

# Album photo ou sandwich chaud

Brichet Raphaël et De Wee Quentin

5 juillet 2022

Élèves de 6<sup>e</sup> secondaire au  
Collège Saint-Benoît de Maredsous

Avec l'aide de leurs enseignants  
Dr Miguël Dhyne et Hélène Gatellier

et des chercheurs

Jean-François de Kemmeter et Christian Mugisho Zagabe de l'université de Namur.

## **Résumé :**

*Pour la coupe du monde, panini à sorti un album de photos à remplir. Pour remplir l'album, il nous faut collectionner 100 cartes différentes qui sont vendues par paquet de 5 cartes dans les magasins. Nous voudrions savoir combien de paquets nous devons acheter pour collectionner les 100 cartes différentes. Pour répondre à cette question, nous avons conçu un programme informatique en python qui nous permet d'obtenir des résultats à partir d'un tirage aléatoire de cartes piochées 5 par 5. Nous avons aussi étudié l'influence de différents facteurs comme des raretés en modifiant le nombre de carte par rareté et la probabilité d'obtenir une certaine rareté.*

## 1 Présentation du problème

Voici notre question de départ :

J'ai acheté un album pour y coller des photos représentant mes sportifs préférés. S'il y a 100 photos à collectionner et qu'on les achète par pochettes de 5, quel nombre de pochettes dois-je acheter pour remplir un album ?

On a premièrement choisi que toutes les cartes avaient autant de chance de tomber (1% par carte car il y a 100 cartes). Pour chaque résultat obtenu, nous avons calculé la moyenne et nous avons regroupé de nos résultats afin d'obtenir une réponse globale.

## 2 Notre hypothèse

Au départ, nous avons réfléchi logiquement et sachant que chaque carte avait la même probabilité de chance d'être choisie, nous savions qu'il faudrait au minimum 20 paquets de 5 cartes pour obtenir les 100 cartes, mais nous pensions aussi qu'il n'y aurait pas de maximum de paquets étant donné qu'il est possible (certe très peu probable mais possible) que nous n'obtenions jamais une ou plusieurs cartes. (Avec notre programme il est impossible de ne jamais obtenir une carte puisqu'un ordinateur n'est jamais totalement aléatoire. De ce fait, nous obtiendrons toujours une réponse autre qu'une infinité).

## 3 Notre démarche pour répondre à la question de départ

Nous avons d'abord conçu un programme informatique écrit en python afin d'obtenir beaucoup plus de résultats et bien plus rapidement qu'à la main pour pouvoir calculer une moyenne et regrouper nos résultats. Nous avons ensuite transféré ces résultats vers Excel ce qui nous a permis d'en faire des graphiques et de déterminer la moyenne très facilement.

### 3.1 Notre programme en python

Notre programme (voir figure 1) contient une double boucle. On commence par déterminer et nommer 100 cartes différentes pour représenter

notre question initiale en leur appliquant une valeur choisie (nous avons choisis la valeur 0 comme valeur initiale pour chaque carte) et on fait de même pour le nombre de paquets générés. Ensuite, le programme regarde si toutes les cartes ont été choisies au moins une fois. Si ce n'est pas le cas, le programme crée un nouveau paquet en augmentant de 1 le nombre de paquets puis il choisit 5 fois 1 carte (il choisit bien 5 fois 1 carte et non pas 5 cartes ce qui permet la répétition d'une carte dans un seul paquet). Chaque fois qu'une carte est choisie, sa valeur est augmentée de 1, lorsqu'une carte ne vaut plus 0, on se rapproche de la condition pour terminer la boucle et lorsqu'une carte vaut plus que 1, c'est un double donc cela augmente le nombre de paquets qui seront générés pour rencontrer la condition de la boucle. Une fois la condition rencontrée (c'est-à-dire que chaque carte à une valeur de 1 ou plus), le programme indique le nombre de paquets qui ont été nécessaires pour obtenir la collection complète puis il recommence le tout 3000 fois (3000 étant un nombre choisi arbitrairement qui nous permet d'avoir assez de données pour calculer une moyenne assez précise).

