

SAÉ 2.04 BDD Stats | Rapport

Table of Contents

1. Les filles n'aiment pas les études scientifiques:	1
2. Les bacs technologiques ont plus de chances que les bacs généraux d'être acceptés en filières sélectives:	2
3. Les boursiers préfèrent les études courtes:	4
4. Détermination des BUT à analyser:	5
5. Taux de filles qui ont candidatées:	6
6. Taux d'admis avec mention:	6
7. Taux de BTN, BG et autres classés:	8
8. Informations complémentaires:	11



****Projet réalisé part FANTUZ CHROBOT Léo, groupe A**

Présentation

Dans un premier temps, nous avons appris à créer une base de données, à y importer des données et à les utiliser. Désormé, nous allons interprété ces données à des fins totalement statistiques nous permettant ainsi d'avoir un réel aperçu concernant certains points principaux de parcoursup (Ici, nous étudierons les données pour les régions Normandie et Pays de la Loire).

[[Première partie]] == Partie 1

1. Les filles n'aiment pas les études scientifiques:

Pour commencer, nous devons trouver le nombre de candidates au total. Ainsi, nous pouvons en déduire qu'il y en a **510 514** à l'aide de la requête suivante:

```
SELECT SUM(candidates) FROM Formations
```

```
WHERE codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --510 514 candidates totales
```

Si nous prenons les matières scientifiques principales (science, santé, chimie, informatique, biotechnologie, classes préparatoires, physique), nous avons 17 001 candidates :

```
SELECT SUM(candidates) FROM Formations
```

```
WHERE (filiereDetailee LIKE '%science%' OR filiereDetailee LIKE '%santé%' OR filiereDetailee
```

```
LIKE '%chimie%' OR filiereDetailee LIKE '%informatique%' OR filiereDetailee LIKE '%biotechnologie%' OR filiereDetailee LIKE '%prépa%' OR filiereDetailee LIKE '%physique%')
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --17 001 candidates
```

En termes de pourcentage, cela équivaut à environ **3.33% des candidates totales** ayant candidaté pour des formations scientifiques (certes, certaines formations scientifiques n'ont pas été mentionnées ici, mais le pourcentage reste approximativement le même, à savoir très faible). Ainsi, nous pouvons en conclure que les filles n'aiment pas les études scientifiques dues à leur faibles candidatures dans ces formations. Pourtant, toutes ces filières correspondent tout de même à **74 formations**. Cela peut être déduit à l'aide de la requête suivante:

```
SELECT count(*) FROM Formations
```

```
WHERE (filiereDetailee LIKE '%science%' OR filiereDetailee LIKE '%santé%' OR filiereDetailee
```

```
LIKE '%chimie%' OR filiereDetailee LIKE '%informatique%' OR filiereDetailee LIKE '%biotechnologie%' OR filiereDetailee LIKE '%prépa%' OR filiereDetailee LIKE '%physique%')
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --74 formations scientifiques
```

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

Pour commencer, nous devons en déduire le nombre total d'étudiants en bac général et technologique. Nous pouvons en déduire qu'il y a au total **31 770 admis généraux** et **8954 admis technologiques** à l'aide des requêtes suivantes:

```
SELECT SUM(a.generaux) FROM Formations AS f, Admis AS a
```

```
WHERE a.formation = f.codeFormation
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --31 770 admis généraux totaux
```

```
SELECT SUM(a.techno) FROM Formations AS f, Admis AS a
```

```
WHERE a.formation = f.codeFormation
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --8954 admis technologiques totaux
```

Ensuite, nous devons en déduire le nombre d'étudiants en filière sélective pour les bacs généraux et technologiques. Nous pouvons en déduire qu'il y a au total **14 433 admis généraux** et **7345 admis technologiques** à l'aide des requêtes suivantes:

```
SELECT SUM(a.generaux) FROM Formations AS f, Admis AS a
```

```
WHERE a.formation = f.codeFormation
```

```
AND selectivite = 'formation sélective'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --14 433 admis généraux en filières sélectives
```

```
SELECT SUM(a.techno) FROM Formations AS f, Admis AS a
```

```
WHERE a.formation = f.codeFormation
```

```
AND selectivite = 'formation sélective'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --7345 admis technologiques en filières sélectives
```

Pour les généraux, après un **produit en croix**, nous avons:

```
14 433 = 46%
```

```
31 770 = 100%
```

