**Applied Data Analysis**

**Milestone 2**

Majda

Julien Burri – SCIPER 296112

Nathan

Albias

Louis Valteau - SCIPER 354926

**Actor Pair Chemistry and Movie Success**

**Abstract:**

1. **Function form of the revenue in function of the number of famous actors.**

Derive the function form of the revenue in function of the number of famous actors. Thus, that analysis could be useful for the movie makers to know whether is affordable to combine the famous actors or not.

We can observe two results:

* No effect – the number of famous actors in a movie has no impact on the revenue of the movie,
* Positive or negative effect – adding additional known actors will increase or decrease the revenue of the movie. Then, we can determine the evolution of this effect (linear, convex, concave, …).

**Mathematics to do:**

1. **Existing effect ?**

* **Method 1 : OLS regression**: , with Nx dummy variables (0 or 1), H0: all β are null except .
* **Method 2 : ANOVA** -> check whether the difference of average revenue in each class (N1, N2, …) are significantly different. -> if yes, it means there is a variation of revenue in function of number of famous actors.

1. **Fitting of the functional form**

* If H0 does not hold (alpha=0.05) -> try to fit the observations (linear, quadratic, polynomial, log transformation) and quantify the error (R2, RSS, MSS)
* Plot the average revenue for each class and draw the functional form. For each class, add the confident interval (95%) -> that will provide an ordinal sequence of box plots:

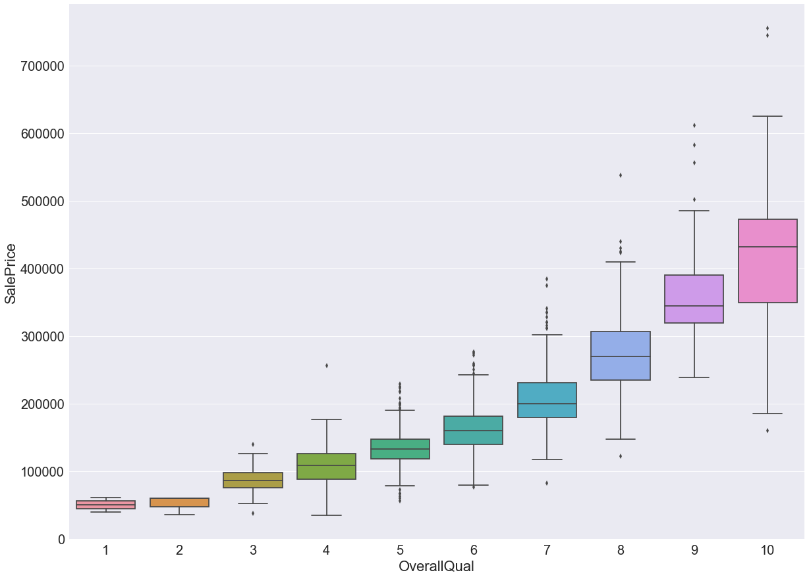
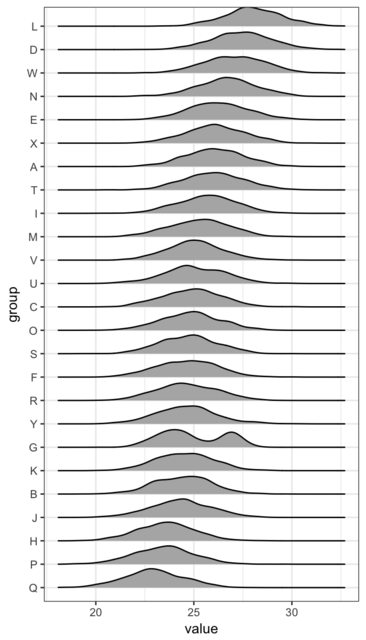


Figure - Just a visual example

We can still use another kind of visualization, other exemples:





1. **Best and worst chemistry between the actor pairs in films**

Quantify the chemistry between the actor pairs in films, how it impacts a movie's financial performance, and whether the pairing of specific actors in multiple films leads to more successful outcomes.



The benchmark will be the average of the actor pairs.

1. **Gender and actor pairs.**

**Remarque :**

Critère de selection des acteurs (# de films, # de connexions, même période 2000-2022)

Groupby “industry” (not the same revenue from Holywood and Bolywood)

Comparaison revenue (inflation, exchange rate, …)

Actor Pairs with a frequence of at least 2/3 times to avoid outlier.