

Cedefop Data Analysis

Claire de Maricourt

28 juin 2019

Data uploading and first infos

Number of offers and number of vacancies

```
load("data/cedefop_ojv.RData") # reload the dataset
```

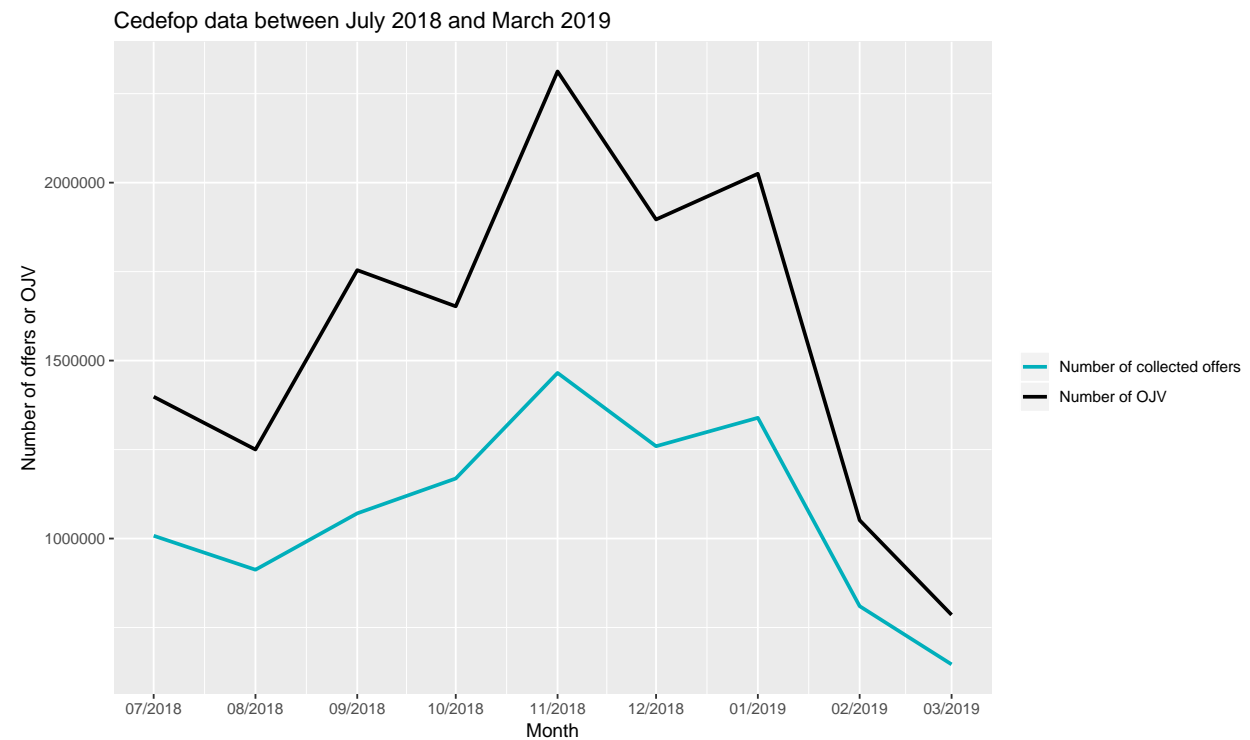
```
dim(cedefop_ojv)[1] # Number of offers
```

```
## [1] 9680329
```

```
sum(cedefop_ojv$NUM_OJV) # Number of OJV
```

```
## [1] 14126212
```

Date distribution



Variable names

```
names(cedefop_ojv)
```

```
## [1] "YEAR_GRAB_DATE"      "MONTH_GRAB_DATE"    "IDESCO_LEVEL_4"  
## [4] "IDREGION"           "IDCONTRACT"         "IDEDUCATIONAL_LEVEL"  
## [7] "IDSECTOR"           "IDSALARY"           "IDWORKING_HOURS"  
## [10] "IDEXPERIENCE"       "SOURCE"             "SITE"  
## [13] "COMPANYNAME"        "NUM_OJV"            "DATE"
```