### 3.2 Les résultats

En implantant les résultats obtenus par le programme informatique dans un tableau Excel, nous avons pu rassembler nos résultats pour en faire des graphiques. (voir figure 2) Nous en avons déterminé le nombre moyen de paquets nécessaires pour obtenir la collection complète, lequel s'élève à 46 paquets. Elle était initialement de 45.45 mais étant donné que nous ne pouvons pas avoir de demi paquet nous devons l'arrondir au nombre de paquet supérieur. L'écart-type est égal à 12.22.

Grâce à la moyenne et l'écart-type, nous pouvons déterminer qu'il y a environ 70% des résultats qui se trouvent entre 34 et 58 (34 et 58 compris) et qu'il y a environ 95% des résultats qui se trouvent entre 22 et 70 (22 et 70 compris).

## 4 Notre dépassement

Après avoir obtenu des résultats pour répondre à la question initiale, nous sommes allés plus loin et nous avons décidé de rajouter des raretés.

Nous avons considéré différents paramètres avec les raretés en faisant varier le nombre de cartes par rareté et la probabilité qu'une rareté soit choisie. Nous avons déterminé 4 raretés différentes par ordre croissant de rareté : Commune, Rare, Epique et Légendaire.

#### **4.1 Nos hypothèses sur la question de dépassement**

En rajoutant des raretés, nous nous attendions à ce que :

- 1) Le nombre moyen de paquets nécessaires pour avoir toutes les cartes devrait augmenter.
- 2) La probabilité qu'une rareté tombe modifie le nombre moyen de paquets nécessaires et c'est la rareté qui a la plus faible probabilité de tomber qui influence le plus le nombre moyen de paquets nécessaires pour avoir toutes les cartes si le nombre de carte par rareté est fixé [sauf si le nombre de cartes présentes dans la rareté ayant la plus faible probabilité d'être choisie est très faible (une seule carte par exemple)] .
- 3) Le nombre de cartes contenues dans chaque rareté influence le nombre moyen de paquets nécessaires pour avoir toutes les cartes et plus particulièrement, si la probabilité de chaque type de rareté est fixe, c'est le nombre de carte de la plus faible probabilité (dont la rareté est grande) qui influencera le nombre moyen de paquets nécessaires pour compléter la collection.

#### **4.2 Notre démarche pour répondre à la question de dépassement**

Nous avons modifié notre programme initial en classant les 100 cartes par rareté et en attribuant à chaque rareté une probabilité d'être sélectionnée. Nous avons par la suite continué à modifier le programme pour modifier le pourcentage de chance de tomber sur une certaine rareté ou modifier le nombre de carte par rareté. Lors de ces modifications, nous avons gardé le même nombre de cartes que précédemment (100 cartes) afin de ne pas trop s'éloigner de la question initiale.

### 4.3 Notre programme en python

Comme pour le programme initial (figure 1) le programme où l'on a implémenté des raretés (voir figure 3) est fait d'une double boucle. On commence par déterminer et nommer 100 cartes différentes pour représenter notre question initiale en leur appliquant une valeur choisie (nous avons choisi la valeur 0 comme valeur initiale) et on fait de même pour le nombre de paquets générés. Ensuite, le programme regarde si toutes les cartes valent un ou plus. En d'autres termes, il regarde si toutes les cartes ont été choisies au moins une fois. Si ce n'est pas le cas, le programme crée un nouveau paquet en augmentant de 1 le nombre de paquets puis il choisit 5 fois 1 carte (il choisit bien 5 fois 1 carte et non pas 5 cartes ce qui permet la répétition d'une carte dans un seul paquet).

