TP1 ATDN2

XIONG Alexis

Etape 1

Nous disposons de 7 variables :

- surface_ha : la surface cultivée en hectares
- type_sol : le type de sol utilisé
- engrais_kg/ha : la quantité d'engrais en kg par hectare
- precipitations_mm : les précipitations moyennes mensuelles en mm
- temperature_C : la température moyenne mensuelle en Celsius
- rendement_t/ha : le rendement obtenu en tonnes/hectare

Le problème métier consiste à optimiser le rendement du maïs en fonction de la température, de la precipitation, du type de sol et de la quantité d'engrais.

La variable cible est rendement_t/ha

Les variables explicatives sont surface_ha, type_sol, engrais_jg/ha, precipitations_mm et temperature_C

La problématique centrale est donc d'optimiser les ressources pour maximiser le rendement par hectare du maïs. Pour cela, il faut analyser les variables explicatives pour comprendre pour comprendre comment optimiser le rendement.

Etape 2

In [3]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('rendement_mais.csv')

moyenne = df['RENDEMENT_T_HA'].mean()
print(f"Moyenne du rendement : {moyenne} T/HA")

mediane = df['RENDEMENT_T_HA'].median()
print(f"Médiane du rendement : {mediane} T/HA")

mode = df['RENDEMENT_T_HA'].mode()
print(f"Mode du rendement : {mode.tolist()} T/HA")

ecart_type = df['RENDEMENT_T_HA'].std()
```

```
print(f"Écart-type du rendement : {ecart_type} T/HA")

variance = df['RENDEMENT_T_HA'].var()
print(f"Variance du rendement : {variance} (T/HA)²")

etendue = df['RENDEMENT_T_HA'].max() - df['RENDEMENT_T_HA'].min()
print(f"Étendue du rendement : {etendue} T/HA")

plt.figure(figsize=(8, 6))
plt.hist(df['RENDEMENT_T_HA'], color='red', edgecolor='black')
plt.title('Distribution du rendement')
plt.show()

plt.figure(figsize=(8, 6))
plt.hist(df['PRECIPITATIONS_MM'], color='red', edgecolor='black')
plt.title('Distribution de la précipitation')
plt.show()

plt.figure(figsize=(8, 6))
plt.hist(df['TEMPERATURE_C'], color='red', edgecolor='black')
plt.title('Distribution de la température')
plt.show()

plt.figure(figsize=(8, 6))
sns.boxplot(y=df['RENDEMENT_T_HA'], color='skyblue')
plt.title('Boxplot du rendement')
plt.ylabel('Rendement (T/HA)')
plt.show()

plt.figure(figsize=(8, 6))
sns.boxplot(y=df['PRECIPITATIONS_MM'], color='skyblue')
plt.title('Boxplot de la précipitation')
plt.ylabel('Précipitation (mm)')
plt.show()

plt.figure(figsize=(8, 6))
sns.boxplot(y=df['TEMPERATURE_C'], color='skyblue')
plt.title('Boxplot de la température')
plt.ylabel('Température (Celsius)')
plt.show()

matrice_de_correlation = df.corr(numeric_only=True)
print(matrice_de_correlation)

