

# ***SUPERHEROES***



Fait par : Reem Al Sobaiee & Rania Ben Ouioua

Projet de Visualisation de Données



## ***LE PLAN :***

***01. PRÉSENTATION***

***02. OUTILS***

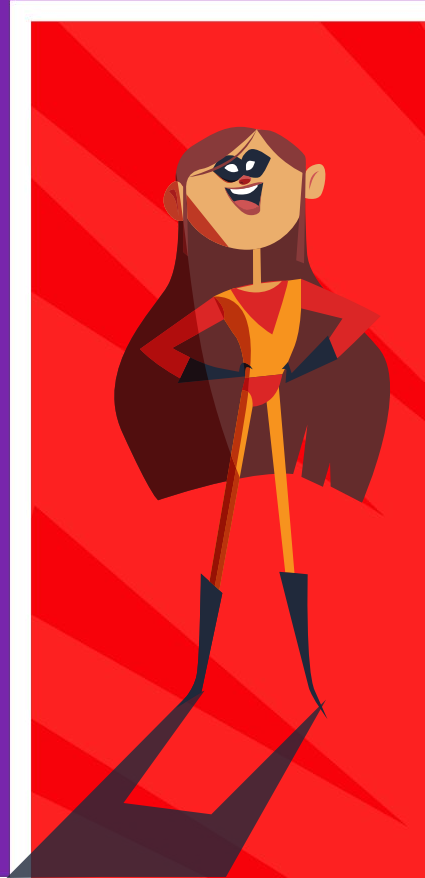
***03. VISUALISATIONS***

***04. PROBLÈMES***

# PRÉSENTATION DU DATASET SUPERHEROES

- Le dataset Superheroes est composé de **283 observations décrites par 18 variables**
- Ce dataset est le fruit de notre projet de data mining
- Données extraites du site Superhero Database :

<https://www.superherodb.com>



# ***NOS VARIABLES :***

## ***BIOGRAPHY***

Name  
Full Name  
Publisher  
Alignment  
Image

## ***POWER STATS***

Class  
Intelligence  
Strength  
Speed  
Durability  
Power  
Combat

## ***APPEARANCE***

Gender  
Race  
Height  
Weight  
Eye Color  
Hair Color

# ***VARIABLES OBSERVABLES QUALITATIVES ET QUANTITATIVES***

## ***QUALITATIVES***

Publisher

Alignement

Gender

Race

Eye color

Hair color

Class  
Intelligence  
Strength  
Speed  
Durability  
Power  
Combat  
Height  
Weight

## ***QUANTITATIVES***

# EXEMPLE D'UN INDIVIDU : THE FLASH

## POWER STATS :

Class : 7  
Intelligence : 88  
Strength : 48  
Speed : 100  
Durability : 60  
Power : 100  
Combat : 60

## APPARENCE

Gender : Male  
Race : Human  
Height : 180  
Weight : 81  
Eye color : Blue  
Hair color : Blond



## BIOGRAPHY

Flash

**NAME**

Barry Allen

**FULL NAME**

DC Comics

**PUBLISHER**

Good

**ALIGNEMENT**



***LA DATAVIZ !***

# **OBJECTIF**

- Étude des profils des superheroes
- Représenter visuellement des données extraite
- Extraire des informations plus facilement
- Mettre en avant les données et relations qui comptent



