# Εργασία στο μάθημα Διαχείριση Μεγάλων Δεδομένων - MongoDB

Στοιχεία Φοιτητών: ΠΑΣΣΟΣ ΔΗΜΗΤΡΙΟΣ 2024201000083 tst10083@go.uop.gr ΜΠΡΑΟΥΝΙ ΓΙΟΑΝΙ 2022201800131 Dit18131@go.uop.gr

#### Λεπτομέριες σχετικά με την υλοποίηση της εργασίας

- Η εργασία υλοποιήθηκε στο MongoDB Compass για την κατασκευή του aggregation pipeline σε συνδιασμό με το πρόγραμμα studio 3T για την εξαγωγή των αποτελεσμάτων. Δεν χρειάστηκε να κάνουμε κάποια μετατροπή των αρχείων εισόδου από csv σε json, διότι και τα 2 προγράμματα δέχονται αρχεία της μορφής csv.
- Επιπρόσθετα, τα αρχεία που χρειάζονται για την υλοποίηση είναι τα GBvideos και USvideos. Ωστόσο, για τα ερωτήματα 2.3 και 2.5, επειδή θέλουμε το πλήθος των βίντεο, έπρεπε να φιλτράρουμε τις διπλοεγγραφές που υπάρχουν στο αρχείο csv με το ίδιο video\_id, για να είναι τα video\_id μοναδικά. Για την δουλειά αυτή, προτιμήσαμε να χρησιμοποιήσουμε κάποιο extension στο libre office αντί να κατασκευάσουμε δικό μας κώδικα.

https://extensions.libreoffice.org/en/extensions/show/remove-duplicatesfast

#### Λεπτομέριες σχετικά με την υλοποίηση της εργασίας

- Για την γραφική απείκονιση των αποτελεσμάτων, χρησιμοποιήσαμε το atlas του MongoDB για το Boxplot.Για το scatter plot χρησιμοποιήσαμε ένα απλό εργαλειό το <a href="https://www.csvplot.com">https://www.csvplot.com</a>. Θα μπορούσαμε να το κάνουμε και στο atlas ωστόσο δεν εμφάνιζε καθαρά τις τιμές.
  - Έχουμε υλοποιήσει και τα 2 bonus όπως αυτά περιγράφονται στην αναφορά στην ενότητα bonus υλοποιήση. Για το 10 bonus, απαραίτητη προυπόθεση αποτελεί η προσθήκη στο Database, τα documents με κατάληξη. csv και. json και μόνο για τα ερωτήματα στα οποίο έχει νόημα. Το bonus 2 έχει υλοποιηθεί ξεχωριστά.

## 2.1 Τι ξέρουμε για τις δημοσιεύσεις του πολύ δημοφιλούς καναλιού Saturday Night Live - Σχολιασμός αποτελεσμάτων

Το κάναλι SNL είναι μια Αμερικάνικη παραγωγή, το οποίο προβάλεται στην τηλεόραση, κομικού χαρακτήρα με σκόπο να σατηρίσει κάποιο δημόσιο πρόσωπο ή οτιδήποτε άλλο το οποίο βρίσκεται στις τάσεις και συζητιέται αρκέτα.

Το video με τίτλο Royal Wedding το οποίο ήταν στην επικαιρότητα για αρκετές ημέρες και συζητήθηκε από πολλούς στα κοινωνικά μέσα , βλέπουμε ότι ειναι πρώτο με τις μεγαλύτερες προβολές.Το 2ο βίντεο που ακολουθεί αφόρα τον Kanye West.

## 2.1 Τι ξέρουμε για τις δημοσιεύσεις του πολύ δημοφιλούς καναλιού Saturday Night Live - Εμφάνιση των αποτελεσμάτων

_id	views	likes	dislikes
"_"Royal Wedding - SNL	<b>⊞</b> 8607264	<u>"</u> " 66559	<u>""</u> 14179
"_"A Kanye Place - SNL	j32 5547578	"_" 110621	<u>"-"</u> 5101
"_"Natalie's Rap 2 - SNL	j32 5156609	"_" 79865	"_" 3195
"_"George W. Bush Returns Cold Open - SNL	5147621	<u>"-"</u> 55076	<b>"_"</b> 6752
"_"Welcome to Hell - SNL	i₃ 4649310	<u>"-"</u> 60641	<b>"_"</b> 8997
"_"Meet the Parents Cold Open - SNL	i32 4548677	<u>"-"</u> 40718	<b>"_"</b> 4666
Morning Joe Michael Wolff Cold Open - SNL	i32 4162540	"_" 31716	<b>"_"</b> 3658
Presidential Address Cold Open - SNL	3763816	<b>""</b> 31791	<b>"_"</b> 3603
What Even Matters Anymore - SNL	3362428	<u>"</u> " 35137	<b>"_" 1</b> 0550
"-"Visit with Santa Cold Open - SNL	i32 2935898	<u>"_"</u> 28777	<b>"_"</b> 4642
White House Tree Trimming Cold Open - SNL	i32 2701442	"_" 27542	<b>"_"</b> 5428
Eminem: Walk on Water, Stan, Love the Way You Lie (ft. Skylar Grey) (Live) - SNL	i32 2200854	"_" 36494	"_" 3044
""Weekend Update: Stefon on St. Patrick's Day - SNL	i32 2189861	<b>""</b> 23694	<b>"_"</b> 761
"-"Weekend Update on Donald Trump's Asia Trip - SNL	i32 1907912	"_" 18641	<u>"-"</u> 1741
""Queer Eye's Tan France Takes Pete Davidson Shopping - SNL	i32 1856381	<b>""</b> 32619	<b>"_" 1483</b>
Taylor Swift:Ready for It? (Live) - SNL	i32 1776127	"_" 35741	<b>"_"</b> 3214
Family Feud: Oscars Edition - SNL	i32 1352472	<u>"</u> 16328	<b>"_"</b> 857
""Handmaids in the City - SNL	்⊞ 1250529	<u>""</u> 9870	<b>"_"</b> 2437
Cut for Time: New Year's Kiss - SNL	i3₹ 830143	"_" 10478	<b>"-"</b> 871

2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις δημοσιεύσεις των βίντεο - Σχολιασμός αποτελεσμάτων

Παρατηρούμε ότι ο κανονάς όσο ποιό πολλά tags τόσες περισσότερες προβολές δεν ισχυεί για τα video στο youtube. Αυτό οφείλεται στο γεγονός ότι τα video με πολλά tags , η πλειοψηφία των tags δεν αντικατροπτίζουν το περιεχόμενο ενός βίντεο αλλά έχουν προστεθεί από τον δημιουργό έτσι ώστε να μπερδέψουν τον αλγόριθμο του youtube και σε κάθε αναζήτηση που κάνει ο χρήστης να εμφανίζεται υψηλότερα στην λίστα, ανεξάρτητα εάν το περιεχόμενο του video είναι σχετικό με αυτό που θέλει να δεί ο χρήστης.Τέλος διακρίνουμε κάποιες εξαιρέσεις, όπως για παράδειγμα ένα βίντεο με id " Ι D 8Z4sJE" έχει 424 εκατομμυρία προβολές

