Le guide est généré par GPT 5o : à vérifier chaque phrase

1. **Chargement & prétraitement**
   * Lecture du CSV (dates parsées, types corrects, nettoyage basique).
   * Ajout d’une colonne **Profit Margin = Profit / Sales**.
   * Vérifications de base et aperçu rapide.
2. **Visualisations & analyses demandées**
   * **États avec le plus de ventes** (bar chart + export CSV).
   * **Comparatif New York vs California** (ventes + profits + marge).
   * **Client “outstanding” à New York** (défini par *plus haut profit*, départage par ventes).
   * **Différences de rentabilité entre états** (Top/Bottom par profit margin, figures + CSV).
3. **Principe de Pareto (80/20) — Clients vs Profit**
   * → Courbe cumulative, calcul de la part minimale de clients expliquant 80% des profits.
   * **Top 20 villes par Ventes & par Profit** (+ comparaison de marges).
   * **Top 20 clients par Ventes** (table + bar chart).
   * **Courbe cumulative — Ventes par clients** (vérification Pareto côté ventes).
4. **Priorisation marketing (États & Villes)**
   * → Score composite basé sur ventes, profits et marge (z-scores), tableaux exportés.
5. **Exports & figures**
6. Tout est sauvegardé dans /mnt/data/outputs/ : (à adapter chez vous)
   * CSV (tables) + PNG (graphiques).

Conseils rapides d’interprétation

* **Client outstanding à NY** : par défaut, “plus haut profit” (tu peux changer la règle si tu préfères CLV ou marge).
* **Pareto** : le 20/80 exact est rare ; vise une **zone** (ex. 18–25% clients ≈ 75–85% profit/ventes).
* **Priorités marketing** : cible d’abord les **états/villes avec ventes élevées et marge saine**.
* Pour les zones **ventes élevées / marge faible**, pense **optimisation des remises**, bundles à meilleure marge, ou **segmentation d’offres**.