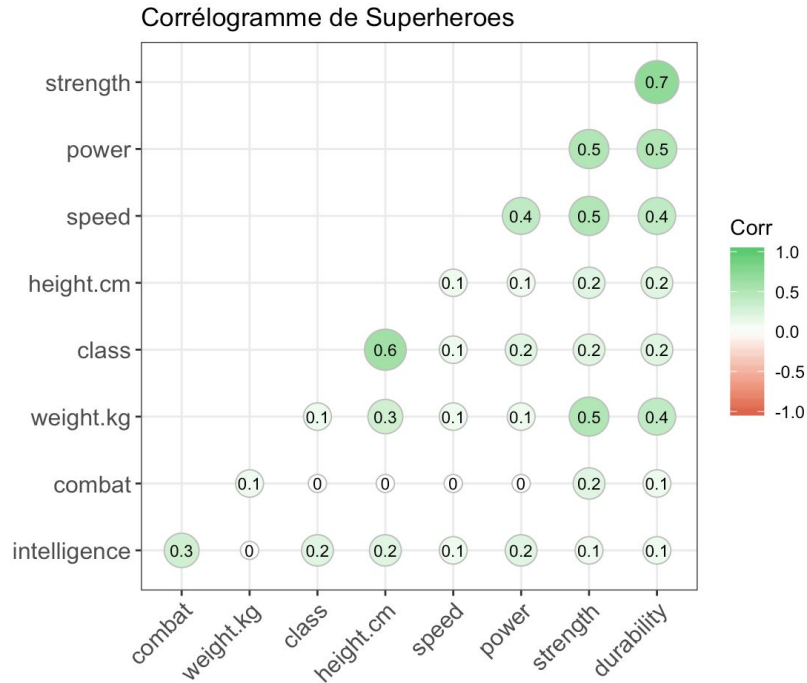
# OUTILS





# ***VISUALISATION DES DONNÉES QUANTITATIVES***

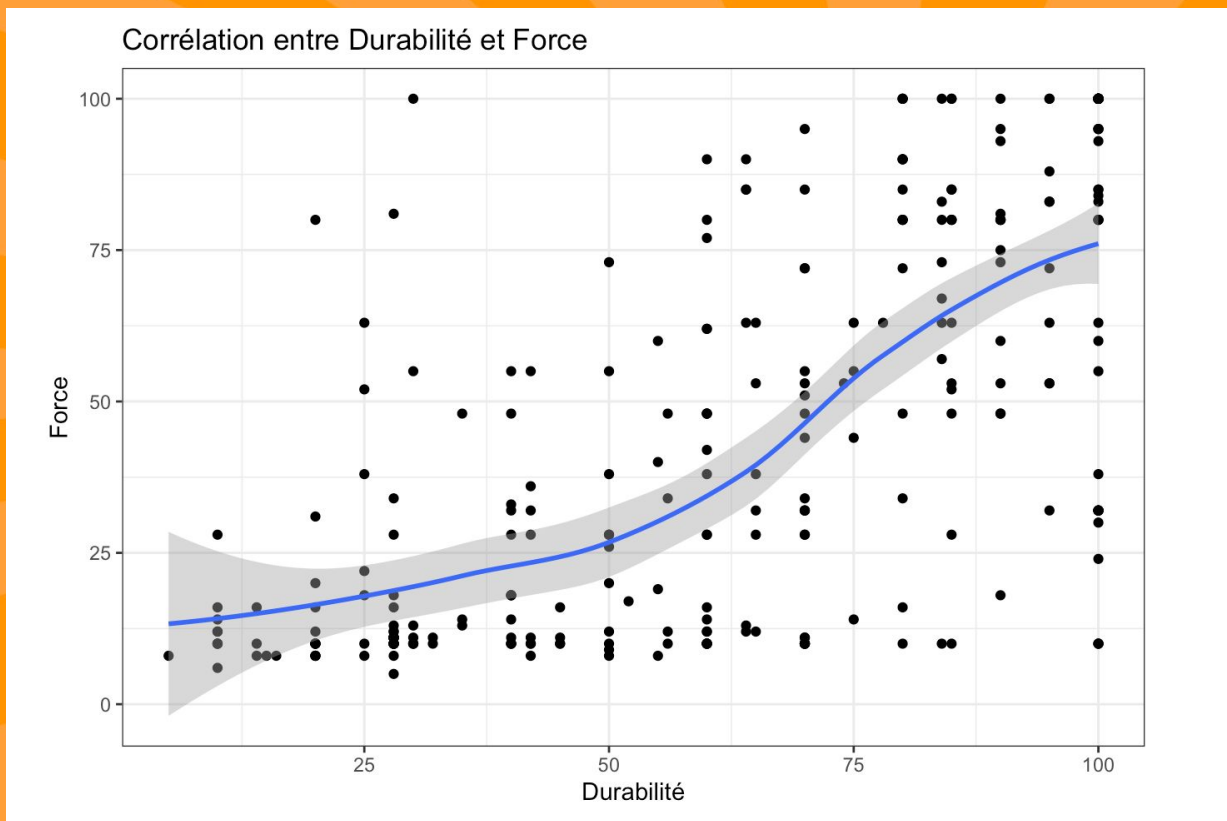
# CORRÉLOGRAMME



→ Création de la matrice de corrélation

→ Affichage du corrélogramme grâce au package 'ggcorrplot'