Pour les technologiques, après un **produit en croix**, nous avons:

```
7345 = 82%
```

```
8954 = 100%
```

Certes, si l'on regarde uniquement les bacs généraux et technologiques admis en filières sélectives, les généraux sont beaucoup plus présent dans ces filières que les technologiques. Cependant, si l'on compare cela au total d'admis pour ces deux bacs, nous pouvons vite nous rendre compte à l'aide d'un produit en croix que les technologiques ont un total de **82%** de leurs candidats admis qui ont été admis en formations sélectives, alors que seulement **46%** des généraux ont été admis en filières sélectives. Ainsi, nous pouvons voir que les technologiques ont bien plus de chance (quasiment le double) d'être acceptés en filières sélectives que les généraux.

3. Les boursiers préfèrent les études courtes:

Dans un premier temps, nous allons déterminer le nombre de formations courtes disponibles sur parcoursup pour les régions Normandie et Pays de la Loire. Grâce à la requête suivante, nous savons qu'il y a **1036 formations courtes** sur parcoursup pour les régions cités précédemment:

```
SELECT count(*) from Formations  
  
WHERE (filiereDetaillee LIKE '%BTS%'  
  
OR filiereDetaillee LIKE '%BUT%'  
  
OR filiereDetaillee LIKE '%Licence%')  
  
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); — 1036 formations courtes
```

Ensuite, nous devons en déduire le nombre total de boursiers en formations courtes. Grâce à la requête suivante, nous savons qu'il y en a **85 430**:

```
SELECT SUM(b.generaux + b.techno + b.professionnels) FROM Formations AS f JOIN Boursiers AS b  
ON (codeFormation = formation)  
  
WHERE codeFormation = formation  
  
AND filiereDetaillee LIKE '%BTS%'  
  
AND filiereDetaillee LIKE '%BUT%'  
  
AND filiereDetaillee LIKE '%Licence%'  
  
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --85 430 Boursiers en formations courtes
```

Par la suite, nous devons en déduire le nombre total de formations longues disponibles sur parcoursup pour les régions Normandie et Pays de la Loire. Grâce à la requête suivante, nous savons qu'il y a **383 formations longues** sur parcoursup pour les régions cités précédemment:

```
SELECT count(*) from Formations  
  
AND filiereDetaillee NOT LIKE '%BTS%'  
  
OR filiereDetaillee NOT LIKE '%BUT%'  
  
OR filiereDetaillee NOT LIKE '%Licence%'  
  
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --383 formations longues
```

Pour finir, nous devons en déduire le nombre total de boursiers en formations longues. Grâce à la

requête suivante, nous savons qu'il y en a **28 413**:

```
SELECT SUM(b.generaux + b.techno + b.professionnels) FROM Formations AS f JOIN Boursiers AS b
ON (codeFormation = formation)

WHERE codeFormation = formation

AND filiereDetaillee NOT LIKE '%BTS%'

AND filiereDetaillee NOT LIKE '%BUT%'

AND filiereDetaillee NOT LIKE '%Licence%'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --28 413 Boursiers en formations longues
```

Ci-dessous, voici le calcul des **moyennes** pour les formations courtes et longues:

Moyennes:

$85\,430 / 1036 = 82$

$28\,413 / 383 = 74$

Ainsi, les Boursiers préfèrent les formations courtes. Cela peut ainsi être déduit en comparant le nombre moyen de candidatures boursières des formations courtes et des formations longues, comme prouvé ici même (tout ça dans les régions Normandie et Pays de la Loire). En effet, nous avons ainsi **82 boursiers en formations courtes** et **74 boursiers en formations longues**. Nous avons ainsi plus de boursiers en formations courtes qu'en formations longues.

