

ÉNONCÉ DU TP

TP1 – Corrélation et Régression linéaire

BTSA ACD 2 – Données réelles : Blé dur dans le Gard (2000–2018)

1. Contexte

On étudie l'évolution du **rendement du blé dur** dans le département du **Gard** entre 2000 et 2018.

Les données proviennent de séries statistiques agricoles départementales.

2. Données

Année	Rendement (q/ha)	Surface (ha)
2000	41	18 200
2002	39	17 500
2004	45	19 100
2006	43	20 300
2008	47	21 000
2010	44	19 800
2012	46	20 100
2014	49	22 400
2016	51	23 100
2018	48	22 900

Variables :

- X_1 : Année
 - X_2 : Surface (ha)
 - Y : Rendement (q/ha)
-

3. Travail demandé

Partie 1 – Observation graphique

1. Représenter le nuage de points :
 - Y en fonction de l'année
 - Y en fonction de la surface
 2. Décrire les tendances observées.
-

Partie 2 – Corrélation

1. Calculer les coefficients de corrélation :
 - r_{X_1Y}
 - r_{X_2Y}
 2. Interpréter le signe et l'intensité de chaque corrélation.
-

Partie 3 – Régression linéaire (année → rendement)

1. Déterminer l'équation de la droite de régression.
 2. Interpréter la pente et l'ordonnée à l'origine.
 3. Donner le coefficient de détermination R^2 .
 4. Commenter la qualité du modèle.
-

Partie 4 – Prédictions

À l'aide du modèle linéaire, estimer le rendement prévu en :

- 2020
- 2030

Puis commenter la pertinence de ces prédictions.

Partie 5 – Modèle quadratique

1. Ajuster un modèle quadratique.
 2. Comparer son R^2 avec celui du modèle linéaire.
 3. Quel modèle semble le plus pertinent ? Justifier.
-

Partie 6 – Conclusion agronomique

Rédiger un paragraphe de synthèse (10–15 lignes) sur :

- l'évolution des rendements,
 - les facteurs explicatifs possibles,
 - l'intérêt de la corrélation/régression pour l'analyse agronomique.
-

TP 2 – Corrélation et Régression linéaire

BTSA ACD 2 – Données réelles : Élevage bovin viande en Aveyron (2008–2022)

1. Contexte

L'Aveyron est l'un des départements français où l'élevage bovin viande est le plus important.

On s'intéresse ici à la relation entre :

- le **nombre de vaches allaitantes** (mères),
- le **nombre de veaux nés vivants**,
- l'évolution de ces deux variables entre 2008 et 2022.

2. Données

Année	Vaches allaitantes (X)	Veaux nés vivants (Y)
2008	182 000	158 500
2010	185 200	160 300
2012	187 400	161 100
2014	189 100	162 000
2016	191 300	163 400
2018	188 700	161 900
2020	186 900	160 800
2022	184 500	159 200

3. Travail demandé

Partie 1 – Observation graphique

1. Représenter le nuage de points :
 - Y en fonction de X (veaux nés en fonction du nombre de mères).
 2. Décrire la forme du nuage :
 - relation croissante ?
 - linéaire ?
 - présence d'un maximum ?
-

Partie 2 – Corrélation

1. Calculer le coefficient de corrélation r_{XY} .
 2. Interpréter :
 - signe,
 - intensité,
 - cohérence zootechnique.
-

Partie 3 – Régression linéaire

1. Déterminer l'équation de la droite de régression :

$$\widehat{Y} = aX + b$$

2. Interpréter la pente.
 3. Donner le coefficient de détermination R^2 .
 4. Commenter la qualité du modèle.
-

Partie 4 – Prédictions

À l'aide du modèle linéaire, estimer le nombre de veaux nés si le département comptait :

- **190 000 vaches allaitantes,**
- **180 000 vaches allaitantes.**

Puis commenter la pertinence de ces prédictions.

Partie 5 – Modèle quadratique

1. Ajuster un modèle quadratique.
 2. Comparer son R^2 avec celui du modèle linéaire.
 3. Quel modèle semble le plus pertinent ? Justifier.
-

Partie 6 – Conclusion zootechnique

Rédiger un paragraphe (10–15 lignes) sur :

- l'évolution du cheptel,
- l'évolution de la production de veaux,
- les facteurs explicatifs possibles (alimentation, sécheresse, démographie agricole...),
- l'intérêt de la corrélation/régression pour analyser un système d'élevage.

TP 3 – Corrélation, régression simple et multiple

Domaine viticole en Languedoc – Irrigation, rendement et qualité du raisin

1. Contexte

Un domaine viticole du Languedoc conduit une parcelle de **Syrah** en AOP. Depuis plusieurs années, l'exploitant suit :

- le **volume d'irrigation** apporté,
- le **rendement** obtenu,
- la **teneur en sucre** des raisins (°Brix),
- l'**acidité totale** (g/L H₂SO₄),
- la **date de vendange**.

L'objectif est de comprendre comment l'irrigation influence à la fois le **rendement** et la **qualité** des raisins, et de voir s'il existe des compromis.