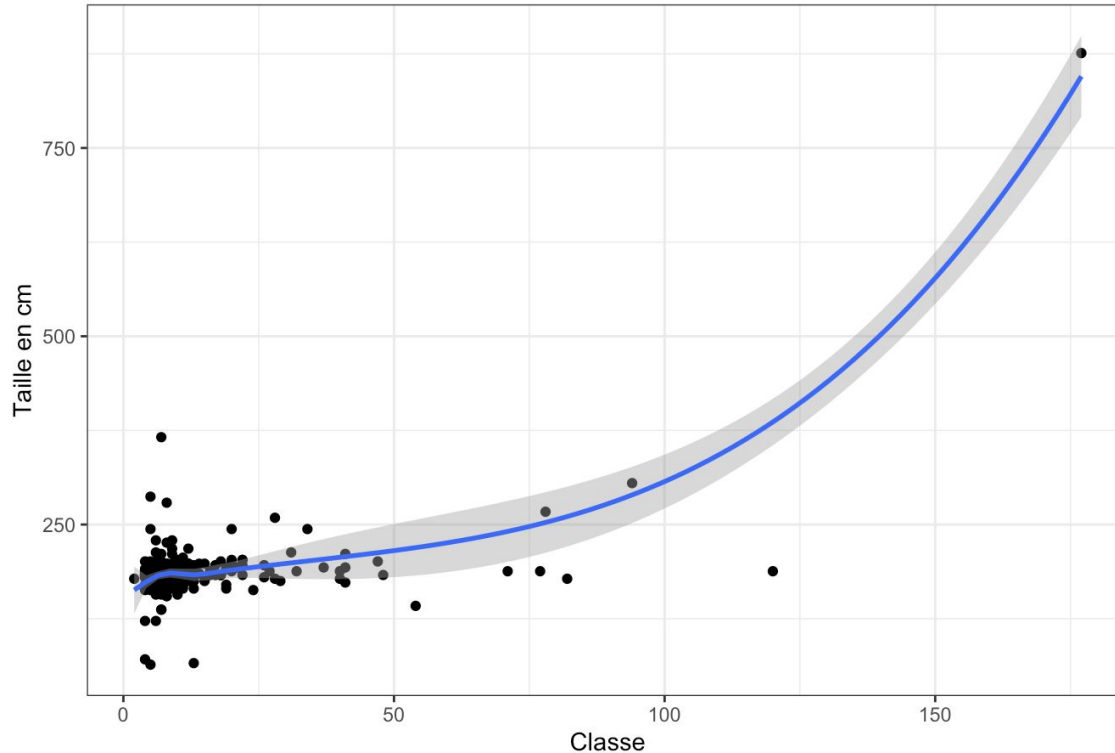
# CORRÉLATION ENTRE DURABILITÉ ET FORCE



- Coefficient de corrélation = 0.7
- Corrélation de qualité moyenne

# CORRÉLATION ENTRE CLASSE ET TAILLE

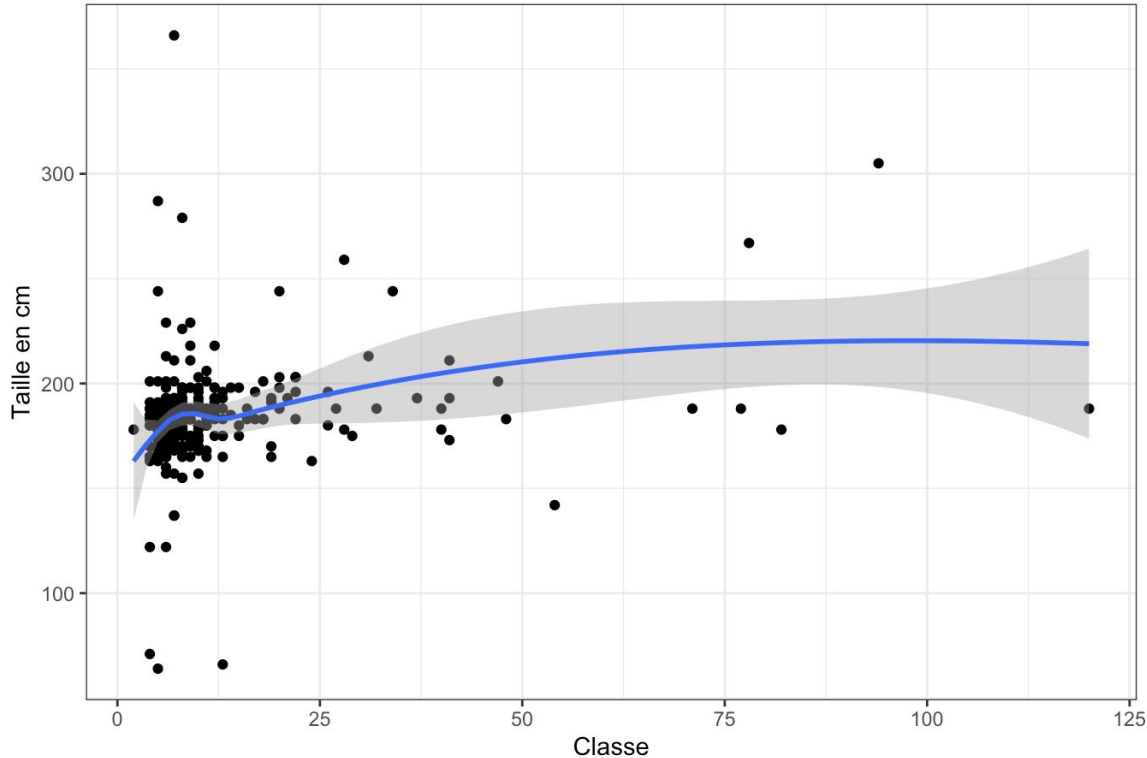
Corrélation entre Classe et Taille



- Coefficient de corrélation = 0.6
- Corrélation de qualité moyenne
- Remarque : un **outlier** !

