**Exemple de présentation fictive**



**Présentation de Léa avec un effet du hasard dans son parcours :**

Bonjour à tous,

Je m'appelle **Léa Dupont**, j'ai **28 ans**, et je travaille comme **consultante en stratégie digitale** depuis 5 ans. Après avoir obtenu un **master en marketing digital** à l'Université de Paris, j'ai commencé ma carrière en tant que chef de projet dans une agence de communication, où j'ai eu l'opportunité de gérer des campagnes pour des marques internationales.

Mon parcours a été influencé par plusieurs facteurs externes. Ayant grandi dans une famille d’entrepreneurs, j'ai toujours été fascinée par la manière dont les entreprises peuvent s'adapter à un environnement changeant. De plus, l'évolution rapide des technologies numériques et leur impact sur notre manière de consommer m'ont motivée à me spécialiser dans ce domaine.

Un tournant inattendu dans ma carrière s'est produit lorsque j'ai participé par hasard à une conférence sur le big data organisée par une connaissance. Ce sujet n’était pas directement lié à mon poste de l’époque, mais il m’a ouvert les yeux sur le rôle crucial des données dans la prise de décisions stratégiques. Cette rencontre m’a poussée à approfondir mes compétences en analyse de données et m’a conduite à décrocher mon poste actuel.

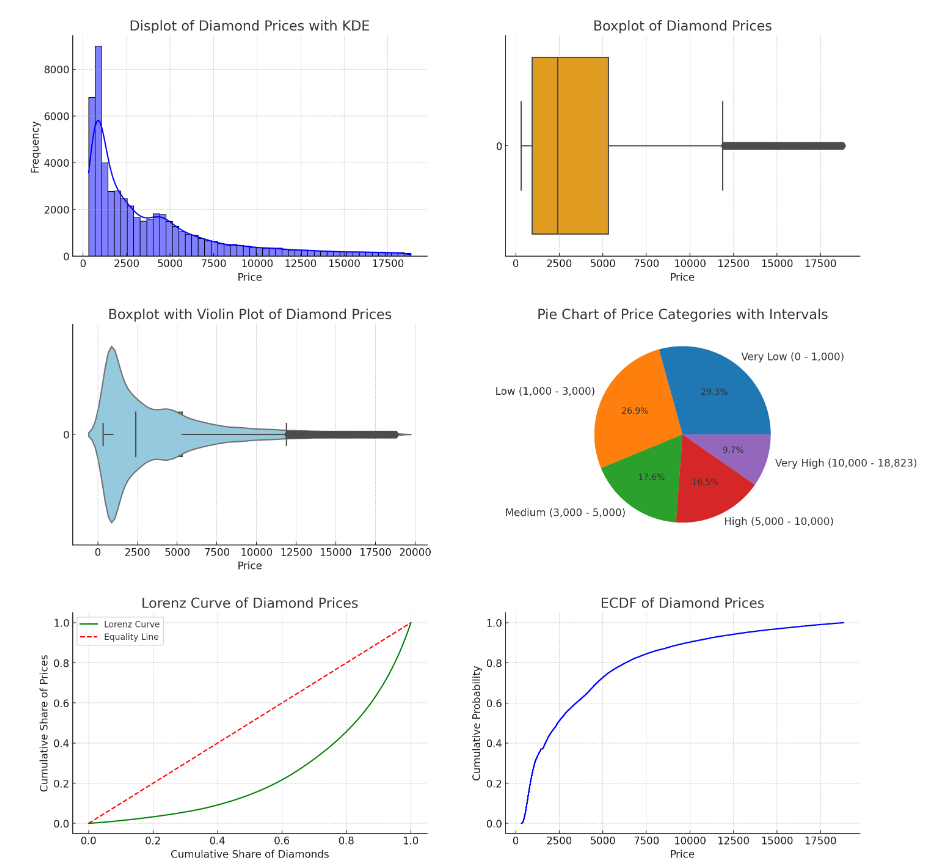
En termes de culture, j'ai été profondément marquée par mes voyages à l’étranger. Vivre quelques mois en Asie pendant mes études a renforcé mon intérêt pour des approches plus collaboratives et inclusives dans la gestion de projets. J’ai également beaucoup appris en observant les différences culturelles dans la manière dont les marques communiquent à travers le monde, ce qui m’a donné une perspective globale précieuse.

