



## Rapport Saé-2.04 - Partie Stats

**Etudiants :**

**[ TURPIN, Samuel ]**

**Groupe : E**

**Mardi 14 mai 2024 – 10h00**



**Université  
de Lille**

Ce rapport porte sur la Saé-2.04 partie stats.

Dans cette partie il nous a été demandé avec les mêmes tables que la partie bdd de regarder en détail les statistiques de parcoussup en fonction d'une région.

La région qui nous a été attribuée est le Pays de la Loire et Le Centre-Val de Loire.

### **Partie 1 :**

Pour cette partie il nous a été demandé de vérifier 3 affirmations.

#### *1) Les filles n'aiment pas les études scientifiques.*

Pour vérifier ceci nous avons eu besoins de 2 requêtes, une qui calcule le nombre de filles du département qui ont candidaté et le nombre de filles qui ont candidaté pour des études scientifiques.

```
SELECT SUM(CAST(n77 AS NUMERIC)) AS NombreFilleScientifique
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN eta-
blissements AS E ON F.eno=E.eno
WHERE filiere_formation LIKE '%scien%' AND (E.region_etablissement='Pays de la
Loire' OR E.region_etablissement='Centre-Val de Loire');

SELECT SUM(CAST(n77 AS NUMERIC)) AS NombreFilleDepartement
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN etablis-
sements AS E ON F.eno=E.eno
WHERE E.region_etablissement='Pays de la Loire' OR E.region_etablisse-
ment='Centre-Val de Loire';
```

La première calcule le nombre de filles parmi les filières de formation avec les noms qui contiennent la chaîne de caractère « scien » dans leur texte qui peut former les mots « science » et « scientifique » étant aussi des départements du Pays de la Loire ou du Centre-Val de Loire.

Pour la deuxième on calcule le nombre total de filles dans les départements.

On obtient pour chacun :

```
--nombrefillescientifique
-----
-- 1325188.0

-- nombrefilledepartement
-----
-- 45230497.0
```

On calcule le nombre de fille en filière scientifique divisé par le nombre total de fille des départements, ce qui nous donne 2,9%.

```
(1325188.0/45230497.0)=0.029
```

Ça pourrait confirmer l'assertion de départ bien qu'on ne puisse pas la confirmer à 100%, en effet nous n'avons pas de données sur les garçons, pour la confirmer pleinement il faudrait comparer le pourcentage des filles par rapport à celui des garçons.

*2) Les bacs technologiques ont plus de chances que les bacs généraux d'être acceptés en filières sélectives.*

Pour vérifier cette assertion nous aurons besoin de calculer deux produits pour les comparer :

```
SELECT Sum(n58) AS BacTechSelect
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN eta-
blissements AS E ON F.eno=E.eno
WHERE F.selectivite='formation sélective' AND (E.region_etablissement='Pays de
la Loire' OR E.region_etablissement='Centre-Val de Loire');

SELECT Sum(n58) AS BacTechTotal
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN eta-
blissements AS E ON F.eno=E.eno
WHERE E.region_etablissement='Pays de la Loire' OR E.region_etablisse-
ment='Centre-Val de Loire';

SELECT Sum(n57) AS BacGenSelect
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN eta-
blissements AS E ON F.eno=E.eno
WHERE F.selectivite='formation sélective' AND (E.region_etablissement='Pays de
la Loire' OR E.region_etablissement='Centre-Val de Loire');

SELECT Sum(n57) AS BacGenTotal
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN eta-
blissements AS E ON F.eno=E.eno
WHERE E.region_etablissement='Pays de la Loire' OR E.region_etablisse-
ment='Centre-Val de Loire';
```

On a comme la première question un calcul des bacs sélectionnés puis un calcul du total dans le but de calculer deux produits que l'on pourra comparer.

Voici ce que ça donne :

```
-- bactechselect
-----
--      851736

-- bactechtotal
-----
--      3669482

-- bacgenselect
-----
--      8206659

-- bacgentotal
-----
--      41845802
```

Pour le ProduitTech ont à 23,2% :

```
--ProduitTech=(851736/3669482)=0.232
```

Pour le ProduitGen ont à 19,6% :

```
--ProduitGen=(8206659/41845802)=0.196
```

ProduitTech étant supérieur à ProduitGen alors l'assertion est vraie bien que la différence ne soit pas énorme et peut être influencer par un grand nombre de BTS ou de DUT dans les départements.