plt.figure(figsize=(8, 6))
sns.heatmap(matrice_de_correlation, annot=True, cmap='coolwarm')
plt.title('Heatmap de la matrice de corrélation')
plt.show()
```

Moyenne du rendement : 7.378418687218944 T/HA
Médiane du rendement : 7.349138167259971 T/HA
Mode du rendement : [3.000276469608442, 3.005880516851696, 3.0100088477656866, 3.037686992651621, 3.057411390659316, 3.058179220107395, 3.061901659976844, 3.0925604634212283, 3.097299752716767, 3.1017228293012025, 3.1160772281057287, 3.12314173284613, 3.124759223215741, 3.1458035760552048, 3.1461909888850834, 3.149332325693381, 3.172399369961276, 3.181240640543608, 3.1884070604405994, 3.198595338140618, 3.2030806053481253, 3.2133697828006653, 3.2155163874343518, 3.216045403604, 3.2247449590534463, 3.226791853389627, 3.2312068038246737, 3.2319263888791605, 3.250670167389669 6, 3.2578718481988904, 3.2709348499018702, 3.274173176780641, 3.2919910144246827, 3.293931547060463, 3.312054338034597, 3.3142039739060145, 3.353424728174871, 3.3545265032353746, 3.3605320714562503, 3.3612497317682988, 3.362322027083089, 3.362569669001986, 3.374210638619344, 3.377236950446318, 3.3965927541334326, 3.4049713911188197, 3.410266421538019, 3.435264399439028, 3.438253635599368, 3.4429882308668542, 3.446529360049537, 3.452187088406839, 3.453145742724179, 3.4537509624198828, 3.45515793441806, 3.4556022504216672, 3.460351850583579, 3.46164464083094, 3.4748873207364235, 3.4910886745442906, 3.500449695124145, 3.510123889492519, 3.511932579998819, 3.5132422260325744, 3.5270877475608047, 3.536283876178293, 3.536555928101931, 3.539489707751774, 3.553016611177088, 3.556276146203004, 3.5583308726590337, 3.5583727526089888, 3.595486699509964, 3.6051076208150494, 3.6160540730941513, 3.6245467731376495, 3.6309732567713127, 3.640401109901188, 3.654614530008016, 3.690822556998642, 3.691227434829794, 3.694687772038212, 3.710454626519, 3.7135225616807257, 3.72095153948642, 3.72168081130902, 3.725538646526416, 3.731066713260054, 3.734174330502293, 3.7343884800435454, 3.7517242321201687, 3.7571489603344457, 3.762195076877761, 3.76550746940317, 3.774289413296489, 3.775189251573123, 3.7831549647250986, 3.79009413352344, 3.808891665625446, 3.822797926354244, 3.8312447530024074, 3.8439887363131486, 3.8497700789443288, 3.854687477055388, 3.858248291457032, 3.8698035085897695, 3.8726910386888234, 3.8820691630266695, 3.885817701912319, 3.9022100735058896, 3.9255833294172895, 3.929385079875436, 3.9309070337443504, 3.9400906629013615, 3.949350847149321, 3.952868863918164, 3.972401034387664, 4.004336999615227, 4.03631560555681, 4.047226366480463, 4.04924767100127, 4.063967600619797, 4.067793891674223, 4.068381859582003, 4.079019841254409, 4.083488209101802, 4.084026059969089, 4.109554013690522, 4.1191125728778335, 4.129376671192356, 4.129948109086294, 4.133596767414224, 4.134018250775949, 4.14141620295376 9, 4.142499723794754, 4.15493275969888, 4.178763686947567, 4.189616919784846, 4.197088190165959, 4.198639131007711, 4.198889317229658, 4.21118640393697, 4.2169558234156925, 4.219135831152217, 4.243127638262225, 4.259886390495866, 4.271011890358345, 4.28300899529804, 4.287010103991538, 4.28953312832377, 4.291225113162181, 4.307940231784537, 4.317926309425345, 4.343245227673265, 4.354250439053978, 4.36289082932967, 4.363009697418216, 4.368437106820426, 4.384053349618945, 4.385099142243963, 4.390429836044332, 4.405813571375393, 4.406852736379747, 4.411314347851251, 4.428630870368786, 4.442030408788557, 4.446116262406607, 4.450969307776103, 4.452306225495448, 4.453120506578161, 4.457307059689923, 4.463947335949366, 4.465118902856064, 4.467172024333976, 4.470300825585782, 4.479649761709909, 4.513832901898663, 4.528983658971516, 4.536375269812321, 4.538427621787635, 4.546465143218926, 4.552714892517258, 4.561464453195198, 4.569683404003495, 4.571824673066966, 4.576108261173179, 4.579424646432217, 4.592072557415856, 4.593740298575609, 4.59438170900363, 4.611284736974254, 4.627066355061416, 4.628027112410085, 4.630649511450745, 4.636009804956258, 4.645540018819354, 4.660131459538951, 4.66169211612632 3, 4.663545581523224, 4.667433198449913, 4.667491653001861, 4.671597648046671, 4.679759458179143, 4.709425228118048, 4.715488497670048, 4.740918086608343, 4.756498654967854, 4.766211450355491, 4.766648720799603, 4.767772942848156, 4.7878011920766514, 4.813755402056816, 4.8206455424238