Aujourd'hui, je me spécialise dans l'accompagnement des entreprises souhaitant renforcer leur présence en ligne et optimiser leurs performances grâce aux outils digitaux. Mon expertise porte notamment sur l'analyse des données clients, le SEO, et la création de stratégies omnicanales. Mon objectif est toujours de trouver des solutions adaptées au contexte économique et culturel de mes clients.

Dans **10 ans**, je me vois à la tête de ma propre agence de conseil en stratégie digitale, travaillant à l’international. Mon ambition est de devenir une actrice majeure dans la transition numérique responsable. Mon rôle dans la société serait d'accompagner des entreprises vers un modèle plus éthique et durable, tout en renforçant l’inclusion dans le monde du numérique.

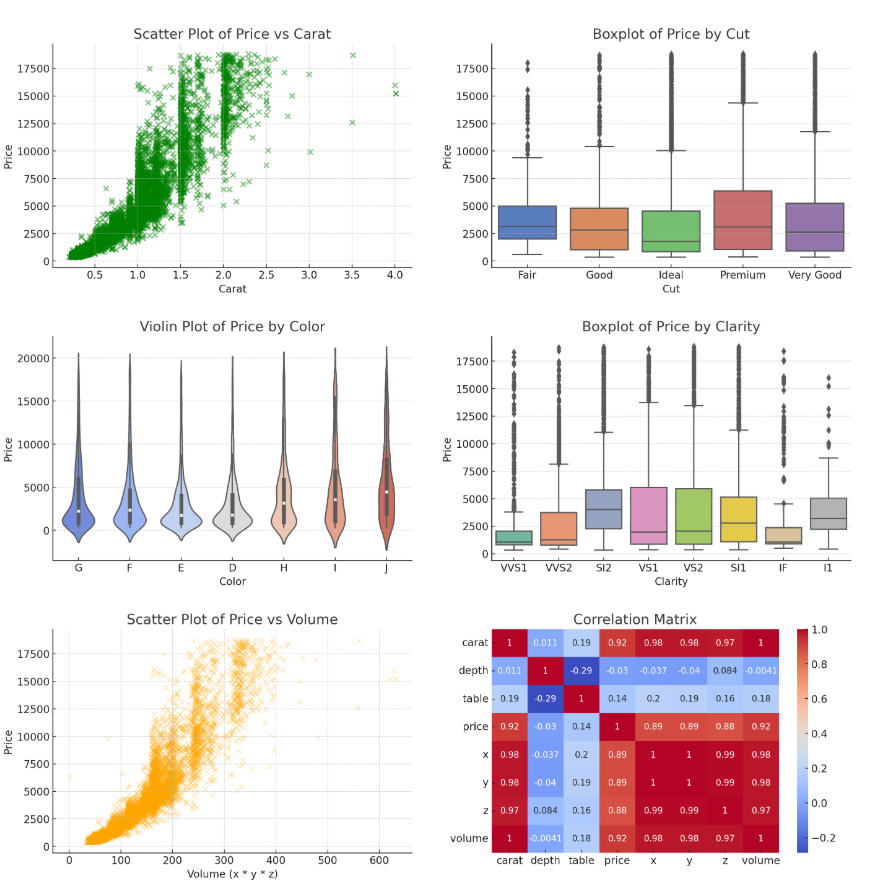
En dehors du travail, je suis passionnée par la photographie et le voyage. Ces expériences nourrissent ma créativité et ma capacité à penser "en dehors des sentiers battus".

Je suis ravie d’être ici aujourd’hui pour échanger avec vous et en apprendre davantage sur vos parcours et vos projets. Merci de m’avoir écoutée !



Dans cette étude sur les prix des diamants, les données révèlent plusieurs éléments essentiels pour le secteur diamantaire.

