



Licence d'Informatique 2ème année

AD Utilisateur (C5-160412-INFO)

CC du 13 février 2020 (1h30)

Carl FRÉLICOT – Dpt Info / Lab MIA

- Autorisés : calculatrice (pas une application sur smartphone), un formulaire **manuscrit** sans **aucun exemple numérique**. Tout **autre document est interdit**.
- Utilisez les cadres réservés pour inscrire vos réponses.
Vous devez **écrire la formule** utilisée avant de donner le détail du moindre calcul.
- Une réponse **non justifiée** sera considérée comme **fausse**.

NOM (en **capitales**), Prénom :

Note : /20

Le tableau ci-contre donne les notes cumulées attribuées par 10 usagers à 8 smartphones pour les caractéristiques suivantes :

- sonnerie (Bell) notée de 1 à 3 (faible, moyen, fort)
- qualité Sonore (Sound) de 1 (exécration) à 10 (excellent)
- qualité de Transmission (Trans.) de 1 à 10
- autonomie (Auto.) de 1 à 10
- facilité d'Utilisation (Use) de 1 à 10

On donne aussi les moyennes et variances de chaque variable :

\bar{x}	18.5	60.5	66.5	67	68
s^2	4.75	14.5	127.5	323.5	32.5

ID.	Bell	Sound	Trans.	Auto.	Use	Avis	Supply
#1	21.0	67.0	88.0	75.0	60.0	1 étoile	AA
#2	19.0	58.0	61.0	40.0	70.0	5 étoiles	BB
#3	20.0	62.0	78.0	69.0	68.0	3 étoiles	AA
#4	14.0	58.0	54.0	86.0	62.0	1 étoile	CC
#5	20.0	66.0	70.0	63.0	79.0	5 étoiles	AA
#6	17.0	56.0	56.0	93.0	67.0	1 étoile	CC
#7	17.0	58.0	56.0	40.0	73.0	5 étoiles	BB
#8	20.0	59.0	69.0	70.0	65.0	3 étoiles	BB

On dispose aussi pour chaque smartphone de l'Avis d'un site web de test et d'un code du fabricant (Supply).

Exercice 1

- 1) Que vaut, pour la variable Bell, l'indicateur de *dispersion* retourné par le widget **Feature Statistics** d'Orange ?

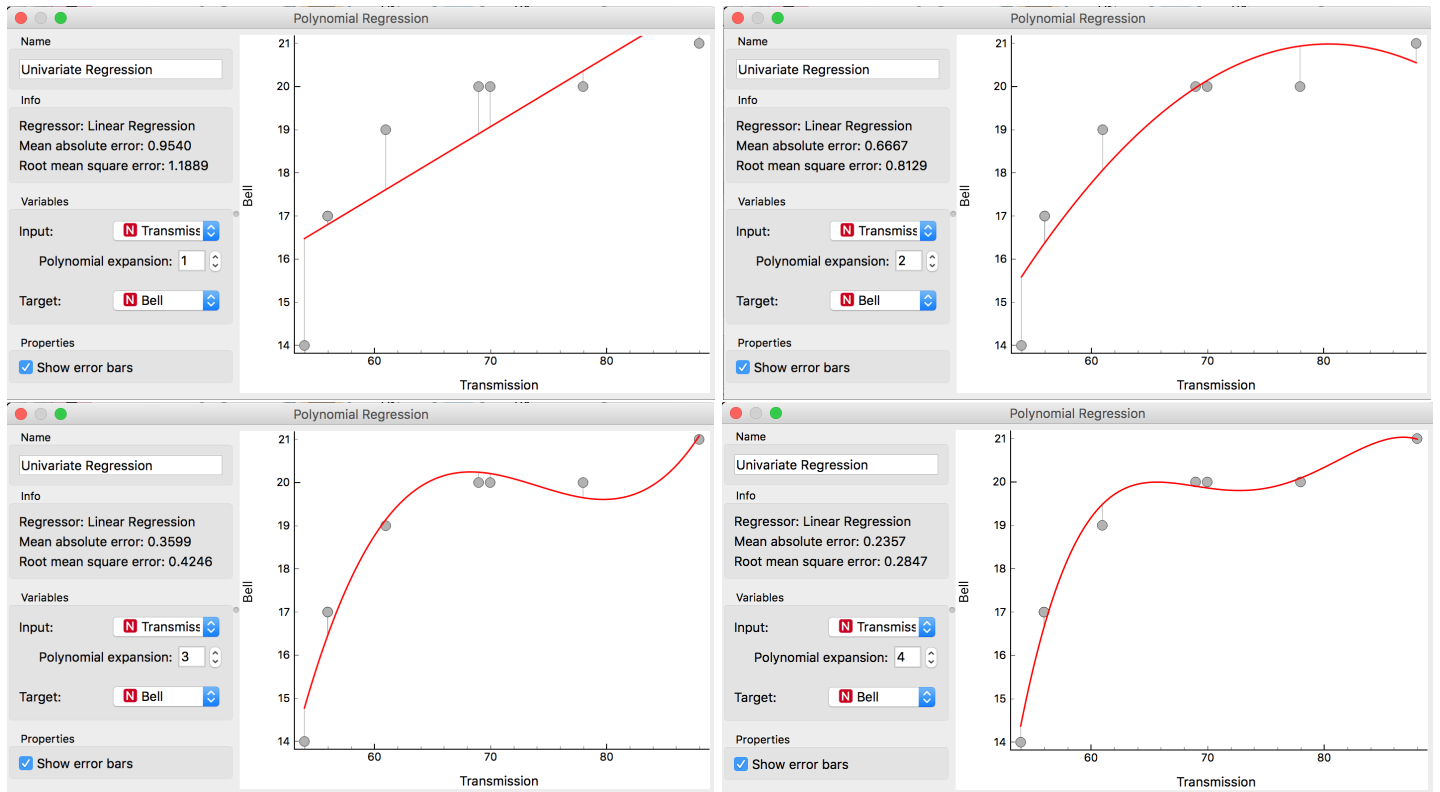
- 2) Quelqu'un qui n'a pas suivi l'EC regarde les statistiques données et a la drôle d'idée de vouloir tracer le boxplot de ces 10 nombres. Donnez à gauche les éléments de calcul et tracez verticalement le boxplot à droite.