Sources

237



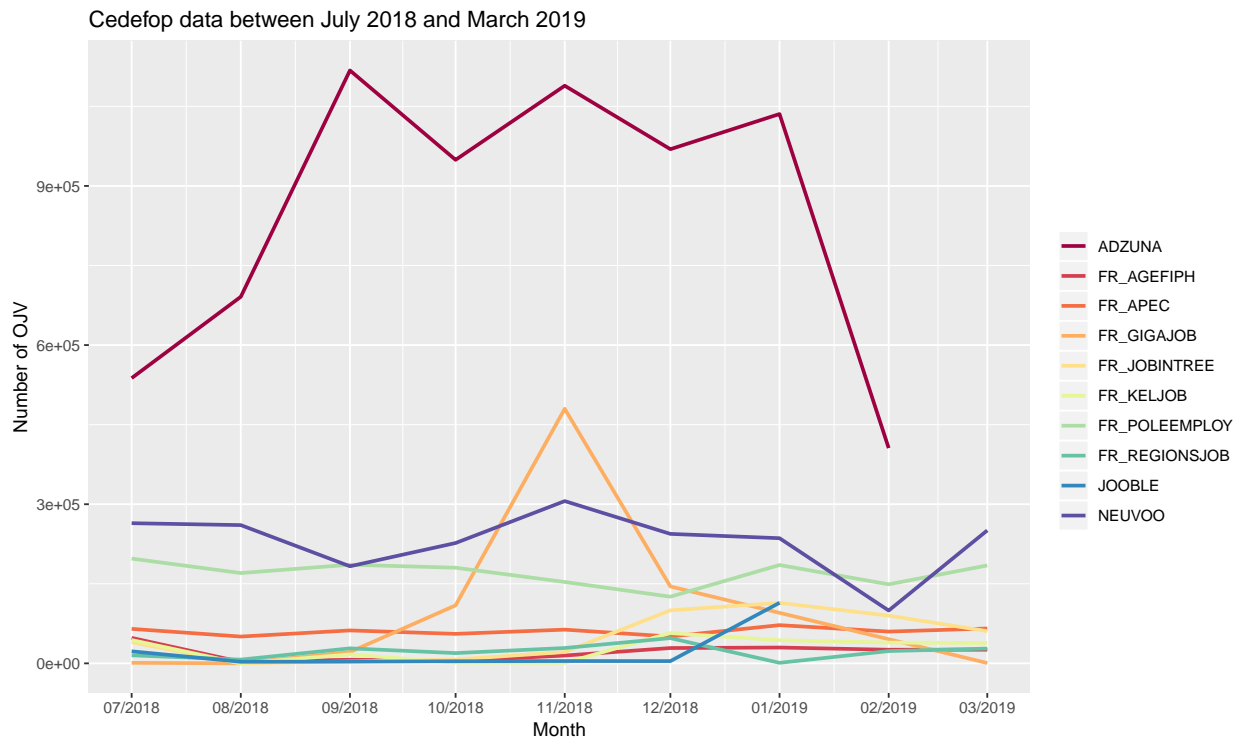
Definition

Sources represents

Most frequent sources

SOURCE	N_OJV	FREQUENCY	SOURCE	N_OJV	FREQUENCY
ADZUNA	6795429	48.11	FR_METEOJOB	147611	1.04
NEUVOO	2070757	14.66	FR_CARRIEREONLINE	112717	0.80
FR_POLEEMPLOY	1531962	10.84	FR_TROVIT	108118	0.77
FR_GIGAJOB	757155	5.36	ADECCO	100873	0.71
FR_APEC	545017	3.86	MONSTER	99606	0.71
FR_JOBINTREE	436556	3.09	FR_CADREMPLOI	93463	0.66
FR_KELJOB	241403	1.71	FR_RECRUT	74303	0.53
FR_REGIONSJOB	198316	1.40	FR_JOBLOBA	69884	0.49
FR_AGEFIPH	187170	1.32	MANPOWER	68786	0.49
JOOBLE	155693	1.10	GIGAJOB	51098	0.36

