

# mini\_rapport\_test

Chloé Blanpain

2025-02-04

## Introduction

Le diamant est un minéral d'une grande pureté, constitué presque exclusivement d'atomes de carbone. Sa formule chimique, notée C, révèle une concentration en carbone variant entre 99,95 % et 99,99 % (Valuae, 2015), faisant de cette gemme l'une des plus pures connues sur Terre.

Cependant, la présence d'éléments traces tels que l'azote, le bore ou l'hydrogène peut influencer ses caractéristiques, notamment sa couleur (Valuae, 2015). Ces variations chimiques confèrent à chaque diamant une identité unique.

## Le prix du diamant

un diamant de 1 carat pourrait coûter entre 2 000 et 25 000 euros. Cela signifie mathématiquement que le prix d'un diamant de milieu de gamme devrait coûter entre 11 000 et 12 000 euros. Cela semble simple mais pas si évident dans la mesure où l'on ne doit pas réfléchir dans ce sens. Un diamant de 1 carat de bonne qualité et de bonne valeur devrait vous coûter entre 4 500 et 7 000 euros. (Celinni, 2024)

## L'Objectif

Notre objectif est de voir comment se comporte le prix des diamants en fonction de ses caractéristiques (couleur, poids, ...) .

## Les données

Les données proviennent de R. NOus les importons grâce à la fonction suivante :

```
data(diamonds)
```

## visualisation du jeu de données

```
head(diamonds) #premier aperçu
```

Il s'agit d'un jeu de données qui repertorie **r** diamants selon ces catégories : carat, cut, color, clarity, depth, table, price, x, y, z. Nous allons surtout nous intéresser au poids, à la couleur et à la qualité de la coupe de ces diamants.

## visualisation statistiques

```
summary(diamonds) # Résumé statistique des variables
```

```
##      carat      cut      color      clarity      depth
## Min.   :0.2000 Fair      : 1610 D: 6775 SI1      :13065 Min.   :43.00
## 1st Qu.:0.4000 Good      : 4906 E: 9797 VS2      :12258 1st Qu.:61.00
## Median :0.7000 Very Good:12082 F: 9542 SI2      : 9194 Median :61.80
## Mean   :0.7979 Premium  :13791 G:11292 VS1      : 8171 Mean   :61.75
## 3rd Qu.:1.0400 Ideal    :21551 H: 8304 VVS2     : 5066 3rd Qu.:62.50
## Max.   :5.0100          I: 5422 VVS1     : 3655 Max.   :79.00
##          J: 2808 (Other): 2531
##      table      price      x      y
## Min.   :43.00 Min.   : 326 Min.   : 0.000 Min.   : 0.000
## 1st Qu.:56.00 1st Qu.: 950 1st Qu.: 4.710 1st Qu.: 4.720
## Median :57.00 Median : 2401 Median : 5.700 Median : 5.710
## Mean   :57.46 Mean   : 3933 Mean   : 5.731 Mean   : 5.735
## 3rd Qu.:59.00 3rd Qu.: 5324 3rd Qu.: 6.540 3rd Qu.: 6.540
## Max.   :95.00 Max.   :18823 Max.   :10.740 Max.   :58.900
##
##      z
## Min.   : 0.000
## 1st Qu.: 2.910
## Median : 3.530
## Mean   : 3.539
## 3rd Qu.: 4.040
## Max.   :31.800
##
```

Nous remarquons que les diamants listés sont majoritairement des diamants avec un poids faible (moyenne autour de 0,8 carat). Le diamant le plus léger est à 0,2 carat et le plus lourd à 5 carat. Nous avons des grandes disparités entre les diamants. Pareil pour les prix. Nous avons des prix variants de 326 à 18 823 euros (grand écart entre les prix) avec une moyenne autour de 4000 euros.