[[Seconde partie]] == Partie 2

4. Détermination des BUT à analyser:

À l'aide de la requête suivante, j'ai décidé de comparer le **BUT Informatique de Nantes** (dans le **Pays de la Loire**) et le **BUT Informatique du Havre** (en **Normandie**):

```
SELECT codeFormation, nom, f.codeDep, f.commune, status, filiereDetaillee FROM Formations AS f,
Etablissements AS e

WHERE (e.UAI = f.UAI AND e.codeDep = f.codeDep AND e.commune = f.commune)

AND filiereDetaillee LIKE '%BUT - Informatique%' AND f.codeDep IN
('14','27','50','61','76','44','49','53','72','85');
```

Ainsi, voici les codes de formation des deux BUT:

Code de formation du BUT de Nantes = 5482

Code de formation du BUT du Havre = 9655

5. Taux de filles qui ont candidatées:

Dans un premier temps, nous allons en déduire le **nombre total de candidates** pour le BUT de Nantes et le BUT du Havre. Ainsi, à l'aide des requêtes suivantes, nous savons qu'il y a **220 candidates pour le BUT de Nantes**, et **202 candidates** pour le BUT du Havre:

Pour Nantes:

```
SELECT candidates FROM Formations
```

```
WHERE codeFormation = '5482'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --220 candidates
```

Pour le Havre:

```
SELECT candidates FROM Formations
```

```
WHERE codeFormation = '9655'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --202 candidates
```

Comme cité précédemment, nous avons ainsi **220 candidates** pour le BUT Informatique de Nantes contre **202** pour celui du Havre. Ainsi, nous avons plus de candidates dans le BUT Informatique de Nantes que dans celui du Havre.

6. Taux d'admis avec mention:

Pour commencer, nous allons en déduire le nombre total d'admis pour le BUT de Nantes et le BUT du Havre. Ainsi, à l'aide des requêtes suivantes, nous savons qu'il y a **82 admis** pour le BUT de Nantes, et **71 admis** BUT pour le BUT du Havre:

Pour Nantes:

```
SELECT totalAdmis FROM Formations AS f, Admis AS a
```

```
WHERE f.codeFormation = a.formation
```

```
AND codeFormation = '5482'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --82 admis totaux
```

Pour le Havre:

```
SELECT totalAdmis FROM Formations AS f, Admis AS a
```

```
WHERE f.codeFormation = a.formation
```

```
AND codeFormation = '9655'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --71 admis totaux
```

Ensuite, nous allons en déduire le nombre total d'admis avec mention pour le BUT de Nantes et le BUT du Havre. Ainsi, à l'aide des requêtes suivantes, nous savons qu'il y a **79 admis** avec mention pour le BUT de Nantes, et **49** pour le BUT du Havre:

Pour Nantes:

```
SELECT mentionAB::int + mentionB::int + mentionTB::int + mentionF::int FROM Formations AS f,  
Admis AS a
```

```
WHERE f.codeFormation = a.formation
```

```
AND codeFormation = '5482'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --79 admis totaux avec mention
```

Pour le Havre:

```
SELECT mentionAB::int + mentionB::int + mentionTB::int + mentionF::int FROM Formations AS f,  
Admis AS a
```

```
WHERE f.codeFormation = a.formation
```

```
AND codeFormation = '9655'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --49 admis totaux avec mention
```

Voici les produits en croix permettant de définir le taux d'admis avec mention pour chaque BUT:

Pour Nantes:

79 = 96%

82 = 100%

Pour le Havre:

49 = 69%

71 = 100%

À l'aide des données trouvées précédemment et à l'aide d'un produit en croix, nous pouvons constater que pour le BUT Informatique de Nantes, **96% des candidates admis** ont eu le bac avec mention contre **69%** pour le BUT - Informatique du Havre. Ainsi, nous pouvons en conclure que Nantes accepte surtout les candidats possèdent un bac avec mentions contrairement au Havre qui en accepte un peu moins. Cela peut probablement être dû au hasard. Quoi qu'il en soit, les chiffres sont là et c'est le BUT - Informatique de Nantes qui bénéficie du plus de candidats ayant reçu une mention au bac.

7. Taux de BTN, BG et autres classés:

Pour information, pour cette question, les colonnes "generauxClasses", "technoClasses", "proClasses" et "autresClasses" (donc n39, n41, n43, n45) ont été ajouté car n'étant pas dans les tables ventilées d'origines.

Pour calculer le taux de BTN, BG et autres classés, nous devons ainsi calculer le nombre total de classés totaux, puis le nombre total de classés pour chaque BUT, puis le nombre total de classés pour chaque bac (BTN, BG, professionnel et autres). Pour finir, nous en déduirons les taux pour chaque formation.