#### 2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις δημοσιεύσεις των βίντεο - Εμφάνιση των αποτελεσμάτων

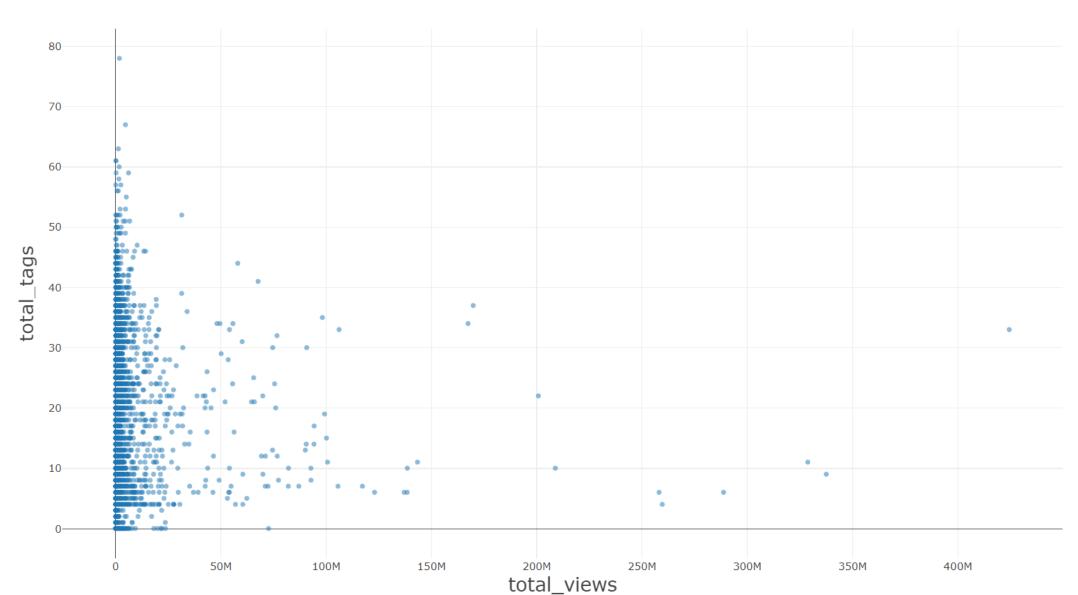
_id	total_tags	total_views
"_"_I_D_8Z4sJE	<b>33</b> 33 33 33 33 33 33 33 33 33 33 33 33	i₃⊋ 424538912
"_"9jI-z9QN6g8	9 SEi	337621571
"_"kLpH1nSLJSs	i32 11	328860380
"_"wfWkmURBNv8	i32 6	3至 288811992
"_"VYOjWnS4cMY	i32 4	32 259721696
"_"xpVfcZ0ZcFM	i32 6	258164991
"_"ffxKSjUwKdU	i32 10	迎 208876887
"_"zEf423kYfqk	i∃2 22	迎 200862743
"_"FIsCjmMhFmw	i∃2 37	169884583
"_"sGIm0-dQd8M	i32 34	⅓ 167456025
"_" TyHvyGVs42U	i32 11	143408235
"_"2Vv-BfVoq4g	i32 10	⅓ 138578860
"_"M4ZoCHID9GI	i32 6	⅓ 138535053
"_" Ck4xHocysLw	i32 6	⅓ 137081637
"_"7C2z4GqqS5E	i32 6	iii 123010920
"_"tCXGJQYZ9JA	i32 7	i32 117270304
"_"U9BwWKXjVaI	i⊞ 33	i32 106147032
"_"au2n7VVGv_c	i32 7	i32 105629911
"_"6ZfuNTqbHE8	<u>₃₃₂</u> 11	ii 100672931

#### 2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις δημοσιεύσεις των βίντεο(Bonus) - Εμφάνιση των αποτελεσμάτων

