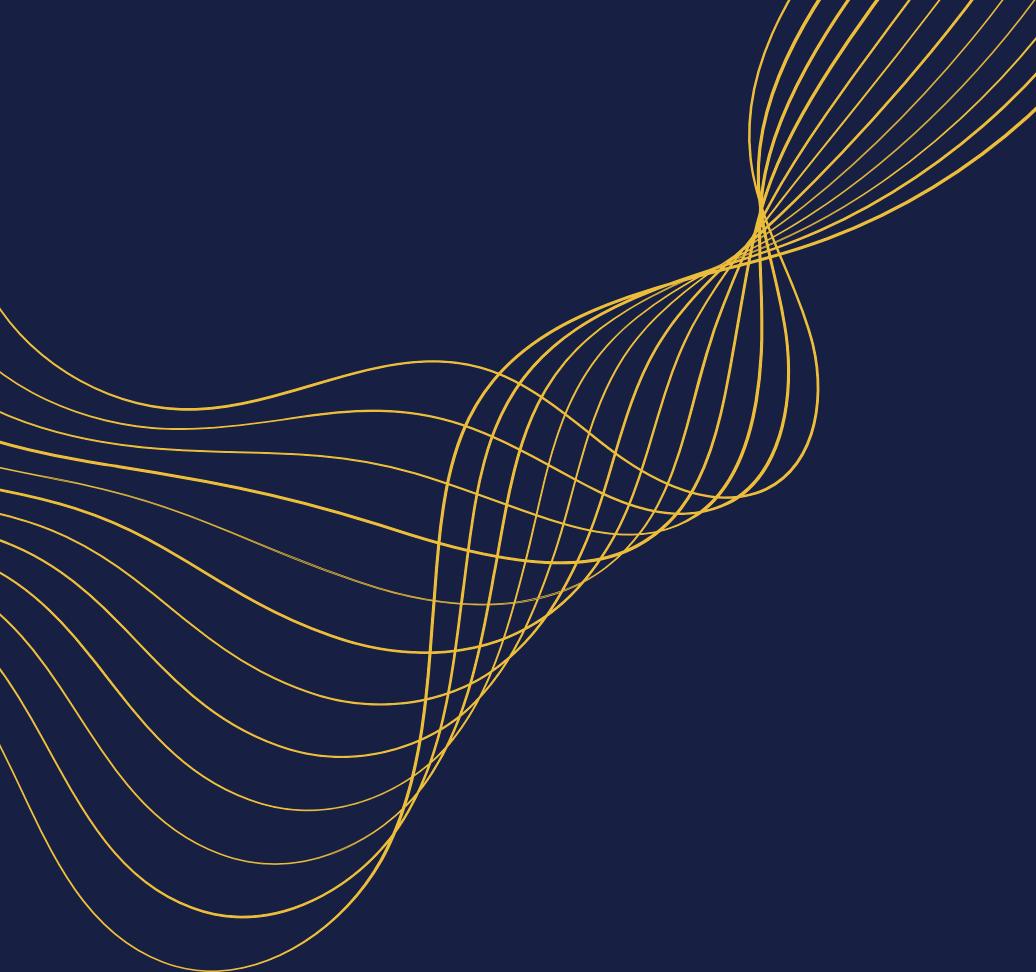


RUSH 1

Data Analyst

Youtube



SOMMAIRE



01

INTRODUCTION

02

PRÉSENTATION DU RUSH

03

DÉMO

04

CONCLUSION

Introduction



YouTube



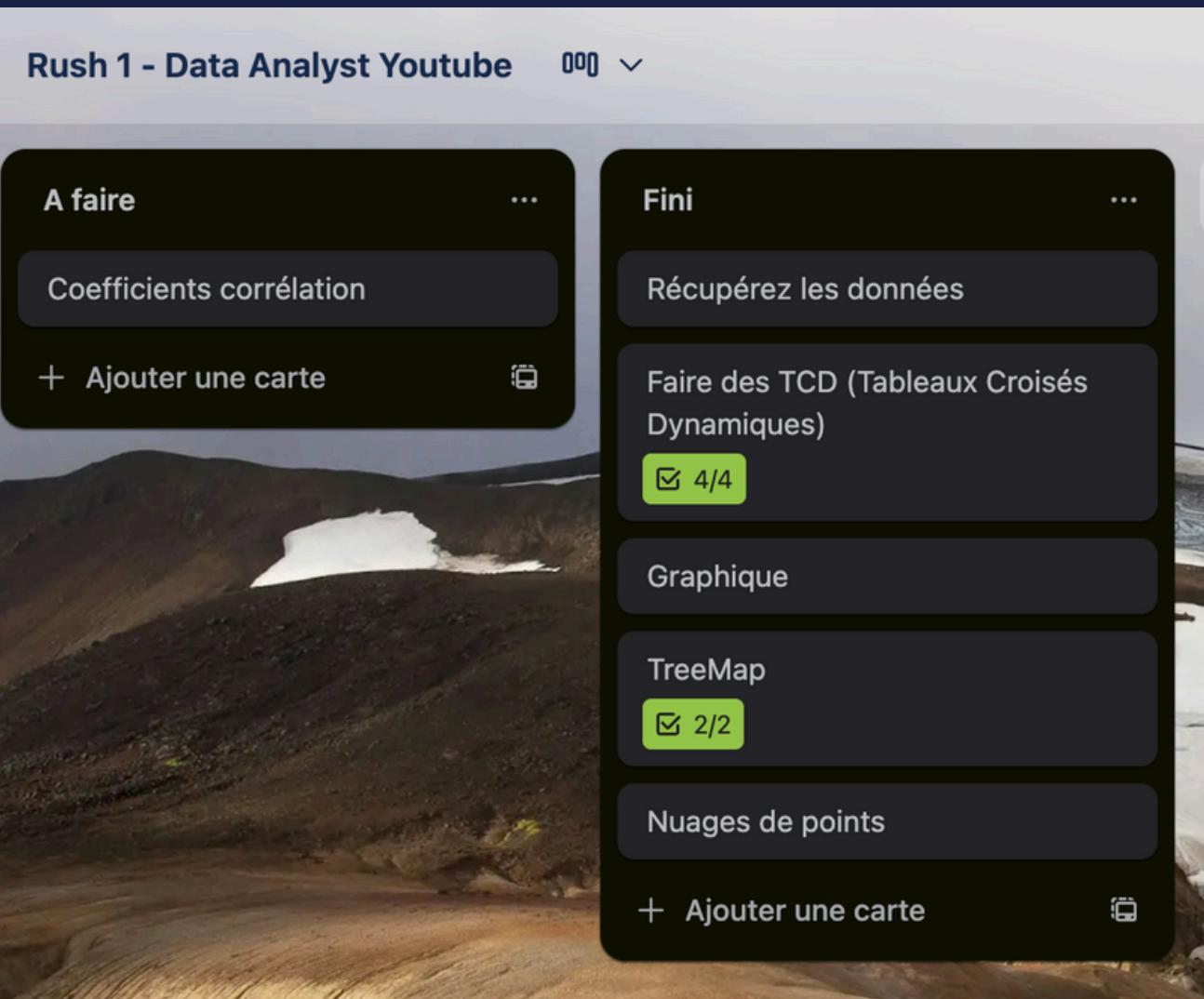
Youtube

Data Analyst



Excel

Introduction



TRELLO

Epitech-MSI-2027 / DATABOOTCAMP_MAINGRE_THOMAS

Issues Pull requests Actions Projects Wiki Security Insights Settings

DATABOOTCAMP_MAINGRE_THOMAS Public

main 1 Branch 0 Tags

ThomasMaingre Initial commit a3cd6d7 · yesterday 1 Commit

README.md Initial commit yesterday

README

DATABOOTCAMP_MAINGRE_THOMAS

GIT HUB

Présentation du projet

Problème

A	B	C	D	E	F	G	H	I	J
1	,Rank,Channel_Name,Subscriber_Count,Video_Views,Video_Count,Genre,Channel_Started								
2	0,1,T-Series,237000000,216495600668,18831,Music,2006								
3	1,2,Cocomelon - Nursery Rhymes,154000000,152638999634,861,Education,2006								
4	2,3,SET India,152000000,140138068504,105649,Film & Animation,2006								
5	3,4,Sony SAB,77500000,92952274861,65028,Film & Animation,2007								
6	4,5,âœ€ Kids Diana Show,108000000,88452629066,1070,People & Blogs,2015								
7	5,6,Like Nastya,104000000,88060349741,762,People & Blogs,2016								
8	6,7,WWE,93900000,74447865775,66901,Sports,2007								
9	7,8,Vlad and Niki,93900000,73333582362,530,Entertainment,2018								
10	8,9,Movieclips,58700000,58923017461,40063,Film & Animation,2006								
11	9,10,Colors TV,60200000,58056997206,104523,Film & Animation,2008								
12	10,11,Zee Music Company,92800000,54295114324,7768,Music,2014								
13	11,12,El Reino Infantil,55600000,54151900416,1434,Music,2011								
