

Rédigé par : l'équipe enseignante	Durée : 1h30
A l'intention de : Elèves d'ING1	Document ou matériel autorisés : Calculatrice de l'EISTI uniquement

Jeu de données

Le jeu de données est constitué des 47 plus grands succès du cinéma français ou américains, en France depuis 1945 jusqu'à 2015.

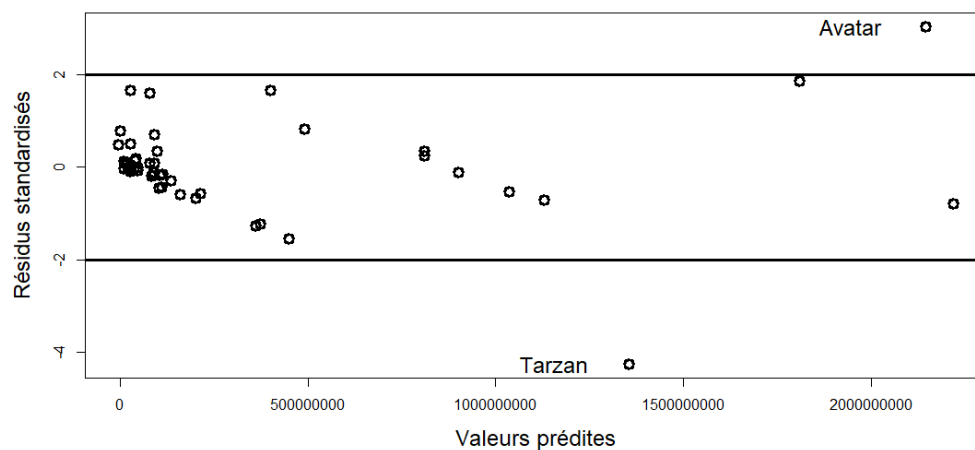
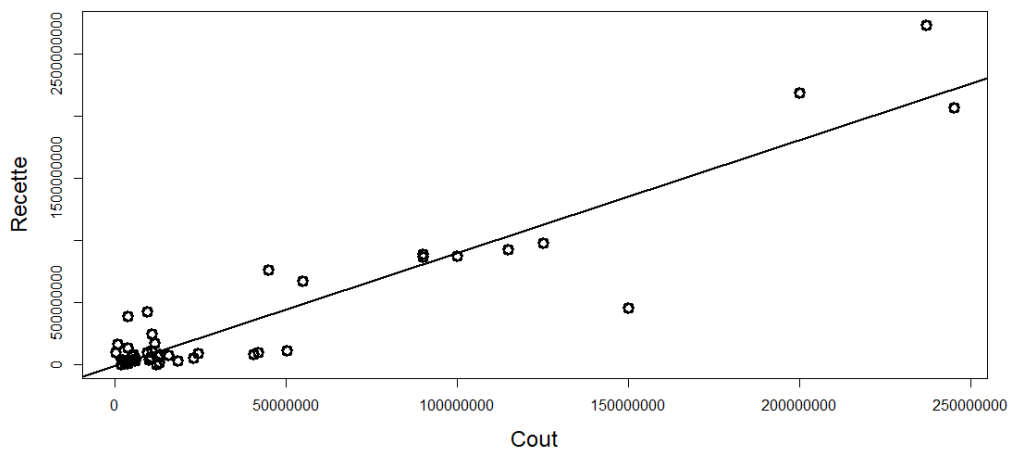
Les films sont caractérisés par

- Le genre (drame, action,...)
- Le cout de production
- La recette
- La durée
- Le nombre d'entrées
- Le nombre de prix remportés

Exercice 1

Nous allons construire un modèle permettant d'évaluer les recettes d'un film en fonction de l'investissement.