Exercice 2 :

On cherche à expliquer les valeurs de Bell (B) par celles de Trans. (T), et on donne : $\sum_k B_k \times T_k = 10\,007$.

- 1) Calculez le coefficient de corrélation entre les deux variables. Commentez le résultat en termes de possibilité d'ajustement par un modèle.

2) Les modèles ci-dessous sont proposés. Lequel retiendriez-vous ?



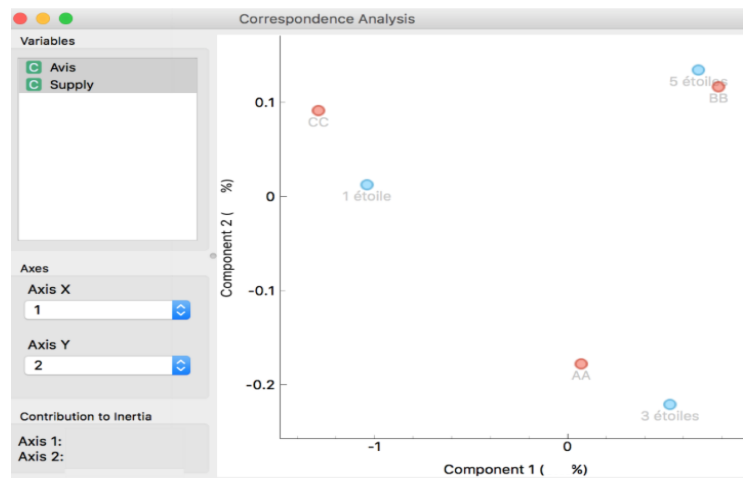
3) Les résidus d'un modèle sont : 2.5, 0.5, 1.5, -4.5, 1.5, -1.5, -1.5, 1.5 ; calculez le *Root mean square error*.

Exercice 3 :

Voici des sorties Orange.

