

Test Data Visualization - Test 1

Sujet: Visualisation de Données

Niveau: Intermédiaire

Nombre de questions: 25

Questions et Réponses

Q1. Quel type de graphique est le plus approprié pour visualiser la distribution d'une variable continue comme les montants de prêts?

R1. L'**histogramme** est le plus approprié car: - Il divise les données en intervalles (bins) - Montre la fréquence de chaque intervalle - Permet de visualiser la forme de la distribution (normale, asymétrique, bimodale) - Révèle les outliers potentiels

Q2. Quelle est la différence fondamentale entre un histogramme et un bar chart?

R2. | Histogramme | Bar Chart | |-----|-----| | Variables **continues** | Variables **catégorielles** | | Barres **adjacentes** (pas d'espace) | Barres **espacées** | | L'axe X représente des intervalles numériques | L'axe X représente des catégories | | Montre une **distribution** | Compare des **quantités** entre groupes |

Q3. Quand utiliseriez-vous un box plot plutôt qu'un histogramme?

R3. Un box plot est préférable quand: - On veut **comparer plusieurs groupes** (ex: scores par agence) - On veut visualiser rapidement les **quartiles et médiane** - On cherche à identifier les **outliers** de manière standardisée - L'espace est limité (plusieurs distributions côte à côte)

Q4. Quels sont les 5 éléments d'un box plot et que représentent-ils?

R4. 1. **Médiane (Q2):** Ligne centrale - 50% des données 2. **Q1 (25ème percentile):** Bord inférieur de la boîte 3. **Q3 (75ème percentile):** Bord supérieur de la boîte 4. **Moustaches:** Min/Max sans outliers (généralement $1.5 \times IQR$) 5. **Points:** Outliers au-delà des moustaches

Q5. Pour visualiser l'évolution du volume de dépôts sur 12 mois, quel graphique choisiriez-vous et pourquoi?

R5. Un **line chart (graphique linéaire)** car: - Montre la **tendance temporelle** - Connecte les points pour visualiser l'évolution - Permet d'identifier **saisonnalité** et **patterns** - L'axe X représente le temps (variable ordonnée)

Q6. Quand un pie chart est-il approprié et quand faut-il l'éviter?

R6. **Approprié:** - Montrer des **parts d'un tout** (100%) - Maximum **5-6 catégories** - Les différences sont **significatives** ($>5\%$)

À éviter: - Plus de 6 catégories - Comparer des valeurs proches - Comparer plusieurs périodes - Les parts ne totalisent pas 100%

Q7. Comment visualiseriez-vous la relation entre le score de crédit et le montant du prêt?

R7. Un **scatter plot (nuage de points)** car: - Montre la relation entre **deux variables continues** - Permet d'identifier la **corrélation** (positive, négative, nulle) - Révèle les **clusters** et **outliers** - On peut ajouter une **ligne de tendance** (régression)

Q8. Qu'est-ce qu'une heatmap et dans quel contexte bancaire l'utiliseriez-vous?

R8. Une **heatmap** représente des données sous forme de matrice colorée où l'intensité de la couleur correspond à la valeur.

Utilisations bancaires: - **Matrice de corrélation** entre variables de risque - **Analyse de concentration** du portefeuille (secteur × région) - **Performance par agence et mois** - **Flux de transactions** entre régions

Q9. Comment le coefficient de variation (CV) influence-t-il le choix de visualisation?

R9. - **CV faible (<30%):** Données homogènes → Histogramme standard - **CV élevé (>30%):** Données dispersées → Considérer échelle log ou box plot - **CV très élevé (>100%):** Probablement des outliers → Box plot ou violin plot

Le CV = (écart-type / moyenne) × 100

Q10. Quelle visualisation utiliseriez-vous pour montrer la répartition du portefeuille par secteur économique au CEO?

R10. Un **treemap** ou un **bar chart horizontal** ordonné: - Treemap: Montre la proportion relative de chaque secteur - Bar chart: Plus précis pour comparer les valeurs exactes

Pour le CEO, préférer le **bar chart horizontal** car: - Plus facile à lire - Labels visibles - Comparaison directe des valeurs

Q11. Qu'est-ce qu'un violin plot et quand l'utiliser?

R11. Un **violin plot** combine: - Un box plot (médiane, quartiles) - Une estimation de densité (forme de la distribution)

Utilisation: - Comparer des distributions entre groupes - Quand la forme de la distribution importe (bimodale, asymétrique) - Plus informatif qu'un box plot quand n est grand

Q12. Comment visualiser l'évolution du NPL ratio avec comparaison à l'année précédente?

R12. Un **line chart avec deux séries:** - Ligne continue: Année actuelle - Ligne pointillée: Année N-1 - Même axe X (mois 1-12)

Ou un **bar chart groupé** si vous voulez montrer les valeurs exactes.

Q13. Quelles couleurs utiliser pour un dashboard bancaire professionnel?