## Analyse et Observation

### Distribution du prix des diamants

Nous allons observer la répartition des prix de ces diamants. Cette distribution suit une loi normale. La majorité de nos diamants se retrouvent dans des prix bas. A l'inverse, peu de diamants ont un prix élevé.

```
ggplot(diamonds, aes(x = price)) +
  geom_histogram(binwidth = 500, fill = "steelblue", color = "lightpink", alpha = 0.5) +
  labs(title = "Distribution du prix des diamants", x = "prix", y = "nombre de diamants") +
  theme_void()
```

### Relation entre le prix et le poids

## Distribution du prix des diamants

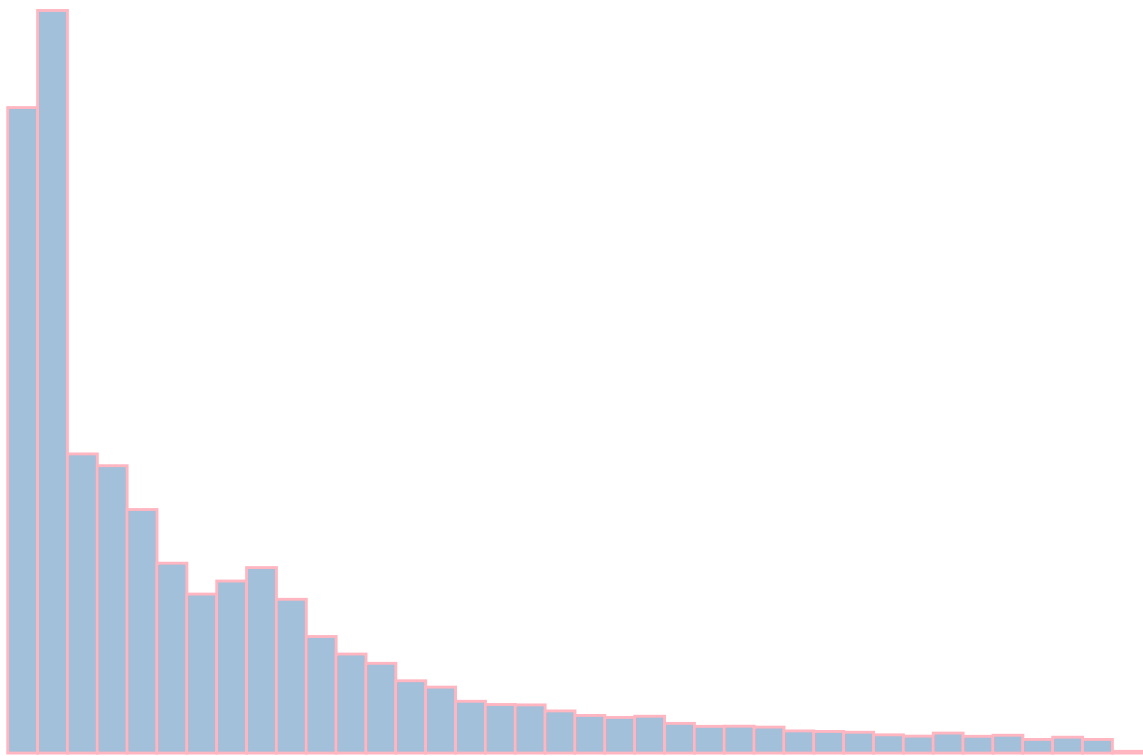


Figure 1: Fig 1 : Distribution du prix des diamants

```