### *3) Les boursiers préfèrent les études courtes.*

Dans nos tables les boursiers sont répartis parmi les bacs techno, généraux et professionnels, ils faut alors les additionner.

Sachant que les études les plus courtes sont généralement les DUT et les BTS durant 2 ans.

```
SELECT CONCAT(Sum(n24)+Sum(n26)+Sum(n28)) AS TotalBoursiers
FROM stats AS S INNER JOIN etablissements AS E ON S.eno=E.eno
WHERE E.region_etablissement='Pays de la Loire' OR E.region_etablissement='Centre-Val de Loire';

SELECT CONCAT(Sum(n24)+Sum(n26)+Sum(n28)) AS CoursEtudeBoursiers
```

```
FROM formations AS F INNER JOIN stats AS S ON F.eno=S.eno INNER JOIN etablis-
sements AS E ON F.eno=E.eno
WHERE (filiere_formation LIKE '%BTS%' OR filiere_formation LIKE '%DUT%') AND
(E.region_etablissement='Pays de la Loire' OR E.region_etablissement='Centre-
Val de Loire');
```

Pour la première le nombre total de boursiers, dans la seconde on regarde parmi les boursiers combien ont choisis soit une formation « BTS » ou « DUT ».

Ce qui donne ça :

```
-- totalboursiers
-----
-- 2049550

-- coursetudeboursiers
-----
-- 960290
```

Comme la première assertion ont calcule pour obtenir le nombre de boursier choisissant des études courtes et on n'obtient 46,8%.

```
ProduitBoursiers=(960290/2049550)=0.468
```

Ce qui n'est pas suffisant pour faire une généralité sur les boursiers mais c'est un chiffre quand même élevé qui s'explique peut-être par une envie d'arriver dans le monde de l'entreprise plus tôt ou un manque de moyen.

## **Partie 2 :**

Pour cette partie il nous à été demandé de choisir dans ces régions 2 IUT qui ont tout deux un BUT Informatique.

Via des recherches internet nous en avons trouvé 2 différents.

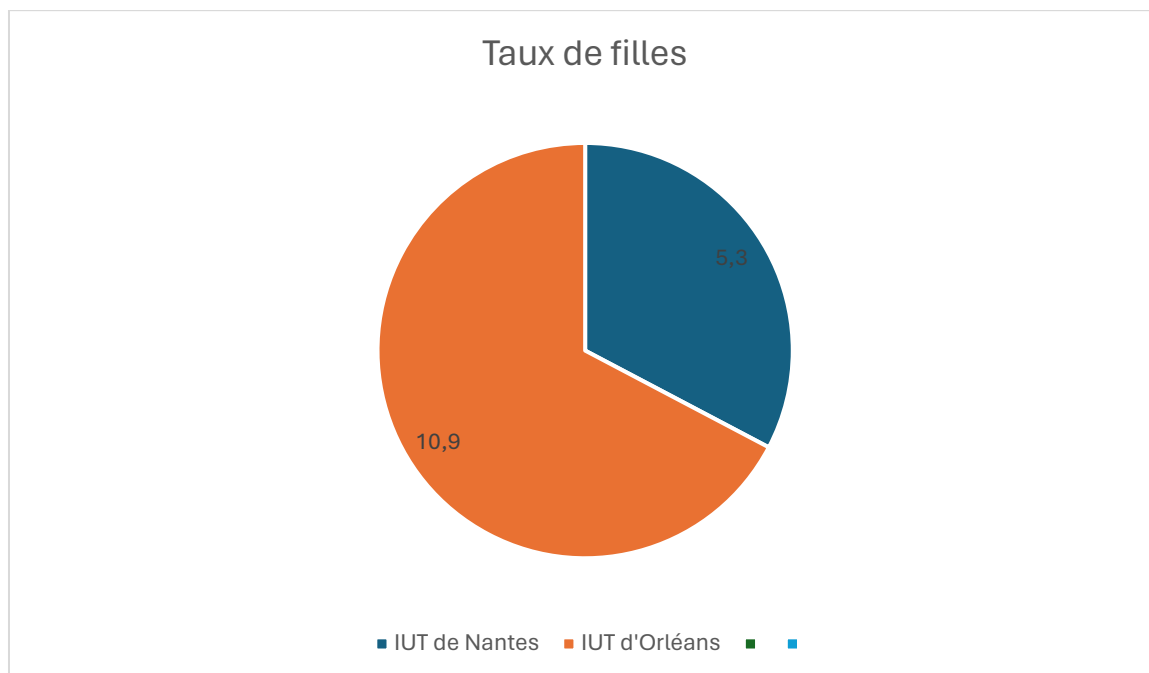
IUT de Nantes pour la région du Pays de la Loire : <https://iutnantes.univ-nantes.fr/formations/bachelor-iut-bac3/but-info>

