

# ping

January 5, 2017

## 1 Exercice 2

- Créez un nouveau notebook “ping.ipynb”
- Créez une série dont, pour chaque ligne : l’index est un code pour chaque dominante de l’Esigelec, la valeur est le nom complet de la dominante
- Ajoutez une colonne avec le département auquel chaque dominante appartient
- Créez une dataframe contenant le nom, prénom et code de dominante pour chaque membre de votre PING
- Créez une dataframe contenant le pourcentage d’élèves par département pour votre PING
- Enregistrez le résultat dans un fichier csv

```
In [1]: import pandas as pd
```

```
dominantes = pd.Series(  
    [  
        "Architecture et Sécurité des Réseaux",  
        "Big Data pour la Transformation Numérique",  
        "Ingénieur d’Affaires : Informatique et Réseaux",  
        "Ingénierie des Services du Numérique",  
        "Ingénieur Finance",  
        "Automatique et Robotique Industrielle",  
        "Énergie et Développement Durable",  
        "Génie Électrique et Transport",  
        "Ingénieur d’Affaires : Distribution Énergie et Signaux",  
        "Mécatronique Génie Électrique",  
        "Ingénierie des Systèmes Embarqués : Véhicule Autonome",  
        "Ingénierie des Systèmes Embarqués : Objets Communicants",  
        "Ingénierie des Systèmes Médicaux",  
        "Electronique des Systèmes pour l’Automobile et l’Aéronautique",  
        "Ingénierie Télécom"  
    ], index = [  
        "ASR",  
        "BDTN",  
        "IA-IR",  
        "ISN",  
        "IF",  
        "ARI",  
    ]  
)
```

```

        "EDD",
        "GET",
        "IA-DES",
        "MCTGE",
        "ISE-VA",
        "ISE-OC",
        "ISYMED",
        "ESAA",
        "ICOM"
    ], name = 'Dominante'
)
dominantes

```

```

Out[1]: ASR                Architecture et Sécurité des Réseaux
BDTN                Big Data pour la Transformation Numérique
IA-IR                Ingénieur d’Affaires : Informatique et Réseaux
ISN                Ingénierie des Services du Numérique
IF                Ingénieur Finance
ARI                Automatique et Robotique Industrielle
EDD                Énergie et Développement Durable
GET                Génie Électrique et Transport
IA-DES            Ingénieur d’Affaires : Distribution Énergie et...
MCTGE                Mécatronique Génie Électrique
ISE-VA            Ingénierie des Systèmes Embarqués : Véhicule A...
ISE-OC            Ingénierie des Systèmes Embarqués : Objets Com...
ISYMED                Ingénierie des Systèmes Médicaux
ESAA                Electronique des Systèmes pour l’Automobile et...
ICOM                Ingénierie Télécom
Name: Dominante, dtype: object

```

```

In [2]: dfDominantes = pd.DataFrame(dominantes)

```

```

dfDepartements = pd.DataFrame(
    [
        "TIC",
        "TIC",
        "TIC",
        "TIC",
        "TIC",
        "GEE",
        "GEE",
        "GEE",
        "GEE",
        "SEI",
        "SEI",
        "SEI",
        "SEI",
        "ET",
    ]
)

```

```

        "ET"
    ], index = [
        "ASR",
        "BDTN",
        "IA-IR",
        "ISN",
        "IF",
        "ARI",
        "EDD",
        "GET",
        "IA-DES",
        "MCTGE",
        "ISE-VA",
        "ISE-OC",
        "ISYMED",
        "ESAA",
        "ICOM"
    ], columns = ['Departement']
)

```

```

In [3]: dfDominantes['Key'] = dfDominantes.index
        dfDepartements['Key'] = dfDepartements.index

```

```

dfDepDom = pd.DataFrame(pd.merge(dfDominantes, dfDepartements, on='Key'))
dfDepDom

```

```

Out [3]:

```

	Dominante	Key	Departement
0	Architecture et Sécurité des Réseaux	ASR	TIC
1	Big Data pour la Transformation Numérique	BDTN	TIC
2	Ingénieur d'Affaires : Informatique et Réseaux	IA-IR	TIC
3	Ingénierie des Services du Numérique	ISN	TIC
4	Ingénieur Finance	IF	TIC
5	Automatique et Robotique Industrielle	ARI	GEE
6	Énergie et Développement Durable	EDD	GEE
7	Génie Électrique et Transport	GET	GEE
8	Ingénieur d'Affaires : Distribution Énergie et...	IA-DES	GEE
9	Mécatronique Génie Électrique	MCTGE	SEI
10	Ingénierie des Systèmes Embarqués : Véhicule A...	ISE-VA	SEI
11	Ingénierie des Systèmes Embarqués : Objets Com...	ISE-OC	SEI
12	Ingénierie des Systèmes Médicaux	ISYMED	SEI
13	Electronique des Systèmes pour l'Automobile et...	ESAA	ET
14	Ingénierie Télécom	ICOM	ET

```

In [4]: member_1 = pd.Series({
        'firstName': 'Mathieu',
        'lastName': 'Cailly',
        'codeDominante': 'ISE-OC'
    })