# ***CORRÉLATION ENTRE CLASSE ET TAILLE SANS OUTLIER***

Corrélation entre Classe et Taille sans le dernier individu

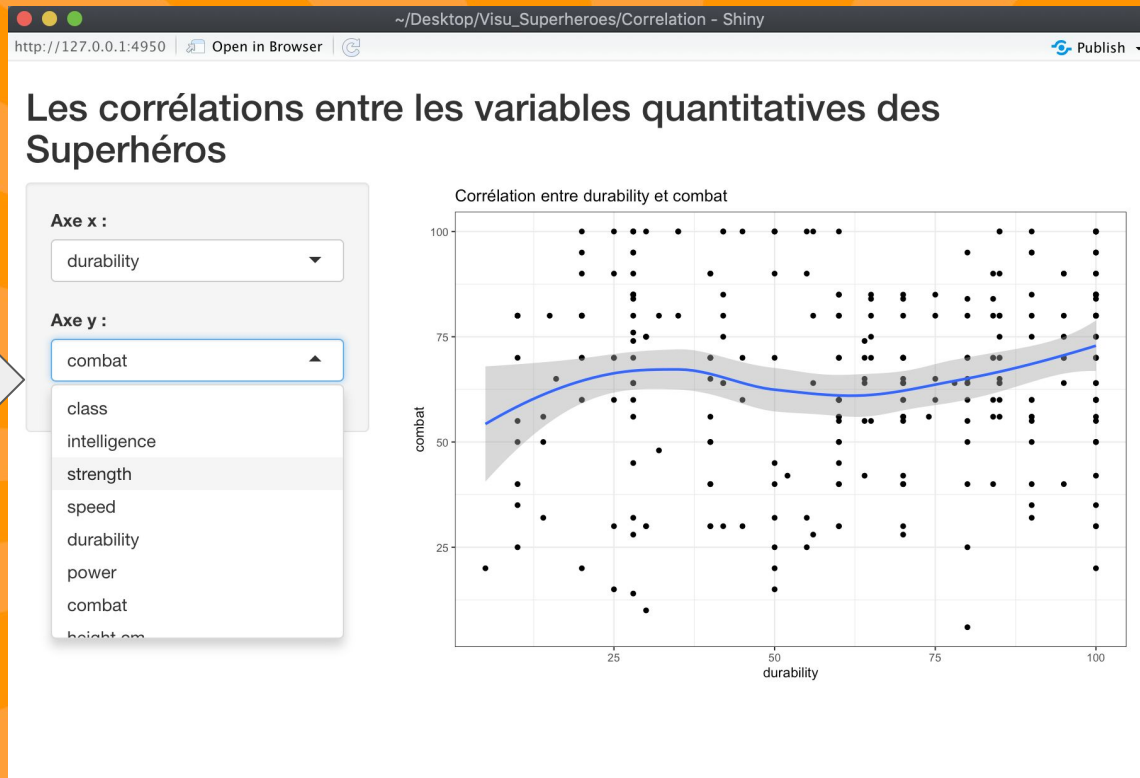


- Change fortement le coefficient de corrélation