Ainsi, à l'aide des requêtes suivantes, nous pouvons en déduire pour le BUT de Nantes qu'il y a **1 117 classés totaux, 726 classés généraux, 314 classés technologiques, 1 classé professionnel et 76 classés pour les autres** (les taux sont également indiquées pour chaque donnée après chaque requête):

Pour Nantes:

```
SELECT generauxClasses + technoClasses + proClasses + autresClasses FROM Formations
```

```
WHERE codeFormation = '5482'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --1117 classés totaux
```

```
SELECT generauxClasses FROM Formations
```

```
WHERE codeFormation = '5482'
```

```
AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --726 classés totaux généraux
```

Taux:

726 = 65%

1117 = 100%

```
SELECT technoClasses FROM Formations
```


WHERE codeFormation = '5482'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --314 classés totaux techno

Taux:

314 = 28%

1117 = 100%

SELECT proClasses FROM Formations

WHERE codeFormation = '5482'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --1 classé total professionnel

Taux:

1 = 0.1%

1117 = 100%

SELECT autresClasses FROM Formations

WHERE codeFormation = '5482'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --76 classés totaux autres

Taux:

76 = 7%

1117 = 100%

Ainsi, à l'aide des requêtes suivantes, nous pouvons en déduire pour le BUT du Havre qu'il y a **980 classés totaux, 722 classés généraux, 206 classés technologiques, 0 classé professionnel et 52 classés pour les autres** (les moyennes sont également indiquées pour chaque donnée après chaque requête):

Pour le Havre:

SELECT generauxClasses + technoClasses + proClasses + autresClasses FROM Formations

WHERE codeFormation = '9655'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --980 classés totaux

SELECT generauxClasses FROM Formations

WHERE codeFormation = '9655'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --722 classés totaux

Taux:

722 = 74%

980 = 100%

SELECT technoClasses FROM Formations

WHERE codeFormation = '9655'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --206 classés totaux

Taux:

206 = 21%

980 = 100%

SELECT proClasses FROM Formations

WHERE codeFormation = '9655'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --0 classé total

Taux:

0 = 0%

980 = 100%

SELECT autresClasses FROM Formations

WHERE codeFormation = '9655'

AND codeDep IN ('14','27','50','61','76','44','49','53','72','85'); --52 classés totaux

Taux:

52 = 5.3%

980 = 100%

À l'aide des données précédentes ainsi qu'un produit en croix, nous pouvons constater que le BUT-Informatique à plus de candidats technologiques, professionnels et autres classés (avec respectivement **28%**, **1%** et **7%** contre **21%**, **0%** et **5.3%** pour le Havre), mais à moins de candidats généraux que le Havre (avec **65%** contre **74%** pour le Havre).

8. Informations complémentaires:

À l'aide des requêtes suivantes, nous pouvons constater que nous avons **89 propositions acceptées** pour l'établissement de Nantes et **78** pour celui du Havre:

Pour Nantes:

```
SELECT propAccep FROM Admis  
  
WHERE formation = '5482';
```

Pour le Havre:

```
SELECT propAccep FROM Admis  
  
WHERE formation = '9655';
```

Cela va sans dire que ces propositions sont ainsi composées de **85 propositions acceptées** avant la fin de la procédure parcoursup dont **34 propositions acceptées** en début de procédure:

Pour Nantes:

```
SELECT propDebProc FROM Admis  
  
WHERE formation = '5482';
```

Pour le Havre:

```
SELECT propAvFinProc FROM Admis  
  
WHERE formation = '5482';
```

Pour le Havre, nous avons ainsi de **72 propositions acceptées** avant la fin de la procédure parcoursup dont **15 propositions acceptées** en début de procédure:

Pour Nantes:

```
SELECT propDebProc FROM Admis
```

```
WHERE formation = '9655';
```

Pour le Havre:

```
SELECT propAvFinProc FROM Admis
```

```
WHERE formation = '9655';
```

Concernant les propositions acceptées en début de procédure, après un calcul en croix, nous pouvons en déduire que cela représente **40%** (comparé aux propositions acceptées avant la fin de la procédure) pour le BUT de Nantes et **20.8%** pour le BUT du Havre.

Pour finir, nous pouvons constater à l'aide des requêtes suivantes que nous avons **89** candidats acceptés en phase principale de parcoursup pour le BUT de Nantes contre **72** pour le BUT du Havre:

Pour Nantes:

```
SELECT admisPhasePrinc FROM Admis
```

```
WHERE formation = '5482';
```

Pour le Havre:

```
SELECT admisPhasePrinc FROM Admis
```

```
WHERE formation = '9655';
```