Contingency Table					
Rows		Supply			
Avis		AA	BB	CC	Σ
1 étoile		1	0	2	3
3 étoiles		1	1	0	2
5 étoiles		1	2	0	3
	Σ	3	3	2	8

ARI: 0.048
AMI: 0.073
 $\chi^2=5.33$, $p=0.255$



- 1) Calculez la contribution du couple de modalités (5 étoiles, BB) à la valeur du $\chi^2 = 5.33$

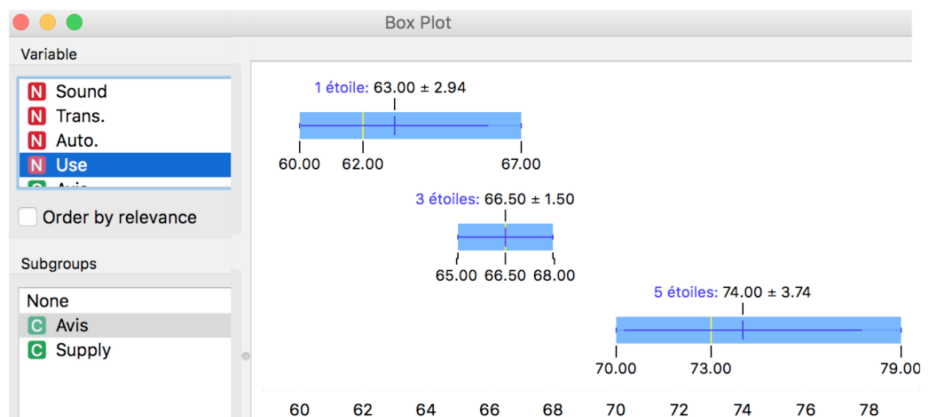
- 2) Peut-on affirmer que Avis et Supply sont indépendants ?

- 3) Donnez quelques éléments pertinents d'analyse des correspondances, et des % crédibles pour les deux composantes.

Exercice 4 :

Voici d'autres sorties Orange. 3

Rank		
Scoring Methods		ANOVA
<input type="checkbox"/> Information Gain	<input type="checkbox"/> Use	
<input type="checkbox"/> Information Gain Ratio	<input type="checkbox"/> Auto.	9.994
<input type="checkbox"/> Gini Decrease	<input type="checkbox"/> Sound	0.004
<input checked="" type="checkbox"/> ANOVA	<input type="checkbox"/> Bell	0.739
<input type="checkbox"/> χ^2	<input type="checkbox"/> Trans.	0.434
<input type="checkbox"/> ReliefF		
<input type="checkbox"/> FCBF		

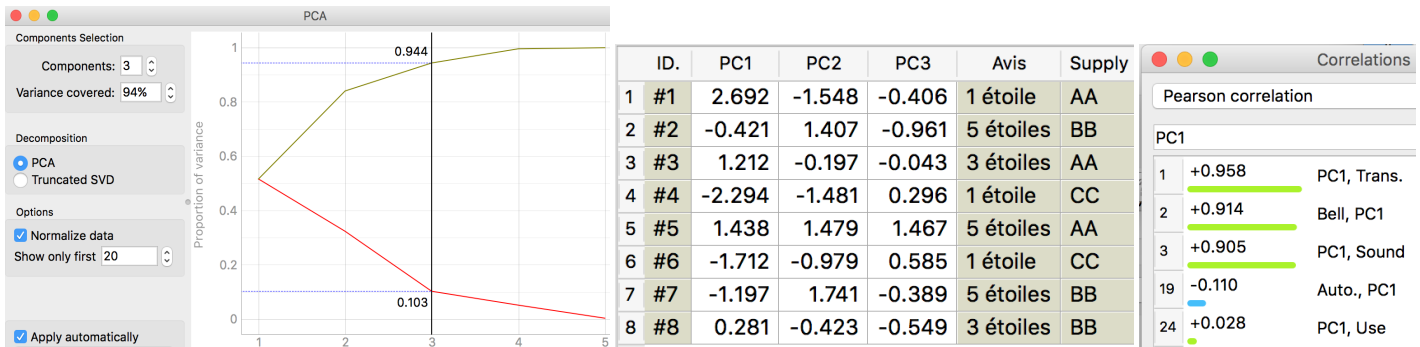


- 1) Calculez la statistique manquante pour la variable Use.

2) Quelle variable parmi est la plus pertinente pour différencier les Avis ?

Exercice 5 :

Une analyse en composantes principales a été réalisée, dont voici les sorties Orange :



1) Auriez-vous fait les mêmes choix que l'analyste ?

- 2) Dessinez le 1er cercle des corrélations, ci-contre.
- 3) Quels que soient vos (dés)accords avec les choix de l'analyste, donnez des éléments d'interprétation des 2 premières composantes afin de leur donner sens.