Collecting dates



Sites

3969



Definition

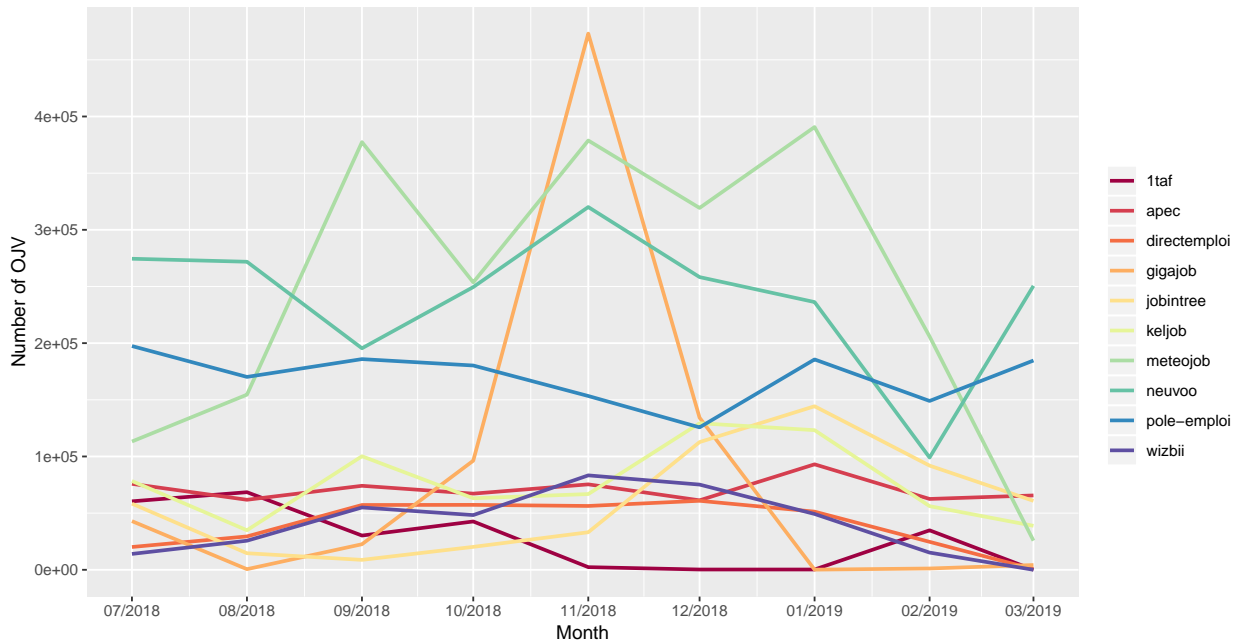
Sites represents

Top 20

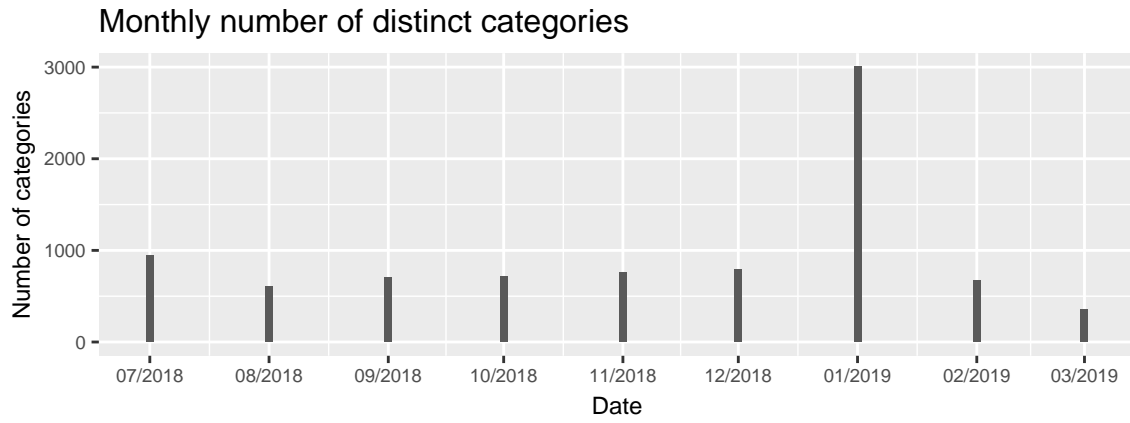
SITE	N_OJV	FREQUENCY	SITE	N_OJV	FREQUENCY
meteojob	2219195	15.71	beep job	239312	1.69
neuvoo	2155488	15.26	regionsjob	235315	1.67
pole-emploi	1532188	10.85	cadreemploi	207036	1.47
gigajob	775243	5.49	bob has a job	197350	1.40
keljob	690949	4.89	multiposting	195462	1.38
apec	636868	4.51	monster	194335	1.38
jobintree	544927	3.86	agefiph	189835	1.34
wizbii	365978	2.59	nounou-top	154380	1.09
directemploi	357878	2.53	carriereonline	133413	0.94
1taf	239522	1.70	jobvitae	130102	0.92