IUT d'Orléans pour la région Centre-Val de Loire : <https://www.univ-orleans.fr/fr/iut-orleans/formation/informatique/but-informatique>

Via divers requêtes sql nous avons calculé les différentes donnée demander :

Les requêtes sont globalement similaires au premières et sont juste plus nombreuse pour comparer les IUT entres eux.

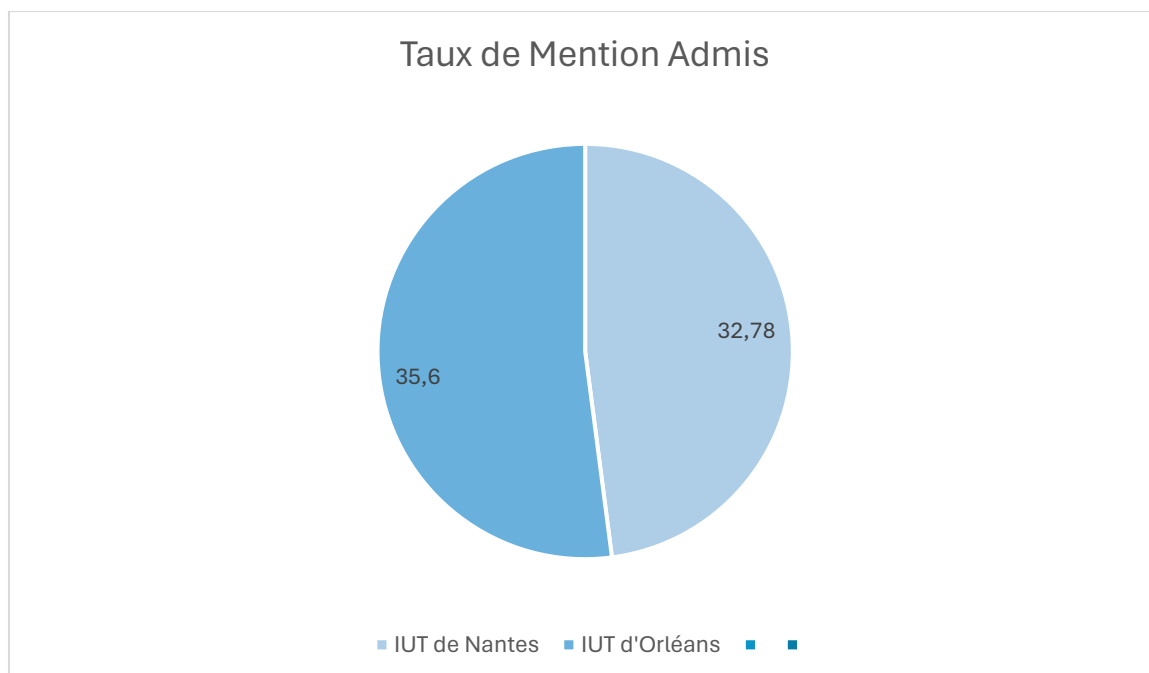
### ***Taux de filles qui ont candidaté :***



Pour ces données j'ai pris le nombre de filles divisé par le nombre de candidats ce qui nous donne ces pourcentages.