```

```

member_2 = pd.Series({
    'fistName': 'Cyrille',
    'lastName': 'Kamdem',
    'codeDominante': 'IA-IR'
})

member_3 = pd.Series({
    'fistName': 'Florent',
    'lastName': 'Bacque',
    'codeDominante': 'BDTN'
})

member_4 = pd.Series({
    'fistName': 'Emmanuel',
    'lastName': 'Veron',
    'codeDominante': 'ISN'
})

member_5 = pd.Series({
    'fistName': 'Aymeric',
    'lastName': 'Descamps',
    'codeDominante': 'ISE-VA'
})

member_6 = pd.Series({
    'fistName': 'Aimy',
    'lastName': 'Carlet',
    'codeDominante': 'BDTN'
})

dfMembers = pd.DataFrame([
    member_1,
    member_2,
    member_3,
    member_4,
    member_5,
    member_6
])
dfMembers

```

```

Out[4]:

```

	codeDominante	fistName	lastName
0	ISE-OC	Mathieu	Cailly
1	IA-IR	Cyrille	Kamdem
2	BDTN	Florent	Bacque
3	ISN	Emmanuel	Veron
4	ISE-VA	Aymeric	Descamps
5	BDTN	Aimy	Carlet

Ajoutons une colonne au dataframe contenant les dominantes et départements

```
In [5]: dfMembers['number'] = 1
dfMembers = dfMembers.groupby('codeDominante').sum()
dfMembers['codeDominante'] = dfMembers.index

dfFinal = dfDepDom.loc[:, ['Key', 'Departement']].copy()
dfFinal = pd.merge(dfFinal, dfMembers, left_on='Key', right_on='codeDominante')
del dfFinal['codeDominante']
dfFinal = dfFinal.fillna(value = 0)
dfFinal = dfFinal.groupby('Departement').sum()
dfFinal['number'] = dfFinal['number']/dfFinal['number'].sum()
```

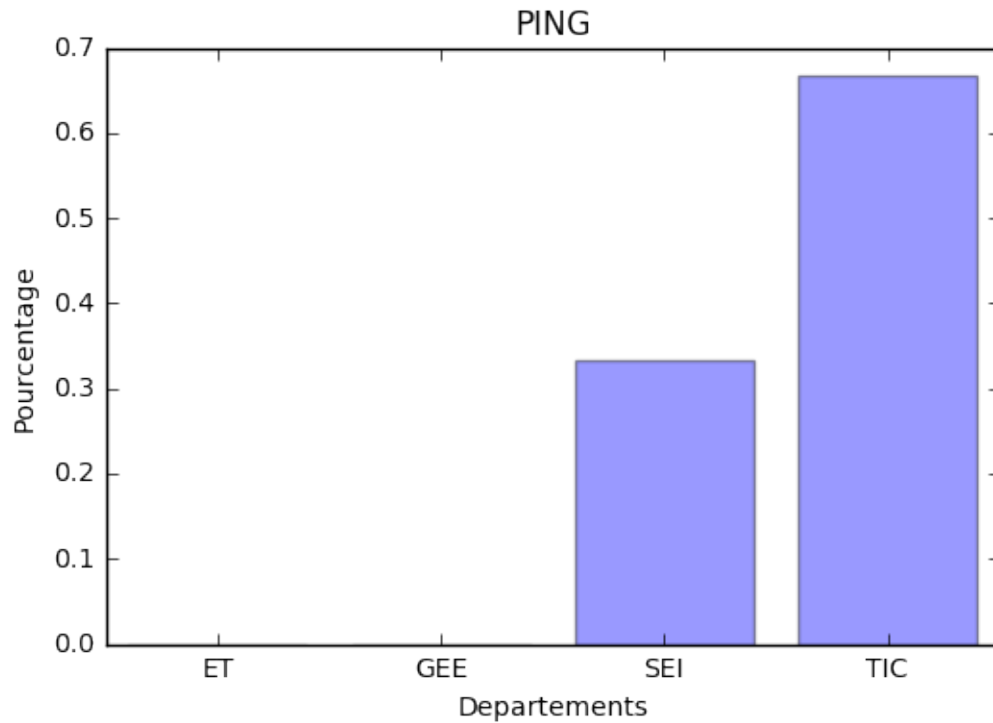
Visualisation des données :

```
In [16]: import matplotlib.pyplot as plt
import numpy as np

# Example data
departements = dfFinal.index.values
x_pos = np.arange(len(departements))
pourcentage = dfFinal['number']

plt.bar(x_pos, pourcentage, align='center', alpha=0.4)
plt.xticks(x_pos, departements)
plt.xlabel('Departements')
plt.ylabel('Pourcentage')
plt.title('PING')

plt.show()
```



```
In [18]: labels = dfFinal.index.values
        sizes = dfFinal['number']*100
        colors = ['yellowgreen', 'gold', 'lightskyblue', 'lightcoral']
        explode = (0, 0, 0, 0.1) # only "explode" the 2nd slice (i.e. 'Hogs')

        plt.pie(sizes, explode=explode, labels=labels, colors=colors,
                autopct='%1.1f%%', shadow=True, startangle=90)
        plt.axis('equal')

        plt.show()
```