Top 10 time series

Cedefop data between July 2018 and March 2019

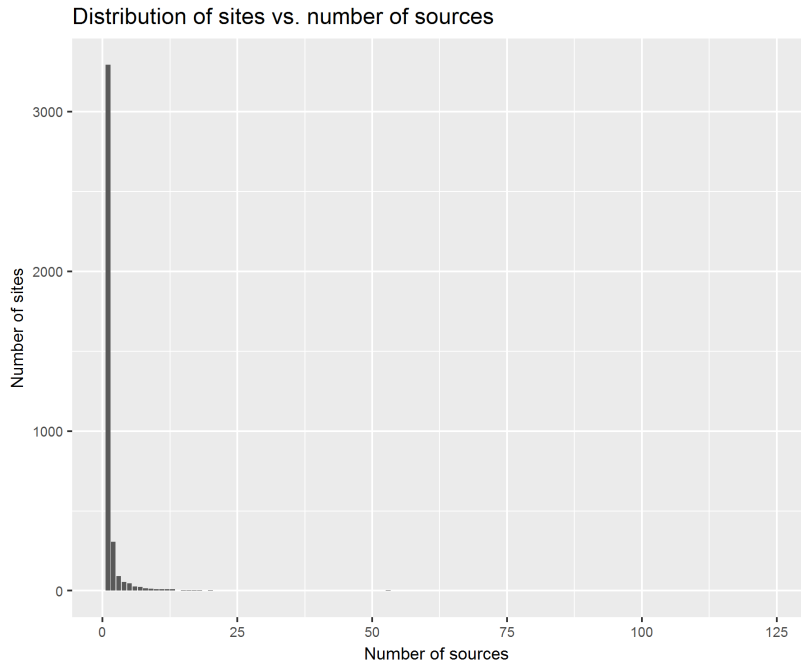


Number of sites per date

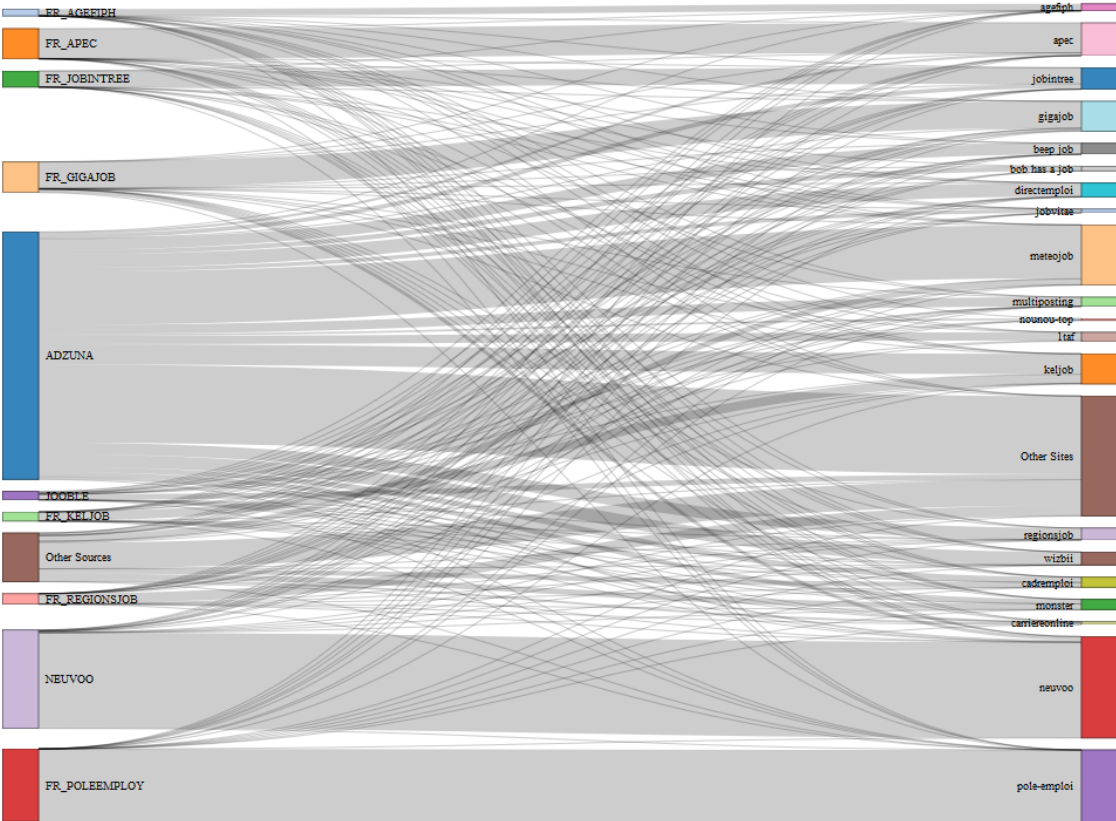


Number of sources per site

Average
1.96
Min / Max
1 / 123



Sankey diagram for top 10 sources and top 20 sites



Companies

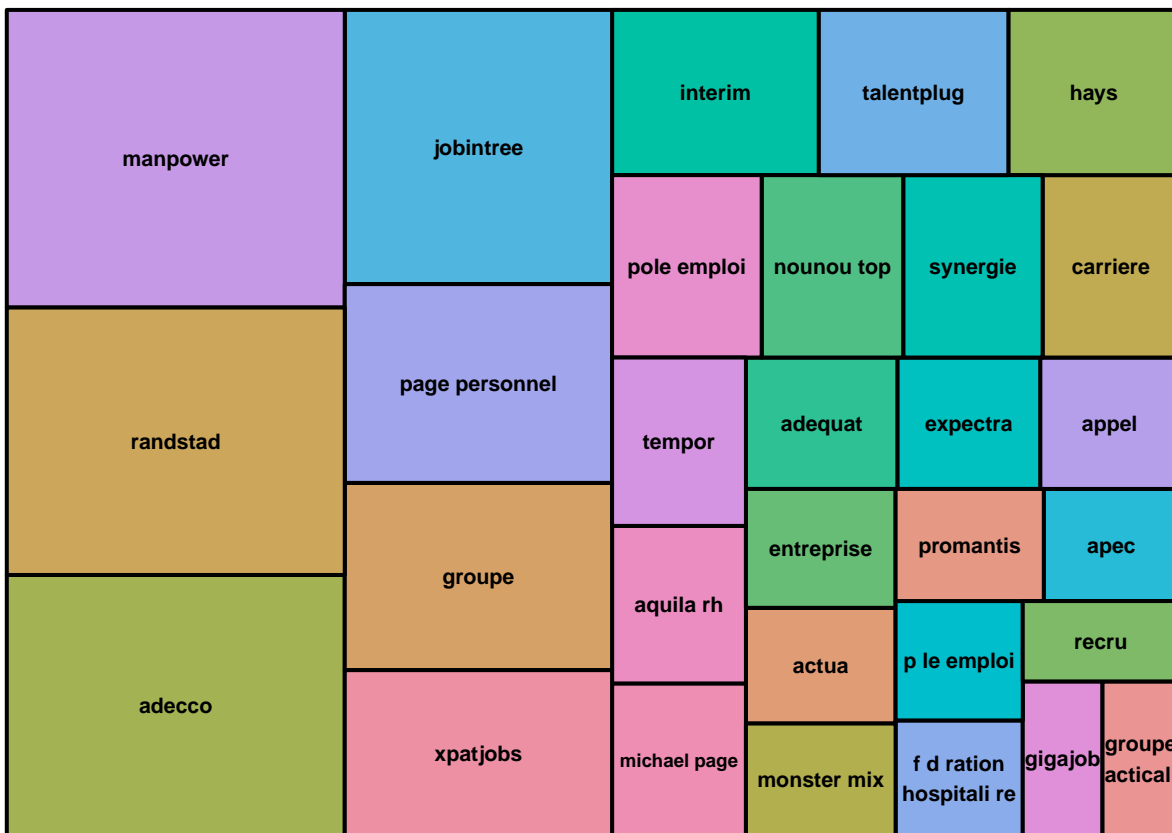
30383



Definition

Identified recruiting company in job offer. Tabulex indicated that companies names are not post processed, which can result in the same company having multiple value names (eg: Pole Emploi could be referred as “pole-emploi” or “pole emploi” in provided data).

