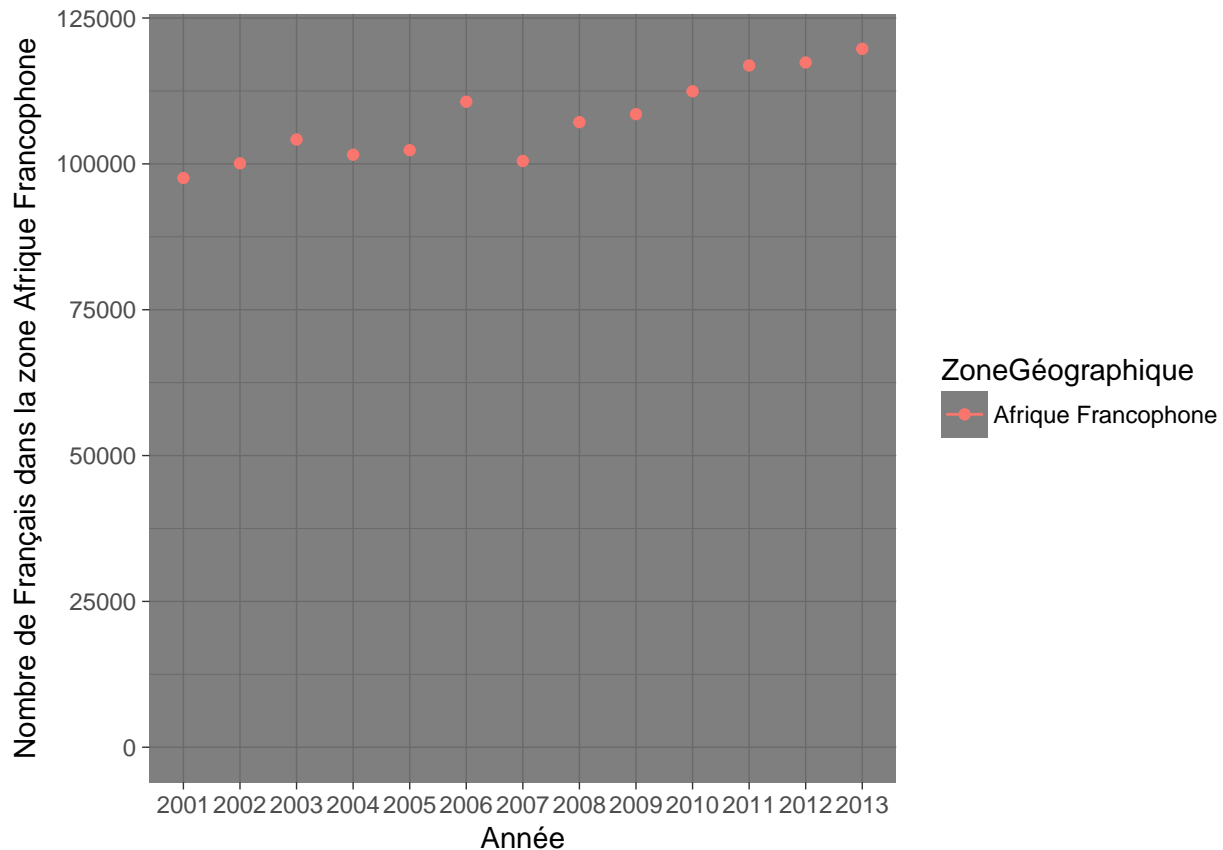
```

library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%

```

```
filter(ZoneGéographique=="Afrique Francophone")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Afrique Francophone")
```

geom_path: Each group consists of only one observation. Do you need to
adjust the group aesthetic?