1. **Description générale des prix** :
   * Nous disposons de **53940 prix unitaires** de diamants, qui varient entre **326 euros** pour les plus abordables et **18823 euros** pour les plus coûteux. La **moyenne des prix** se situe autour de **3933 euros**, ce qui nous donne une première idée des valeurs typiques dans cette gamme de produits. Les prix les plus fréquents sont **605 euros**.
2. **Analyse des quartiles et de la répartition des prix** :
   * Les quartiles de distribution nous montrent que :
   * 25% des prix des diamants sont inférieurs à **950 euros**, indiquant une catégorie de diamants plus accessible pour les consommateurs.
   * 50% des diamants coûtent moins de **2401 euros** (médiane), et les prix peuvent monter jusqu'à **5324.25 euros** pour 75% des produits.
   * **50% des diamants** étudiés se situent entre **950** et **5324** euros, ce qui représente une gamme assez large et représentative de l'offre.
   * **10 à 90% des diamants** étudiés se situent respectivementre entre **[646.0, 837.0, 1087.0, 1698.0, 2401.0, 3465.0, 4662.0, 6301.0, 9821.0]** euros
3. **Dispersion et variabilité des prix** :
   * La **dispersion par rapport à la moyenne** est élevée (101%), indiquant une large gamme de prix autour de la moyenne de **3933 euros**. Par ailleurs, le **coefficient de variation par rapport à la médiane** (75%) met en évidence une grande variabilité entre les diamants de prix moyen et ceux de très haute valeur. Ces diamants de valeur exceptionnelle, atteignant jusqu’à **18823 euros**, influencent fortement les statistiques et justifient une offre de niche dans la diamantaire.
4. **Identification des valeurs aberrantes** :
   * Plusieurs diamants, qualifiés de valeurs atypiques ou aberrantes, se trouvent dans les prix supérieurs à **11886 euros**. Ces prix incluent des spécimens exclusifs, probablement rares, avec une qualité de coupe, de couleur, et de clarté exceptionnelle.
5. **Structure de la distribution** :
   * Les prix ne semblent pas être générés de manière aléatoire. La courbe de densité des prix montre un pic autour des prix moyens, ce qui peut refléter des stratégies commerciales visant à rendre certains diamants plus accessibles. Cette structure pourrait également correspondre aux choix d'approvisionnement et de ciblage du marché.
6. **Analyse de la distribution (Skewness et Kurtosis)** :
   * **Skewness (Asymétrie)** : La skewness est de **2**, indiquant une distribution légèrement asymétrique à droite. Cela signifie que les valeurs élevées sont plus fréquentes.
   * **Kurtosis (Aplatissement)** : La kurtosis est de **2**, ce qui montre une distribution avec plus de valeurs concentrées autour de la moyenne (leptokurtique).
7. **Analyse de la distribution (Loi de probabilité)** :
   * Ces résultats suggèrent que les prix **ne semblent pas aléatoires** car la statique de Jarques Bera **34205.09 depasse largenement 5.99 (seuil)**. Une distribution asymétrique et/ou avec un aplatissement élevé indique une structuration des prix influencée par des facteurs tels que la qualité, le poids en carats et la clarté des diamants, qui déterminent la valeur perçue de chaque pièce.
8. **Analyse de la Contribution des Diamants au Chiffre d'Affaires**
   * Les **1%** des diamants les plus chers génèrent **5%** du chiffre d'affaires total
   * Les **10%** des diamants les plus chers génèrent **34%** du chiffre d'affaires total, montrant une forte concentration de la valeur dans le haut de gamme.
   * Les **25%** des diamants les plus chers génèrent **62%** du chiffre d'affaires, illustrant l'importance des gammes premium.
   * Les **50%** des diamants génèrent **86%** du chiffre d'affaires, indiquant une répartition plus équilibrée à ce niveau.
   * Les **75%** des diamants contribuent à **96%** du chiffre d'affaires total, soulignant l'importance des diamants abordables pour le volume global des ventes.
   * Ces données confirment que les diamants haut de gamme dominent le marché en termes de chiffre d'affaires, tandis que les segments intermédiaires et abordables soutiennent les ventes globales en volume.



**Analyse des Corrélations avec les Prix des Diamants :**

**Analyse des Corrélations entre le Prix et les Attributs Quantitatives :**

Les calculs de corrélation montrent les relations entre les différentes caractéristiques des diamants et leur prix, permettant de dégager les facteurs ayant le plus d'impact sur la valorisation.