Pour chaque carte le programme doit choisir une rareté. Pour rajouter les raretés, nous avons dû attribuer une valeur comprise entre 0 et 1 à chacune d'entre elles sachant que les valeurs ne pouvaient pas être les mêmes. Les intervalles sont construites en additionnant le nombre décimal précédent plus le pourcentage de chance que l'on veut attribuer à la rareté. Par exemple, on veut attribuer à la rareté la plus difficile à obtenir 5% de chance, on va donc lui attribuer le nombre 0.05 (ce qui signifie que le nombre généré doit se trouver entre 0 et 0.05 pour que le programme pioche une carte dans cette rareté) mais si après on veut donner un pourcentage de chance de 30% à la deuxième rareté, il faudra faire le calcul  $0.05+0.30$  ce qui vaut 0.35 (le nombre généré doit être égal ou plus petit que 0.35 pour être pioché dans la deuxième rareté mais s'il est égal ou plus petit que 0.05, alors le programme piochera une carte de la première rareté.), 0.35 étant la valeur que l'on attribuera à la deuxième rareté et ainsi de suite jusqu'à la dernière rareté qui aura le reste des nombres décimaux comme intervalle. Si une carte dans le paquet était déjà tombée, alors il s'agit d'un doublon qui va forcer le programme à créer plus de paquets pour compléter la collection. Si par contre elle n'était pas encore tombée, sa valeur n'est plus égale à 0 et cela agit directement sur la condition de la boucle.

Une fois la condition rencontrée (c'est à dire que toutes les cartes ont une valeur égale ou supérieure à 1), le programme indique le nombre de paquets qui ont été nécessaires puis il recommence le tout 3000 fois (3000 étant un nombre choisit arbitrairement et qui nous permet d'avoir assez de données

pour calculer une moyenne relativement correcte et précise).

A chaque fois que nous avons terminé avec certaines données, nous changeons les paramètres comme le nombre de cartes par raretés et le pourcentage de chance que chaque rareté tombe.

#### 4.4 Les résultats

Premièrement, nos hypothèses (cf. sous section 4.1) sont vérifiées. La moyenne a bel et bien augmentée dès que nous avons rajouté les raretés et nous avons remarqués après quelques tests que la rareté possédant le moins de chances de tomber est la rareté qui influence le plus les résultats. Nous avons pu remarquer aussi que dans tous les cas de figure le graphique se comporterait de la même manière (ils ont plus ou moins la même forme) bien que nous n'ayons pas pu déterminer la cause de ce phénomène. Nous avons aussi pu remarquer que les deux paramètres (nombre de cartes par rareté et la probabilité de choisir chaque rareté) influencent le graphique à leur manière mais que les deux paramètres sont liés car ils s'influencent l'un et l'autre. Par exemple si l'on n'a que 0.1% de chance de devoir choisir une carte légendaire et que l'on a 1 carte ou bien 10 cartes, la moyenne et le graphique qui découleront des résultats auront des valeurs bien plus différentes que si on a 10% de chance de choisir dans les légendaires avec 1 carte ou 10 cartes.

## 5 Nos perspectives

Nous avons commencé à modifier le nombre de cartes par rareté et la probabilité de tomber sur chaque rareté. Par ailleurs, nous aurions aussi voulu trouver le pourcentage de chance de trouver toutes les cartes en 20, 25, 30, ... paquets en fonction de chaque paramètre (nombre de carte par rareté et pourcentage de chance de tomber sur chaque rareté) et essayer de trouver une formule générale pour trouver la probabilité de tomber sur toutes les cartes en un certain nombre de paquets en fonction de chaque paramètre.

## 6 Conclusion

Grâce au programme en python, nous avons pu obtenir un grand nombre de résultats rapidement. En utilisant Excel, nous en avons fait des graphiques que nous avons comparés et nous avons calculés les moyennes ce qui nous a permis de répondre à la question. Avec les raretés, nous avons pu modifier plusieurs paramètres et trouver des liens dans les différents paramètres que nous avons modifiés. Notre programme combiné à Excel nous permet donc de répondre à la question initiale.