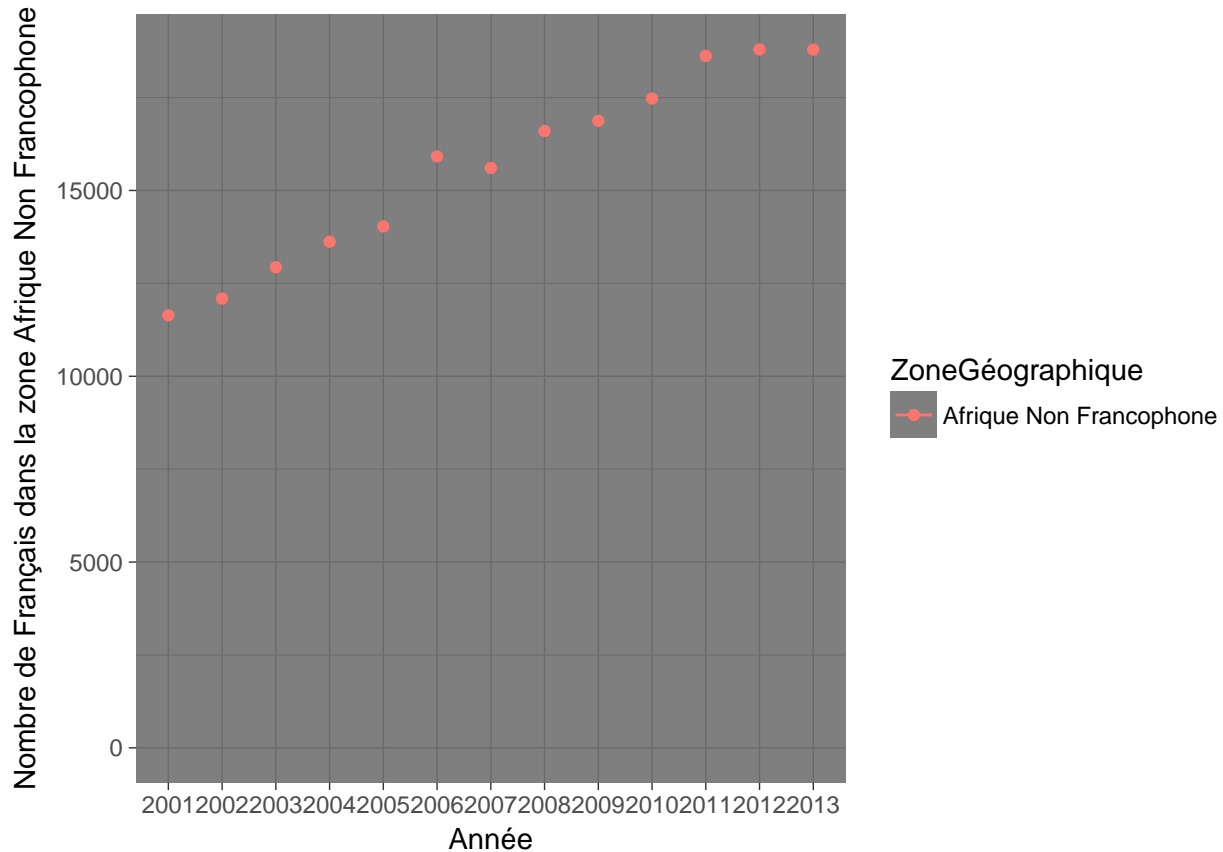
Annexe 3 (Évolution du nombre de Français dans la Zone Afrique non Francophone) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Afrique Non Francophone")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
```