1. **Corrélations positives significatives** : Les caractéristiques avec une corrélation positive supérieure à 0.0 avec le prix sont : ['carat', 'x', 'y', 'z', 'color\_encoded', 'table', 'cut\_encoded']. Ces facteurs augmentent le prix du diamant lorsqu'ils sont élevés, ce qui est attendu pour des caractéristiques comme le poids en carats, qui influence directement la perception de la valeur.
2. **Corrélations négatives significatives** : Les caractéristiques avec une corrélation négative supérieure à -0.0 avec le prix sont : ['depth', 'clarity\_encoded', 'Unnamed: 0']. Cela indique que certains attributs peuvent réduire la valeur lorsque leurs niveaux sont plus élevés, comme certaines couleurs ou clartés moins prisées.

**Conclusion professionnelle :** Ces corrélations confirment que le **poids (carat)** et d'autres caractéristiques de qualité comme la **clarté** et la **coupe** ont un impact direct sur le prix, ce qui justifie leur importance dans l'industrie de la diamantaire. Les bijoutiers et négociants peuvent ainsi ajuster les prix en fonction de ces caractéristiques pour optimiser la valorisation de chaque diamant.

**Analyse des Corrélations entre le Prix et les Attributs Catégoriels :**

Les variables catégorielles ayant été encodées en valeurs numériques, nous pouvons maintenant interpréter les corrélations obtenues.

1. **Corrélation entre price et cut** : La corrélation est de 0.04. Une valeur positive suggère que les diamants de meilleure coupe tendent à être plus chers.
2. **Corrélation entre price et color** : La corrélation est de 0.17. Une valeur positive indique que les diamants avec une couleur plus pure (proche du transparent) tendent à être plus chers.
3. **Corrélation entre price et clarity** : La corrélation est de -0.07. Une valeur négative indique que les diamants avec une meilleure clarté sont généralement plus abordables.

**Conclusion professionnelle :** Ces résultats confirment que des attributs de qualité supérieure, tels qu'une coupe idéale, une couleur incolore et une clarté élevée, sont associés à des prix plus élevés dans le marché des diamants. Les bijoutiers peuvent ainsi utiliser ces informations pour ajuster les prix en fonction de ces caractéristiques afin de maximiser la valorisation de chaque pierre précieuse.

Le lien entre la description de Léa et l'approche utilisée pour analyser les données dans des bases de données en statistiques réside dans **l'organisation et l'interprétation des informations pour obtenir un profil ou une vision globale cohérente.**

Voici quelques points de convergence :

1. **Structuration des données :**
   * Dans la présentation de Léa, on organise les informations en catégories (parcours, influences, aspirations, etc.).
   * En statistiques, les données sont également structurées en variables (colonnes) et observations (lignes) pour décrire des entités ou événements.
2. **Analyse descriptive :**
   * La description de Léa permet de résumer des informations pertinentes sur son parcours, ses influences et ses objectifs.
   * En statistiques, une analyse descriptive synthétise les caractéristiques des données (moyennes, médianes, distributions) pour donner un aperçu clair.
3. **Corrélation et explication :**
   * La présentation de Léa met en lien ses choix et ses influences (par exemple, ses voyages ou sa famille d’entrepreneurs expliquent ses décisions professionnelles).
   * En statistiques, on cherche à identifier les relations entre les variables (corrélation, causalité) pour expliquer les phénomènes observés.
4. **Projection et prédiction :**
   * La vision de Léa à 10 ans est une projection basée sur son parcours et ses ambitions actuelles.
   * En statistiques, les modèles prédictifs utilisent les données passées et présentes pour anticiper les résultats futurs.
5. **Utilité sociale :**
   * Léa décrit son impact futur dans la société, en montrant comment ses compétences pourront résoudre des problèmes ou améliorer des situations.
   * En statistiques, on utilise les résultats des analyses pour prendre des décisions éclairées et créer de la valeur (par exemple, optimiser des processus ou améliorer des politiques publiques).

En résumé, tout comme une base de données, la description de Léa est un ensemble structuré de données qualitatives et quantitatives qu'on peut analyser pour comprendre son passé, interpréter ses choix présents, et projeter son avenir. En statistique, ces mêmes principes sont appliqués à grande échelle pour explorer et prédire des tendances.