- Plus aucune corrélation

***ET SI L'ON VOULAIT VISUALISER LA  
CORRÉLATION ENTRE TOUTES LES  
VARIABLES ?***

# SOLUTION ? UNE APPLICATION INTERACTIVE !

Choix de la  
variables pour  
chaque axe



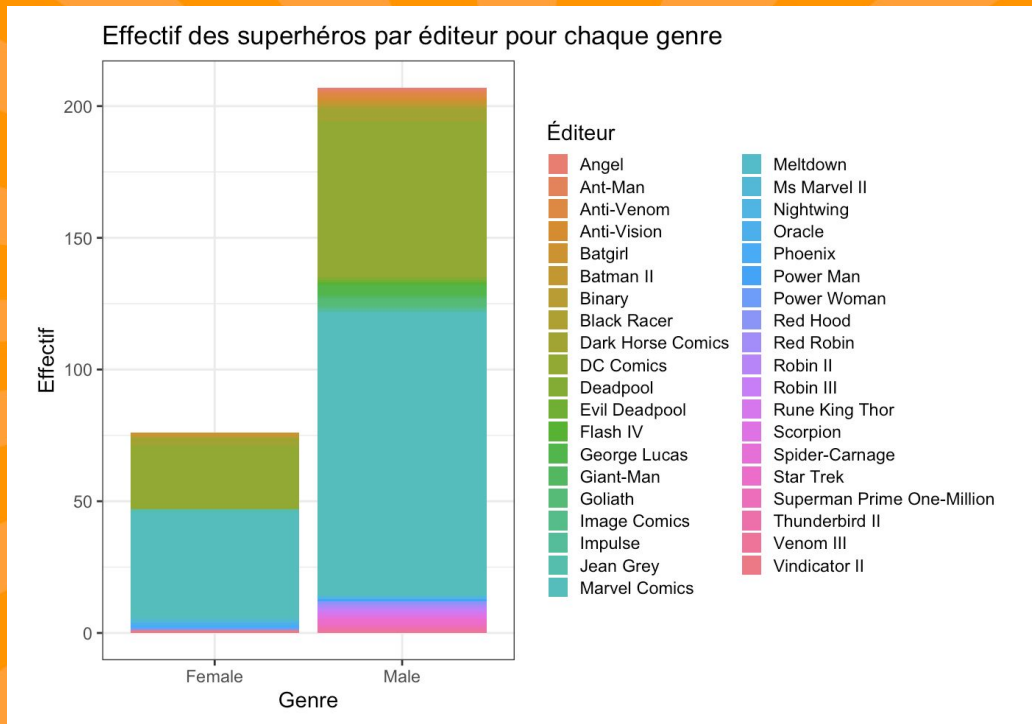
Shiny



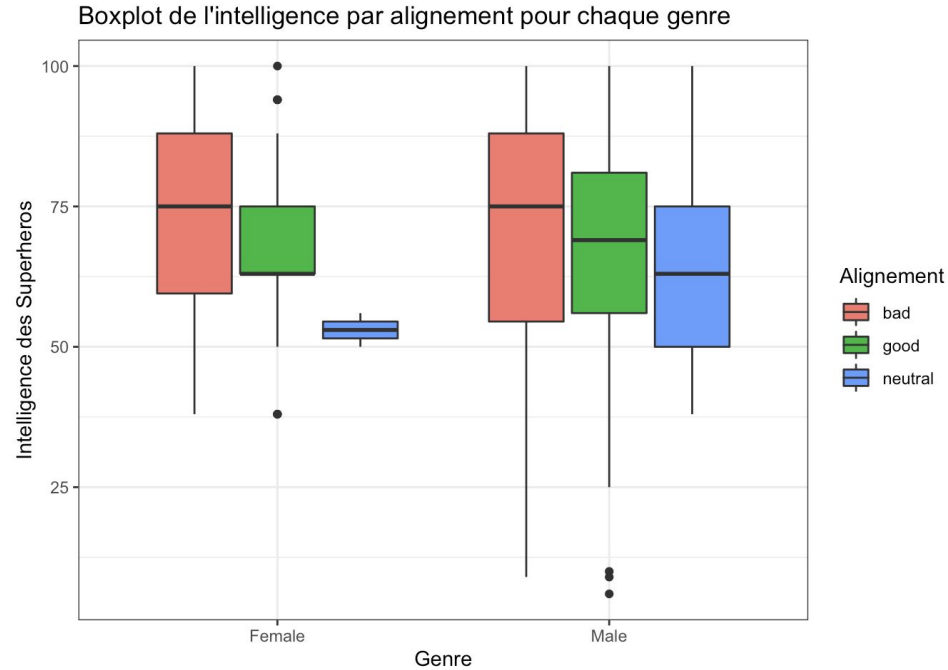


# ***VISUALISATION DES DONNÉES QUALITATIVES***

# QUI EST PLUS CONNU ? SUPERHÉROS OU SUPERHÉROINES ?

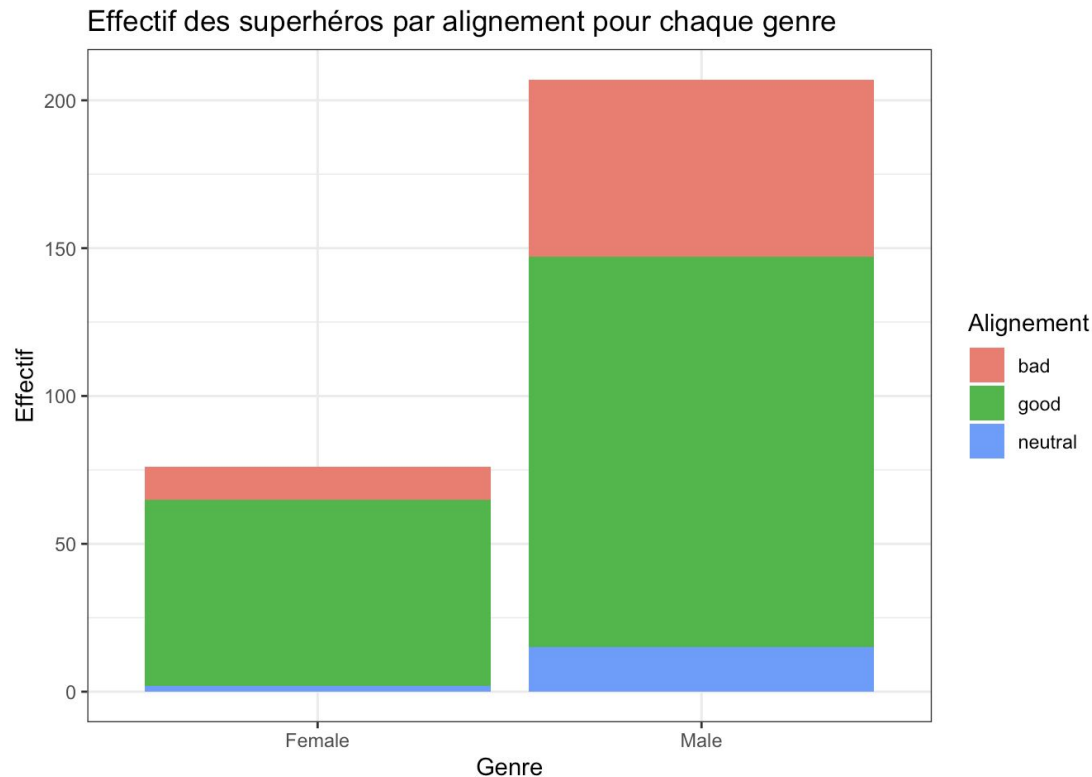