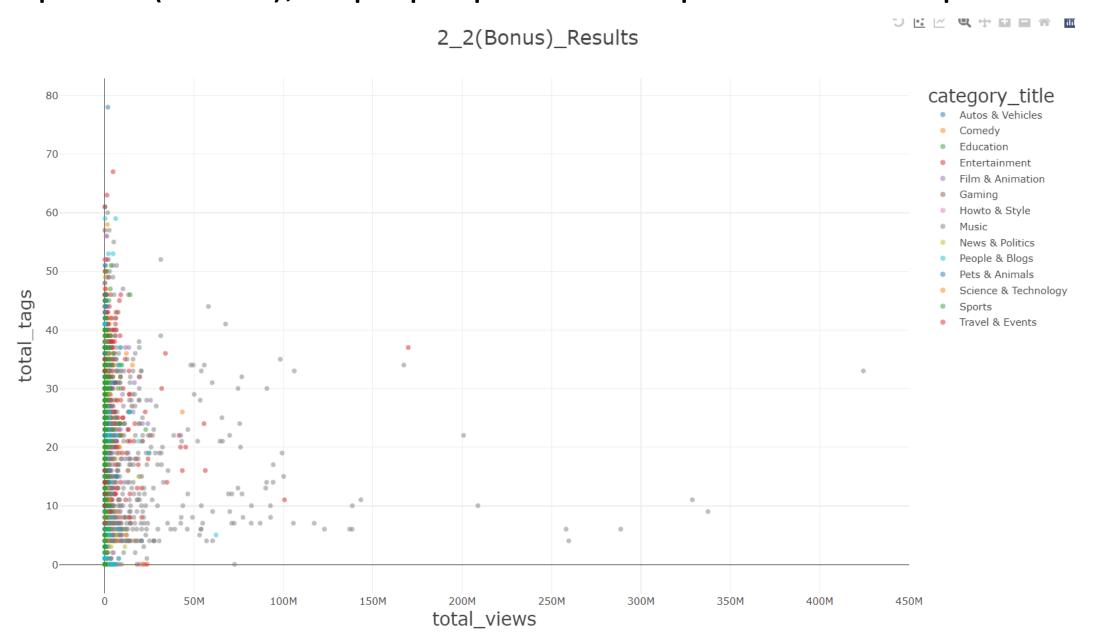
IIII kLpH1nSLJSs         i32         11         i32         328860380         IIII Music           IIII wfWkmURBNv8         i32         6         i32         288811992         IIII Music           IIII vYOjWnS4cMY         i32         4         i32         259721696         IIII Music           IIII xpVfcZ0ZcFM         i32         6         i32         258164991         IIII Music           IIII xpVfcZ0ZcFM         i32         10         i32         208876887         IIII Music           IIII zpf423kYfqk         i32         22         i32         200862743         IIII Music           IIII zpf423kYfqk         i32         32         169884583         IIII Music           IIII zpf456025         IIII Music         IIII Music           IIII zpf456025         IIII Mu				
IIII klpH1nSLJSs         j32         11         j32         328860380         IIII Music           IIII wfWkmURBNv8         j32         6         j32         288811992         IIII Music           IIII VYOjWnS4cMY         j32         4         j32         259721696         IIII Music           IIII xpVfcZ0ZcFM         j32         6         j32         258164991         IIII Music           IIII xpVfcZ0ZcFM         j32         10         j32         208876887         IIII Music           IIII xpVfqXx4Si         j32         22         j32         200862743         IIII Music           IIII xpVfqXx4Si         j32         37         j32         169884583         IIII Entertainment           IIII xpVfqXyGVs42U         j32         34         j32         167456025         IIII Music           IIII xpVfqXyGVs42U         j32         11         j32         143408235         IIII Music           IIII xpVfqXyGVs42U         j32         10         j32         138578860         IIII Music           IIII xpVfqXyGVs4         j32         6         j32         137081637         IIII Music           IIII xpVfqXyGVs4         j32         6         j32         137081637         IIII Music <td>"_"_I_D_8Z4sJE</td> <td>i32 33</td> <td>i₃⊋ 424538912</td> <td>"_" Music</td>	"_"_I_D_8Z4sJE	i32 33	i₃⊋ 424538912	"_" Music
III. wfWkmURBNv8         i32 6         i32 288811992         III. Music           III. VYOjWnS4cMY         i32 4         i32 259721696         III. Music           III. xpVfcZ0ZcFM         i32 6         i32 258164991         III. Music           III. ffxKSjUwKdU         i32 10         i32 208876887         III. Music           III. zEf423kYfqk         i32 22         i32 200862743         III. Music           III. zef423kYfqk         i32 37         i32 169884583         III. Entertainment           III. sGIm0-dQd8M         i32 34         i32 167456025         III. Music           III. yHvyGVs42U         i32 11         i32 143408235         III. Music           III. 2Vv-BfVoq4g         i32 10         i32 138578860         III. Music           III. AVACCHID9GI         i32 6         i32 137081637         III. Music           III. Ck4xHocysLw         i32 6         i32 137081637         III. Music           III. TCZ24GqqS5E         i32 6         i32 123010920         III. Music           III. TCXGJQYZ9JA         i32 7         i32 117270304         III. Music           III. U9BwWKXjVal         i32 3         i32 106147032         III. Music           III. U9BwWKXjVal         i32 7         i32 105629911         III. Music <td>9jI-z9QN6g8</td> <td>i32 9</td> <td>₃ 337621571</td> <td>"_" Music</td>	9jI-z9QN6g8	i32 9	₃ 337621571	"_" Music
************************************	"_"kLpH1nSLJSs	i32 11	i₃ 328860380	"_" Music
***ExpVfcZ0ZcFM         i32 6         i32 258164991         ***Music           ***EffxKSjUwKdU         i32 10         i32 208876887         ***Music           ***Ezf423kYfqk         i32 22         i32 200862743         ***Music           ***EffsCjmMhFmw         i32 37         i32 169884583         ***Entertainment           ***SGIm0-dQd8M         i32 34         i32 167456025         ***Music           ***TyHvyGVs42U         i32 11         i32 143408235         ***Music           ***2Vv-BfVoq4g         i32 10         i32 138578860         ***Music           ***M4ZoCHID9GI         i32 6         i32 138535053         ****Music           ****Ck4xHocysLw         i32 6         i32 137081637         ****Music           ****TC2z4GqqS5E         i32 6         i32 123010920         ****Music           ****TCXGJQYZ9JA         i32 7         i32 117270304         ****Music           *****U9BwWKXjVal         i32 3         i32 106147032         ****Music           ************************************	"_"wfWkmURBNv8	i32 6	<b>32</b> 288811992	"_" Music
Image: Control of the contro	"_"VYOjWnS4cMY	i32 4	<b>32</b> 259721696	"_" Music
""zEf423kYfqk         332         22         332         200862743         "" Music           ""FISCjmMhFmw         332         37         332         169884583         "" Entertainment           ""sGlm0-dQd8M         332         34         332         167456025         "" Music           ""TyHvyGVs42U         332         11         332         143408235         "" Music           ""2Vv-BfVoq4g         332         10         332         138578860         "" Music           ""M4ZoCHID9GI         332         6         332         138535053         "" Music           ""Ck4xHocysLw         332         6         332         137081637         "" Music           ""TC2z4GqqS5E         332         6         332         123010920         "" Music           ""tCXGJQYZ9JA         332         7         332         117270304         "" Music           ""U9BwWKXjVal         332         33         332         106147032         "" Music           ""au2n7vvGv_c         332         7         332         105629911         "" Music	"_"xpVfcZ0ZcFM	<u>i32</u> 6		"_" Music
Image: Sign of the content o	"_"ffxKSjUwKdU	j32 10	<b>32</b> 208876887	"_" Music
"""sGIm0-dQd8M       332 34       332 167456025       """ Music         """TyHvyGVs42U       332 11       332 143408235       """ Music         """2Vv-BfVoq4g       332 10       332 138578860       """ Music         """M4ZoCHID9GI       332 6       332 138535053       """ Music         """Ck4xHocysLw       332 6       332 137081637       """ Music         """7C2z4GqqS5E       332 6       332 123010920       """ Music         """tCXGJQYZ9JA       332 7       332 117270304       """ Music         """U9BwWKXjVal       332 33       332 106147032       """ Music         """au2n7VVGv_c       332 7       332 105629911       """ Music	"_"zEf423kYfqk	<u>i3₹</u> 22	32 200862743	"_" Music
"""TyHvyGVs42U       i32       11       i32       143408235       "" Music         """2Vv-BfVoq4g       i32       10       i32       138578860       "" Music         """M4ZoCHID9GI       i32       6       i32       138535053       "" Music         """Ck4xHocysLw       i32       6       i32       137081637       "" Music         """7C2z4GqqS5E       i32       6       i32       123010920       "" Music         """tCXGJQYZ9JA       i32       7       i32       117270304       "" Music         """U9BwWKXjVal       i32       33       i32       106147032       "" Music         """au2n7VVGv_c       i32       7       i32       105629911       "" Music	"_" FIsCjmMhFmw	i3₹ 37	<b>169884583</b>	"_" Entertainment
""2Vv-BfVoq4g       iii 10       iii 138578860       "" Music         ""M4ZoCHID9GI       iii 6       iii 138535053       "" Music         ""Ck4xHocysLw       iii 6       iii 137081637       "" Music         ""7C2z4GqqS5E       iii 6       iii 123010920       "" Music         ""tCXGJQYZ9JA       iii 7       iii 117270304       "" Music         ""U9BwWKXjVal       iii 32 33       iii 106147032       "" Music         ""au2n7VVGv_c       iii 105629911       "" Music	"_"sGlm0-dQd8M	i32 34	<b>32</b> 167456025	"_" Music
"-" M4ZoCHID9GI       jii 6       jiii 138535053       "-" Music         "-" Ck4xHocysLw       jii 6       jiii 137081637       "-" Music         "-" 7C2z4GqqS5E       jii 6       jiii 123010920       "-" Music         "-" tCXGJQYZ9JA       jii 7       jii 117270304       "-" Music         "-" U9BwWKXjVal       jii 33       jii 106147032       "-" Music         "-" au2n7VVGv_c       jii 7       jii 105629911       "-" Music	"_"TyHvyGVs42U	i32 11	<b>143408235</b>	"_" Music
"_"Ck4xHocysLw       32 6       32 137081637       "_" Music         "_"7C2z4GqqS5E       32 6       32 123010920       "_" Music         "_"tCXGJQYZ9JA       32 7       32 117270304       "_" Music         "_"U9BwWKXjVal       32 33       32 106147032       "_" Music         "_"au2n7VVGv_c       32 7       32 105629911       "_" Music	"_"2Vv-BfVoq4g	j32 10		"_" Music
"-"7C2z4GqqS5E       32 6       32 123010920       "-" Music         "-"tCXGJQYZ9JA       32 7       32 117270304       "-" Music         "-"U9BwWKXjVal       32 33       32 106147032       "-" Music         "-"au2n7VVGv_c       32 7       32 105629911       "-" Music	"_"M4ZoCHID9GI	i32 6	138535053	"_" Music
"_"tCXGJQYZ9JA       332 7       332 117270304       "_" Music         "_"U9BwWKXjVal       332 33       332 106147032       "_" Music         "_"au2n7VVGv_c       332 7       332 105629911       "_" Music	"_"Ck4xHocysLw	i32 6	<b>137081637</b>	"_" Music
"_"U9BwWKXjVal       33       32       106147032       "_" Music         "_"au2n7VVGv_c       32       7       32       105629911       "_" Music	"_"7C2z4GqqS5E	i32 6	iii 123010920	"_" Music
"_"au2n7VVGv_c	"_"tCXGJQYZ9JA	i32 7	iii 117270304	"_" Music
	"_"U9BwWKXjVaI	i∃2 33	₃ 106147032	"_" Music
"_"6ZfuNTqbHE8	"_"au2n7VVGv_c	i32 7	105629911	"_" Music
	"_"6ZfuNTqbHE8	i32 11	iii 100672931	"_" Entertainment