```
theme_dark() +
xlab("Année") +
ylab("Nombre de Français dans la zone Afrique Non Francophone")
```

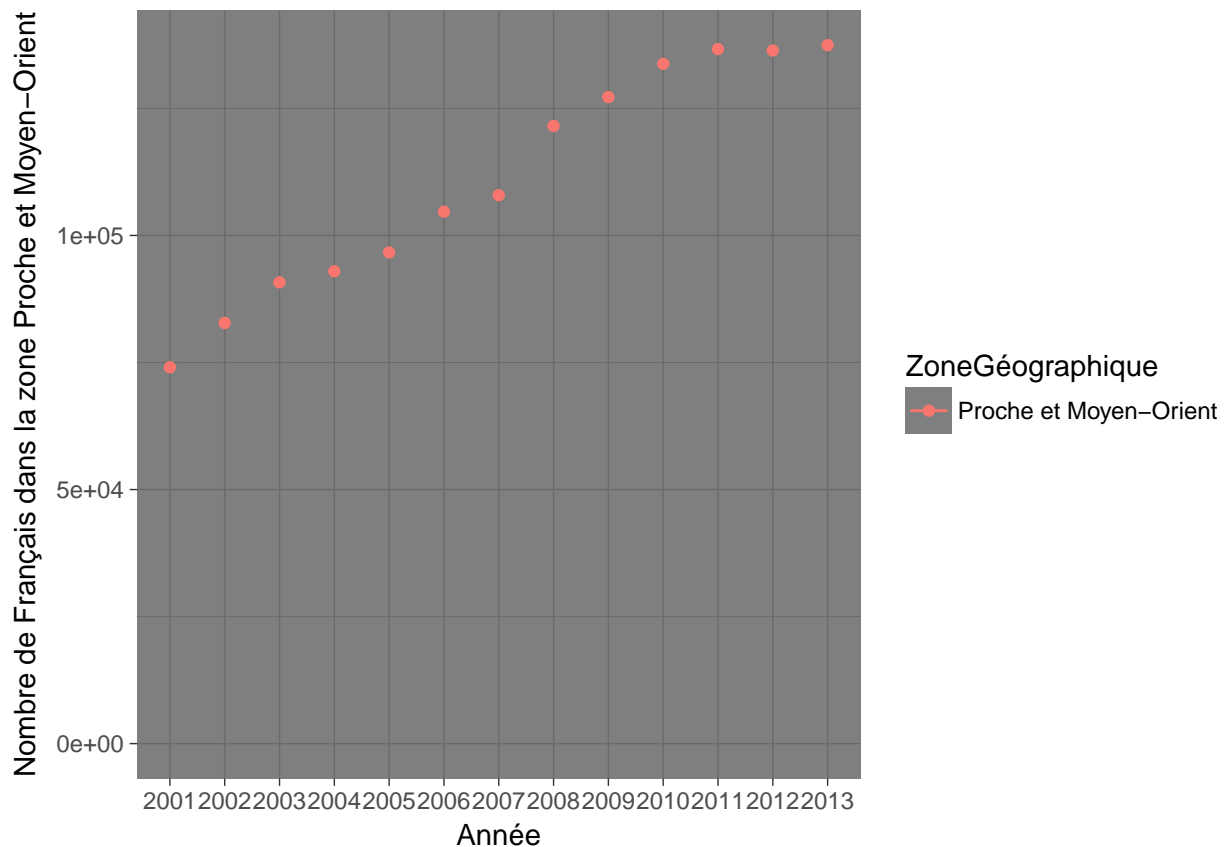
geom_path: Each group consists of only one observation. Do you need to
adjust the group aesthetic?



Annexe 4 (Évolution du nombre de Français dans la Zone Proche et Moyen-Orient) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Proche et Moyen-Orient")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Proche et Moyen-Orient")
```

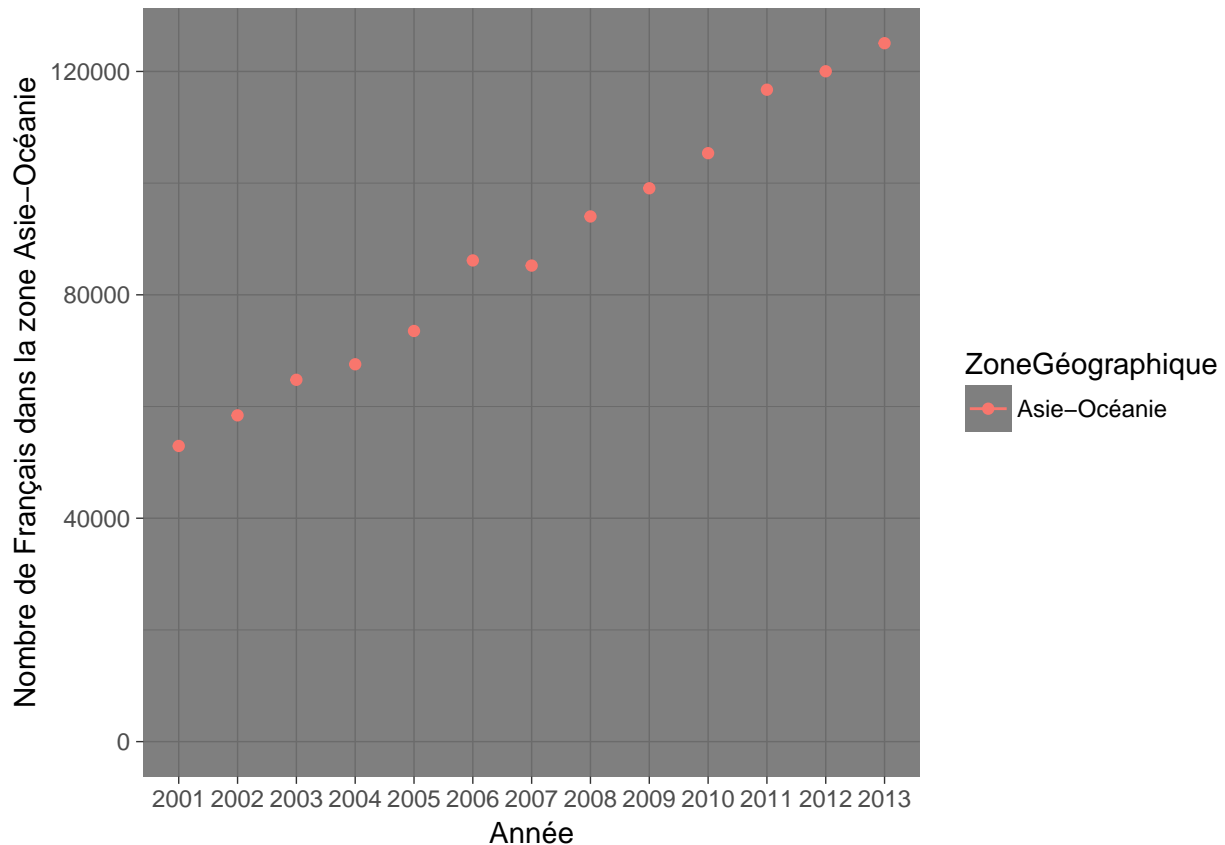
geom_path: Each group consists of only one observation. Do you need to
adjust the group aesthetic?