# QUI EST PLUS INTELLIGENT ?





# ***QUI EST GENTIL & QUI EST MÉCHANT ?***

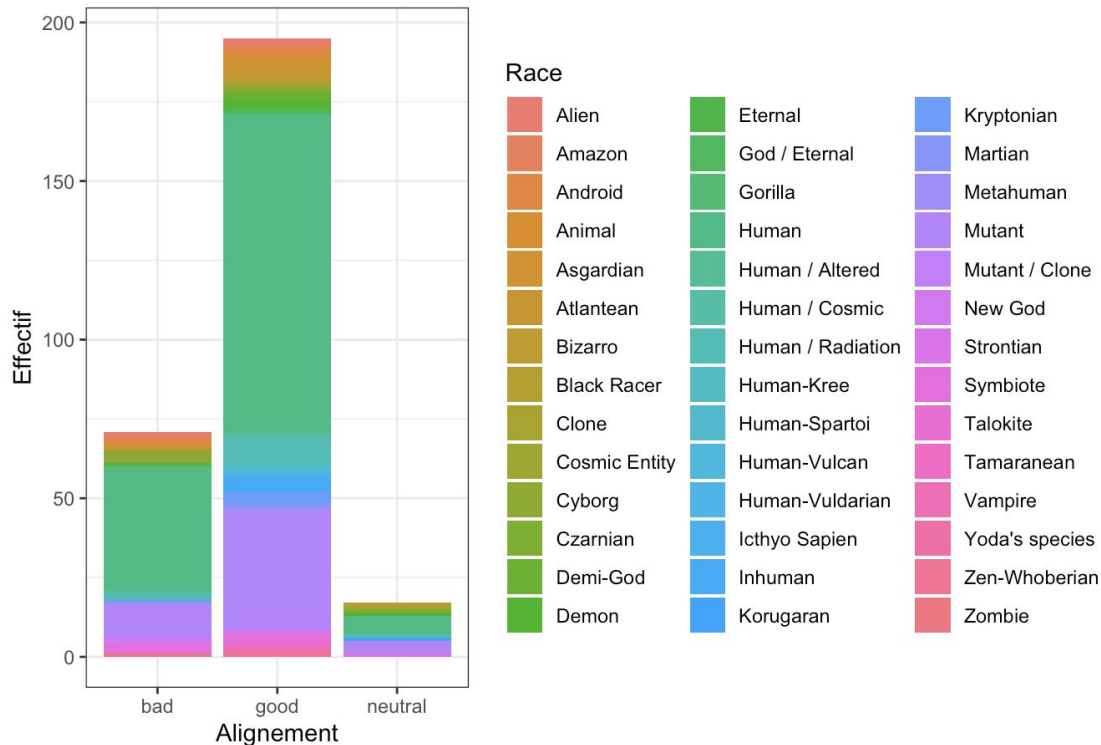




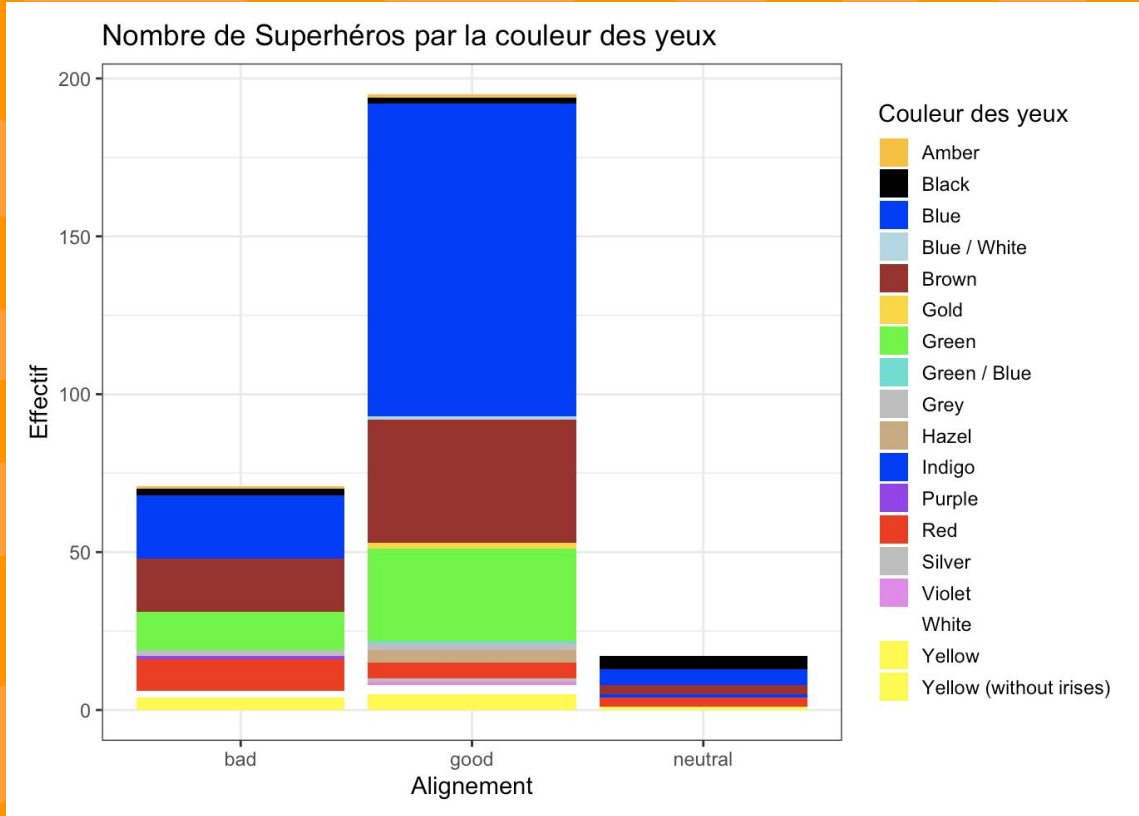
# QUI EST GENTIL & QUI EST MÉCHANT ?