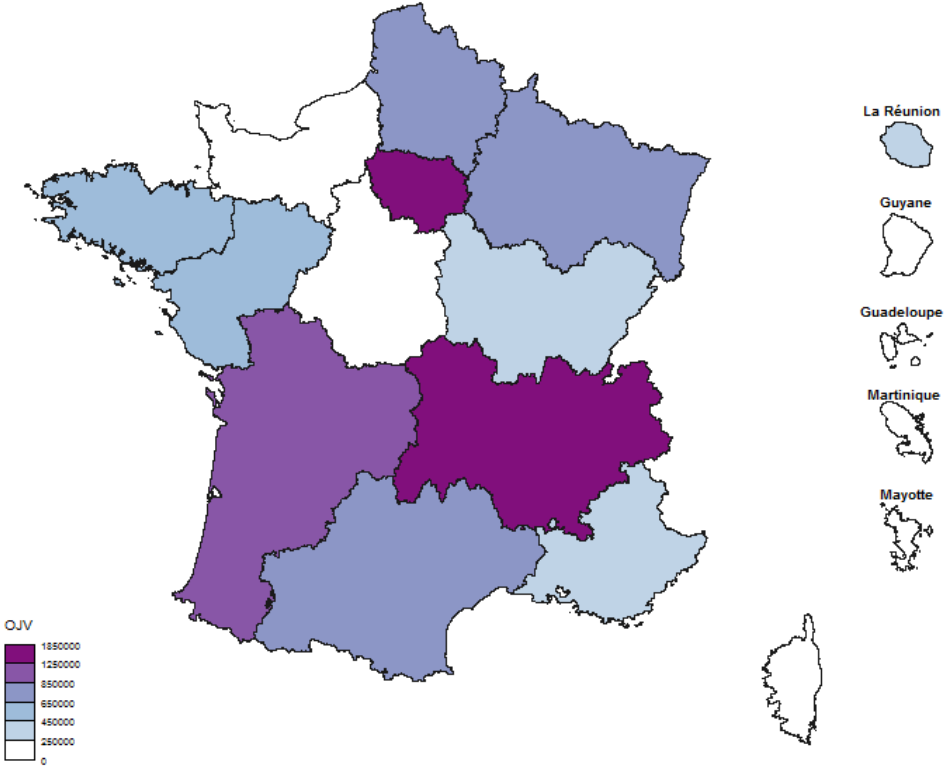
Top 30 identified companies



Regions



Number of oju per region



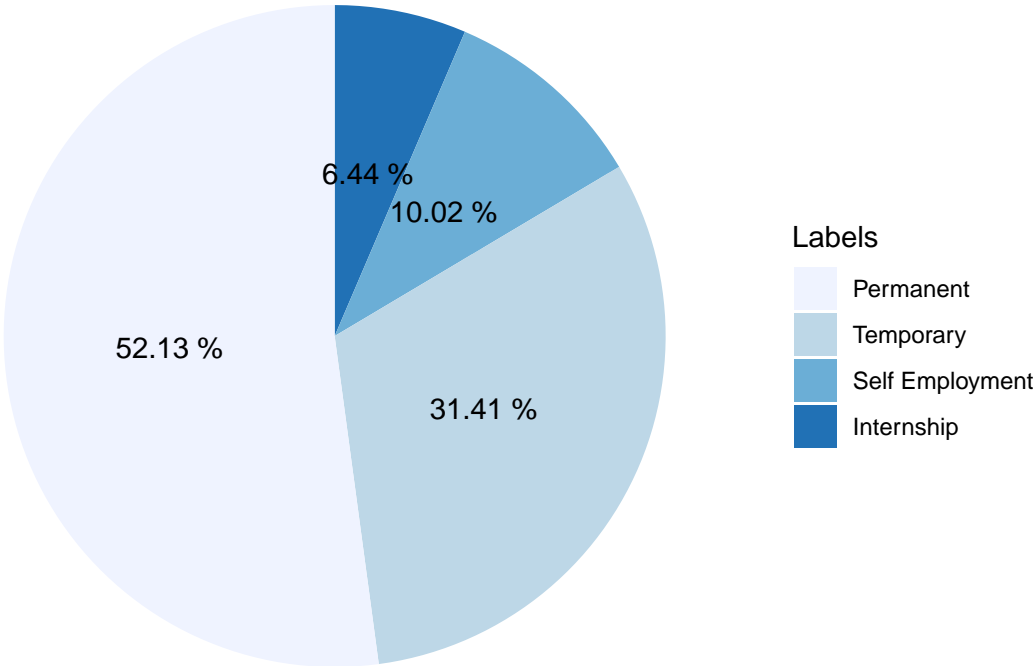
Contracts



```
unique(cedefop_ojv$IDCONTRACT)
```

```
## [1] 1 2 3 4  
## Levels: 1 2 3 4
```

Contract type (repartition per OJV, for non null value)