14	12,13,Ryan's World,34400000,54008224558,2321,Entertainment,2015								
15	13,14,netd mÃ½zik,23600000,53867289619,22108,Music,2014								
16	14,15,ABS-CBN Entertainment,42600000,49164270097,187424,People & Blogs,2008								
17	15,16,Toys and Colors,38600000,44642348659,910,Entertainment,2016								
18	16,17,ChuChu TV Nursery Rhymes & Kids Songs,61900000,42589514748,546,Education,2013								
19	17,18,Zee TV,65900000,42064672138,98369,Entertainment,2005								
20	18,19,Super Simple Songs - Kids Songs,37800000,41715976696,667,Education,2006								
21	19,20,Wave Music,56300000,39258909308,19031,Music,2014								
22	20,21,Little Baby Bum - Nursery Rhymes & Kids Songs,40800000,38840192263,2287,Education,2011								
23	21,22,Canal KondZilla,66400000,37297721449,2393,Music,2012								
24	22,23,Pinkfong Baby Shark - Kids' Songs & Stories,65700000,36688824621,2647,Education,2011								
25	23,24,WorkpointOfficial,38300000,35265515153,68255,Entertainment,2012								
26	24,25,ĐœĐ°ÑĐ° Đ, ĐœĐµĐ’ÐµĐ’ÑŒ,41600000,35041524012,1316,Film & Animation,2011								
27	25,26,HAR PAL GEO,38000000,33884922133,84696,Entertainment,2008								
28	26,27,ARY Digital HD,39800000,33692884745,92357,Entertainment,2016								
29	27,28,YRF,52200000,33001702137,3528,Film & Animation,2006								
30	28,29,Tips Official,57000000,31392740249,4182,Music,2007								
31	29,30,LooLoo Kids - Nursery Rhymes and Children's Songs,52900000,31261734882,621,Music,2014								
32	30,31,LankyBox,23300000,31155421572,7307,Comedy,2016								
33	31,32,Shemaroo Filmi Gaane,64000000,30308486987,8535,Music,2010								