ggplot(data=diamonds, mapping= aes(x=price, y= carat, color = color))+
  geom_point(alpha=0.5, size=5)+
  scale_color_brewer(palette = "Set1")+
  geom_smooth( method=lm , color = "yellow") +
  theme_dark()+
  labs(title = "Prix des diamants en fonction du poids (carat)",
        x="Prix (en Euro)", y="Poid (en carat)")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

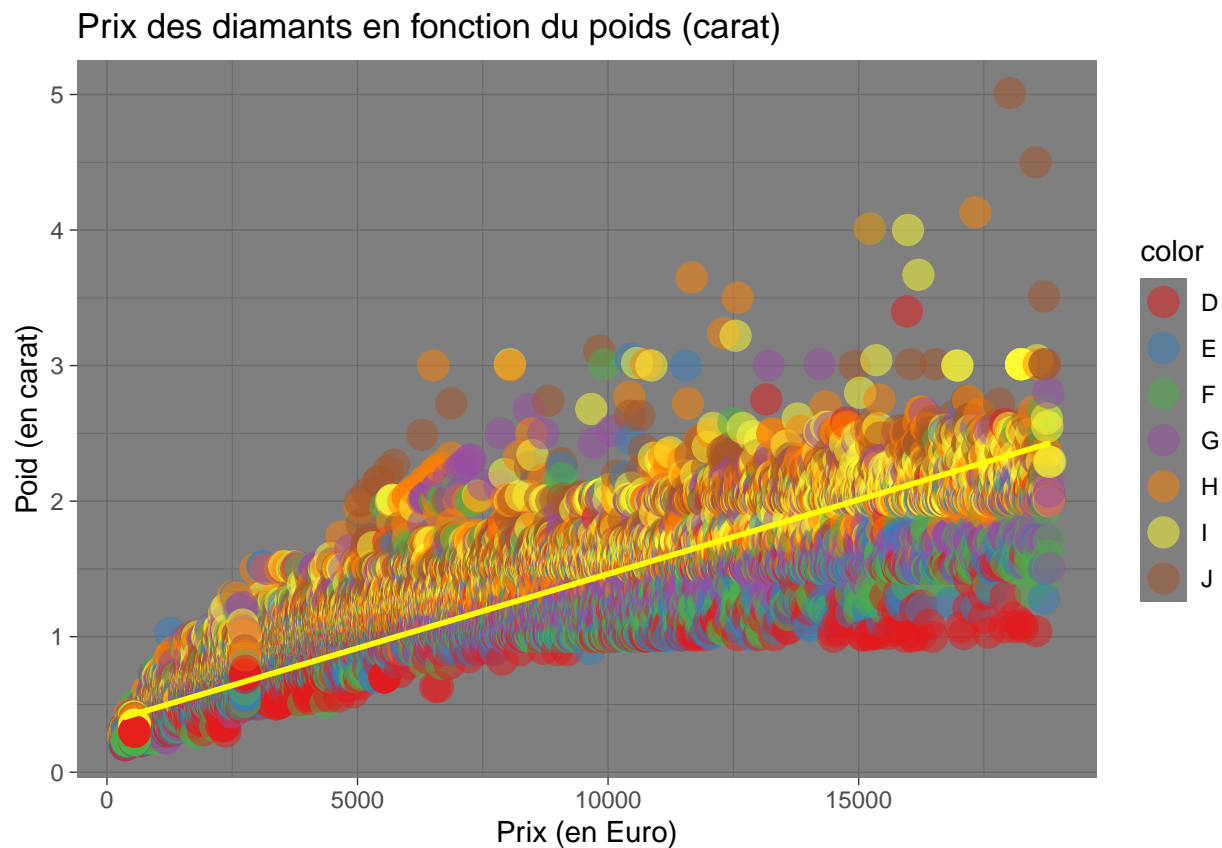


Figure 2: Fig 2 : Répartition du prix en fonction du poids des diamants

Nous remarquons une corrélation entre le prix et le poids des diamants. En effet, il y a une tendance : plus un diamant est lourd plus son prix tend à être élevé.

## Relation entre le prix et la couleur