Pour répondre à la question initiale, nous avons fait un programme en python qui nous a permis de rapidement obtenir de vraies réponses et en très grand nombre. Avec ces données nous avons pu faire une moyenne qui nous permet de donner une réponse globale du nombre de paquets nécessaires à acheter et nous avons aussi représenté nos résultats sous forme de graphiques afin que cela soit plus lisible. Une fois que nous avons répondu à la question de recherche, nous avons décidés de nous pencher sur l'ajout de raretés, ce qui devait modifier nos résultats, et ce fut le cas, nos résultats étaient différents, les nombres augmentaient et diminuaient en fonction des paramètres que nous déterminions, mais les moyennes qui en découlaient restaient toujours supérieures à la moyenne obtenue pour la question de recherche initiale.

### 6.1 Remerciements

Nous souhaitons remercier Dr Miguël Dhyne, Hélène Gatellier, Jean-François de Kemmeter et Christian Mugisho Zagabe pour nous avoir suivi dans notre recherche tout au long de l'année scolaire.

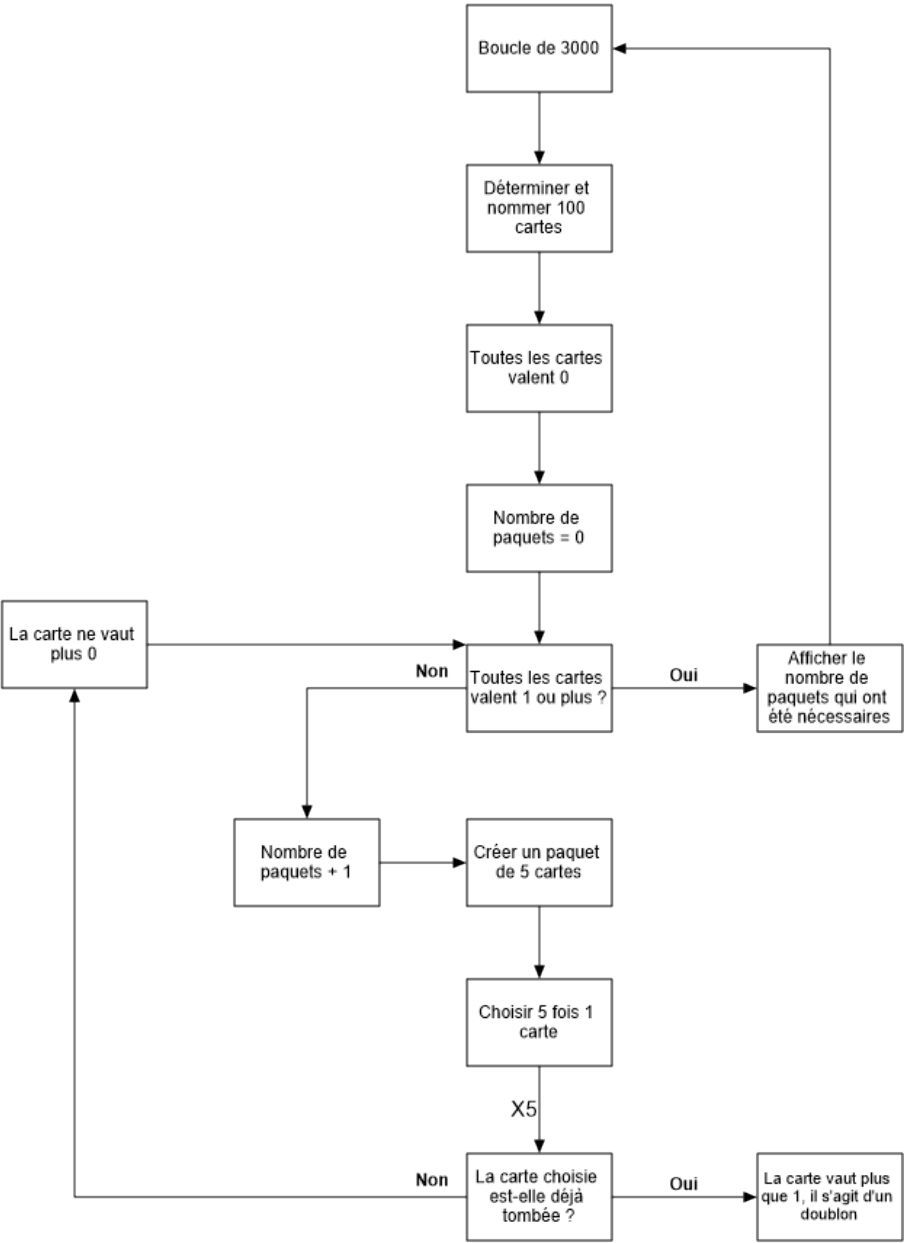


FIGURE 1 – Cette figure représente notre programme pour notre question initiale.