- 1) Quelle est l'équation du model ?
- 2) Pour un investissement du 160000000 USD, quelle recette peut-on espérer récolter ?
- 3) A quoi correspond le « Multiple R-squared » ?
- 4) Ce modèle vous paraît-il valide ?



```
> summary(model)
```

```
Call:
lm(formula = Recette ~ Cout, data = tab)

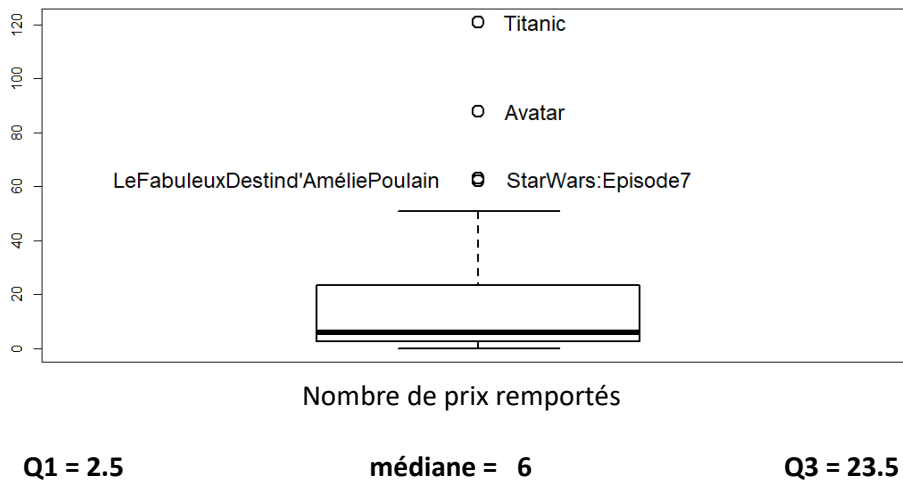
Residuals:
    Min       1Q   Median       3Q      Max
-899729597 -79424089  -8165021  65015174  584516608

Coefficients:
            Estimate      Std. Error t value      Pr(>|t|)
(Intercept) -8016997.8306  38586172.4751  -0.208      0.836
Cout          9.0910       0.5309   17.123 <0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 221700000 on 45 degrees of freedom
Multiple R-squared:  0.8669,    Adjusted R-squared:  0.864
F-statistic: 293.2 on 1 and 45 DF,  p-value: < 0.00000000000000022
```

Exercice 2

Nous allons maintenant nous intéresser aux récompenses remportées par les films.

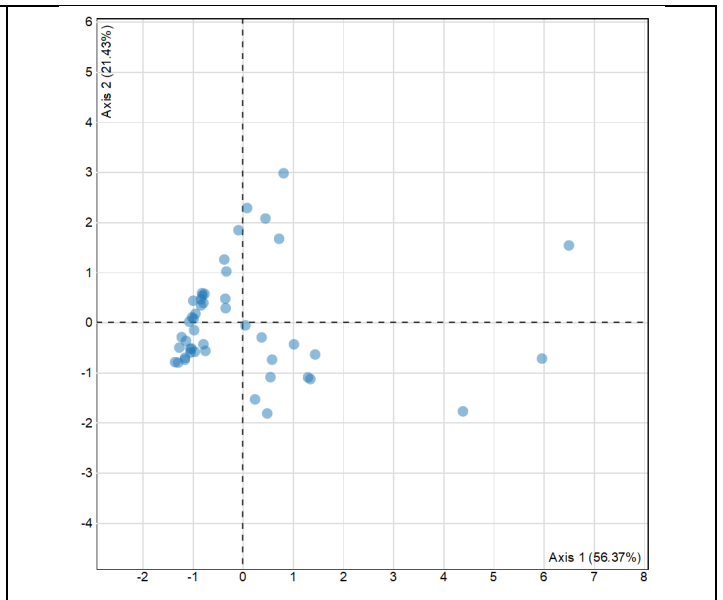
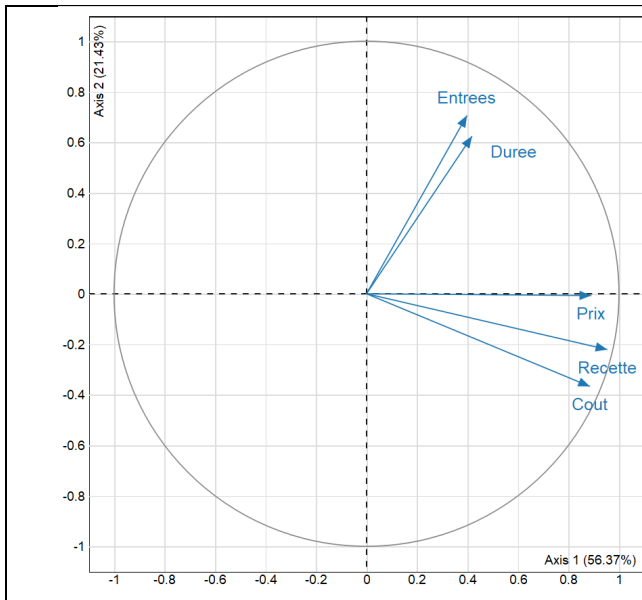


- 1) Expliquez pourquoi Q1 et Q3 sont des réels et non des entiers.
- 2) Exprimez dans une phrase Q1.
- 3) Les films cités sur le graphique ont reçu plus de ??? prix. Calculez ce nombre à partir des quantiles.
- 4) Que pouvez-vous dire sur le nombre de prix moyen.