Le taux de filles est bien inférieur pour l'IUT de Nantes cela peut s'expliquer par la communication de l'IUT d'Orléans par rapport à celui de Nantes mais aussi de la proximité d'Orléans par rapport à Paris.

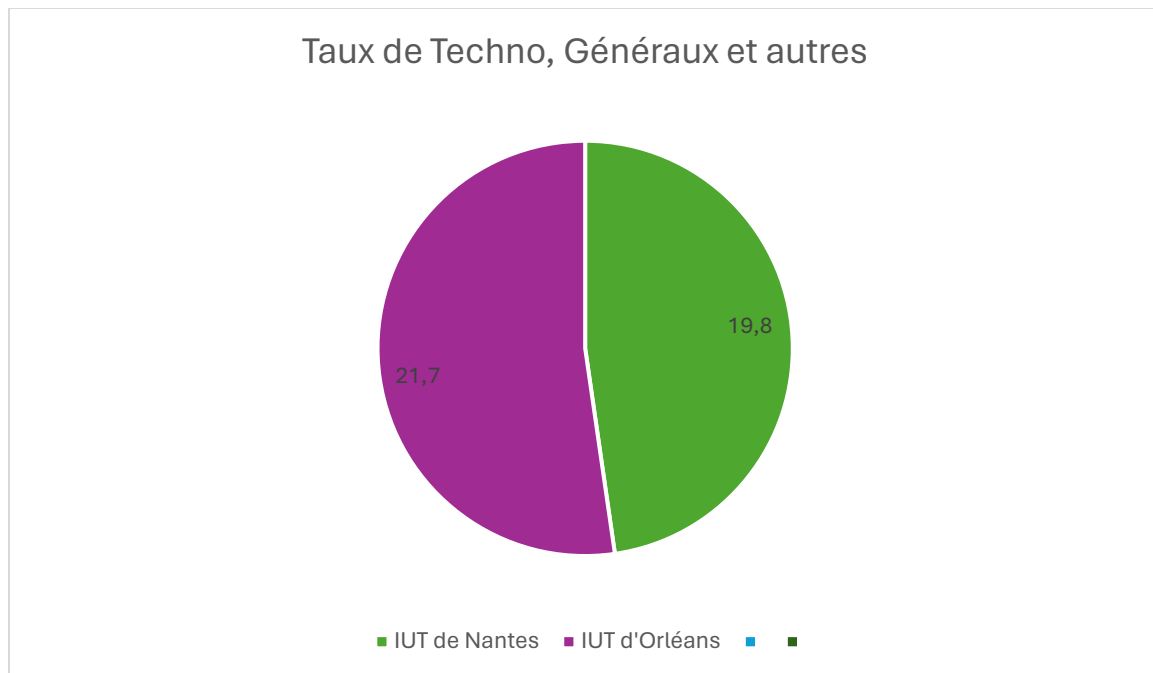
### ***Taux d'admis ayant une mention :***



Pour ces données j'ai additionné ensemble le nombre de personnes de filière X ayant eu une mention.

L'IUT d'Orléans à encore un pourcentage supérieur ce qui peut s'expliquer encore une fois par la proximité d'Orléans par rapport à Paris et par ce facteur là les néo-bacheliers s'inscrivent à cette formation et ceci peut être suivis par un temps de transport moins long que Nantes et tout ça pour la même formation

### ***Taux de BTN(Techno), BG(Généraux) et autres classés :***



Rien de particulier au niveau des requêtes sql un simple calcul des taux d'admis par filières, c'est finalement le seul graphique où les deux IUT sont à quasi-égalité puisque les BUT Informatique ont tout deux les mêmes cibles.

### **Conclusion**

A travers cette partie de la Saé-2.04 nous avons appris à utiliser les données jusque là mis de côté dans la partie bdd, en faisant des requêtes qui ne concernent que des régions spécifiques ceci nous a permis de voir d'énormes différences entre les régions de France.

La chose qui pourrait améliorer ce rapport est l'ajout d'une colonne avec le total de certains effectifs ainsi que le taux de garçons.