#### 2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις δημοσιεύσεις των βίντεο; - Γραφική απεικόνιση των αποτελεσμάτων

2.2\_Results



#### 2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις δημοσιεύσεις των βίντεο(Bonus); - Γραφική απεικόνιση των αποτελεσμάτων

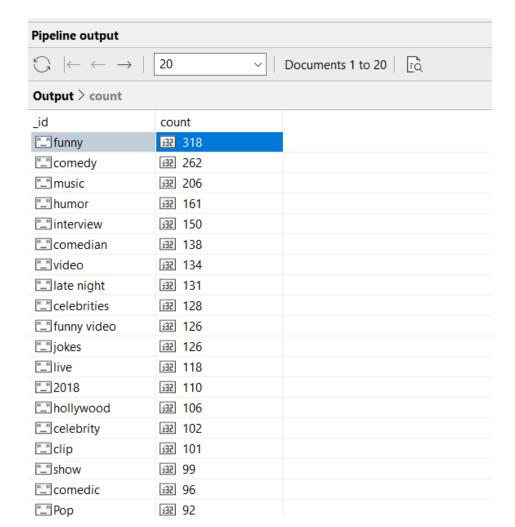


2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Σχολιασμός αποτελεσμάτων

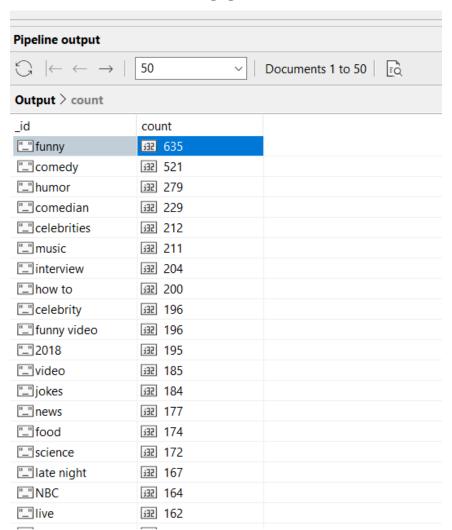
Οι πιο δημορφιλείς ετικέτες στα βίντεο GB και US είναι τα funny ,comedy και music, δηλαδή στα βίντεο όπου έχουν στόχο να διασκεδάσουν τον χρήστη.

#### 2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Εμφάνιση των αποτελεσμάτων

GB



US



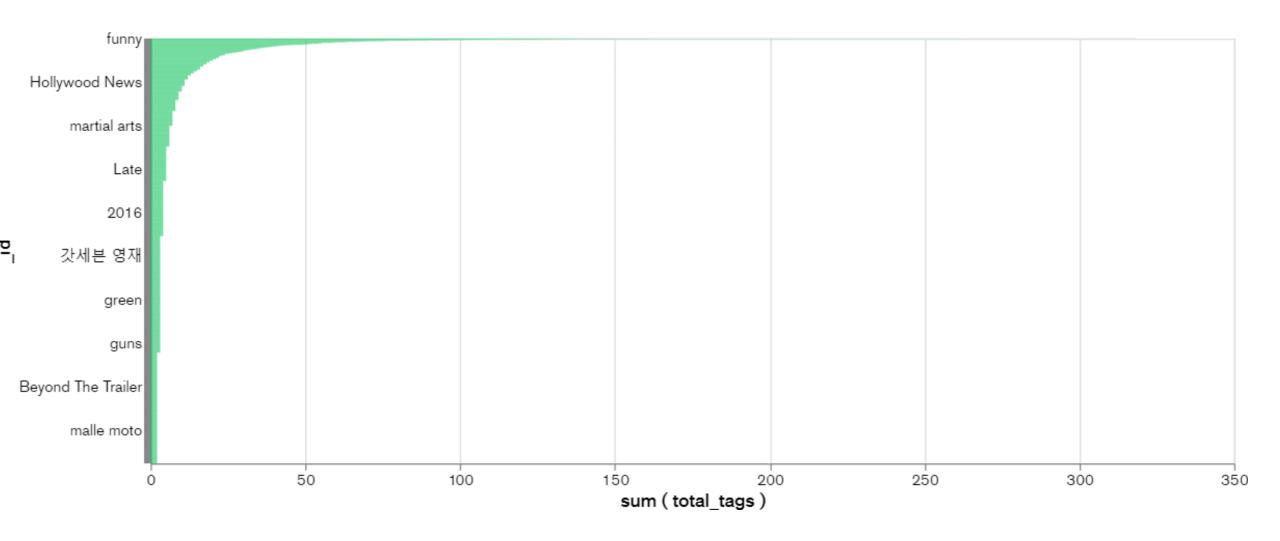
### 2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Εμφάνιση των αποτελεσμάτων GBvideos(Bonus)