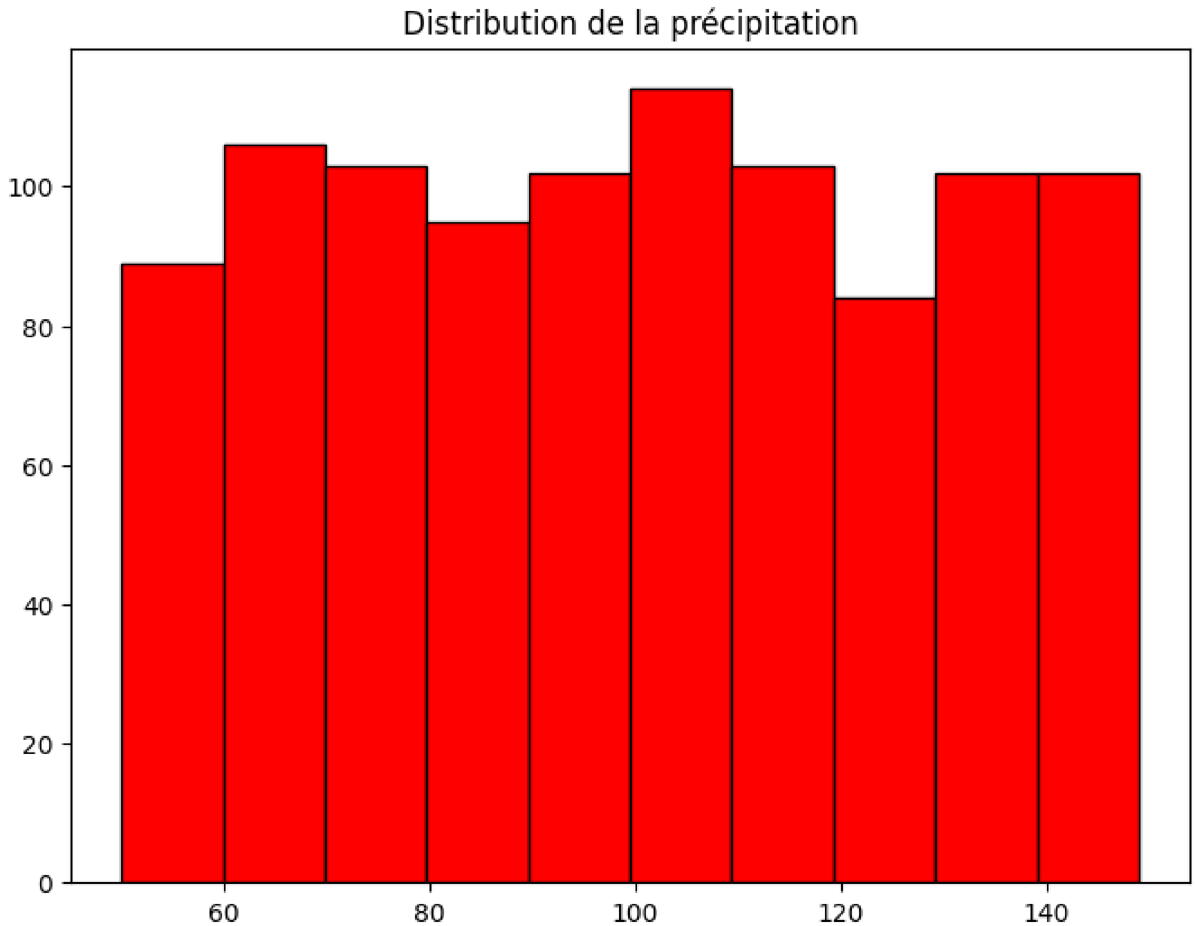
33, 4.840035352178457, 4.867868425851331, 4.874838510076661, 4.8794784194513
1, 4.881534363361614, 4.890964736211166, 4.892867006572143, 4.90013348302233
7, 4.9120254947592565, 4.925952108284822, 4.9322629871702, 4.94699499867744
3, 4.961022416640696, 4.96386781096955, 4.968388202912249, 4.96842500015885
2, 4.9840140382087705, 4.984388754732743, 4.984538370920824, 4.9858595365957
36, 4.991148268014576, 5.0040847121077645, 5.018867029681736, 5.021916578775
899, 5.034711639095065, 5.036451870028676, 5.038694634324951, 5.040953627876
812, 5.068440532178566, 5.069943785669165, 5.0797276758704655, 5.08242572739
1632, 5.0832769238876825, 5.09999540384394, 5.103667128330699, 5.10431549783
7577, 5.1098304001036325, 5.132315945070511, 5.132399924159838, 5.1373070107
70318, 5.143319882991267, 5.162692111053301, 5.16478426621338, 5.1660346328
14378, 5.212626982440603, 5.219927542248659, 5.2217132483615964, 5.247020574
662072, 5.257911899016302, 5.260523407935612, 5.270807343627183, 5.275723064
251, 5.277237068282279, 5.290331984564342, 5.298027363869197, 5.305723629954
194, 5.327324320213056, 5.337037550553726, 5.338453241462695, 5.349436736113
633, 5.350461892597743, 5.351997160968454, 5.369226118662101, 5.389467666861
651, 5.393356055689069, 5.41356930970422, 5.419178535271783, 5.4357954778358
91, 5.443428022884402, 5.444698631469327, 5.447591034583788, 5.4479874480992
04, 5.477152629333075, 5.483304249424922, 5.493623247420597, 5.4971821175245
86, 5.507787568035328, 5.523012772962712, 5.525407319141244, 5.5294477515020
02, 5.534544378950685, 5.549559631433196, 5.560872471092404, 5.5706338970601
88, 5.571368108948013, 5.584990882846357, 5.591508084083932, 5.5936607336357
65, 5.604955694575391, 5.623148221444145, 5.640180792969089, 5.646607159591,
5.662416877468824, 5.663388264929545, 5.666932217098201, 5.667218521619596,
5.669512339413839, 5.675584838531938, 5.699478627993203, 5.729573017643546,
5.75169561702562, 5.757739044159328, 5.757895404330345, 5.7649477531914375,
5.769938147638342, 5.776214533405357, 5.806914921873528, 5.824879470502235,
5.8249026886352455, 5.834266001300062, 5.837154911845057, 5.838273689395057
5, 5.841958361245464, 5.857947302318296, 5.911773712369424, 5.91896439616366
3, 5.926900766485902, 5.941152352719996, 5.942469219030078, 5.95361534517850
8, 5.954097951105055, 5.97256200751876, 5.982169968626853, 5.98774354262260
9, 6.007734427557989, 6.009610610898738, 6.018344455957162, 6.03708851067806
05, 6.042759684296352, 6.0537474057861065, 6.059746774139233, 6.063704005028
336, 6.0686271429200245, 6.071867813550469, 6.078968587772346, 6.08457832885
1165, 6.090480319598798, 6.09561917901917, 6.105753709619103, 6.109327683511
012, 6.112254885693825, 6.127669157823791, 6.149440682136427, 6.157371984015
141, 6.160974486139028, 6.177926000445505, 6.191995844865746, 6.192557010571
914, 6.193486699888464, 6.194079151378986, 6.198954036276556, 6.209365583875
851, 6.220301244020337, 6.2232574487623005, 6.227594237582323, 6.23287872284
4213, 6.2442705462739845, 6.249672537069335, 6.252634011604677, 6.2630381690
31851, 6.283402231870188, 6.306516558273634, 6.343205363619392, 6.3495592972
22074, 6.352580383981193, 6.361121506194348, 6.370934190416231, 6.3724977345
66387, 6.40055292354468, 6.411364368447315, 6.417763706273857, 6.41981171068
8552, 6.428566605211275, 6.428921463015599, 6.439255273493123, 6.44505690362
7365, 6.4489382935252655, 6.461499683808913, 6.480363513395801, 6.4854697028
23572, 6.485902661777827, 6.49853713088892, 6.525682055060504, 6.52719506649
5182, 6.528842192141154, 6.536055980038611, 6.537935168034609, 6.53832416544
6137, 6.551150837870314, 6.556397677250745, 6.55733383888042, 6.563435035763
549, 6.564008697470537, 6.566813629483985, 6.572929459867267, 6.579159729897
