

## Solution Summary :

### Question 1

1) How would you define most powerful superhero from the information available in dataset?

a> Who is the most powerful superhero from each creator?

- The field provided in the data set 'overall\_score' is a good estimate of the power of the superhero.
- However, the challenge is that there are :

18 incidents with ' $\infty$ ' as overall\_score  
107 incidents with '-' as overall\_score

- To overcome this challenge we built a linear regressor model to predict the score given the features
- Then we predict the score for

18 incidents with ' $\infty$ ' as overall\_score  
107 incidents with '-' as overall\_score

- Who ever gets the maximum score is the strongest superhero

*Note : More sophisticated model than linear regressor can be used. However, to provide an intuition of the :*

*solution we stuck to the base model.*

> Who is the most powerful superhero from each creator ?

- Top 10 creator details provided
- Incidents for which creator information is not provided is grouped as nan

name
golden master mech
black alice
true form oblivion
devilman
dracula
unicron
golden ninja
man miracle
skeletor
link

## Question 2

2) Find the top 5 superpowers in descending order.

- Intuition

We observe the contribution of various powers to the 'overall\_score' field.  
Larger the positive coefficient greater the contribution.  
Hence sorting and picking top 5 gives the result.

column_name	coeff
has_immortality	3.4241
has_matter_manipulation	2.5903
intelligence_score	1.8214
has_mind_control	1.5616
has_teleportation	1.5106

## Question 3

Which race has the most immortal superheroes?

nan = undefined group

nan	38
god eternal	36
human	29
demon	18
new god	11

## Question 4

4)Name the creator having most superheroes of type “Parademon”.

```
data[data['type_race']=='parademon']['creator']
```

991      dc comic

## Question 5

- Which comic creator has most superhero teams?
  - a. Find names, real names and alias of superhero who is part of most teams.
  - b. Are there any crossovers between creators and teams?

- Which comic creator has most superhero teams?

```
df1.sort_values('team_count',ascending=False).head(2)
```

	creator	teams	team_count
2	dc comic	['Blue Lantern Corps', 'Green Lantern Corps', ...	73
8	marvel comic	['Annihilators', 'Asgardians', 'Avengers', 'Ne...	56

a> Answer

```
df2.sort_values('team_count',ascending=False).head(5)
```

	name	teams	team_count
142	deadpool	['Deadpool Corps', 'Agency X', 'X-Force', 'Dee...	14
530	wolverine	['X-Force', 'X-Men', 'Weapon X', 'Secret Defen...	14

**Find names, real names and alias of superhero who is part of most teams.**

```
data[(data['name']=='deadpool') | (data['name']=='wolverine')][['name','real_name','aliases']]
# ['name','real_name','aliases']
```

	name	real_name	aliases
370	deadpool	wade wilson	wade wilson jack chiyonosake wolf rice wine rh...
1415	wolverine	logan	weapon x weapon ten death mutate jim logan emi...

**b. Are there any crossovers between creators and teams?**

yes

```
agg_by_team.sort_values('creator_count',ascending=False).head(5)[[
```

	clean_team	creator_count	clean_creator
289	incredible family	2	disney ,dark horse comic
186	titans	2	marvel comic ,dc comic
0	a force	1	marvel comic

## Question 6

- What are the characteristics that can predict a superhero alignment.

The traditional way would be to build a classifier and observe the factord that contribute most to it's success.

Considering the time constraint I couldn't follow that approach

However by analysis/profiling of data following observations were made :

1> Superheros with negative alignment have higher strength, speed , durability,power score

2>Superheros with positive alignment have higher intelligence and combat score

Question 7

- From history of superheroes,
  - a. Find list of superheroes having negative past but now aligned positively.  
(Negativity of past can be decided by multiple methods, please use which is familiar to you)
  - b. Extract patterns from superhero history for each creator.

My consideration of negative past is if history\_text sentiment is negative.

- predicting past\_alignment from history\_text took too much time
- Hence, we limited ourself to picking just 50 samples

a. Find list of superheroes having negative past but now aligned positively

```
man
abomb
aa
aaron cash
abe sapien
abin sur
absorb man mcu
adam strange
agent
agent bob
agent coulson destroyer gun
agent coulson mcu
agent may mcu
agent zero
akita
alfred pennyworth
allan quatermain
ancient one mcu
ando masahashi
angel dust
angel salvadore
angel
animal man
antman ii
antman
aquababy
aquagirl
aquaman injustice
ardina
ariel
armor
```

b. Extract patterns from superhero history for each creator.

- Creators distribution with superheroes having negative past but now aligned positively

```
past[past['past_alignment']=='bad']['creator'].value_counts()
```

```
marvel comic      17
dc comic          9
dark horse comic  2
lego              1
wildstorm         1
nbc hero          1
Name: creator, dtype: int64
```

- Creators distribution with superheroes having positive past and now aligned positively

```
past[past['past_alignment']=='good']['creator'].value_counts()
```

```
dc comic      6
marvel comic  5
ubisoft       2
george lucas  1
nbc hero      1
j r r tolkien 1
Name: creator, dtype: int64
```

## Question 8

- Report on the 10 superheroes with most relatives,
- status of those relatives where possible,
- and the alignment of those superheroes.

	name	relatives_cleaned	alignment	relatives_count
936	namor	Elanna (maternal ancestor), Tanas (maternal an...	good	53
574	havok	Oscar Summers (adoptive paternal distant ances...	good	48
75	aquaman	Koryak (son), Arthur Curry, Jr. (son), A.J. (s...	good	37
1100	robin v	Talia al Ghul (mother),\nBruce Wayne (Batman, ...	good	36
121	baron zemo	Harbin Zemo (distant ancestor, deceased),\nHad...	bad	33
515	ghost rider king hell	Illyana Kale (maternal ancestor, deceased), De...	good	33
1251	supergirl	Zor-El (father), Allura In-Ze (mother), Jor-El...	good	33
340	cyclops	Oscar Summers (adoptive paternal distant ances...	good	33
1328	toxin	Carl Brock (father, estranged), Janine Brock (...)	good	28
312	colossus	Grigory Efimovich Rasputin (great-grandfather,...)	good	28

## Question 9

- Find out any other interesting insights from given data.
- Which 3 comic characters can you recommend to your friends to read or watch?

- Super heros are product of an imaginative mind with escapist tendencies.
- Failing to deal with the reality we resort to imagination of super power or super being to seek comfort in the thought that a miracle would set us free.
- But miracles rarely do happen and even less to those waiting for it.
- Hence, I would recommend watching superheros with the least 'over\_all' score.
- It doesn't take a lot of courage to stand before the bullet when you know it would bounce off your skin. but takes a mighty heart to do so knowing full well that it would pierce through them.
- Hence I believe superheros with the least 'over\_all' score rank higher when it come to attitude.
- Sends the message a sharp blade don't defeat the dragon, but a valiant warrior does.
- But knowing my friends how shallow they are they would rather enjoy watching the strongest superheros.
- Because I care for their happiness I would recommend superheros with the highest 'over\_all' score.

name	
508	golden master mech
173	black alice
1285	true form oblivion