				1					
• _id	• total_tags	• categ	• 0 (category_titles.0)	• 1 (category_titles.1)	• 2 (category_titles.2)	• 3 (category_titles.3)	• 4 (category_titles.4)	• 5 (category_titles.5)	6 (category_titles.)
"_" funny	1≥3 318.0	[ 12	"_" Howto & Style	"_" Comedy	" Sports	Pets & Animals	"_" Education	People & Blogs	Science & Tech
"_" comedy	123 262.0	[] [9	"_" Howto & Style	"_" News & Politics	" Comedy	"_" Sports	"_" People & Blogs	"_" Music	" Film & Animati
"_" music	123 206.0	[] [9	"_" Howto & Style	"_" Comedy	"_" Sports	"_" Education	"_" Music	"_" People & Blogs	"_" Film & Animati
"_" humor	161.0	[] [6	"_" Music	"_" Comedy	"_" News & Politics	"_" Gaming	"_" Entertainment	" People & Blogs	
"_" interview	150.0	[] [10	"_" Howto & Style	"_" Comedy	"_" News & Politics	"_" Sports	"_" Education	" People & Blogs	"_" Music
"_" comedian	138.0	[] [6	"_" Music	"_" Entertainment	People & Blogs	" Film & Animati	"_" Comedy	" News & Politics	
"_" video	134.0	[] [10	"_" Sports	"_" Comedy	"_" News & Politics	"_" Howto & Style	"_" Pets & Animals	"_" Music	People & Blogs
"_"late night	131.0	[] [4	"_" Comedy	"_" Entertainment	"_" Music	People & Blogs			
"_" celebrities	128.0	[] [5	"_" Comedy	"_" Entertainment	People & Blogs	"_" Music	"_" Howto & Style		
"_" funny video	126.0	[] [6	"_" Comedy	"_" Pets & Animals	Entertainment	"_" Music	"_" People & Blogs	"_" Howto & Style	
"_"jokes	126.0	[] [4	"_" Comedy	" Gaming	"_" Entertainment	People & Blogs			
"_" live	118.0	[] [7	"_" Science & Tech	" News & Politics	"_" Sports	"_" Entertainment	"_" Music	"_" Travel & Events	"_" Gaming
"_"2018	110.0	[] [10	"_" Howto & Style	"_" News & Politics	"_" Sports	"_" Comedy	"_" Education	"_" Film & Animati	People & Blogs
"_"hollywood	106.0	[] [5	"_" Entertainment	"_" People & Blogs	"_" Music	"_" Comedy	" News & Politics		
"_" celebrity	102.0	[] [6	"_" Howto & Style	"_" Comedy	"_" Music	"_" Entertainment	"_" People & Blogs	" News & Politics	
"_"clip	101.0	[] [5	"_" Comedy	"Film & Animati	"_" Music	"_" Entertainment	" Gaming		
"_" show	123 99.0	[] [8	"_" Sports	"_" Comedy	" Howto & Style	"_" People & Blogs	"_" Music	" Film & Animati	" Gaming
"_" comedic	123 96.0	[] [3	"_" Music	"_" Comedy	"_" Entertainment				
"_" Pop	123 92.0	[] [3	"Film & Animati	"_" Music	" Comedy				
"_"NBC	123 90.0	[] [4	"_" Music	"_" Entertainment	"_" Comedy	"_" News & Politics			

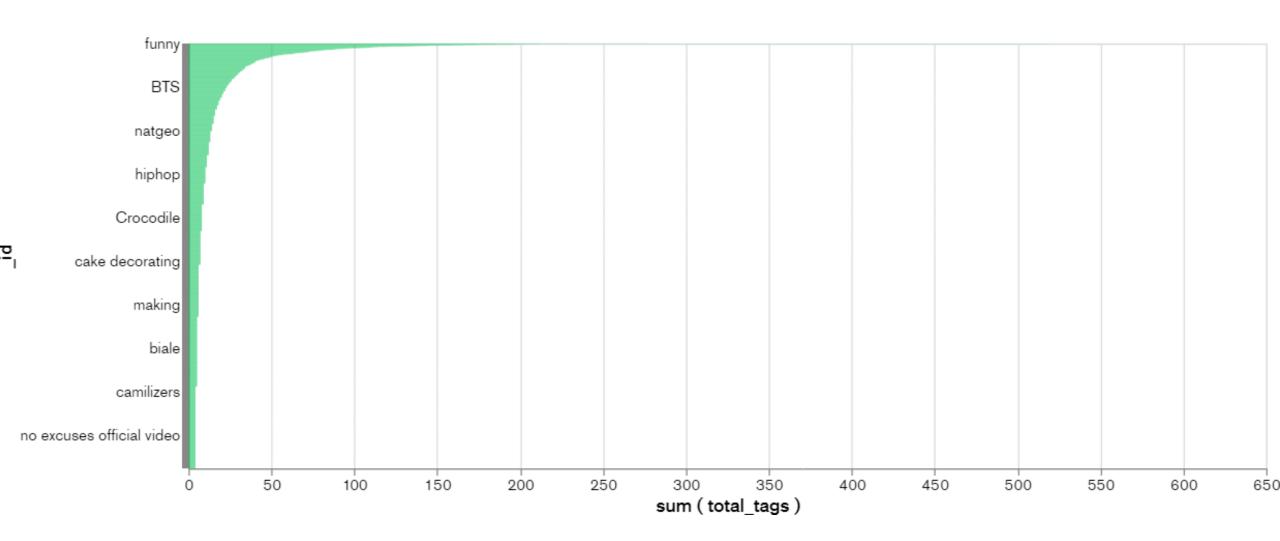
#### 2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Εμφάνιση των αποτελεσμάτων USvideos(Bonus)