492, 6.580140968361883, 6.580791327156845, 6.593993221793003, 6.595669061282
413, 6.601910482647048, 6.602815525556947, 6.6063375015102, 6.61315601180138
9, 6.618244951994166, 6.621656333693806, 6.6243334889510415, 6.6246553627214
77, 6.62886980519927, 6.628895698654405, 6.648805426252395, 6.6593148362352
5, 6.661243208565512, 6.666289717915491, 6.669732945024591, 6.67610023267634
4, 6.687055823632672, 6.697099289996506, 6.70071428282542, 6.70077739994291
2, 6.701735748596081, 6.71051755048499, 6.718552654869476, 6.72726961394099

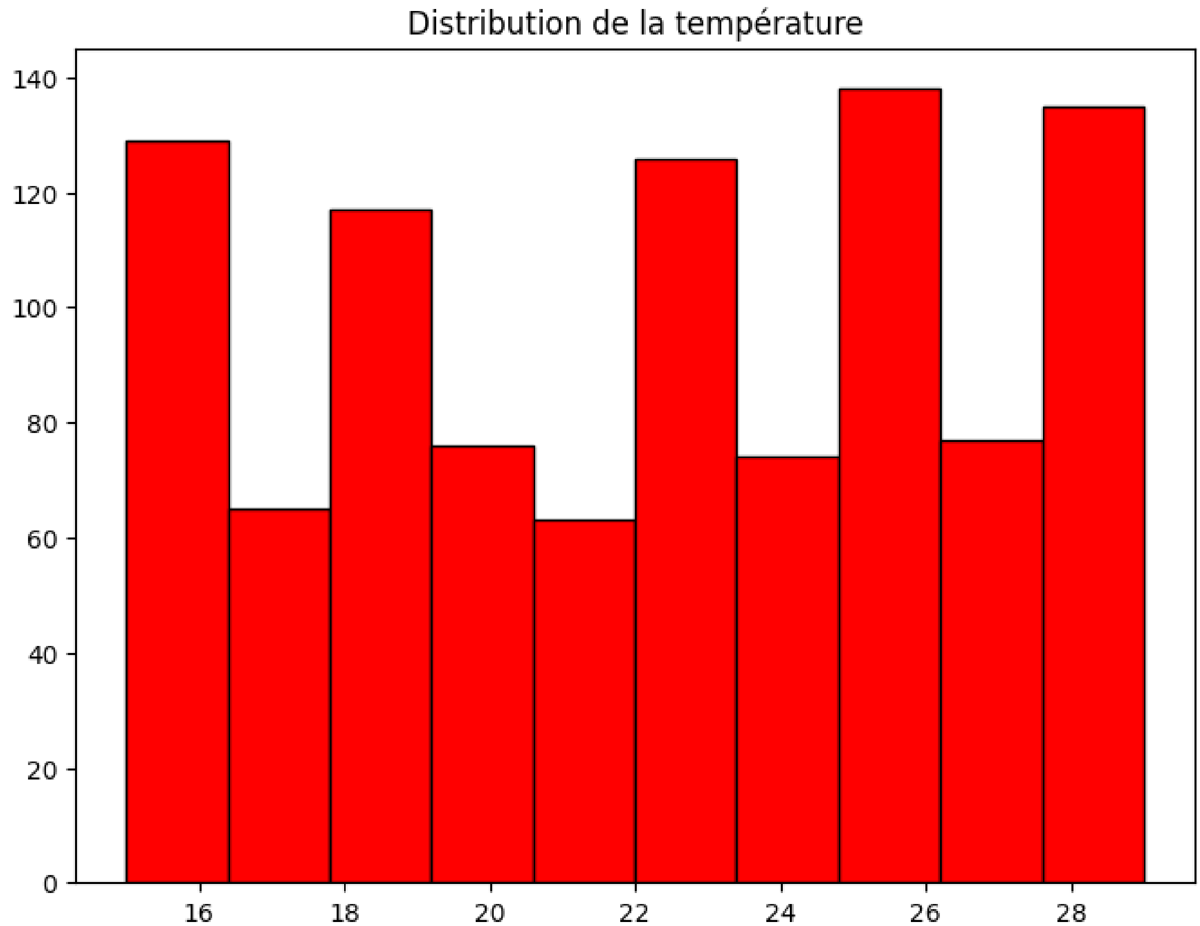
6, 6.733430899692168, 6.738987005328807, 6.742367112165699, 6.75474076708118
8, 6.759002962352953, 6.763303093465197, 6.765505800405002, 6.77189525625376
7, 6.772645883119091, 6.776646445261293, 6.780099038660966, 6.83509114157007
7, 6.837585211214603, 6.83849772517675, 6.847378086812412, 6.85612113463915
9, 6.861561323723993, 6.867060270196332, 6.869618578018395, 6.86986692708953
8, 6.881104828124021, 6.881715357069131, 6.89494335907637, 6.90459530569719
8, 6.9436075192351, 6.9587007574245145, 6.976503343217591, 6.97830078838011
2, 6.9789633121174806, 6.983196113395339, 6.9851832288936375, 6.987303357038
575, 6.996932012159568, 7.020430745326683, 7.0413489113711325, 7.05052775483
6027, 7.065655283832994, 7.085909200797291, 7.0876529646303545, 7.1213609499
06251, 7.124912351288179, 7.126667059683138, 7.132116209580128, 7.1758352835
019865, 7.188032405455375, 7.195634320148013, 7.196694716458369, 7.200238087
959676, 7.209130912363893, 7.212621488108611, 7.238642271912601, 7.244845416
755358, 7.245658537665951, 7.275808855800912, 7.280365726398075, 7.291549899
598119, 7.296153583101612, 7.296745705323183, 7.306419630964703, 7.306755329
275703, 7.331501560973133, 7.339188890278577, 7.348627769823016, 7.349648564
696927, 7.354222858745826, 7.356185152584637, 7.375174366934317, 7.385388186
900457, 7.393599077797326, 7.404532836659299, 7.406620158452203, 7.406796684
288293, 7.432818426428604, 7.433121636779679, 7.435376260795421, 7.466152997
107326, 7.483169491932381, 7.499630927683096, 7.502506816869843, 7.511480801
631564, 7.513661643151497, 7.521845584497388, 7.551270694032365, 7.553384888
876026, 7.588686085954119, 7.595226817136039, 7.605367024032308, 7.623370696
937743, 7.626858982961338, 7.630151949659048, 7.646549692779066, 7.662697482
363103, 7.667055197991606, 7.6765050822270675, 7.692393193263671, 7.71061233
8298074, 7.722620738952889, 7.7292327662030385, 7.769465282372097, 7.7816461
65158388, 7.783094227703748, 7.784149374180581, 7.786845898545149, 7.7906268
8356309, 7.842204348275127, 7.862510038686789, 7.862668786112765, 7.86269466
08534976, 7.870024423909602, 7.871430998475103, 7.872369788166944, 7.8977879
6253332, 7.89910398870889, 7.918126875553119, 7.920261138649171, 7.922359883
869293, 7.933306824552198, 7.937295735492196, 7.939855213883398, 7.948817064
985847, 7.954785125366031, 7.955970414967663, 7.960439657247349, 7.971295240
355109, 7.973960409990734, 7.981245985888649, 7.992183223095305, 7.992586229
541189, 7.996204080778696, 8.005716233670105, 8.008181736139399, 8.009205856
51938, 8.011064198353692, 8.013364619475317, 8.020863080940659, 8.0596992995
80725, 8.075036289498621, 8.075276100800615, 8.078086350160824, 8.0813157504
0117, 8.091704555519602, 8.092033993568968, 8.09265504281922, 8.094165823990
387, 8.099447330149168, 8.106518497404661, 8.109545861379491, 8.111768675844
726, 8.132582448582514, 8.137896751555285, 8.141344229433876, 8.144468845346
495, 8.147962904552335, 8.149106673499857, 8.157379805571141, 8.164792647093
867, 8.165548794621094, 8.178232558161888, 8.180340857760738, 8.191716430194
301, 8.191930811800876, 8.192834511759138, 8.194053741491159, 8.198685789099
804, 8.200263452259762, 8.201020014246794, 8.20523786804886, 8.2064325740639
71, 8.21383019995744, 8.225807033186706, 8.23562982679068, 8.24443004410213
5, 8.248171971896252, 8.249271530813012, 8.260888019335123, 8.26345030175211
9, 8.270073559139718, 8.270289154853511, 8.279615264552964, 8.29192466517449
5, 8.292004170866655, 8.295169988848379, 8.341152444342514, 8.34402568749106