Exercice 3

Une analyse en composantes principales sur les variables quantitatives donne les résultats ci-après. L'axe 1 explique 56.37% de l'inertie du nuage de points et l'axe 2 explique 21.43%.

- 1) Combien y-a-t-il d'axes en tout ?
- 2) Quelles variables contribuent le plus à la constitution de l'axe 1 ? de l'axe 2 ?
- 3) Sur quel(s) axe(s) peut-on interpréter le film Avatar ? Comment peut-on le qualifier ?
- 4) Sur quel(s) axe(s) peut-on interpréter le film « Autant en emporte le vent » ? Comment peut-on le qualifier ?
- 5) Quelle est la contribution moyenne d'un film ? Y-a-t-il des films qui ne respectent pas cette valeur moyenne ? Quel impact cela peut-il avoir ?



Résultats concernant les variables sur les 3^{ers} axes

\$cos2	Dim.1			Dim.2			Dim.3		
	Dim.1	Dim.2	Dim.3	Dim.1	Dim.2	Dim.3	Dim.1	Dim.2	Dim.3
Entrees	0,159	0,499	0,330	5,63	46,54	42,96	27,78	12,58	0,20
Cout	0,783	0,135	0,002	32,26	4,57	0,24	6,20	36,31	56,34
Recette	0,909	0,049	0,002	28,12	0,00	0,27			
Duree	0,175	0,389	0,433						
Prix	0,793	0,000	0,002						

Résultats concernant les individus sur les 3^{ers} axes

	\$scoord			\$cos2			\$scontrib		
	Dim.1	Dim.2	Dim.3	Dim.1	Dim.2	Dim.3	Dim.1	Dim.2	Dim.3
LeDictateur	-0,97	-0,16	0,56	0,667	0,018	0,225	0,709	0,050	0,878
LeGrandBleu	-0,83	0,34	0,91	0,420	0,069	0,503	0,522	0,226	2,290
LaGrandeEvasion	-0,81	0,58	1,34	0,233	0,120	0,643	0,494	0,672	5,005
Les101Dalmatiens	-0,99	0,43	-1,45	0,293	0,056	0,631	0,734	0,369	5,811
Qu'est-cequ'onafaitaubondieu?	-0,94	0,17	-0,67	0,625	0,020	0,321	0,670	0,057	1,261
Pourquisonneleglas?	-0,83	0,46	1,40	0,242	0,074	0,680	0,524	0,420	5,405
Vingtmillielieussouslesmers	-1,01	0,10	0,38	0,865	0,008	0,120	0,770	0,019	0,392
LesCanonsdeNavarone	-0,80	0,53	0,68	0,460	0,201	0,327	0,488	0,562	1,275
L'Ours	-1,02	-0,52	-0,08	0,780	0,205	0,005	0,788	0,546	0,019
Cendrillon	-0,98	0,07	-1,28	0,368	0,002	0,629	0,722	0,010	4,531
LaBelleetleclochard	-1,22	-0,30	-0,85	0,647	0,038	0,313	1,120	0,174	1,986
LesAristochats	-1,06	0,01	-1,09	0,486	0,000	0,511	0,855	0,000	3,306
Pinocchio	-1,29	-0,81	0,01	0,688	0,269	0,000	1,255	1,289	0,000
LeJourlepluslong	-0,37	1,25	0,83	0,055	0,645	0,287	0,102	3,118	1,931
WestSideStory	-0,34	0,28	0,95	0,067	0,046	0,528	0,087	0,157	2,497
E.T.l'extra-terrestre	0,05	-0,06	0,16	0,001	0,001	0,010	0,002	0,007	0,068
LesCouloirsdutemps	-1,15	-0,75	0,18	0,689	0,289	0,016	1,004	1,108	0,087
Taxi2	-1,13	-0,37	-0,45	0,786	0,085	0,123	0,966	0,274	0,553
LeDînerdecons	-1,15	-0,71	-0,39	0,663	0,252	0,078	0,998	0,996	0,432
UnIndiendanslaville	-1,35	-0,79	0,05	0,731	0,254	0,001	1,373	1,254	0,007
BenHur	0,46	2,07	1,03	0,037	0,756	0,189	0,156	8,526	2,968
LesDixCommandements	0,09	2,28	1,14	0,001	0,770	0,194	0,006	10,39	3,626
LesChoristes	-0,96	-0,59	0,03	0,616	0,234	0,001	0,689	0,688	0,003
LesBronzés3	-0,78	-0,44	-0,25	0,602	0,189	0,063	0,463	0,383	0,178
Rienàdéclarer	-1,04	-0,61	0,33	0,683	0,231	0,071	0,817	0,728	0,310