Annexe 5 (Évolution du nombre de Français dans la Zone Asie-Océanie) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Asie-Océanie")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Asie-Océanie")
```

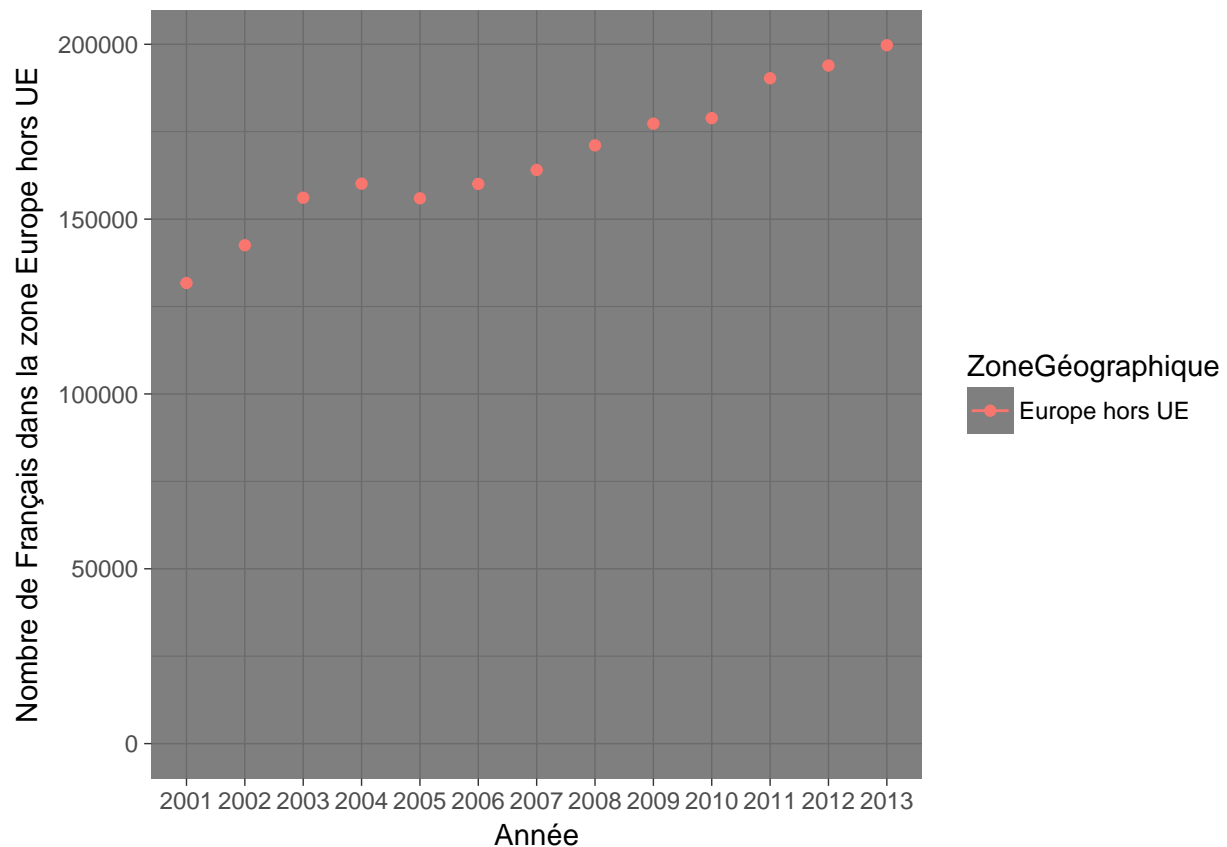
```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```