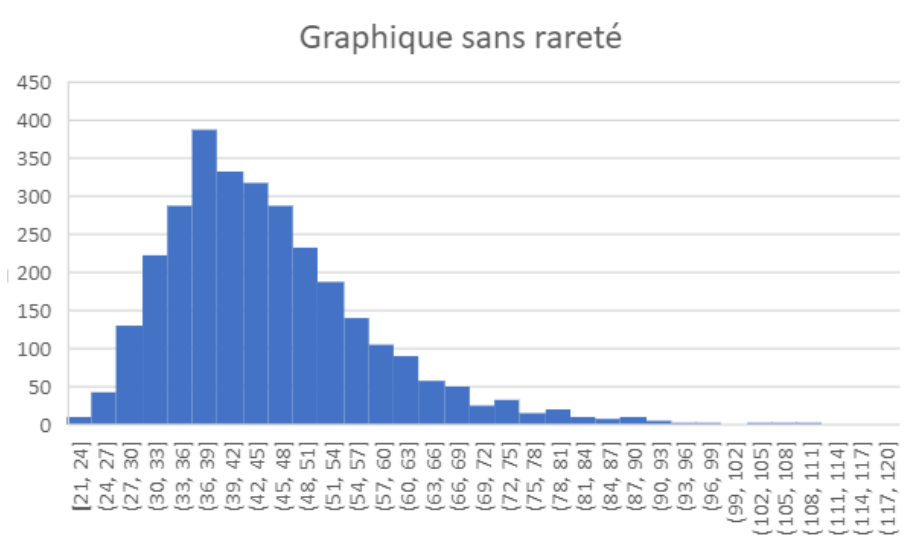
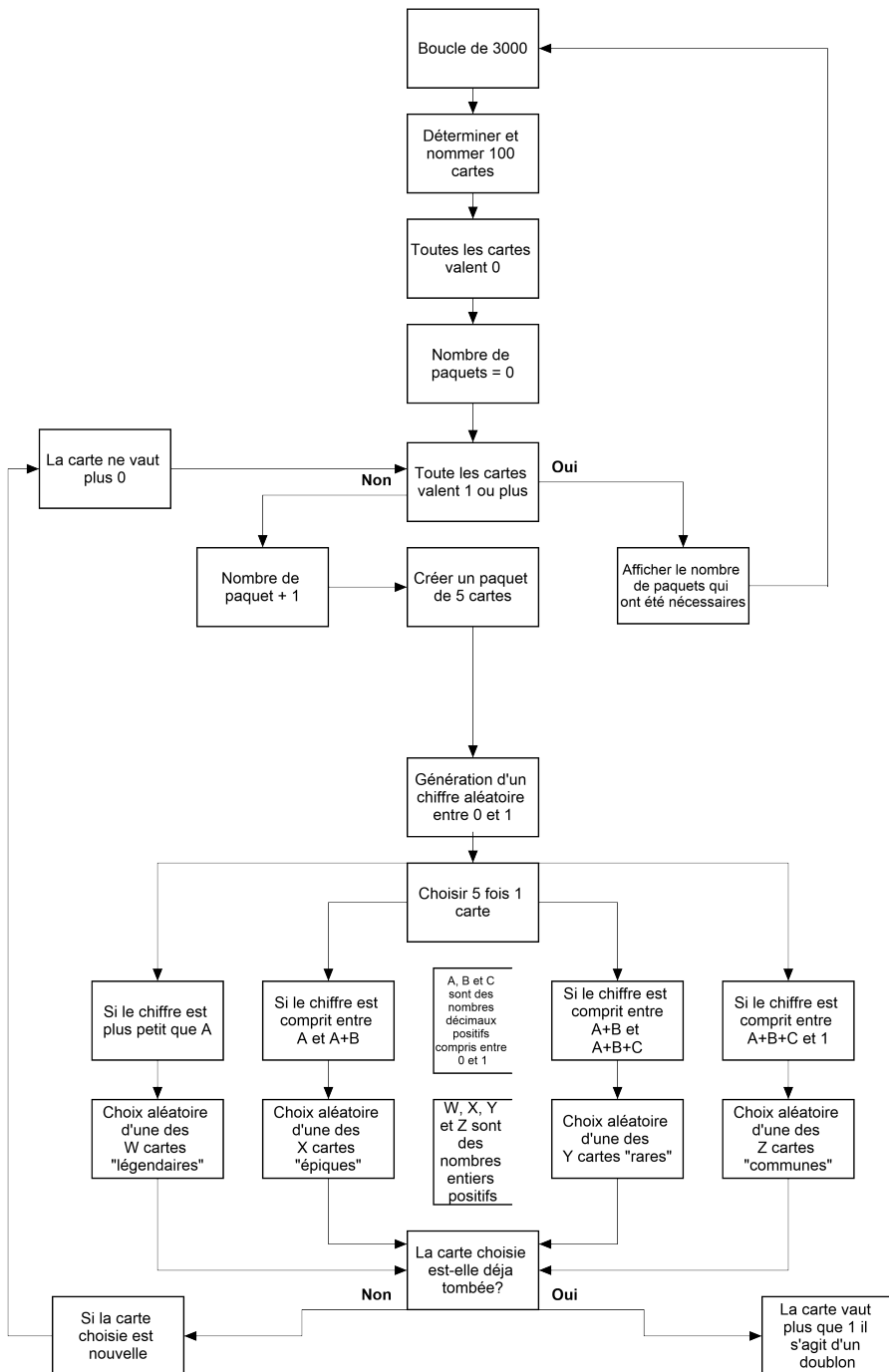


FIGURE 2 – Cette figure représente le graphique que nous avons obtenu pour la question initiale. L'axe des abscisses représente le nombre de paquets qu'il faut acheter pour obtenir une collection complète. Pour en faciliter la lecture, nous avons regroupés les paquets en classes de 3 unités. L'axe des ordonnées représente le nombre d'occurrences pour chaque classe de paquets (la somme des occurrences vaut donc 3000).



10

FIGURE 3 – Cette figure représente notre programme pour notre question de dépassement