USvideos > categor	ry_titles								
•_id	• total_tags	• category	• 0 (category_titles.0)	• 1 (category_titles.1)	• 2 (category_titles.2)	• 3 (category_titles.3)	• 4 (category_titles.4)	• 5 (category_titles.5)	• 6 (category_t
"_"funny	123 635.0	[] [ 13 ele	"_" Gaming	"_" Comedy	"_" Howto & Style	"_" Science & Tech	"_" News & Politics	"_" People & Blogs	"_" Education
"_"comedy	123 521.0	[] [ 10 ele	"_" Entertainment	"_" Gaming	"_" Sports	"_" Shows	People & Blogs	"Film & Animati	"_" Music
"_"humor	123 279.0	[] [7 ele	People & Blogs	"_" Education	"_" Comedy	" Howto & Style	" Film & Animati	"_" Music	"_" Entertainm
"_"comedian	123 229.0	[] [ 5 ele	"_" Comedy	" Howto & Style	"_" Entertainment	"Film & Animati	People & Blogs		
"_" celebrities	123 212.0	[] [ 5 ele	"_" Comedy	"-" Howto & Style	People & Blogs	"_" Entertainment	"_" Sports		
"_" music	123 211.0	[] [ 11 ele	"_" Education	"_" Music	People & Blogs	" Film & Animati	" Howto & Style	" Science & Tech	"_" Comedy
"_"interview	123 204.0	[] [ 9 ele	" News & Politics	"_" Comedy	" Howto & Style	" Science & Tech	"_" Music	"Film & Animati	"-" People & E
"_"how to	123 200.0	[] [ 12 ele	" Entertainment	"_" Howto & Style	"_" Comedy	Science & Tech	Pets & Animals	"_" Education	"-" People & E
"_"funny video	123 196.0	[] [ 5 ele	" Entertainment	"_" People & Blogs	"_" Howto & Style	"_" Comedy	Pets & Animals		
"_"celebrity	123 196.0	[] [8 ele	"_" Comedy	" News & Politics	"_" Sports	"_" Entertainment	"Film & Animati	People & Blogs	"_" Music
"_"2018	123 195.0	[] [ 13 ele	"_" Entertainment	"-" Howto & Style	"_" Comedy	"_" News & Politics	Pets & Animals	"_" Education	" Film & Ani
"_"video	185.0	[] [ 12 ele	" Entertainment	"_" Howto & Style	"_" Comedy	" News & Politics	Science & Tech	"_" Education	"_" Music
"_"jokes	184.0	[] [ 5 ele	"_" Sports	" Howto & Style	" Entertainment	" People & Blogs	"_" Comedy		
"_" news	123 177.0	[] [ 11 ele	" News & Politics	" Howto & Style	" Science & Tech	" Pets & Animals	"_" Music	People & Blogs	"_" Education
"_" food	174.0	[] [8 ele	"_" Education	People & Blogs	" Howto & Style	"_" Travel & Events	"_" Comedy	" News & Politics	"_" Science &
"_" science	172.0	[] [ 11 ele	" News & Politics	"_" Travel & Events	Pets & Animals	Science & Tech	"Film & Animati	"_" Education	"- People & E
"_"late night	123 167.0	[] [3 ele	"_" Entertainment	"_" Comedy	People & Blogs				
"_"NBC	164.0	[] [4 ele	"_" Comedy	"_" Sports	"_" Entertainment	" News & Politics			
"_"live	162.0	[] [7 ele	"_" Comedy	"_" News & Politics	People & Blogs	Science & Tech	"_" Entertainment	"_" Music	"_" Sports
"_"hollywood	159.0	[] [7 ele	People & Blogs	"_" Comedy	" News & Politics	"Film & Animati	" Gaming	"_" Music	"_" Entertainm
"_" show	158.0	[] [7 ele	"_" Entertainment	":" Film & Animati	"_" Howto & Style	"_" Comedy	Science & Tech	"_" Music	"-" People & E
"_"clip	152.0	[] [6 ele	" Film & Animati	People & Blogs	"_" Comedy	"_" Howto & Style	"_" Music	"_" Entertainment	

#### 2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Γραφική απεικόνιση των αποτελεσμάτων για τα GB videos



#### 2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στα ανερχόμενα βίντεο - Γραφική απεικόνιση των αποτελεσμάτων για τα US videos



2.4 Τι αντίκτυπο έχει στο κοινό η απενεργοποίηση των σχολίων - Σχολιασμός αποτελεσμάτων.

Τα βίντεο όπου τα σχόλια είναι απενεργοποιήμενα, παρατηρούμε ότι Είναι λιγότερο πιθανόν οι χρήστες να πατήσουν το κουμπί Like,χώρις Όμως να δημιουργούνται αντιπάθειες προς τον δημιουργό του βίντεο. Η απενεργοποίηση των σχολίων είναι μια καλή προσέγγιση που μπορεί να ακολουθήσει κάποιος, σε περίπτωση όπου το περιεχόμενο του βίντεο μπορεί να προκαλέσει διαφωνίες, διαφορετικές ερμηνείες ή στις περιπτώσεις όπου θίγει ένα θέμα taboo

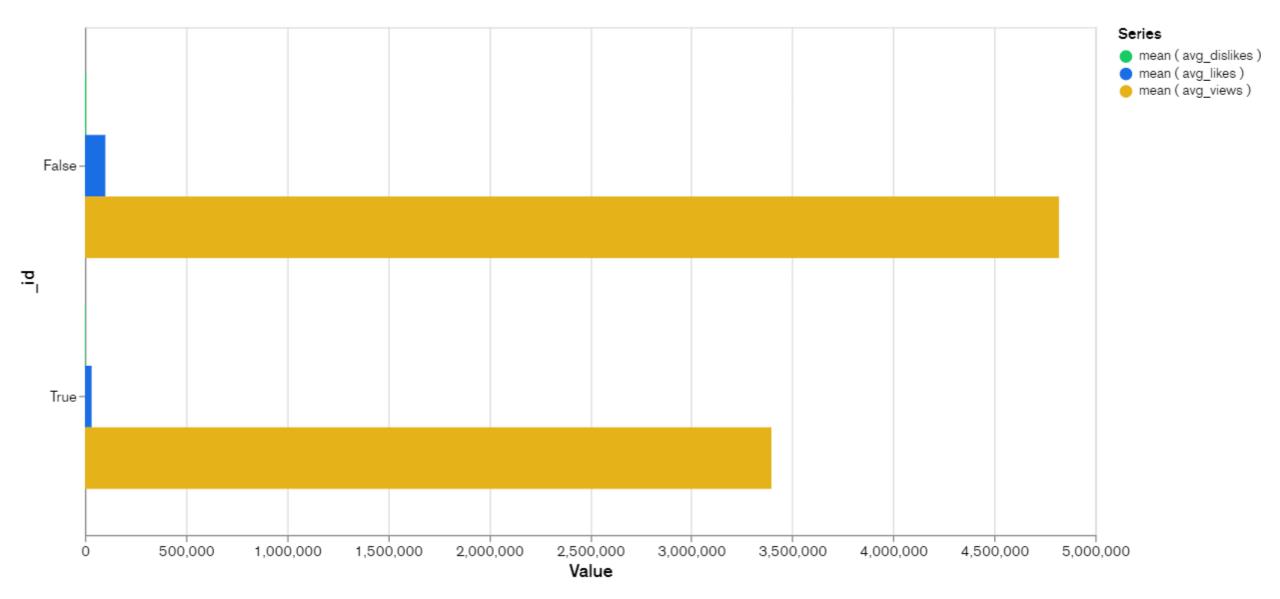
_id	avg_views	avg_likes	avg_dislikes
"_"True	13397406.16981	№ 33029.1698113	№ 3531.81132075
"_" False	1≥3 4831839.63497	100906.713575	№ 5921.66853059