Annexe 6 (Évolution du nombre de Français dans la Zone Europe hors UE) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Europe hors UE")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Europe hors UE")
```

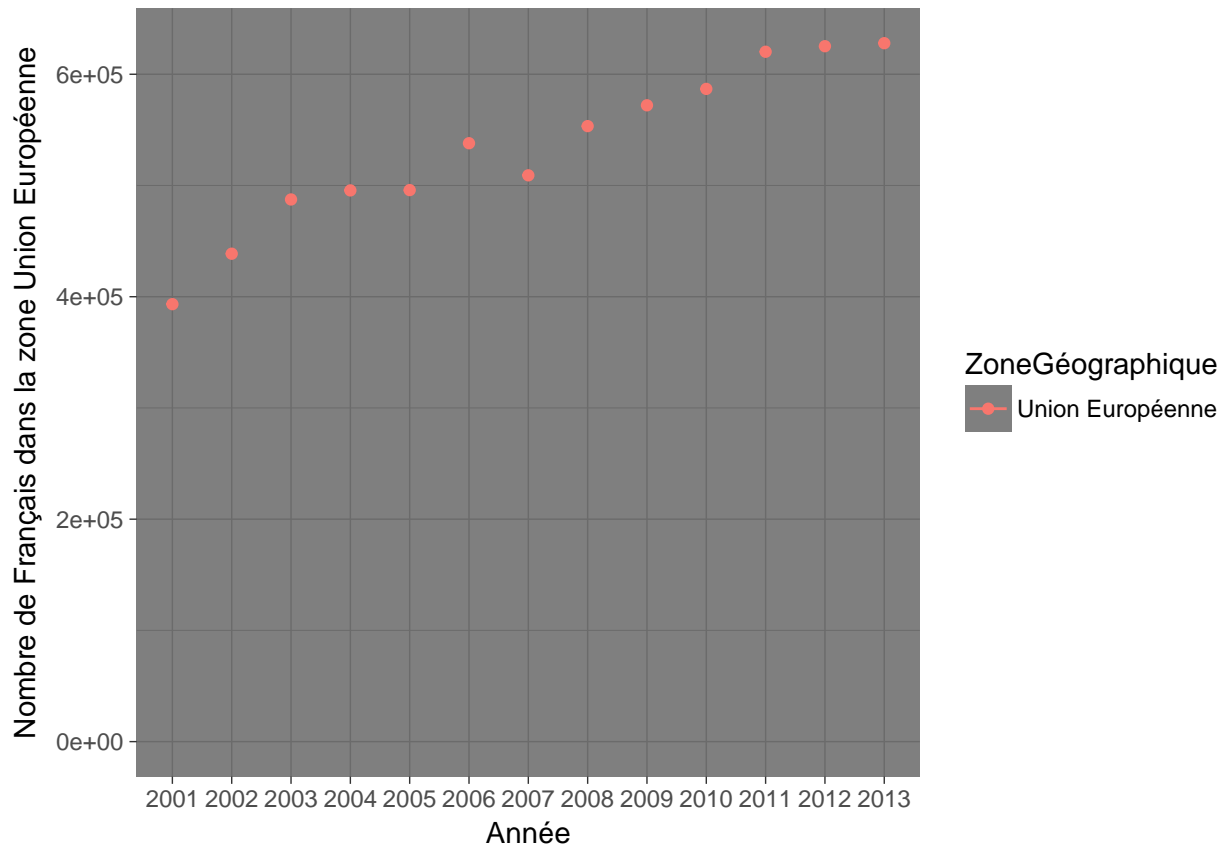
```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```