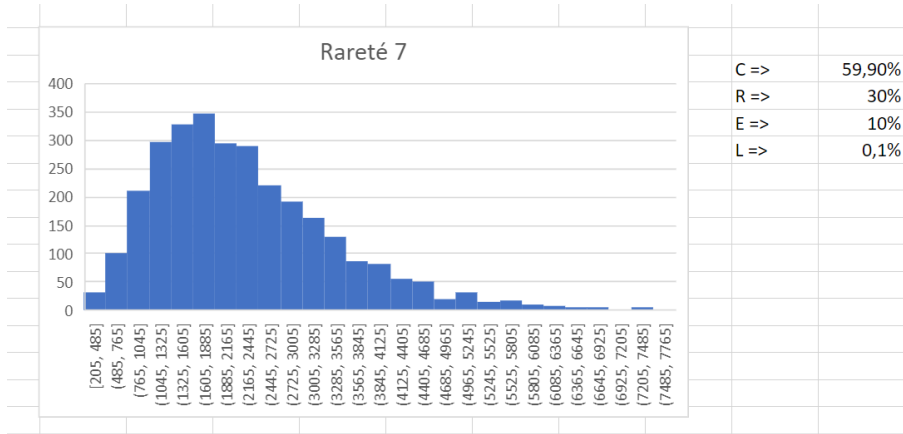


FIGURE 4 – Cette figure représente le graphique que nous avons obtenu pour la question initiale. L’axe des abscisses représente le nombre de paquets qu’il faut acheter pour obtenir une collection complète. Pour en faciliter la lecture, nous avons regroupés les paquets en intervalles. L’axe des ordonnées représente le nombre d’occurrences pour chaque intervalle de paquets (la somme des occurrences vaut donc 3000).

Ce graphique à été fait avec des raretés et on peut remarquer que les intervalles de l’axe des abscisses ont des valeurs plus grandes que celles du graphique initial. Les données qui sont écrites à côté sont tout simplement le pourcentage de chance d’obtenir chaque rareté.

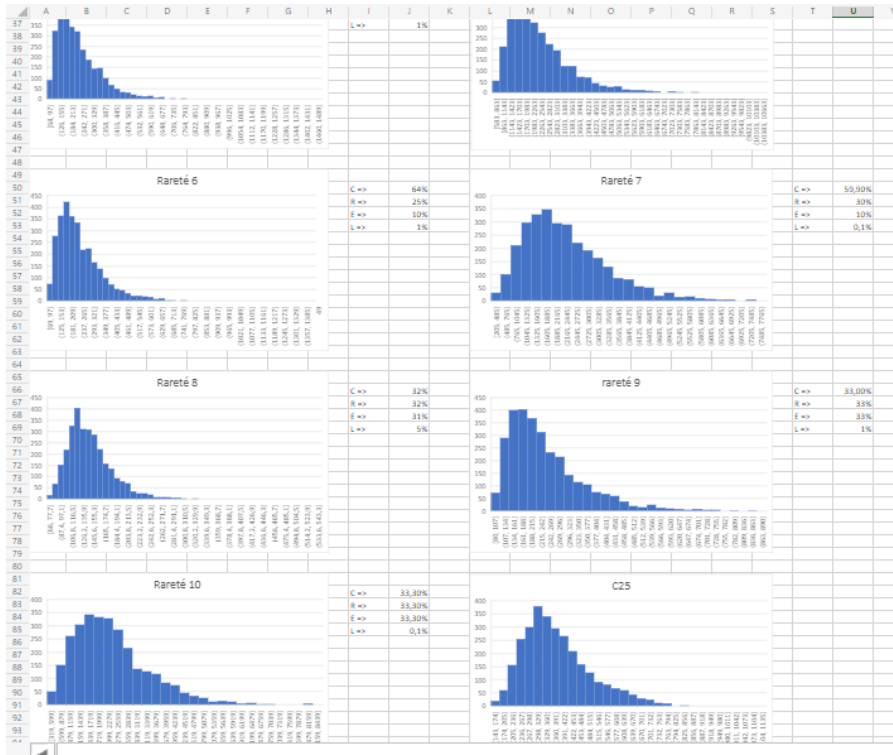


FIGURE 5 – Cette figure représente quelques résultats que nous avons obtenus en jouant avec les raretés. L'important ici est de montrer que même avec des modifications, la courbe des graphiques reste globalement la même.