9, 8.348176445777375, 8.353626415517187, 8.356915507141872, 8.35706759447461
4, 8.361118417453767, 8.364737597144753, 8.372196214919905, 8.37250701071855
4, 8.38706693895984, 8.393693583057498, 8.394237525379491, 8.39515025692845,
8.408208050931774, 8.416260605509025, 8.42204497769929, 8.428351081106523,
8.431759258510088, 8.43923387743294, 8.445805666864846, 8.453972705439229,
8.464217527952194, 8.474875500069945, 8.490971298326112, 8.503172501030948,
8.506479180928993, 8.515162519828593, 8.522741126904453, 8.525732756927095,
8.536162264870798, 8.53878701838752, 8.553953499857908, 8.57255461829253, 8.
574509724298517, 8.604205868548178, 8.604780629158569, 8.627971529590084, 8.
63085710484549, 8.631137210352174, 8.645960415816102, 8.668385889912368, 8.6

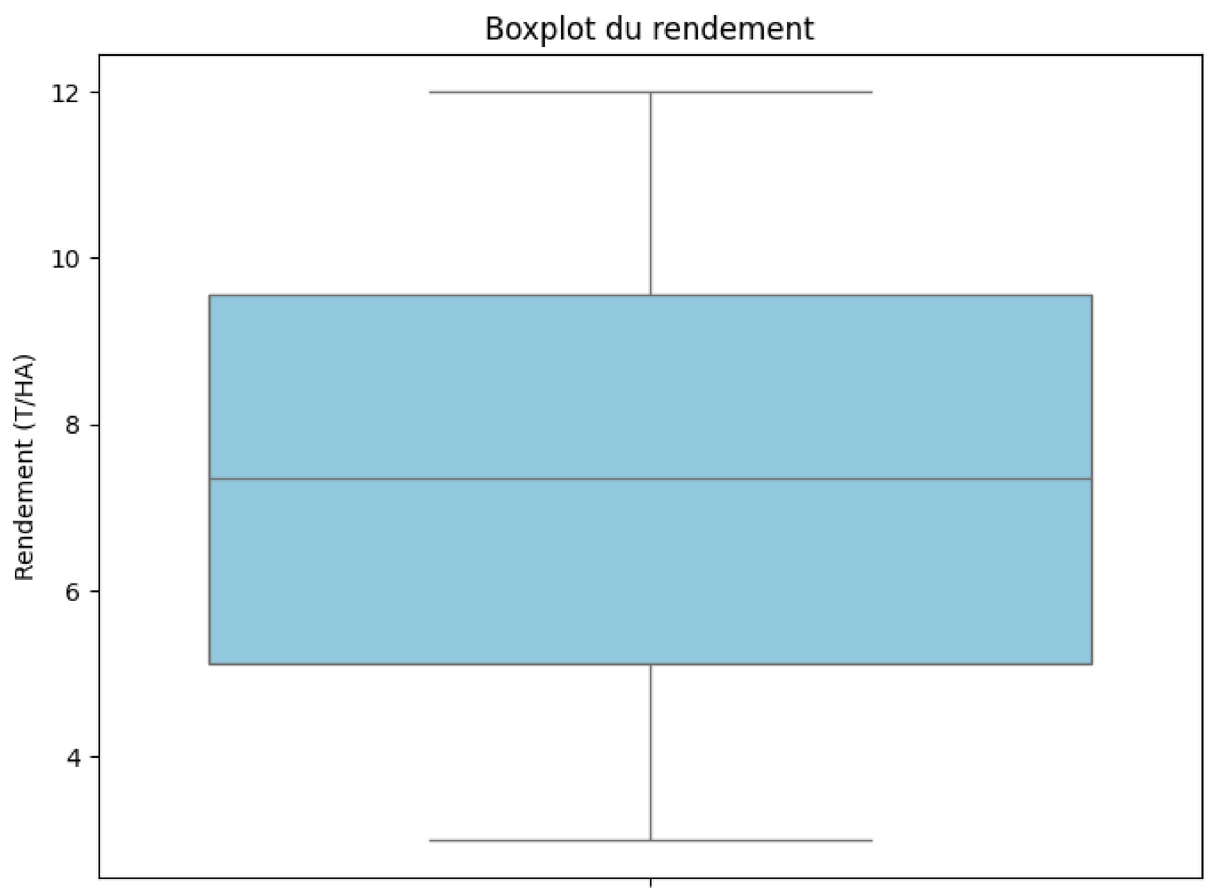
7083814175497, 8.682377934434259, 8.683129133247185, 8.691566308189577, 8.70
3404397889761, 8.71872921175719, 8.739967276610324, 8.76868975288882, 8.7735
40331306265, 8.776375651212918, 8.795609215028522, 8.802707237191367, 8.8101
49632846851, 8.834101219270615, 8.838888475434473, 8.844409382129601, 8.8480
92420878697, 8.907686431505912, 8.913274484106966, 8.922919962508429, 8.9234
38418678161, 8.927952194457902, 8.9284587485432, 8.934806658737978, 8.935553
834861944, 8.937212456374269, 8.955005336662701, 8.95546814379561, 8.9559995
92244748, 8.956511266025984, 8.957222693733105, 8.967244833508326, 8.9717123
45221071, 8.974420273674525, 8.98550390105968, 8.998697790237316, 9.01081116
030732, 9.019888779366168, 9.0229712693238, 9.03390900417226, 9.039838413773
785, 9.041437427062192, 9.050256185647752, 9.051454589625331, 9.052565364062
843, 9.0548815297122, 9.057829178200953, 9.064550084430042, 9.07910440802783
1, 9.080966119863657, 9.082406325936663, 9.087249610854515, 9.0883537456854
4, 9.096538168999857, 9.09694264105465, 9.097912246653983, 9.10649971463374,
9.111599440460736, 9.116698641556292, 9.117274951500107, 9.139725203295887,
9.142792135839494, 9.15479806651809, 9.158828240328102, 9.19241001366716, 9.
201788883734068, 9.203793708924213, 9.219502314653356, 9.228101303519892, 9.
228454977357703, 9.231188755389727, 9.24079915250678, 9.25457807408743, 9.27
4473229241824, 9.28049506954653, 9.305503467550064, 9.315644464895314, 9.318
967087024134, 9.356215462997358, 9.384949781705773, 9.409233412925444, 9.432
469501518325, 9.432779487995512, 9.4444497591793, 9.481417148378462, 9.49859
4577834067, 9.52138458162866, 9.555337689056811, 9.557906558129424, 9.559002
490055828, 9.560452800946416, 9.571995459815245, 9.57224270576614, 9.5778331
11985443, 9.583377237328042, 9.583468176388251, 9.597000965003607, 9.6073613
31374817, 9.609291088429966, 9.611474706394816, 9.614332626554049, 9.6152266
7570841, 9.64155852657833, 9.649800094197294, 9.651602783454353, 9.674773066
52438, 9.68044253027058, 9.72370440416126, 9.74873938159442, 9.7523918754495
8, 9.7569268428995, 9.760594232586383, 9.77158094073409, 9.77221503827845,
9.78124899593637, 9.784812744389216, 9.796202557006763, 9.806284122801909,
9.813433032287628, 9.815546804479506, 9.84144836811143, 9.84518078209754, 9.
862022628073817, 9.873230424152885, 9.888270274722366, 9.898745845994643, 9.
91723197821882, 9.925040046738424, 9.954942776985712, 9.977196466514544, 10.
015650312816572, 10.018387956090612, 10.0219138368207, 10.04183230991509, 1
0.049681846604429, 10.05122604044341, 10.059649258915176, 10.08634762334345
7, 10.129720753215086, 10.133353318882213, 10.144362935172357, 10.1633374842
9152, 10.166192485107484, 10.177626407668374, 10.179590370133926, 10.1879020
23209729, 10.192421739943038, 10.198050280012051, 10.208826303630284, 10.213
011986006585, 10.227106933736884, 10.237490084854958, 10.248870387413811, 1
0.248890963699877, 10.250839248270076, 10.25143073101396, 10.27565165004322,
10.287075326393005, 10.289319782726094, 10.291449809637324, 10.3067951160934
2, 10.312264389397544, 10.323510093500047, 10.329481013418851, 10.3320390423
89903, 10.35574780991913, 10.356670625754983, 10.365423823015815, 10.3686620