→ Données illisibles

→ Tout dans la colonne A

→ 300 lignes à décoder

Présentation du projet

Solution

A	B	C	D	E	F	G	H	I	J
1	,Rank,Channel_Name,Subscriber_Count,Video_Views,Video_Count,Genre,Channel_Started								
2	0,1,T-Series,237000000,216495600668,18831,Music,2006								
3	1,2,Cocomelon - Nursery Rhymes,154000000,152638999634,861,Education,2006								
4	2,3,SET India,152000000,140138068504,105649,Film & Animation,2006								
5	3,4,Sony SAB,77500000,92952274861,65028,Film & Animation,2007								
6	4,5,âœ¿ Kids Diana Show,108000000,88452629066,1070,People & Blogs,2015								
7	5,6,Like Nastya,104000000,88060349741,762,People & Blogs,2016								
8	6,7,WWE,93900000,74447865775,66901,Sports,2007								
9	7,8,Vlad and Niki,93900000,73333582362,530,Entertainment,2018								
10	8,9,Movieclips,58700000,58923017461,40063,Film & Animation,2006								
11	9,10,Colors TV,60200000,58056997206,104523,Film & Animation,2008								
12	10,11,Zee Music Company,92800000,54295114324,7768,Music,2014								
13	11,12,El Reino Infantil,55600000,54151900416,1434,Music,2011								
14	12,13,Ryan's World,34400000,54008224558,2321,Entertainment,2015								
15	13,14,netd mÃ½zik,23600000,53867289619,22108,Music,2014								
16	14,15,ABS-CBN Entertainment,42600000,49164270097,187424,People & Blogs,2008								
17	15,16,Toys and Colors,38600000,44642348659,910,Entertainment,2016								
18	16,17,ChuChu TV Nursery Rhymes & Kids Songs,61900000,42589514748,546,Education,2013								
19	17,18,Zee TV,65900000,42064672138,98369,Entertainment,2005								
20	18,19,Super Simple Songs - Kids Songs,37800000,41715976696,667,Education,2006								
21	19,20,Wave Music,56300000,39258909308,19031,Music,2014								
22	20,21,Little Baby Bum - Nursery Rhymes & Kids Songs,40800000,38840192263,2287,Education,2011								
23	21,22,Canal KondZilla,66400000,37297721449,2393,Music,2012								
24	22,23,Pinkfong Baby Shark - Kids' Songs & Stories,65700000,36688824621,2647,Education,2011								
25	23,24,WorkpointOfficial,38300000,35265515153,68255,Entertainment,2012								
26	24,25,ĐœĐ°ÑÐ° Đ, ĐœÐµÐ’Ð²ÐµÐ’ÑŒ,41600000,35041524012,1316,Film & Animation,2011								
27	25,26,HAR PAL GEO,38000000,33884922133,84696,Entertainment,2008								
28	26,27,ARY Digital HD,39800000,33692884745,92357,Entertainment,2016								
29	27,28,YRF,52200000,33001702137,3528,Film & Animation,2006								
30	28,29,Tips Official,57000000,31392740249,4182,Music,2007								
31	29,30,LooLoo Kids - Nursery Rhymes and Children's Songs,52900000,31261734882,621,Music,2014								
32	30,31,LankyBox,23300000,31155421572,7307,Comedy,2016								
33	31,32,Shemaroo Filmi Gaane,64000000,30308486987,8535,Music,2010								