#### 2.4 Τι αντίκτυπο έχει στο κοινό η απενεργοποίηση των σχολίων - Εμφάνιση των αποτελεσμάτων(Bonus).

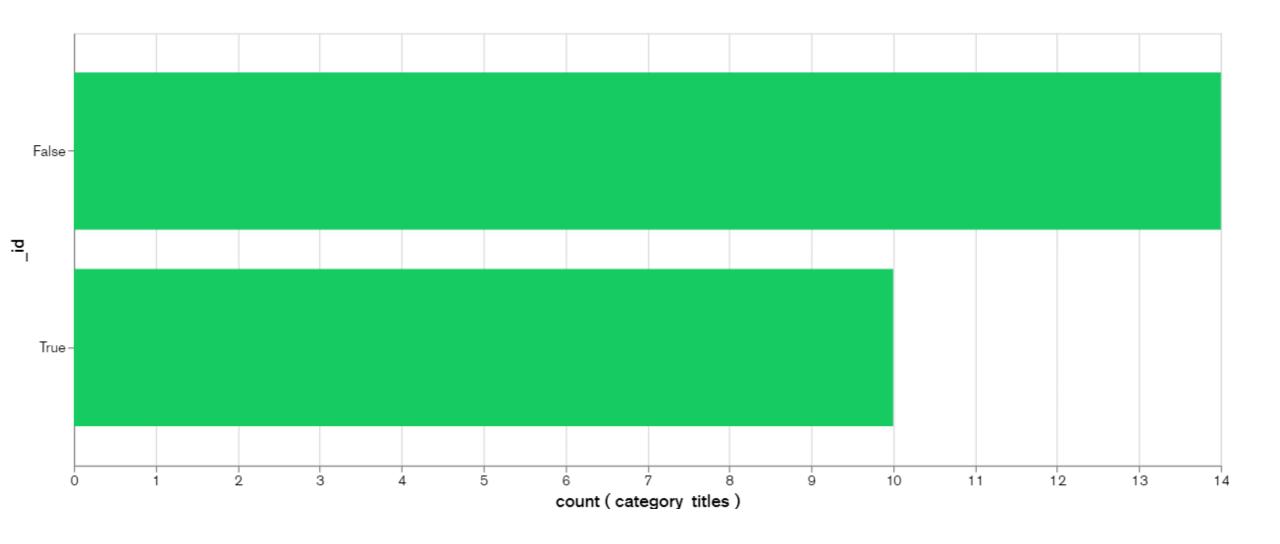
```
"_id" : "True",
"avg views" : 3396680.2264150945,
"avg likes" : 33014.56603773585,
"avg dislikes" : 3531.5471698113206,
"category titles" : [
    "Music",
    "Film & Animation",
    "News & Politics",
    "People & Blogs",
    "Sports",
    "Science & Technology",
    "Entertainment",
    "Autos & Vehicles",
    "Gaming",
    "Education"
```

```
"_id" : "False",
"avg views": 4807800.924680984,
"avg likes": 100230.5284780579,
"avg_dislikes" : 5749.71864301276,
"category titles" : [
    "Science & Technology",
    "Howto & Style",
    "Entertainment",
    "Travel & Events",
    "People & Blogs",
   "Education",
    "Music",
    "Film & Animation",
    "Gaming",
    "Comedy",
    "Autos & Vehicles",
    "News & Politics",
    "Dets & Animals"
```

### 2.4 Τι αντίκτυπο έχει στο κοινό η απενεργοποίηση των σχολίων - Γραφική Απεικόνιση των αποτελεσμάτων.



2.4 Τι αντίκτυπο έχει στο κοινό η απενεργοποίηση των σχολίων - Γραφική Απεικόνιση των τίτλων κατηγορίας.



#### 2.5 Ποιες ήταν οι πιο δημοφιλείς ημερομηνίες για δημοσίευση βίντεο - Σχολιασμός αποτελεσμάτων.

Παρατηρούμε ότι υπάρχει ένα μοτίβο ως προς το ρυθμό που δημοσιέυονται τα βίντεο.Ειδικότερα, 2 εβδομάδες πριν τις γιορτές των Χριστουγέννων και της Πρωτοχρονιάς ,οι δημοσιέυσεις των βίντεο αυξάνονται έτσι ωστέ να υπάρχει μια επάρκεια των βίντεο πρός παρακολούθηση στις γιορτές, ενώ κατά την διάρκεια των εορτών ότι οι δημιουργοί των βίντεων κάνουν μια πάυση για διακοπές και μειώνονται οι δημοσιεύεσεις.Επιπρόσθετα, μετά την Πρωτοχρονιά υπάρχει μια άυξηση για να καλήψει το κένο ώς προς το πλήθος των βίντεων που υπάρχουν διαθέσιμα για παρακολούθηση ,το οποίο είχε δημιουργηθεί στις γιορτές.Τέλος, το μήνα Φεβρουάριο ο ρυθμός των δημοσιεύσεων επανέρχεται πάλι στο κανονικό.

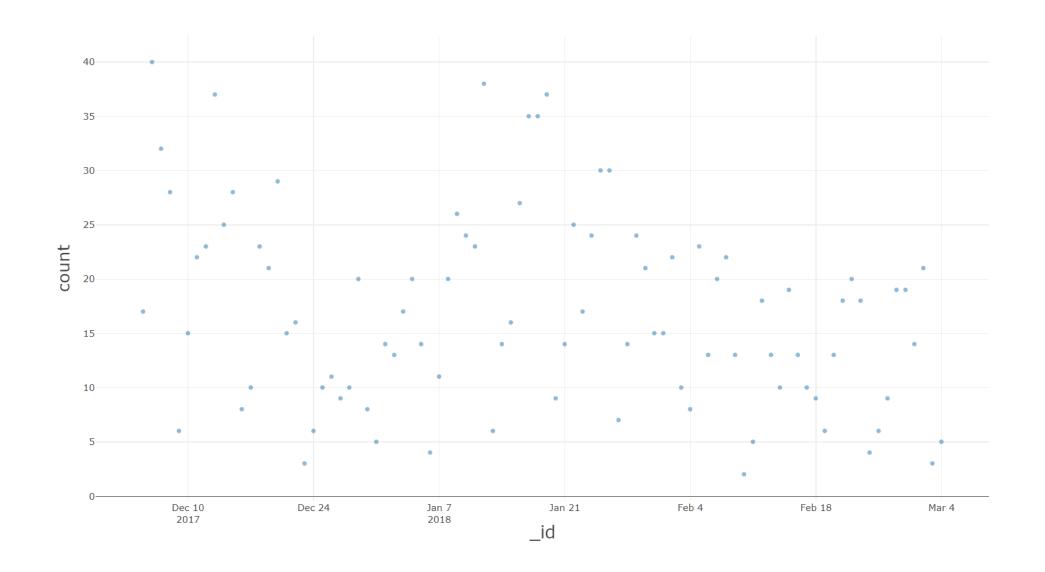
#### 2.5 Ποιες ήταν οι πιο δημοφιλείς ημερομηνίες για δημοσίευση βίντεο - Εμφάνιση των αποτελεσμάτων GBvideos.