65356457, 10.380607518500844, 10.384388260322474, 10.386265446901014, 10.388
829155692267, 10.409458914120272, 10.411970039506372, 10.412612496740762, 1
0.419154025244143, 10.420243658440942, 10.431398564046004, 10.43339077695014
6, 10.43458216093457, 10.437067298551682, 10.43895966772318, 10.441727048754
778, 10.445719347635212, 10.447977470466425, 10.455948038450916, 10.45787615
5064604, 10.47560484388612, 10.478542539495615, 10.483664998010012, 10.49127
1856948956, 10.49350595993518, 10.496367440168052, 10.502796094276553, 10.51
1146112300649, 10.514524851249632, 10.544073789736206, 10.552319472636771, 1
0.552623554822688, 10.5609727059528, 10.568178393626164, 10.579488389853427,
10.597910927575288, 10.60199318181202, 10.606087663464312, 10.62198501986057
4, 10.632983115140672, 10.64501765603162, 10.649633974528848, 10.65517075213
6044, 10.660356618833836, 10.73815207330481, 10.75498396229191, 10.759826762
923698, 10.790061448549832, 10.794913908400511, 10.797808651857302, 10.81568
5715522406, 10.83602130712966, 10.836196099116954, 10.84721755152455, 10.855

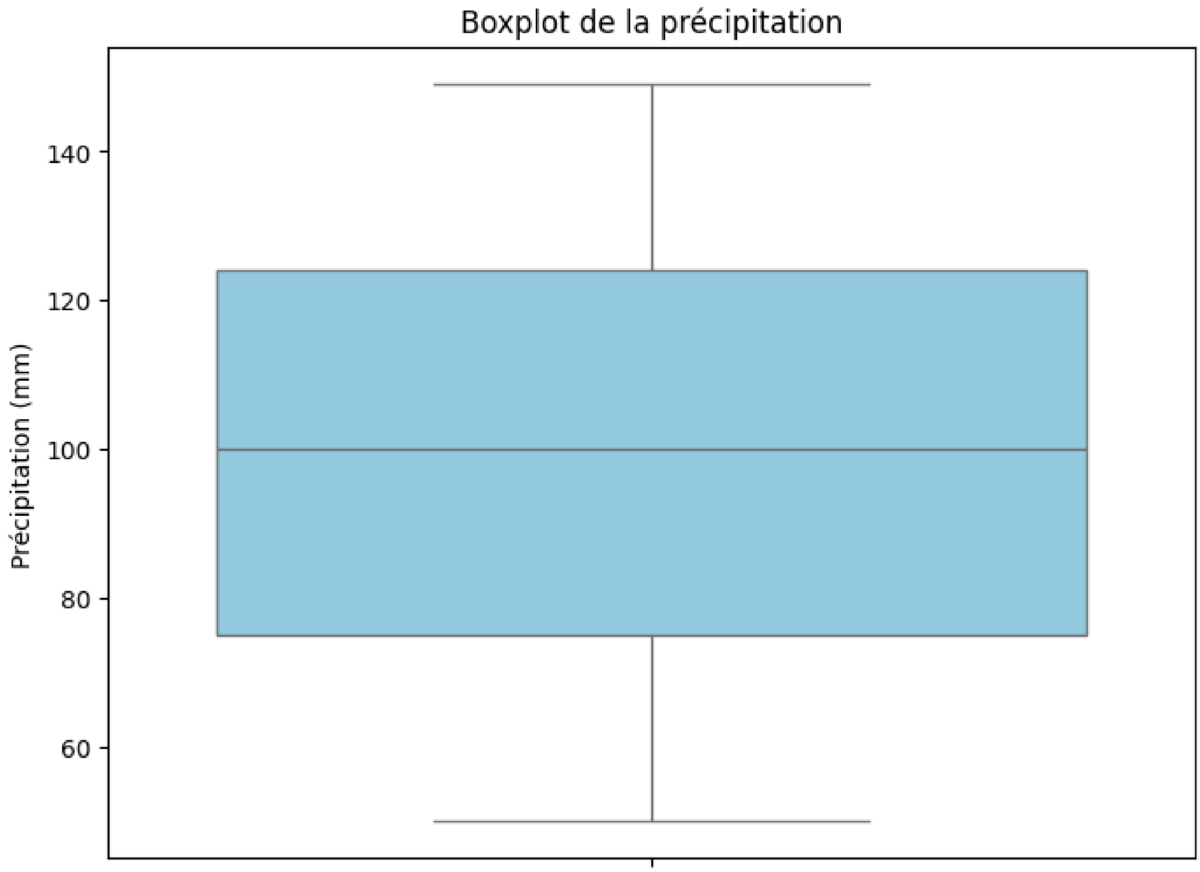
73479409098, 10.862566758634935, 10.866645959746666, 10.868766007233726, 10.
876202567590369, 10.902301615600113, 10.913971941106263, 10.91641347040586,
10.917771620649091, 10.949462412836011, 10.961281994066525, 10.9636960592640
38, 10.974774979321648, 10.975541646503938, 10.975635949164106, 11.005673799
393868, 11.009220387265762, 11.013458887697396, 11.015926099156582, 11.06006
5633186548, 11.061209498399547, 11.069171926221092, 11.076471457173431, 11.0
81417640993632, 11.081681821630095, 11.085478199256317, 11.102377000818633,
11.121629484996072, 11.12171729957493, 11.138189622668088, 11.13923797494471
7, 11.163377643735195, 11.172166908794962, 11.175910966914378, 11.1825773764
82271, 11.195451454929795, 11.202052710707676, 11.217373202008703, 11.235546
592655002, 11.247784900788597, 11.255412599199126, 11.26557787873402, 11.286
656853083718, 11.30190855290512, 11.338174181605664, 11.344114036780043, 11.
357675059038549, 11.35973594861009, 11.362302454905189, 11.37358578152393, 1
1.38399422878533, 11.385663629848574, 11.392993905773809, 11.39477645154282,
11.398845789833755, 11.40374480701323, 11.408737027507412, 11.40989734878097
1, 11.412569239394708, 11.419360621747462, 11.423435234237036, 11.4448415578
61844, 11.446339423927393, 11.446768051258172, 11.456117301226492, 11.464535
575914624, 11.46792152603284, 11.473268137554587, 11.513321796288778, 11.513
738008610394, 11.520806202230167, 11.52445316610223, 11.540826003423012, 11.
543457956210553, 11.55479538858968, 11.563003430967829, 11.591093359729296,
11.610483905004148, 11.61868481645329, 11.621512739335238, 11.62306208027490
4, 11.638119183643653, 11.65343092116701, 11.67024288401434, 11.671319527268
947, 11.6877245126736, 11.691724888377728, 11.7126874330566, 11.723896310557
61, 11.727512695181794, 11.729338728106718, 11.738462246441568, 11.740358166
314572, 11.764203720214343, 11.772215364436963, 11.795519393098026, 11.80398
6016739024, 11.80519308351171, 11.81236399590554, 11.825114043546511, 11.842
691450304349, 11.852104858622967, 11.886025427215912, 11.89578894567133, 11.
905060178685508, 11.905548000458902, 11.910979614781793, 11.912990734142834,
11.92379498244648, 11.927664725147787, 11.93263158145774, 11.93323615938987
3, 11.960120104416193, 11.968736779187571, 11.972544756705028, 11.9797445010
91805, 11.996019329253947] T/HA
Écart-type du rendement : 2.569990985326707 T/HA
Variance du rendement : 6.6048536646605385 (T/HA)²
Étendue du rendement : 8.995742859645505 T/HA

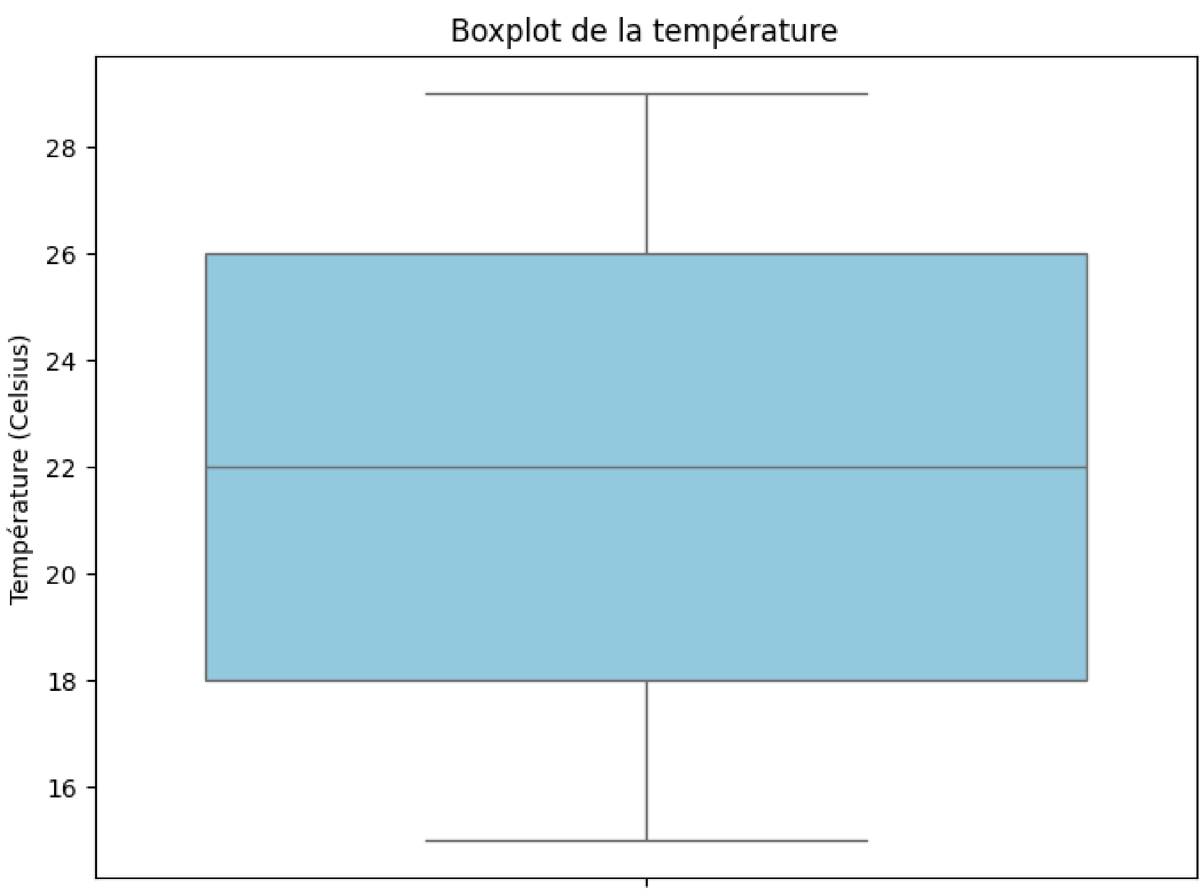

Distribution du rendement
120
100
80
60
40
20
0
4
6
8
10
12


Distribution de la précipitation
0
20
40
60
80
100
60
80
100
120
140


Distribution de la température
0
20
40
60
80
100
120
140
16
18
20
22
24
26
28

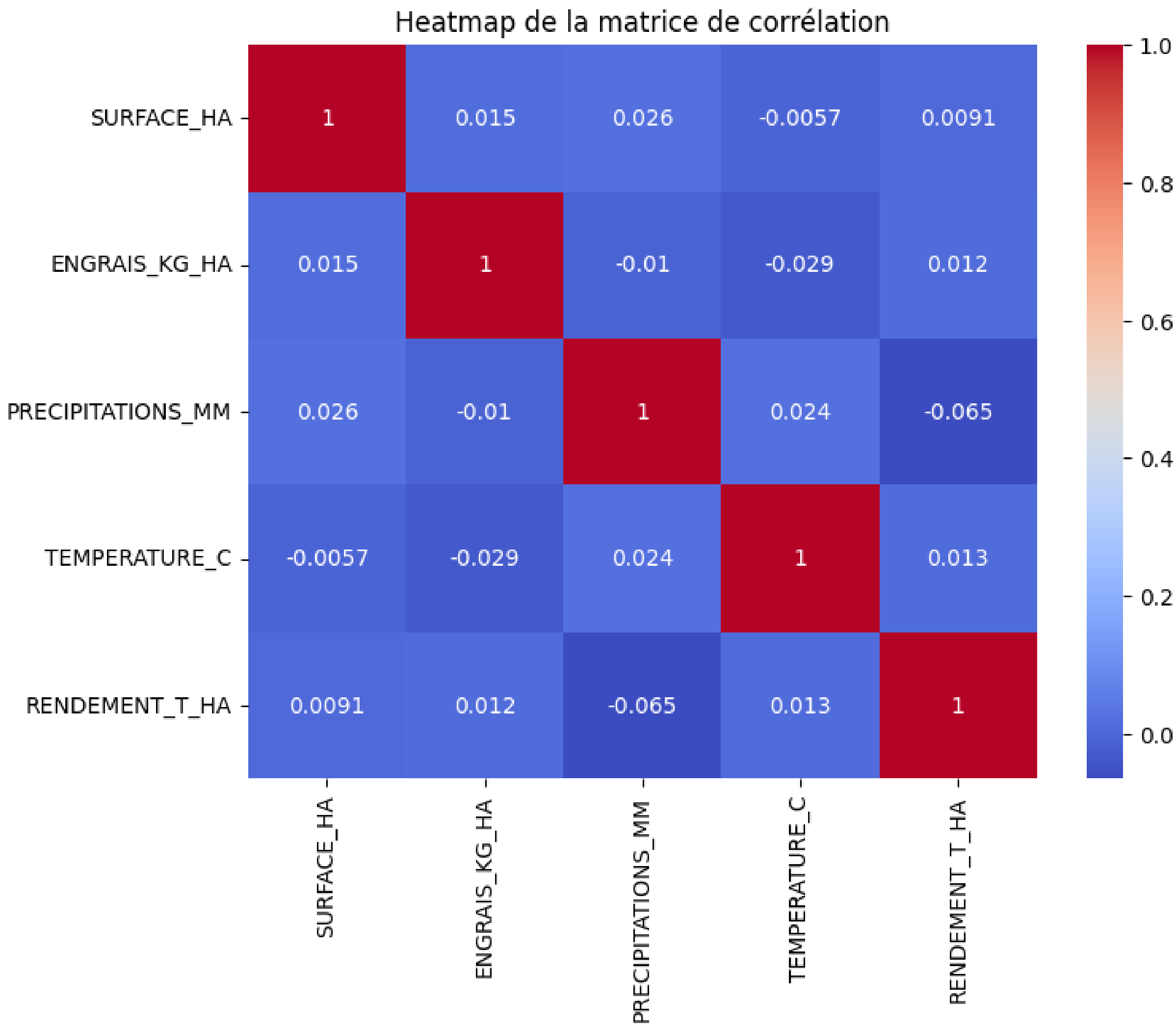

Boxplot du rendement
Rendement (T/HA)
4
6
8
10
12


Boxplot de la précipitation
Précipitation (mm)
140
120
100
80
60


Boxplot de la température
Température (Celsius)
28
26
24
22
20
18
16