Diviser chaque catégories de données dans une colonne

Utiliser le séparateur de millier

Limiter à 10-20 lignes pour une meilleure visibilité

Présentation du projet

Création de TCD général

	A	B	C	D	E	F	G
1	Genre (Tous)						
2	Channel_Started (Tous)						
3							
4	Étiquettes de lignes	Somme de Subscriber_Count	% de Subscriber_Count	Somme de Video_VIEWS	% de Video_VIEWS	Somme de Video_Count	% de Video_Count
5	⑤ T-Series	237 000 000	18,50%	216 495 600 668	23,04%	18 831	9,06%
6	1	237 000 000	18,50%	216 495 600 668	23,04%	18 831	9,06%
7	⑥ Cocomelon - Nursery Rhymes	154 000 000	12,02%	152 638 999 634	16,24%	861	0,41%
8	2	154 000 000	12,02%	152 638 999 634	16,24%	861	0,41%
9	⑦ SET India	152 000 000	11,86%	140 138 068 504	14,91%	105 649	50,84%
10	3	152 000 000	11,86%	140 138 068 504	14,91%	105 649	50,84%
11	⑧ âœ; Kids Diana Show	108 000 000	8,43%	88 452 629 066	9,41%	1 070	0,51%
12	5	108 000 000	8,43%	88 452 629 066	9,41%	1 070	0,51%
13	⑨ Like Nasty	104 000 000	8,12%	88 060 349 741	9,37%	762	0,37%
14	6	104 000 000	8,12%	88 060 349 741	9,37%	762	0,37%
15	⑩ WWE	93 500 000	7,30%	74 447 865 775	7,92%	66 901	32,19%
16	7	93 500 000	7,30%	74 447 865 775	7,92%	66 901	32,19%
17	⑪ Vlad and Niki	93 900 000	7,33%	73 333 582 362	7,80%	530	0,26%
18	8	93 900 000	7,33%	73 333 582 362	7,80%	530	0,26%
19	⑫ Zee Music Company	92 800 000	7,24%	54 295 114 324	5,78%	7 768	3,74%
20	11	92 800 000	7,24%	54 295 114 324	5,78%	7 768	3,74%
21	⑬ PewDiePie	111 000 000	8,66%	28 920 114 696	3,08%	4 700	2,26%
22	38	111 000 000	8,66%	28 920 114 696	3,08%	4 700	2,26%
23	⑭ MrBeast	135 000 000	10,54%	22 825 311 283	2,43%	735	0,35%
24	76	135 000 000	10,54%	22 825 311 283	2,43%	735	0,35%
25	Total général	1 281 200 000	100,00%	939 607 636 053	100,00%	207 807	100,00%