Le parallèle entre l’utilisation des **lois en droit pénal** par les avocats et les **lois de probabilité** en statistiques peut être établi à travers leur rôle dans la prise de décisions rationnelles et fondées sur des principes bien définis.

### 1. ****Structure des lois et des règles :****

* **Avocats et lois pénales :** Les avocats utilisent les lois pénales (ex. : articles du Code pénal) comme cadre pour interpréter les situations, défendre leurs clients, ou prouver la culpabilité ou l'innocence. Par exemple, un avocat peut s’appuyer sur l’article 121-3 du Code pénal français pour argumenter sur l'absence d'intention criminelle.
* **Statistiques et lois de probabilité :** En statistiques, les lois de probabilité (ex. : loi normale, loi de Poisson) sont utilisées pour modéliser et interpréter les comportements d’un phénomène. Par exemple, la loi normale est souvent appliquée pour décrire la répartition des notes des élèves ou les tailles dans une population.

### 2. ****Analyse des cas ou des données :****

* **Droit pénal :** Lorsqu’un avocat défend une cause, il analyse les faits, les circonstances atténuantes, ou les éléments de preuve pour construire son argumentaire. Par exemple, l’analyse des témoignages ou des preuves ADN peut être décisive dans une affaire criminelle.
* **Statistiques :** Les statisticiens analysent les données en s’appuyant sur des lois pour tirer des conclusions. Par exemple :
  + Utiliser la loi binomiale pour calculer la probabilité qu’un échantillon de 10 produits contienne au moins 3 produits défectueux.
  + Utiliser la loi de Poisson pour prédire le nombre d'accidents dans une ville donnée pendant une semaine.

### 3. ****Prise de décision :****

* **Avocats et juges :** Les décisions en droit sont basées sur une interprétation des lois et sur la jurisprudence. Par exemple, un juge peut décider d’accorder une peine avec sursis en tenant compte des précédents judiciaires et de la gravité des actes.
* **Statistiques :** Les décisions basées sur des lois de probabilité permettent d'évaluer des risques ou de prédire des résultats. Exemples :
  + Calculer la probabilité qu’un médicament ait un effet secondaire grave (loi normale ou loi de Bernoulli).
  + Décider de l’implantation d’un hôpital dans une zone en fonction des données de population (loi exponentielle pour modéliser les temps d’attente).

### 4. ****Exemples concrets de lois de probabilité utilisées comme base de décision :****

* **Loi normale (gaussienne) :**
  + Exemple : Dans un test d’évaluation des élèves, la moyenne est 12/20 avec un écart-type de 3. La loi normale permet de déterminer qu’environ 68 % des élèves auront une note entre 9 et 15.
* **Loi de Poisson :**
  + Exemple : Un gestionnaire d’un service d’urgence veut savoir combien de patients arrivent en moyenne toutes les heures. S’il observe qu’en moyenne 3 patients arrivent par heure, la loi de Poisson peut prédire la probabilité d’avoir plus de 5 patients en une heure.
* **Loi binomiale :**
  + Exemple : Une entreprise fabrique 100 produits par jour. Si 95 % des produits respectent les normes, on peut calculer la probabilité d’avoir au moins 97 produits conformes dans une journée.

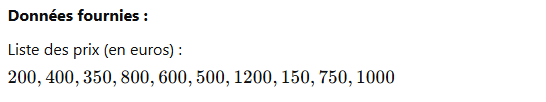
### Conclusion :

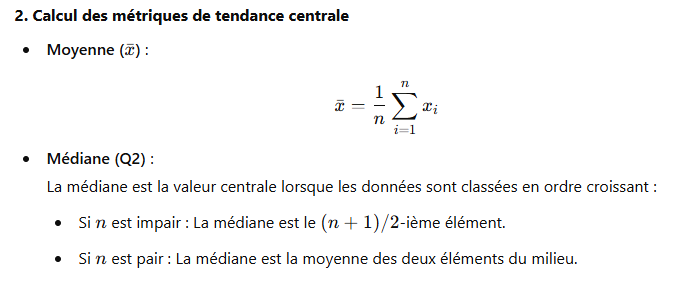
Tout comme un avocat applique des lois juridiques pour défendre une cause ou guider une décision, les statisticiens utilisent les lois de probabilité pour modéliser des phénomènes, interpréter des données et prendre des décisions éclairées. Dans les deux cas, les lois fournissent une **structure rationnelle et rigoureuse** pour analyser des situations complexes et parvenir à une conclusion logique.