```
                   SURFACE_HA  ENGRAIS_KG_HA  PRECIPITATIONS_MM  \
SURFACE_HA           1.000000       0.014786           0.025900   
ENGRAIS_KG_HA        0.014786       1.000000          -0.010348   
PRECIPITATIONS_MM    0.025900      -0.010348           1.000000   
TEMPERATURE_C       -0.005717      -0.028524           0.023692   
RENDEMENT_T_HA       0.009141       0.012368          -0.065479   

                   TEMPERATURE_C  RENDEMENT_T_HA  
SURFACE_HA             -0.005717        0.009141  
ENGRAIS_KG_HA          -0.028524        0.012368  
PRECIPITATIONS_MM       0.023692       -0.065479  
TEMPERATURE_C           1.000000        0.013166  
RENDEMENT_T_HA          0.013166        1.000000  
```



D'après la matrice de corrélation, toutes les variables ont une corrélation nulle ou très légerement positive ou négative, ce qui montre qu'aucune variable ne semble avoir d'impact sur le rendement. La variable qui semblent avoir le plus d'impact sur le rendement est la température (0,013).

La diagonale vaut 1 car une variable x aura forcément une corrélation parfaite avec elle même. Dans la matrice, les valeurs :

- qui tendent vers 1 montre que si la variable x augmente, y augmente.

- qui tendent vers 0 montre que si la variable x augmente, y ne varie pas.
- qui tendent vers -1 montre que si la variable x augmente, y diminue.

Etape 3

In [4]:
```
import statsmodels.api as sm
from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm

formule = 'RENDEMENT_T_HA ~ C(TYPE_SOL)'
model = ols(formule, data=df).fit()
tableAnova = anova_lm(model)

print(tableAnova)
```

```
                 df       sum_sq   mean_sq         F    PR(>F)
C(TYPE_SOL)     2.0    17.900287  8.950143  1.356052  0.258151
Residual      997.0  6580.348524  6.600149       NaN       NaN
```

Selon les résultats obtenus, on peut voir que F est inférieur à 19 (table de Fisher). De plus, la P value est également supérieure à 0,05. On ne rejette donc pas l'hypothèse nulle. On peut donc dire que le type de sol n'a pas d'influence significative sur le rendement. (ce qui est totalement différent avec les calculs effectués à la main).

Etape 4

In [5]:
```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import GradientBoostingRegressor, RandomForestRegresso
from sklearn.neighbors import KNeighborsRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_scor
import numpy as np

df = pd.read_csv('rendement_mais.csv')

# Séparation des données
df = pd.get_dummies(df, columns=['TYPE_SOL'], drop_first=True) # Utilisation

X = df.drop('RENDEMENT_T_HA', axis=1)
y = df['RENDEMENT_T_HA']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, rar

# Création des modèles : Regression Linéaire, Random Forest et les K plus pr

modele_lineaire = LinearRegression()
modele_gradient_boosting = GradientBoostingRegressor(random_state=42)
modele_knn = KNeighborsRegressor(n_neighbors=4)
modele_rf = RandomForestRegressor(random_state=42)

modele_lineaire.fit(X_train, y_train)
modele_gradient_boosting.fit(X_train, y_train)
```

```
modele_knn.fit(X_train, y_train)
modele_rf.fit(X_train,y_train)

predictions_lineaire = modele_lineaire.predict(X_test)
predictions_gradient_boosting = modele_gradient_boosting.predict(X_test)
predictions_knn = modele_knn.predict(X_test)
predictions_rf = modele_rf.predict(X_test)

mae_lineaire = mean_absolute_error(y_test, predictions_lineaire)
mae_gradient_boosting = mean_absolute_error(y_test, predictions_gradient_boo
mae_knn = mean_absolute_error(y_test, predictions_knn)
mae_random_forest = mean_absolute_error(y_test, predictions_rf)

rmse_lineaire = np.sqrt(mean_squared_error(y_test, predictions_lineaire))
rmse_gradient_boosting = np.sqrt(mean_squared_error(y_test, predictions_grac
rmse_knn = np.sqrt(mean_squared_error(y_test, predictions_knn))
rmse_random_forest = np.sqrt(mean_squared_error(y_test, predictions_rf))

r2_gradient_boosting = r2_score(y_test, predictions_gradient_boosting)
r2_lineaire = r2_score(y_test, predictions_lineaire)
r2_knn = r2_score(y_test, predictions_knn)
r2_random_forest = r2_score(y_test, predictions_rf)

print(f"  MAE Linéaire : {mae_lineaire}")
print(f"  RMSE Linéaire : {rmse_lineaire}")
print(f"  R² Linéaire : {r2_lineaire}")

print(f"  MAE Gradient Boosting : {mae_gradient_boosting}")
print(f"  RMSE Gradient Boosting : {rmse_gradient_boosting}")
print(f"  R² Gradient Boosting : {r2_gradient_boosting}")

print(f"  MAE K plus proches voisins : {mae_knn}")
print(f"  RMSE K plus proches voisins : {rmse_knn}")
print(f"  R² K plus proches voisins : {r2_knn}")

print(f"  MAE K Random Forest : {mae_random_forest}")
print(f"  RMSE K Random Forest : {rmse_random_forest}")
print(f"  R² K Random Forest : {r2_random_forest}")
```

```
 MAE Linéaire : 2.0959068723189698
 RMSE Linéaire : 2.4621318034215456
 R² Linéaire : -0.02795424759225562
 MAE Gradient Boosting : 2.1368719722291654
 RMSE Gradient Boosting : 2.552499419169278
 R² Gradient Boosting : -0.10479701797140728
 MAE K plus proches voisins : 2.176323224530716
 RMSE K plus proches voisins : 2.627464495700877
 R² K plus proches voisins : -0.1706441573519648
 MAE K Random Forest : 2.062098979473595
 RMSE K Random Forest : 2.499626906235486
 R² K Random Forest : -0.0595014882624707
```

Random Forest semble être le modèle le plus performant car son R^2 est le plus haut. Je suppose qu'il est plus performant car il est plus adaptés aux variables non linéaire, De plus, il réduit le surapprentissage.

Etape 5

In [6]:
```
importance_variables = modele_rf.feature_importances_

for i, importance in enumerate(importance_variables):
    print(f"{X.columns[i]}: {importance}")
```

```
SURFACE_HA: 0.15330184381275322
ENGRAIS_KG_HA: 0.31739862503925254
PRECIPITATIONS_MM: 0.2895961305268603
TEMPERATURE_C: 0.17944367059984936
TYPE_SOL_Limoneux: 0.03148171371912523
TYPE_SOL_Sableux: 0.028778016302159325
```

D'après les résultats obtenus, on peut remarquer l'importance de la valeur des variables malgré le fait qu'il ne nous permettent pas de savoir exactement quel est la variable qui pourrait nous permettre d'avoir un rendement plus élevé. Il est simplement possible que malheureusement aucun de ses modèles ne soient adaptés à ce problème. Cependant, au vu des résultats obtenus, on peut supposer que les variables les plus importantes serait les engrais, la précipitation ainsi que la température.

Je recommanderais donc concrètement d'utiliser des des engrais adaptés ainsi que la bonne quantité d'engrais pour le maïs, il faudrait également hydrater le sol au maximum par des arrosages automatiques réguliers ou mettre en place un système d'irrigation et également adapter la culture des maïs et donc choisir un type de maïs plus adapté à la température ou alors les cultiver dans de grandes serres afin de mieux gérer la température.

Pour améliorer le modèle il faudrait utiliser des modèles plus complexes capable de gérer des variables ayant peu de corrélation entre elles.

Pour optimiser sa production, la ferme pourrait faire des tests avec différentes types de maïs, faire de l'irrigation ou utiliser des arroseurs automatiques, tester des engrais.