

Exercice :

Maladie cardiaque	Sexe	Date de naissance	Douleur thoracique	Tension artérielle au repos	Taux de cholestérol dans le sang mg/dl	Limite supérieure de la fréquence cardiaque
Oui	0	21/01/1964	0	100	248	122
Non	1	20/06/1964	0	114	318	140
Non	1	11/06/1967	0	160	289	145
Non	1	11/08/1976	0	120	249	144
Non	1	23/12/1968	0	122	286	116
Oui	0	6/07/1951	0	112	149	125
Non	0	8/02/1979	0	132	341	136
Oui	0	12/10/1988	1	118	210	192
Non	1	19/11/1971	0	140	298	122
Non	1	13/05/1970	0	128	204	156
Oui	0	2/06/1988	1	118	210	192
Oui	0	11/04/1971	2	140	308	142
Non	1	3/04/1968	0	124	266	109
Oui	0	13/09/1972	1	120	244	162
Oui	1	23/01/1964	2	140	211	165
Non	1	23/06/1962	2	140	185	155
Oui	0	19/12/1955	0	106	223	142
Oui	1	13/10/1977	0	104	208	148
Oui	0	10/10/1959	2	135	252	172
Oui	0	4/10/1980	2	120	209	173
Non	0	23/07/1961	0	145	307	146

Questions :

1. Saisir les données et coder les variables selon leurs types : Nominale, Ordinale, Echelles.
2. Créer une nouvelle variable Age en utilisant la date de naissance.
3. Calculer : la moyenne, l'écart type, la médiane, le max et le min des variables de types échelle.
4. Tracer un camembert des effectifs selon le sexe.
5. Tracer un camembert des gens qui ont une maladie cardiaque.
6. Tracer des diagrammes de dispersions des trois dernières variables.
7. Si on considère la statistique F (test de Fisher/Snedecor), laquelle parmi les trois dernières variables est expliquée par l'âge ? tracer la droite des moindres carrés (Régression).
8. Créer un modèle linéaire qui explique l'existence d'une maladie cardiaque selon les informations présentes dans le tableau.
9. Quelles sont les facteurs qui peuvent être la cause d'une maladie cardiaque ?
10. Recalculer le modèle avec seulement les facteurs qui peuvent causer une maladie cardiaque.

Remarque : on prend le risque 10%