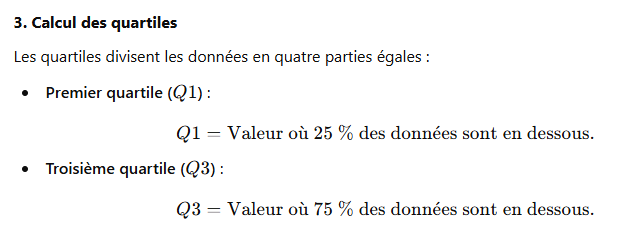
### ****Exercice : Analyse Statistique des Prix des Diamants****

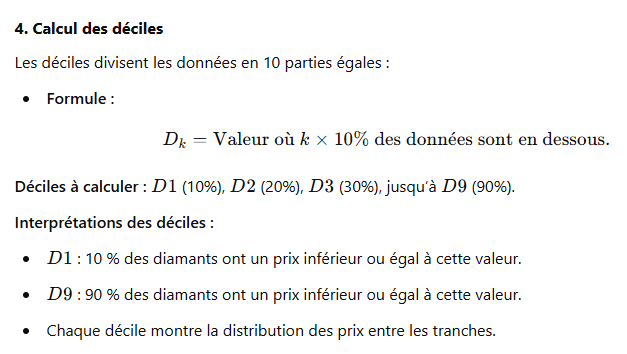
Vous avez passé un entretien pour le poste de data analyst ou data scientist. L'entreprise a constaté que de nombreux candidats ne maîtrisent plus les fondamentaux statistiques. Afin de valider votre candidature, vous êtes sélectionné parmi les 5 autres finalistes pour réaliser l'exercice suivant.

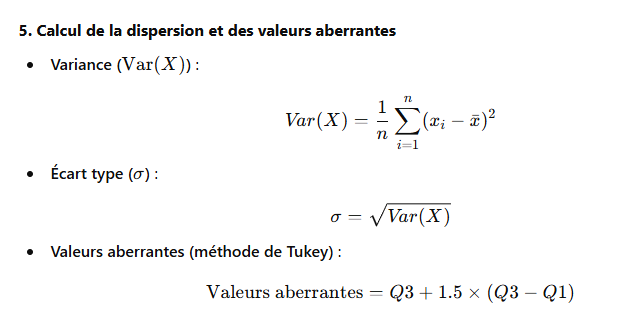
Vous êtes un data scientist travaillant pour une entreprise spécialisée dans les diamants. On vous a fourni une liste de prix de 10 diamants. Votre mission est de réaliser une analyse statistique complète en suivant les étapes suivantes.

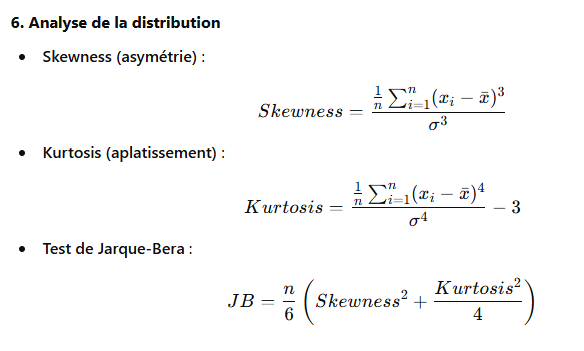


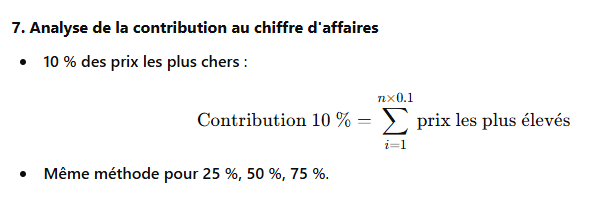


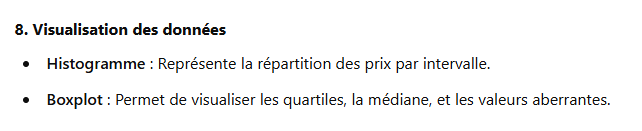












**Questions**

