**Medical Appointment No Shows**

**Pourquoi 30 % des patients manquent-ils leur rendez-vous médical ?**

**À propos du jeu de données**

**Contexte**

Une personne prend rendez-vous chez le médecin, reçoit toutes les instructions… et ne se présente pas. Qui blâmer ?  
Ce jeu de données est une référence à **Joni Hoppen** et **Aquarela Analytics**.

**Contenu**

Le jeu de données contient **110 527 rendez-vous médicaux** et leurs **14 variables associées** (caractéristiques).  
La variable la plus importante est celle indiquant si le patient **s’est présenté** ou **a manqué** son rendez-vous (**No-show**).  
Les noms des variables sont explicites, mais voici le dictionnaire complet :

**Dictionnaire des données**

1. **PatientId**  
   Identifiant unique d’un patient.
2. **AppointmentID**  
   Identifiant unique de chaque rendez-vous.
3. **Gender**  
   Sexe du patient : Male ou Female.  
   Remarque : la proportion de femmes est plus élevée, elles prennent davantage soin de leur santé que les hommes.
4. **DataMarcacaoConsulta**  
   Date réelle du rendez-vous, jour où le patient doit voir le médecin.
5. **DataAgendamento**  
   Date à laquelle le rendez-vous a été pris (par appel ou enregistrement), forcément avant la date de consultation.
6. **Age**  
   Âge du patient.
7. **Neighbourhood**  
   Lieu où a lieu le rendez-vous.
8. **Scholarship**  
   True ou False – Indique si le patient bénéficie du programme **Bolsa Família** (voir <https://en.wikipedia.org/wiki/Bolsa_Fam%C3%ADlia>).
9. **Hipertension**  
   True ou False – Indique si le patient souffre d’hypertension.
10. **Diabetes**  
    True ou False – Indique si le patient est diabétique.
11. **Alcoholism**  
    True ou False – Indique si le patient souffre d’alcoolisme.
12. **Handcap**  
    True ou False – Indique si le patient présente un handicap.
13. **SMS\_received**  
    Nombre de SMS envoyés au patient (1 ou plus).
14. **No-show**  
    True ou False – Indique si le patient **n’est pas venu** au rendez-vous.

**Inspiration**

Peut-on prédire si un patient ne se présentera pas à son rendez-vous ?  
Ce jeu de données permet d’explorer cette question grâce à l’analyse statistique et au machine learning.