▶ Filtrage selon les Genres et Années

▶ Ajout des %, pour une autre approche de lecture

▶ Limiter à 10 lignes pour une meilleure visibilité

Présentation du projet

Création de TCD de chaque catégories

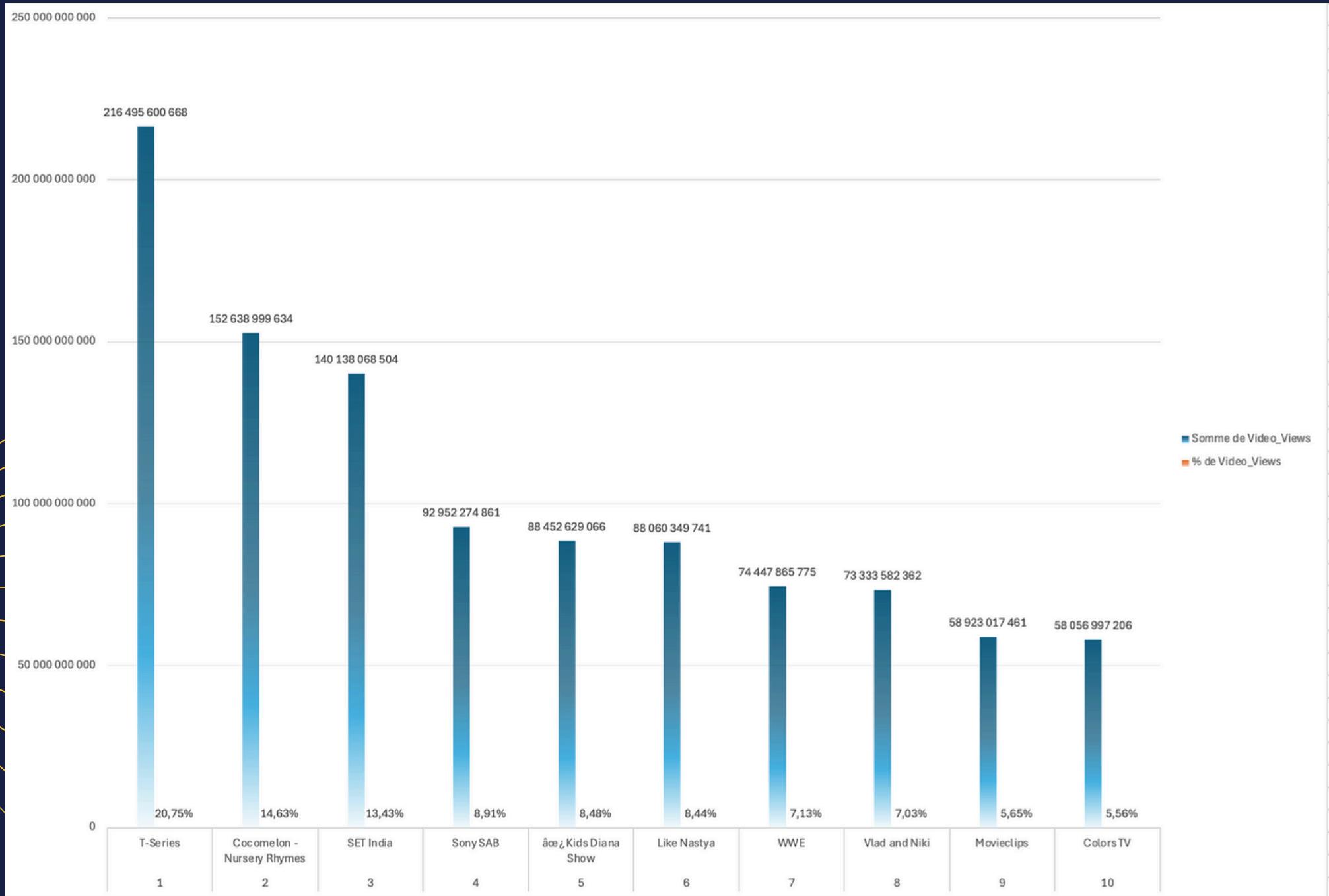
A	B	C
Zone Nom	(Tous)	
2 Channel_Started	(Tous)	
3		
4 Étiquettes de lignes	▼ Somme de Video_VIEWS % de Video_VIEWS	
5 1	216 495 600 668	20,75%
6 T-Series	216 495 600 668	20,75%
7 2	152 638 999 634	14,63%
8 Cocomelon - Nursery Rhymes	152 638 999 634	14,63%
9 3	140 138 068 504	13,43%
10 SET India	140 138 068 504	13,43%
11 4	92 952 274 861	8,91%
12 Sony SAB	92 952 274 861	8,91%
13 5	88 452 629 066	8,48%
14 âœ¿ Kids Diana Show	88 452 629 066	8,48%
15 6	88 060 349 741	8,44%
16 Like Nastya	88 060 349 741	8,44%
17 7	74 447 865 775	7,13%
18 WWE	74 447 865 775	7,13%
19 8	73 333 582 362	7,03%
20 Vlad and Niki	73 333 582 362	7,03%
21 9	58 923 017 461	5,65%
22 Movieclips	58 923 017 461	5,65%
23 10	58 056 997 206	5,56%
24 Colors TV	58 056 997 206	5,56%
25 Total général	1 043 499 385 278	100,00%

- ▶ Filtrage selon les Genres et Années
- ▶ Ajout des %, pour une autre approche de lecture
- ▶ Limiter à 10 lignes pour une meilleur visibilité
- ▶ Focus sur une seule catégorie

Exemple ici Vues par vidéo

Présentation du projet

Générer des graphiques par TCD



- ▶ Filtrage selon les Genres et Années
- ▶ Ajout des %, pour une autre approche de lecture
- ▶ Focus sur une seule catégorie