Output > count		
_id	count	
"_"2017-12-05	17.0	
"_"2017-12-06	123 40.0	
"_"2017-12-07	123 32.0	
"_"2017-12-08	123 28.0	
<u>"</u> 2017-12-09	123 6.0	
"_"2017-12-10	15.0	
"_"2017-12-11	123 22.0	
"_"2017-12-12	1≥3 23.0	
"_"2017-12-13	123 37.0	
"_" 2017-12-14	123 25.0	
"_"2017-12-15	123 28.0	
"_"2017-12-16	123 8.0	
"_"2017-12-17	123 10.0	
"_"2017-12-18	123 23.0	
"_"2017-12-19	123 21.0	
"_"2017-12-20	123 29.0	
"_"2017-12-21	123 15.0	
"_"2017-12-22	16.0	
"_" 2017-12-23	123 3.0	

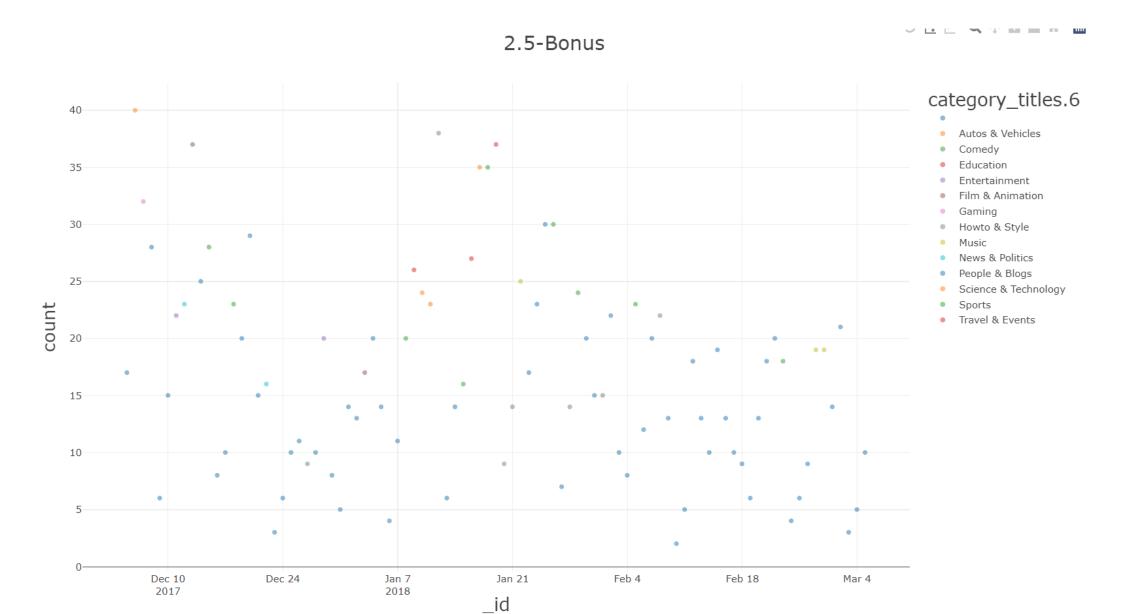
#### 2.5 Ποιες ήταν οι πιο δημοφιλείς ημερομηνίες για δημοσίευση βίντεο - Εμφάνιση των αποτελεσμάτων GBvideos(Bonus).

• _id	• count	category_titles	• 0 (category_titles.0)	• 1 (category_titles.1)	• 2 (category_titles.2)	• 3 (category_titles.3)	• 4 (category_titles.4)	• 5 (category_titles.5)
<u>"_"</u> 2017-12-05	17.0	[] [ 9 elements ]	"_" People & Blogs	"_" Howto & Style	"_" News & Politics	"_" Film & Animati	"_" Entertainment	"_" Music
"_"2017-12-06	123 40.0	[] [ 9 elements ]	"_" Music	" Gaming	" Howto & Style	"_" News & Politics	" Autos & Vehicles	People & Blogs
"_"2017-12-07	123 32.0	[] [ 10 elements ]	" People & Blogs	"_" Sports	" News & Politics	"_" Howto & Style	"_" Comedy	"_" Entertainment
"_"2017-12-08	123 28.0	[7 elements]	" Film & Animati	"_" Gaming	"_" Music	"_" People & Blogs	"_" Comedy	"_" Autos & Vehicles
"_" 2017-12-09	123 6.0	[] [5 elements]	"_" People & Blogs	"_" News & Politics	"_" Sports	"_" Gaming	"_" Entertainment	
"_" 2017-12-10	15.0	[ 5 elements ]	"_" Howto & Style	"_" Sports	"_" People & Blogs	"_" Entertainment	"_" Gaming	
"_" 2017-12-11	133 22.0	[] [8 elements]	"_" People & Blogs	"_" News & Politics	"_" Howto & Style	"_" Sports	"_" Film & Animati	"_" Music
"_" 2017-12-12	13 23.0	[] [8 elements]	"_" People & Blogs	" News & Politics	"_" Sports	"_" Music	"_" Entertainment	"_" Comedy
"_" 2017-12-13	123 37.0	[] [ 11 elements ]	"_" News & Politics	"_" Comedy	"_" Gaming	"_" Entertainment	"_" Music	" Film & Animati
"_" 2017-12-14	25.0	[] [7 elements]	"_" Sports	Entertainment	"_" Music	"_" Comedy	"_" Gaming	People & Blogs
"_" 2017-12-15	123 28.0	[] [ 10 elements ]	"_" Entertainment	"_" Music	"_" Comedy	" Autos & Vehicles	Science & Tech	"_" Gaming
"_" 2017-12-16	123 8.0	[] [ 4 elements ]	"_" Entertainment	Gaming	People & Blogs	"_" News & Politics		
"_" 2017-12-17	10.0	[] [ 6 elements ]	"_" Entertainment	Gaming	"_" Music	People & Blogs	Howto & Style	"_" News & Politics
"_" 2017-12-18	1≥3 23.0	[] [8 elements]	"_" Gaming	"_" Music	" Film & Animati	"_" Howto & Style	" News & Politics	People & Blogs
"_" 2017-12-19	123 20.0	[] [ 6 elements ]	"_" Comedy	Entertainment	"_" Music	"_" Gaming	Education	"_" Sports
"_" 2017-12-20	123 29.0	[] [ 9 elements ]	People & Blogs	News & Politics	"Film & Animati	"_" Entertainment	"_" Music	"_" Comedy
"_" 2017-12-21	15.0	[] [8 elements]	People & Blogs	" News & Politics	"_" Gaming	"_" Entertainment	"_" Music	"_" Comedy
"_" 2017-12-22	16.0	[] [8 elements]	"_" People & Blogs	"_" Howto & Style	"_" Sports	"_" News & Politics	"_" Music	"_" Entertainment
"_" 2017-12-23	123 3.0	[] [ 3 elements ]	"_" People & Blogs	"_" Howto & Style	"_" Comedy			
"_" 2017-12-24	123 6.0	[] [3 elements]	People & Blogs	Entertainment	"_" Music			
"_" 2017-12-25	10.0	[] [4 elements]	People & Blogs	" News & Politics	"_" Entertainment	"_" Music		
"_" 2017-12-26	11.0	[] [ 5 elements ]	"_" People & Blogs	"_" Sports	" Film & Animati	"_" Entertainment	"_" Music	

2.5 Ποιες ήταν οι πιο δημοφιλείς ημερομηνίες για δημοσίευση βίντεο - Γραφική απεικόνιση των αποτελεσμάτων GBvideos.



#### 2.5 Ποιες ήταν οι πιο δημοφιλείς ημερομηνίες για δημοσίευση βίντεο - Γραφική απεικόνιση των αποτελεσμάτων GBvideos (Bonus).



#### Bonus 2

Ο μέσος όρος ημερών που περνάει από τη μέρα που θα δημοσιευτεί ένα βίντεο μέχρι να τη μέρα που θα βρεθεί στη λίστα με τα πιο δημοφιλή βίντεο του YouTube είναι 36.85 ημέρες