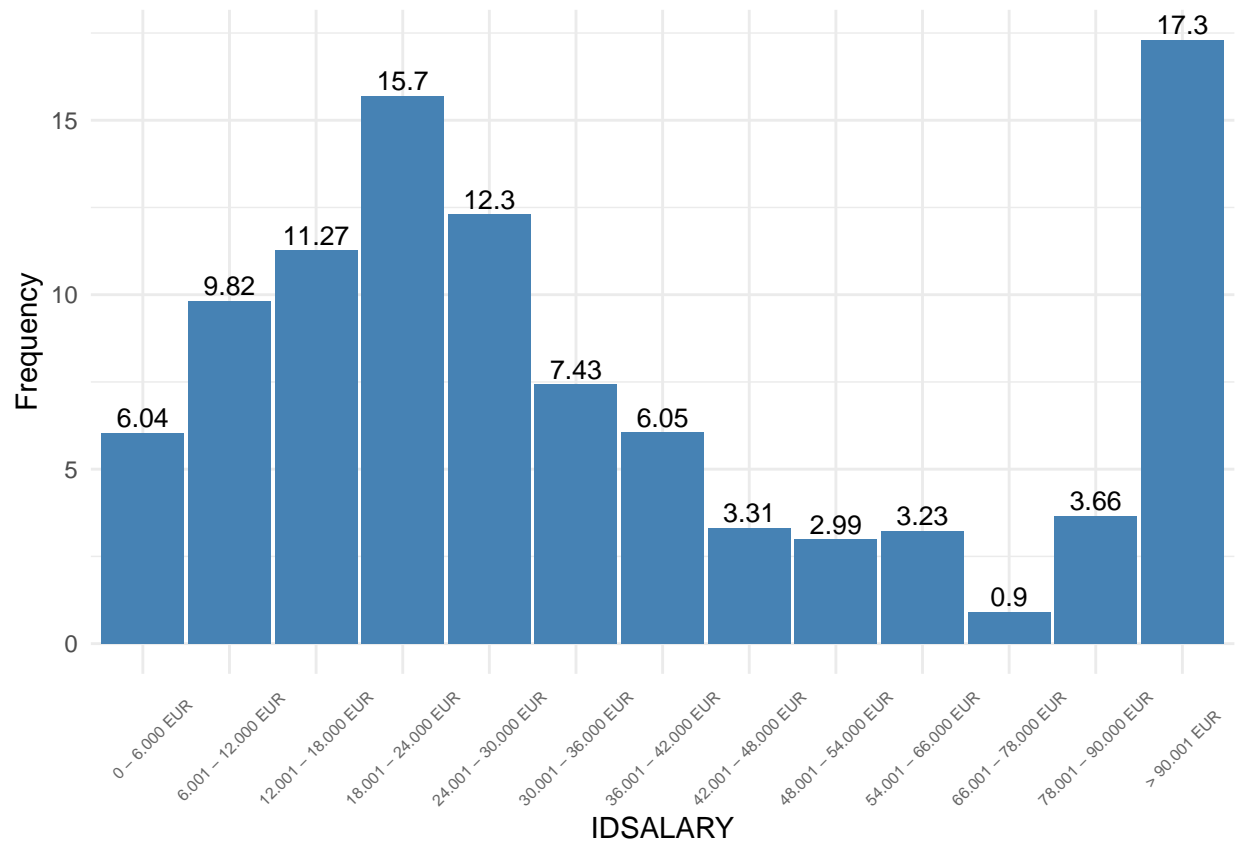


Salary

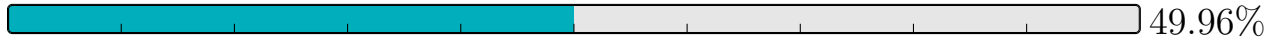


Salary (repartition per OJV, for non null value)

```
## [1] 0 - 6.000 EUR 6.001 - 12.000 EUR 12.001 - 18.000 EUR
## [4] 18.001 - 24.000 EUR 24.001 - 30.000 EUR 30.001 - 36.000 EUR
## [7] 36.001 - 42.000 EUR 42.001 - 48.000 EUR 48.001 - 54.000 EUR
## [10] 54.001 - 66.000 EUR 66.001 - 78.000 EUR 78.001 - 90.000 EUR
## [13] > 90.001 EUR
## 13 Levels: 0 - 6.000 EUR 6.001 - 12.000 EUR ... > 90.001 EUR
```