```
ggplot(diamonds, aes(x = price, fill = color)) +
  geom_histogram(binwidth = 250, alpha = 0.7, position = "identity", color = "pink") +
  facet_wrap(~ color) +
  labs(title = "Prix des diamants en fonction de la couleur", x = "prix", y = "nombre de diamants") +
  theme_minimal()
```

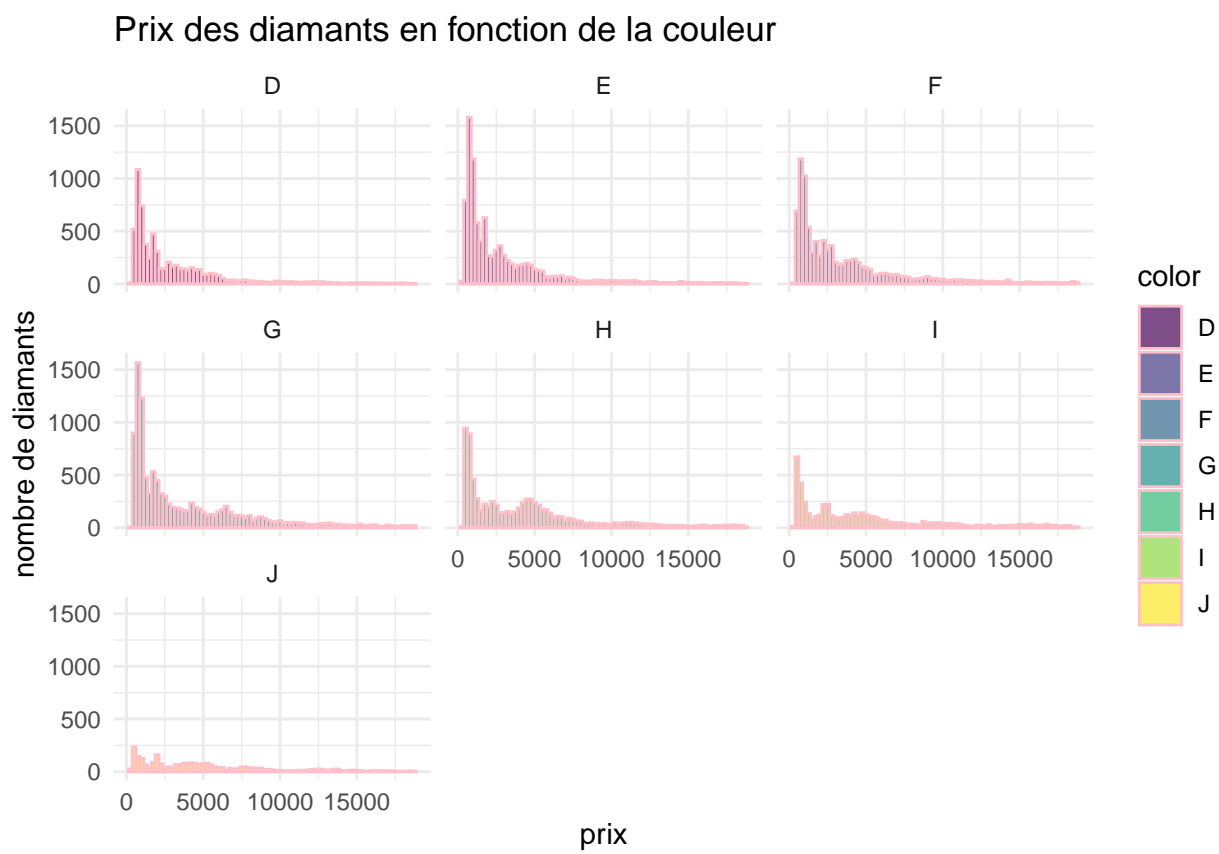


Figure 3: Fig 3 : Répartition du prix en fonction de la couleur des diamants

Ce graphique nous montre que la couleur n'est pas un facteur déterminant mais qu'il a une influence sur le prix du diamant.

## Conclusion

Comme nous l'avons vu avec nos références bibliographiques, la pureté du diamant dépend de son poids. Cependant, c'est cette pureté qui influence le prix : plus un diamant est pur, plus il est lourd, plus il est cher.

## Références

### Bibliographie

Liste des publications :

CELINNI, 2024. Prix du diamant (2024) : Comment évaluer sa valeur ? [en ligne]. 2024. [Consulté le 4 février 2025]. Disponible à l'adresse : <https://www.celinni.com/fr/content/category/10-prix-du-diamant>

VALUAE, 2015. D'où viennent les diamants - la géologie du diamant. [en ligne]. 2015. [Consulté le 4 février 2025]. Disponible à l'adresse : <https://valuae.com/guide/quest-ce-quun-diamant/douviennent-les-diamants-la-geologie-du-diamant/>