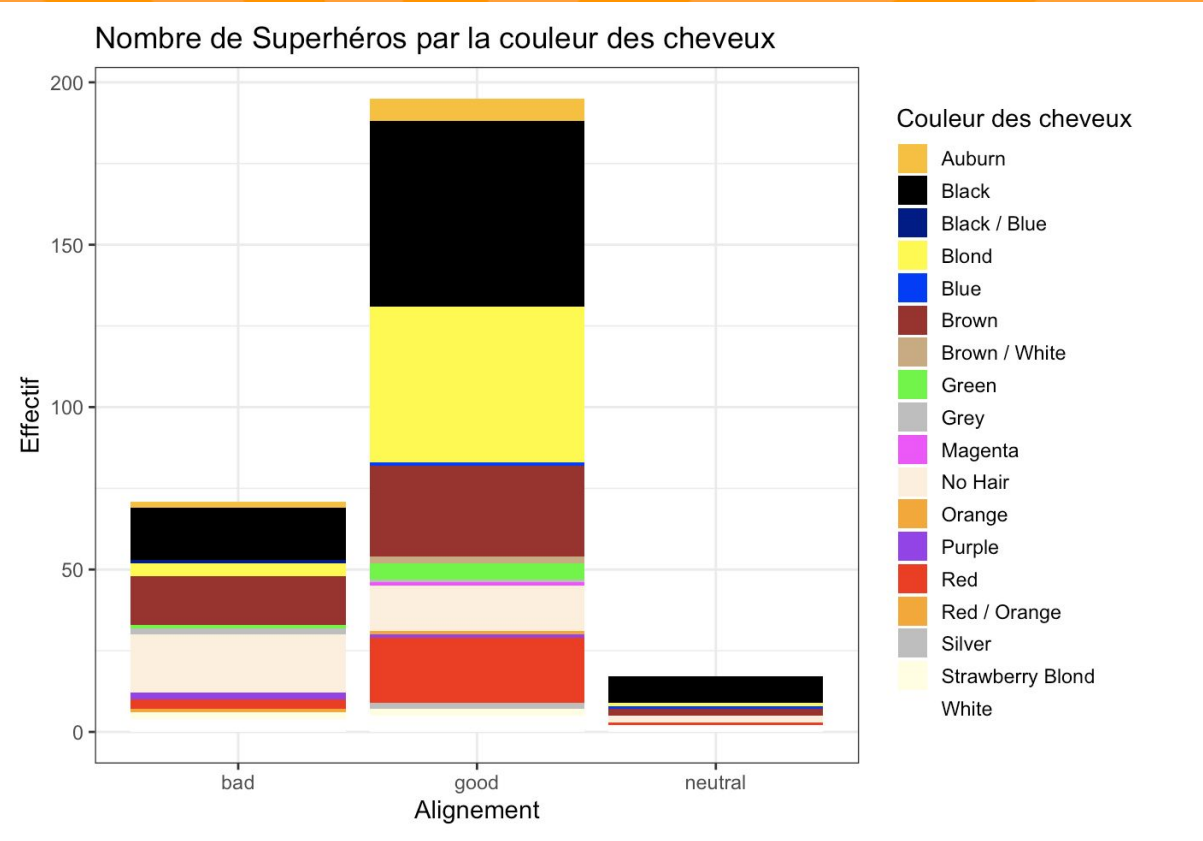
Effectif des superhéros par race pour chaque alignement



# PEUT-ON EN DÉDUIRE UN STÉRÉOTYPE BASÉ SUR L'APPARENCE DES SUPERHÉROS ?



# PEUT-ON EN DÉDUIRE UN STÉRÉOTYPE BASÉ SUR L'APPARENCE DES SUPERHÉROS ?





***CONCLUSION ?***





# PROBLÈMES RENCONTRÉS

- Choix des graphes
- Extractions des informations pertinentes
- Analyse des graphes





**THANK  
YOU!**