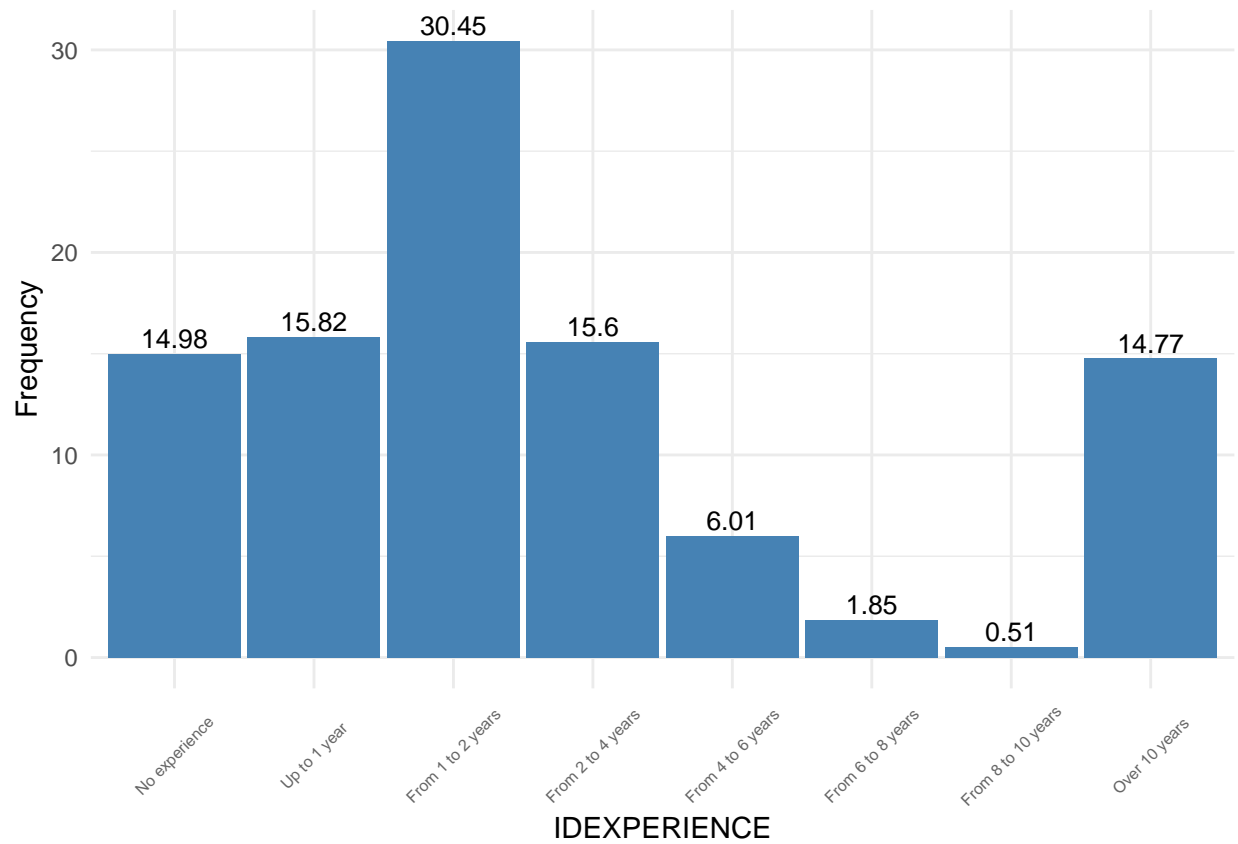


Experience



Experience (repartition per OJV, for non null value)

```
## [1] No experience      Up to 1 year      From 1 to 2 years
## [4] From 2 to 4 years    From 4 to 6 years From 6 to 8 years
## [7] From 8 to 10 years    Over 10 years
## 8 Levels: No experience Up to 1 year ... Over 10 years
```

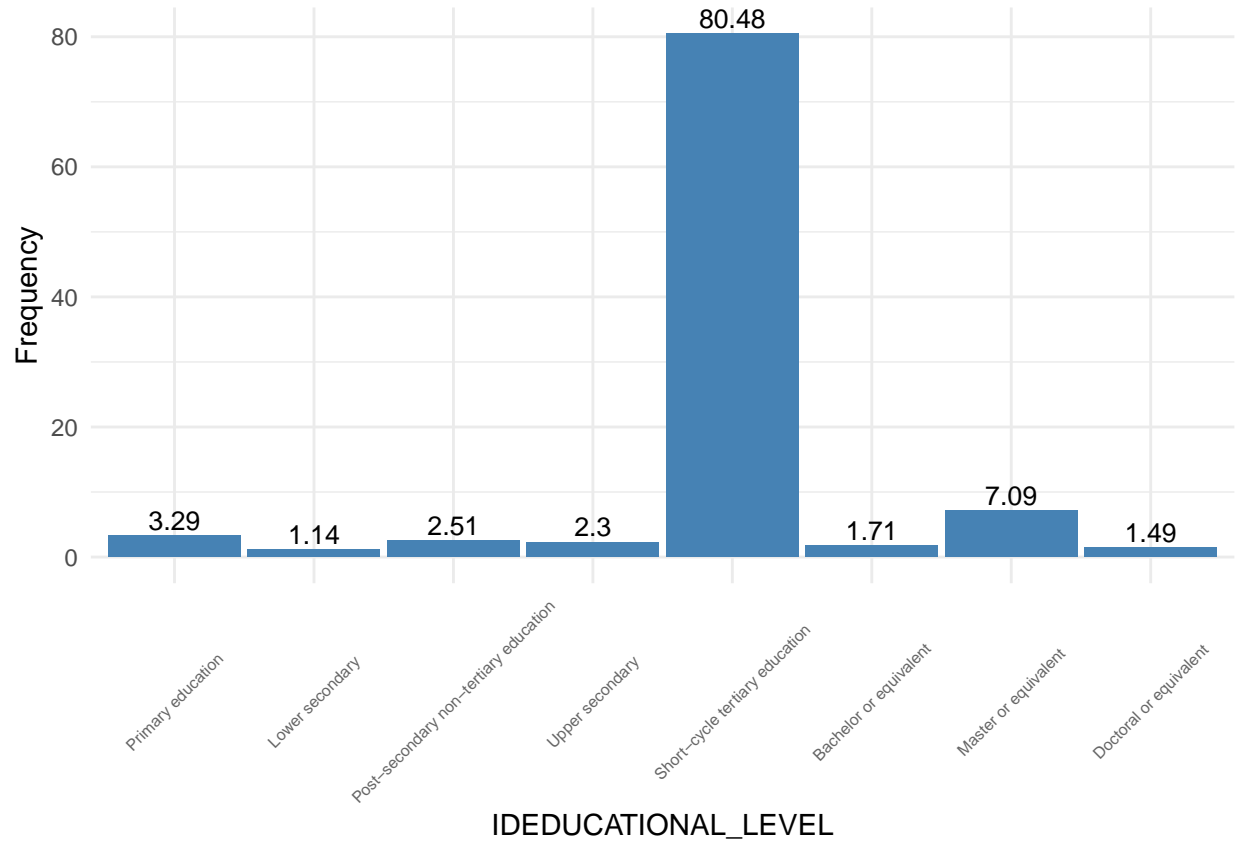


Education

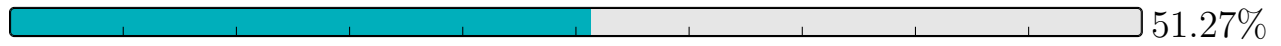


Education (repartition per OJV, for non null value)