Annexe 7 (Évolution du nombre de Français dans la Zone Union Européenne) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Union Européenne")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Union Européenne")
```

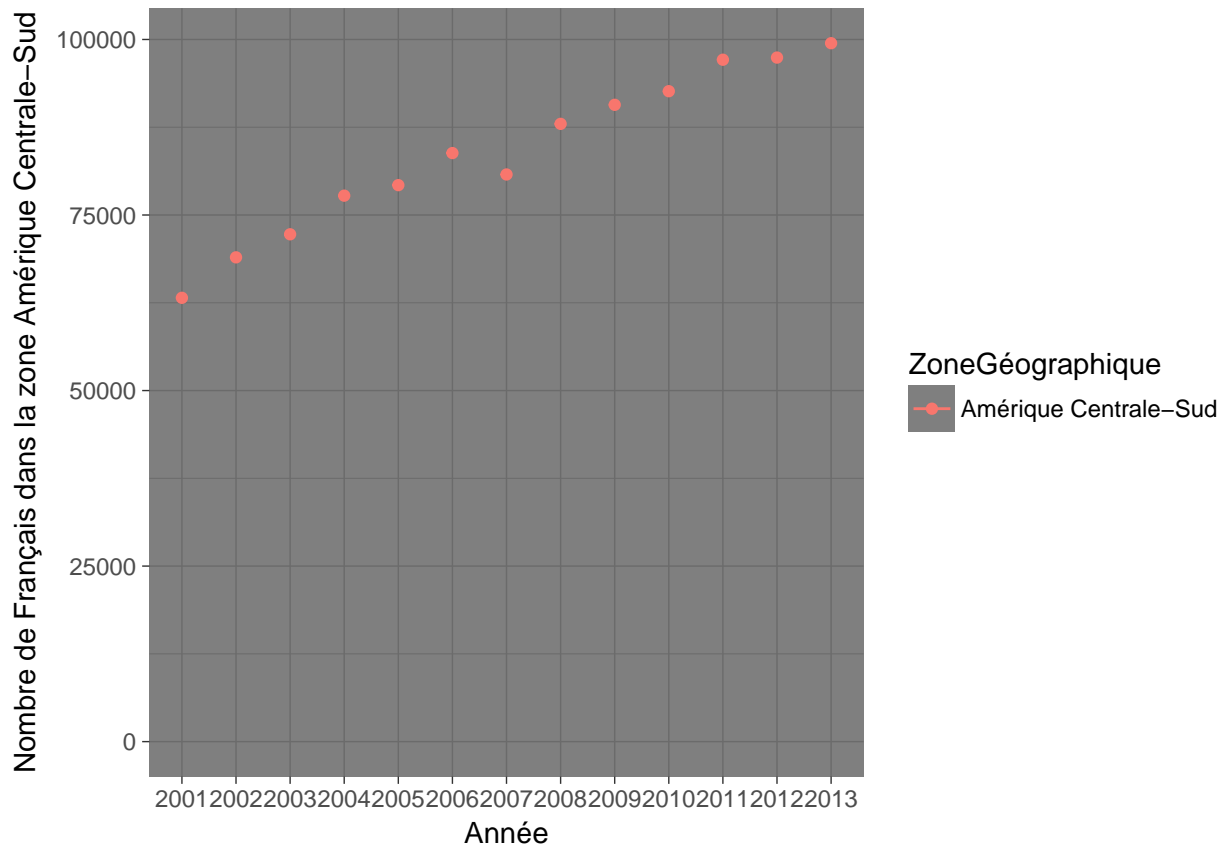
```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```



Annexe 8 (Évolution du nombre de Français dans la Zone Amérique Centrale-Sud) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Amérique Centrale-Sud")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Amérique Centrale-Sud")
```

```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```



Annexe 9 (Évolution du nombre de Français dans la Zone Amérique du Nord) :

```
library(dplyr);
library(reshape2);
library(ggplot2);
dataf<-read.csv("FrancaisE.csv",sep = ";")
melt(dataf,id.vars = "ZoneGéographique.Année")%>%
mutate(Année = as.integer(gsub("X","",variable)))%>%
select(-variable)%>%
rename(ZoneGéographique = ZoneGéographique.Année)%>%
filter(ZoneGéographique=="Amérique du Nord")%>%
ggplot(aes(x=as.factor(Année),y=value, colour = ZoneGéographique))+
  geom_point()+
  ylim(0,NA)+
  geom_line() +
  theme_dark() +
  xlab("Année") +
  ylab("Nombre de Français dans la zone Amérique du Nord")
```

```
## geom_path: Each group consists of only one observation. Do you need to
## adjust the group aesthetic?
```