R13. Bonnes pratiques: - **Bleu foncé:** Couleur principale (confiance, professionnalisme) - **Vert:** Performance positive, croissance - **Rouge:** Alertes, déclin, risque - **Gris:** Données secondaires, comparaisons

À éviter: - Trop de couleurs (max 5-6) - Couleurs saturées - Rouge/vert ensemble (daltonisme)

Q14. Comment visualiser la migration entre buckets de retard (current → 30j → 60j → 90j)?

R14. Un **Sankey diagram** ou un **flow chart**: - Montre les flux entre états - Épaisseur proportionnelle au volume - Permet de voir les transitions positives (régularisation) et négatives (dégradation)

Alternative: **Matrice de transition** avec heatmap

Q15. Quelle est l'importance du ratio données/encre (data-ink ratio)?

R15. Concept d'Edward Tufte:

Data-ink ratio = Encre utilisée pour les données / Encre totale

Objectif: Maximiser ce ratio en éliminant: - Grilles excessives - Bordures inutiles - Effets 3D - Légendes redondantes

Un bon graphique: Simple, épuré, focus sur les données.

Q16. Comment représenter un objectif vs réalisé sur un dashboard?

R16. Un **bullet chart** est idéal: - Barre principale: Valeur réalisée - Marqueur: Objectif (target) - Zones de fond: Seuils (mauvais/moyen/bon)

Alternative: **Gauge chart** (moins précis mais plus visuel)

Q17. Quelle visualisation pour montrer la décomposition du résultat net?

R17. Un **waterfall chart (graphique en cascade)**:

Revenus d'intérêts
+ Commissions
- Charges d'intérêts
- Charges opérationnelles
- Provisions
- Impôts
= Résultat Net

Montre la contribution positive/négative de chaque composante.

Q18. Comment gérer les outliers dans une visualisation?

R18. Options: 1. **Identifier explicitement:** Points distincts sur scatter plot 2. **Échelle logarithmique:** Comprime les extrêmes 3. **Axe tronqué:** Avec indication claire de la troncature 4. **Visualisation séparée:** Un graphique principal + zoom sur outliers 5. **Box plot:** Affiche naturellement les outliers

Ne jamais: Les supprimer silencieusement

Q19. Qu'est-ce qu'un sparkline et où l'utiliser?

R19. Un **sparkline** est un mini-graphique intégré dans une cellule de tableau: - Sans axes ni labels - Montre la tendance rapidement - Idéal pour les tableaux de bord exécutifs

Exemple: Dans un tableau d'agences, ajouter un sparkline de l'évolution du PNB sur 12 mois.

Q20. Comment visualiser la concentration du portefeuille (top clients)?

R20. Une **courbe de Lorenz** ou un **Pareto chart**: - Axe X: % cumulé des clients (triés par exposition) - Axe Y: % cumulé de l'exposition totale

Interprétation: Si 20% des clients représentent 80% de l'exposition → forte concentration (risque)

Q21. Quelle visualisation pour un funnel de conversion (demandes → approbations → décaissements)?

R21. Un **funnel chart**:

Demandes reçues	1000
Éligibles	800
Approuvées	500
Décaissées	450

Montre les taux de conversion à chaque étape.

Q22. Comment choisir entre un graphique interactif et statique?

R22. Interactif (Power BI, Tableau): - Exploration de données - Plusieurs dimensions à filtrer - Utilisateur technique - Dashboard self-service

Statique (rapport PDF): - Message clair et unique - Présentation exécutive - Archivage/audit - Communication externe

Q23. Quelle est la règle des 5 secondes pour un graphique?

R23. Un bon graphique doit communiquer son message principal en **5 secondes** ou moins: - **Titre clair** et actionnable - **Axes labellisés** clairement - **Données mises en évidence** (couleur, annotation) - **Pas de surcharge** d'informations

Test: Montrer le graphique 5 secondes, demander le message → doit être compris.

Q24. Comment annoter efficacement un graphique pour le management?

R24. Bonnes pratiques d'annotation: - **Titre actionnable**: "Le NPL a augmenté de 0.5% ce trimestre" (pas juste "NPL") - **Callouts**: Flèches vers les points importants - **Valeurs clés**: Afficher les chiffres importants directement - **Contexte**: "Objectif: 5%" sur le graphique - **Source**: Mentionner la date et source des données

Q25. Créez une recommandation de graphique pour chaque KPI bancaire suivant:

R25. | KPI | Graphique Recommandé | **ROE mensuel** | Line chart (tendance) | **Répartition secteur** | Treemap ou Bar horizontal | **NPL vs objectif** | Bullet chart | **Évolution dépôts** | Area chart (volume) | **Score crédit distribution** | Histogram + box plot | **Corrélation variables** | Heatmap | **Migration buckets** | Sankey diagram | **Performance agences** | Bar chart ranking |

Scoring

Score	Niveau
0-10	À améliorer
11-17	Intermédiaire
18-22	Avancé
23-25	Expert