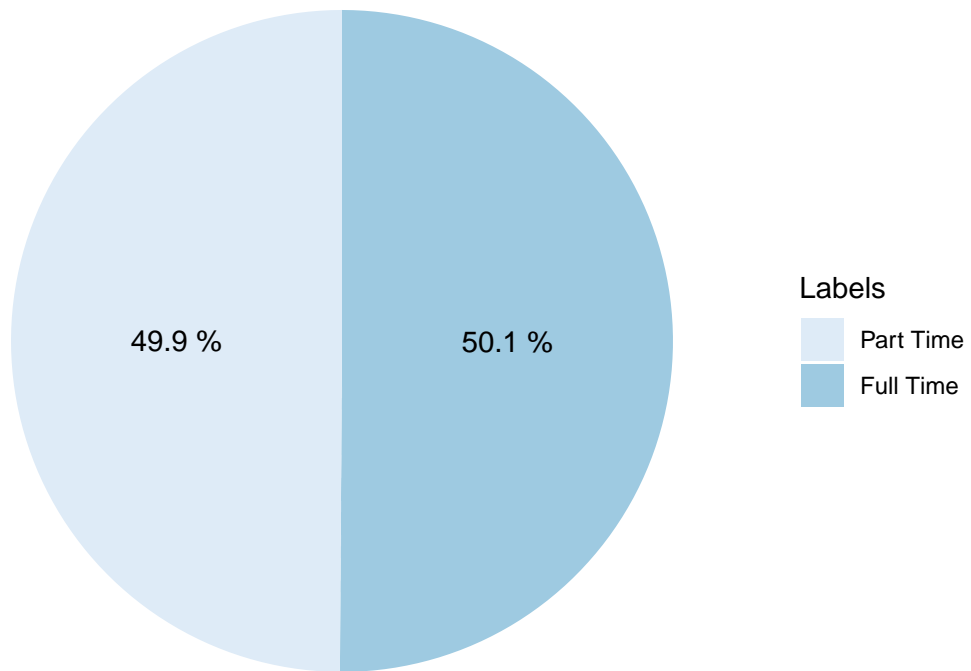
```
## [1] Primary education
## [2] Lower secondary
## [3] Post-secondary non-tertiary education
## [4] Upper secondary
## [5] Short-cycle tertiary education
## [6] Bachelor or equivalent
## [7] Master or equivalent
## [8] Doctoral or equivalent
## 8 Levels: Primary education ... Doctoral or equivalent
```



Working hours



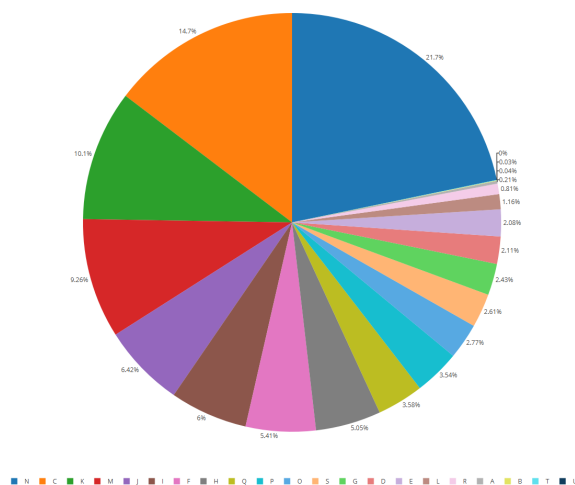
Working hours (repartition per OJV, for non null value)



Sector



Visualisation of sectors repartition of OJV with 10 levels



IDSECTOR	N_OJV	FREQUENCY	LABELS
A	28924	0.21	Agriculture, sylviculture et pêche
B	5525	0.04	Industries extractives
C	2067705	14.67	Industrie manufacturière
D	297245	2.11	Production et distribution d'électricité, de gaz, de vapeur et d'air conditionné
E	292638	2.08	Production et distribution d'eau; assainissement, gestion des déchets et dépollution
F	762894	5.41	Construction
G	342458	2.43	Commerce; réparation d'automobiles et de motocycles
H	712508	5.05	Transports et entreposage
I	846154	6.00	Hébergement et restauration
J	905214	6.42	Information et communication
K	1421008	10.08	Activités financières et d'assurance
L	163511	1.16	Activités immobilières
M	1305304	9.26	Activités spécialisées, scientifiques et techniques
N	3065516	21.75	Activités de services administratifs et de soutien
O	389760	2.77	Administration publique
P	498376	3.54	Enseignement
Q	504818	3.58	Santé humaine et action sociale
R	114333	0.81	Arts, spectacles et activités récréatives
S	367656	2.61	Autres activités de services
T	4380	0.03	Activités des ménages en tant qu'employeurs; activités indifférenciées des ménages en tant que producteurs de biens et services pour usage propre
U	72	0.00	Activités extraterritoriales

ESCO Classification

421



Visualisation of Esco repartition of OJV with 10 levels