LesVisiteurs	-0,76	0,57	-0,76	0,366	0,202	0,360	0,439	0,637	1,585
Emmanuelle	-1,26	-0,51	-0,06	0,834	0,134	0,002	1,204	0,507	0,010
AstérixetObélixcontreCésar	-0,74	-0,57	0,23	0,545	0,322	0,054	0,417	0,647	0,152
AstérixetObélix:missionCléopâtre	-0,35	0,47	-0,87	0,069	0,128	0,434	0,091	0,439	2,084
DocteurJivago	-0,33	1,01	1,50	0,031	0,303	0,662	0,080	2,045	6,233
LeLivredelajungle	-0,78	0,39	-1,50	0,201	0,049	0,738	0,463	0,297	6,231
Bambi	-1,05	-0,53	-0,88	0,490	0,126	0,343	0,827	0,559	2,127
LeFabuleuxDestind'AméliePoulain	0,38	-0,30	0,36	0,034	0,022	0,032	0,108	0,181	0,367
BienvenuechezlesCh'tis	-0,08	1,84	-2,09	0,001	0,410	0,529	0,005	6,708	12,06
Autantenemportelevé	0,82	2,97	0,93	0,063	0,836	0,082	0,503	17,59	2,415
Intouchables	0,73	1,67	-1,88	0,076	0,399	0,507	0,398	5,520	9,772
Tarzan	0,49	-1,82	0,08	0,048	0,669	0,001	0,180	6,579	0,017
SixièmeSens	0,55	-1,10	0,29	0,146	0,569	0,039	0,232	2,387	0,231
LeRoilion	0,59	-0,75	-0,63	0,204	0,330	0,237	0,259	1,105	1,106
LeMondedeNemo	1,35	-1,13	-0,17	0,514	0,362	0,009	1,377	2,555	0,084
HarryPotteretlachambredessecre	1,02	-0,44	1,12	0,346	0,064	0,412	0,788	0,385	3,450
L'Âgedeglace3	0,25	-1,54	0,12	0,019	0,744	0,005	0,046	4,705	0,043
StarWars:épisode1	1,30	-1,10	0,83	0,464	0,332	0,189	10,27	2,404	1,915
HarryPotteràl'écoledessorciers	1,45	-0,64	0,85	0,550	0,109	0,190	11,57	0,820	1,993
StarWars:Episode7	4,39	-1,78	0,28	0,838	0,137	0,004	14,56	6,274	0,225
Titanic	6,51	1,53	-1,03	0,913	0,051	0,023	31,95	4,671	2,960
Avatar	5,97	-0,72	-0,21	0,975	0,014	0,001	26,89	1,043	0,122

Exercice 4

L'AFC permet de représenter qualitatives. Notons m le nombre de modalités de la 1ère et p le nombre de modalités de la deuxième.

- 1) Pourquoi y-a-t-il deux nuages de points sur le graphique de l'AFC ? A quoi correspondent-ils ? Combien y-a-t-il de points en tout sur le graphique ?
- 2) Quelle est la dimension (le nombre d'axes) de ce nuage ?
- 3) Chaque axe est associé à un pourcentage mesurant l'information retranscrite par cet axe. A quoi correspond ce pourcentage ?
- 4) On transforme la variable Nombre de prix remportés en variable qualitative avec 4 modalités choisies en fonction des quartiles. On la croise avec la variable Genre et on effectue une analyse factorielle des correspondances. Analysez le graphique obtenu.