Exemple ici Vues par vidéo

Présentation du projet

Création des TreeMap

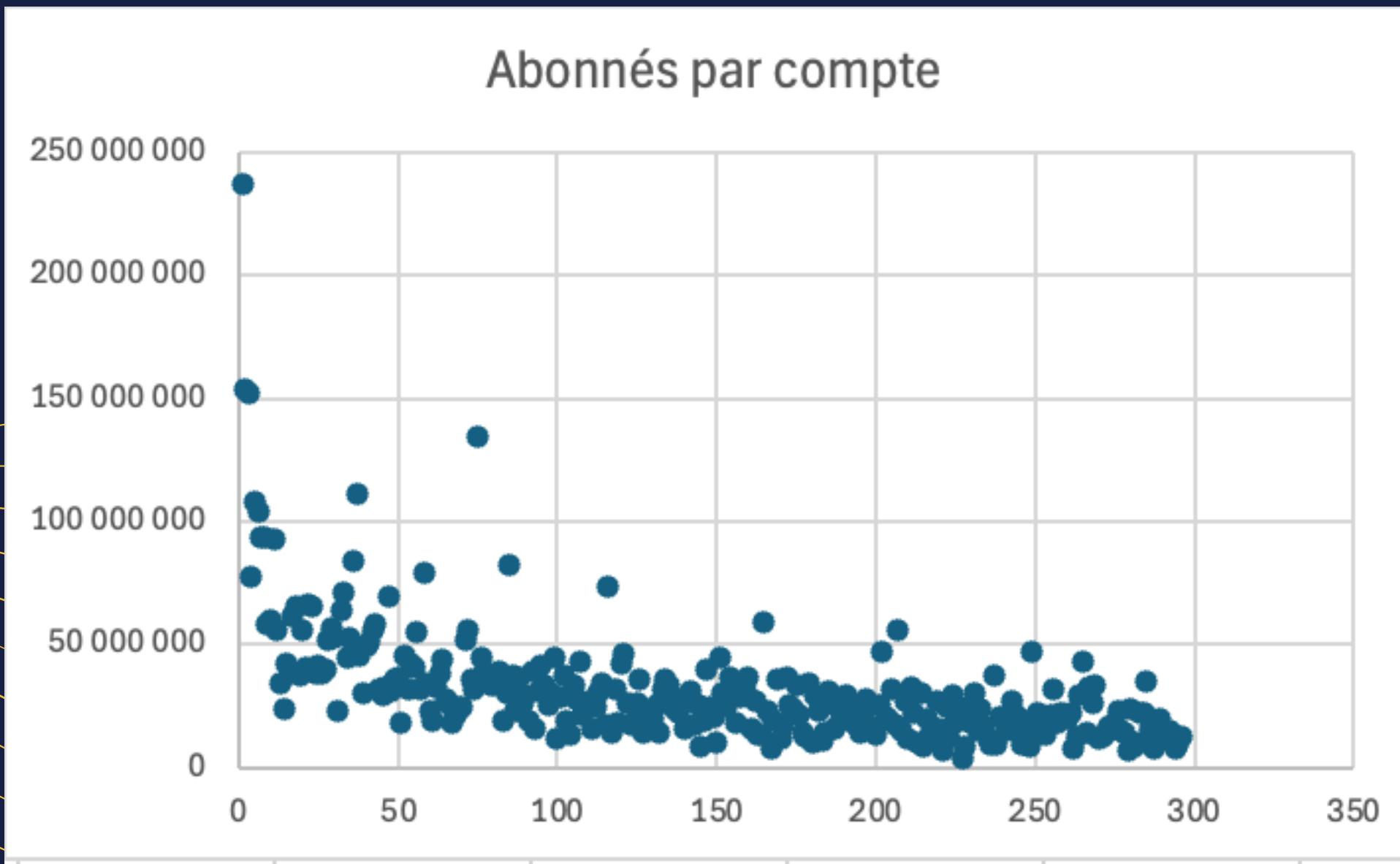


▶ Lecture colorée des proportions

Exemple ici Nombre de vues par rapport aux différents genre

Présentation du projet

Création des Nuages de points



▶ Lecture par points des proportions

Exemple ici Abonnés par compte

Présentation du projet

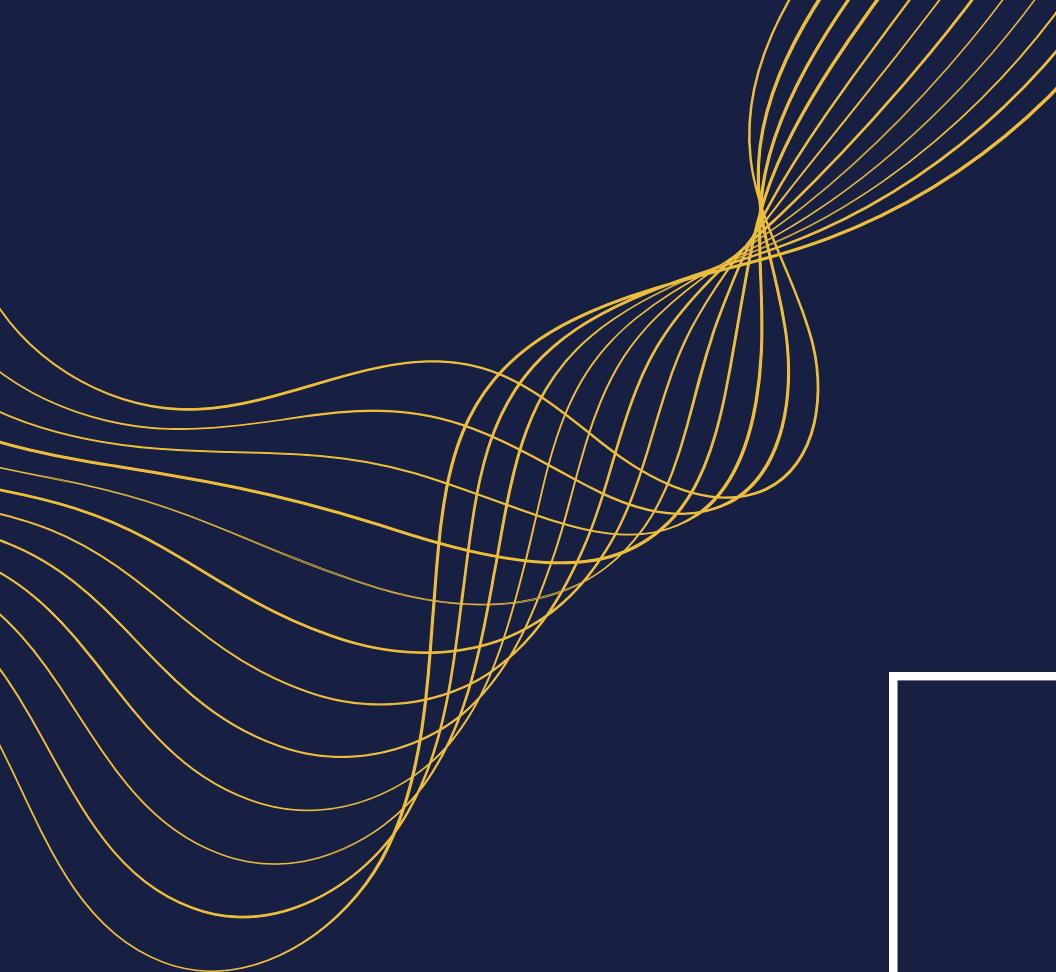
Coefficients de corrélation

Exemple d'une formule pour calculer le coefficient de corrélation entre le nombre d'abonnés par compte et le nombre de Vues par vidéo

=COEFFICIENT.CORRELATION(C1:C297;E1:E297)

**Mon résultat = 1
Bon résultat = 0,8...**

- +1 : relation linéaire parfaite
- 0 : pas de relation linéaire
- -1 : relation inverse parfaite
- Règle de pouce : ~0.2 faible ~0.5 moyenne
≥0.7 forte.