2. Données

Parcelle de Syrah – Languedoc – 10 millésimes

Année	Irrigation (mm)	Rendement (hL/ha)	Sucre (°Brix)	Acidité (g/L)	Date vendange (jour julien)
2012	0	38	25,0	5,8	270
2013	20	42	24,5	5,6	268
2014	40	48	24,0	5,3	266
2015	30	45	24,8	5,1	267
2016	50	52	23,8	5,0	265
2017	10	40	25,5	5,9	271
2018	60	54	23,5	4,8	264
2019	0	36	26,0	6,1	272
2020	25	44	24,7	5,4	268
2021	45	50	23,9	5,2	266

Variables :

- X_1 : Irrigation (mm)
 - Y_1 : Rendement (hL/ha)
 - Y_2 : Sucre (°Brix)
 - Y_3 : Acidité (g/L)
 - X_2 : Date de vendange (jour julien)
-

3. Partie 1 – Exploration graphique

1. Saisir les données dans un tableur ou R.
 2. Tracer les nuages de points suivants :
 - Y_1 (rendement) en fonction de X_1 (irrigation)
 - Y_2 (sucre) en fonction de X_1
 - Y_3 (acidité) en fonction de X_1
 3. Décrire qualitativement :
 - L'irrigation semble-t-elle **augmenter** le rendement ?
 - L'irrigation semble-t-elle **diluer** le sucre ?
 - L'irrigation semble-t-elle **modifier** l'acidité ?
-

4. Partie 2 – Corrélations

1. Calculer les coefficients de corrélation linéaire :

- r_{X_1, Y_1} (irrigation / rendement)
- r_{X_1, Y_2} (irrigation / sucre)
- r_{X_1, Y_3} (irrigation / acidité)

2. Interpréter pour chaque r :

- signe,
- intensité,
- cohérence agronomique.

3. Calculer aussi :

- r_{Y_1, Y_2} (rendement / sucre)
- r_{Y_1, Y_3} (rendement / acidité)

→ Question : observe-t-on un **compromis** entre rendement et qualité ?

5. Partie 3 – Régression linéaire simple

5.1. Modèle 1 : Rendement en fonction de l'irrigation

1. Ajuster un modèle linéaire :

$$\widehat{Y}_1 = aX_1 + b$$

2. Relever :

- la pente a ,
- l'ordonnée à l'origine b ,
- le coefficient de détermination R^2 .

3. Interpréter :

- De combien de hL/ha le rendement augmente-t-il pour 10 mm d'irrigation en plus ?
- Le modèle explique-t-il bien la variabilité du rendement ?

5.2. Modèle 2 : Sucre en fonction de l'irrigation

1. Ajuster :

$$\widehat{Y}_2 = cX_1 + d$$

2. Interpréter le signe de c :

- L'irrigation augmente-t-elle ou diminue-t-elle la teneur en sucre ?

3. Commenter R^2 .

6. Partie 4 – Régression multiple (niveau plus avancé)

On suppose que le rendement dépend à la fois :

- de l'irrigation X_1 ,
- de la date de vendange X_2 .

On cherche un modèle du type :

$$\widehat{Y}_1 = \alpha X_1 + \beta X_2 + \gamma$$

1. Ajuster ce modèle avec un logiciel (R ou tableur avec régression multiple).
 2. Relever :
 - α, β, γ ,
 - le nouveau R^2 .
 3. Comparer ce R^2 avec celui du modèle simple (rendement ~ irrigation).
 4. Interpréter :
 - Effet de l'irrigation à date de vendange fixée,
 - Effet de la date de vendange à irrigation fixée.
-

7. Partie 5 – Modèle non linéaire (optionnel mais intéressant)

En observant le nuage de points **rendement / irrigation**, on peut suspecter une **saturation** :

→ au-delà d'un certain niveau d'irrigation, le gain de rendement devient faible.

1. Ajuster un modèle **quadratique** :

$$\widehat{Y}_1 = aX_1^2 + bX_1 + c$$

2. Comparer son R^2 avec le modèle linéaire.
 3. Discuter :
 - Le modèle quadratique est-il plus cohérent agronomiquement ?
 - Met-il en évidence un **optimum d'irrigation** ?
-

8. Partie 6 – Synthèse agronomique

Rédiger une synthèse (15–20 lignes) répondant aux questions :

1. Quel est l'effet global de l'irrigation sur :
 - le rendement,
 - la teneur en sucre,
 - l'acidité ?
 2. Observe-t-on un **compromis rendement / qualité** ?
 3. L'ajout de la date de vendange dans le modèle améliore-t-il la compréhension du système ?
 4. Quel type de modèle (linéaire, quadratique, multiple) te semble le plus pertinent pour :
 - piloter l'irrigation,
 - décider de la date de vendange ?
 5. Comment un viticulteur pourrait-il utiliser ce type d'analyse pour :
 - adapter sa stratégie d'irrigation,
 - viser un certain niveau de qualité (vin de garde vs vin plus productif) ?
-