DÉMO

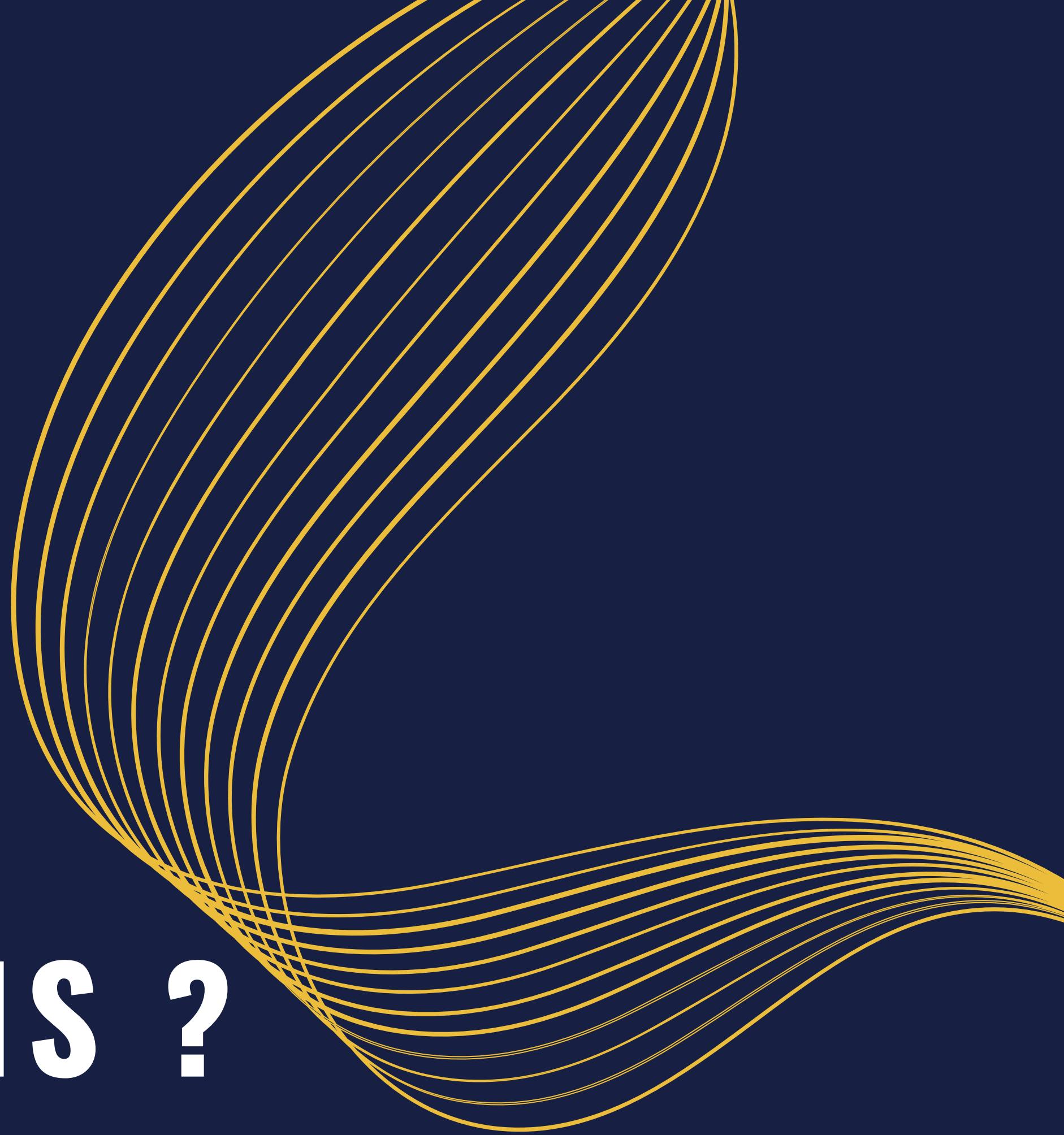
CONCLUSION

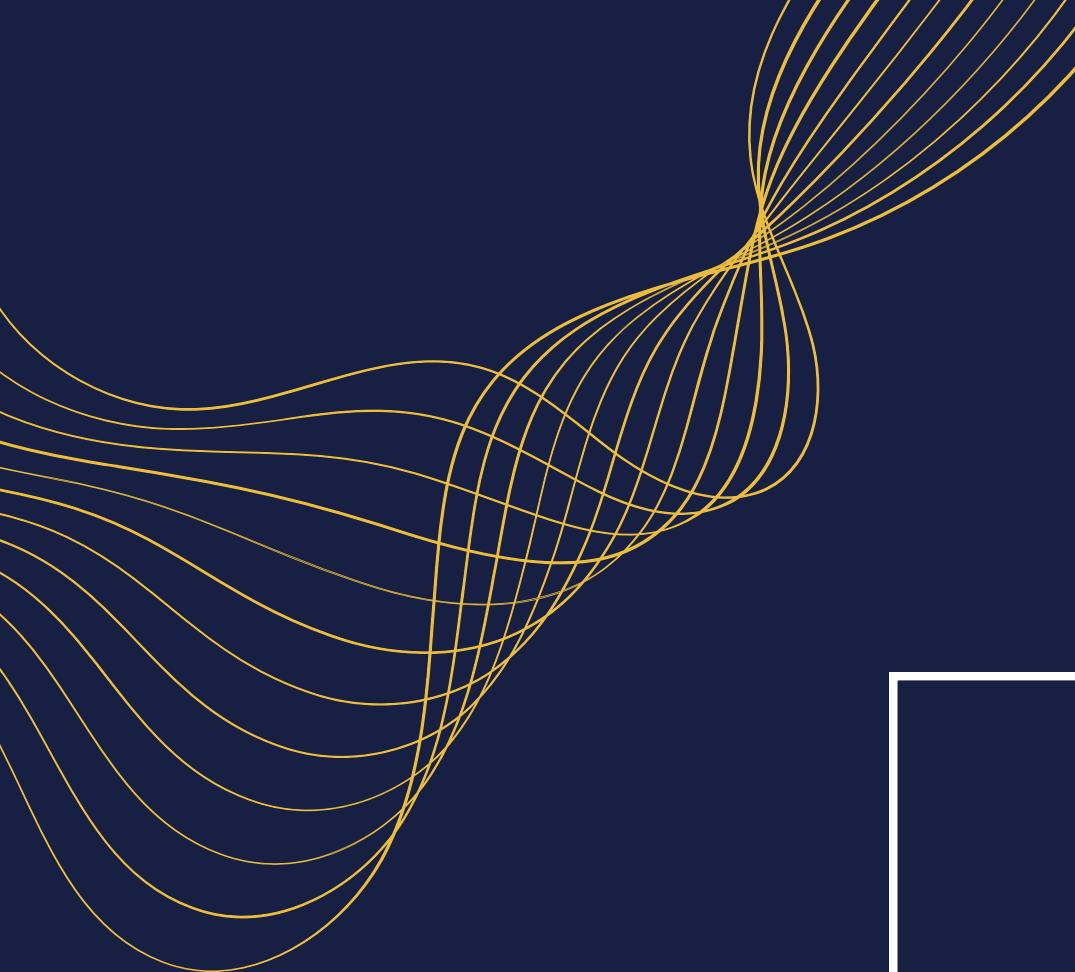


Rush 1

- Difficulté : ★
- Apprentissage : ★★★★★
- Compréhension : ★★★★★
- Fun ★★★★★

**AVEZ-VOUS
DES QUESTIONS ?**





RUSH 1

Data Analyst

Youtube

MERCI POUR VOTRE ATTENTION