



NTNU

Kunnskap for en bedre verden

Analysing Tidal's listening logs

Jan William Johnsen

Ph.D. student

Norwegian University of Science and Technology

May 7th 2019

COINS Winter School

Logs

- Received 74,1 GB
 - 65 CSV files
- 65 days in a 110 days period
- Two distinct periods with consecutive days

Period 1: Jan 21st – Mar 3rd

1

Period 2: Apr 18th – May 9th

2

January 2016

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
53					1	2	3
1	4	5	6	7	8	9	10
2	11	12	13	14	15	16	17
3	18	19	20	21	22	23	24
4	25	26	27	28	29	30	31

February 2016

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
5	1	2	3	4	5	6	7
6	8	9	10	11	12	13	14
7	15	16	17	18	19	20	21
8	22	23	24	25	26	27	28
9	29						

March 2016

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
13		1	2	3	4	5	6
14	7	8	9	10	11	12	13
15	14	15	16	17	18	19	20
16	21	22	23	24	25	26	27
17	28	29	30	31			

April 2016

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
13					1	2	3
14	4	5	6	7	8	9	10
15	11	12	13	14	15	16	17
16	18	19	20	21	22	23	24
17	25	26	27	28	29	30	

May 2016

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
17							1
18	2	3	4	5	6	7	8
19	9	10	11	12	13	14	15
20	16	17	18	19	20	21	22
21	23	24	25	26	27	28	29
22	30	31					

Content of logs

```
timestamp;countrycode;systemuserid;trackid;offlineplay  
0028-02-12 18:31:19.000;NO;13356188;20659857;Y  
0028-02-12 20:25:54.000;NO;13866374;422113;Y  
0028-02-12 22:51:47.000;NO;15380556;1647477;Y  
0028-02-12 22:58:55.000;NO;15380556;6640790;Y  
0028-02-12 23:03:43.000;NO;15380556;6640791;Y  
0028-02-12 23:10:23.000;NO;15380556;6640784;Y  
0028-02-13 00:19:50.000;NO;13356188;20659854;Y  
0028-02-13 00:23:18.000;NO;13356188;20659855;Y  
0028-02-13 00:37:23.000;NO;15380556;6640790;Y
```

Content of logs

timestamp	countrycode	systemuserid	trackid	offlineplay
0028-02-12 18:31:19.000	NO	13356188	20659857	Y
0028-02-12 20:25:54.000	NO	13866374	422113	Y
0028-02-12 22:51:47.000	NO	15380556	1647477	Y
0028-02-12 22:58:55.000	NO	15380556	6640790	Y
0028-02-12 23:03:43.000	NO	15380556	6640791	Y
0028-02-12 23:10:23.000	NO	15380556	6640784	Y
0028-02-13 00:19:50.000	NO	13356188	20659854	Y
0028-02-13 00:23:18.000	NO	13356188	20659855	Y
0028-02-13 00:37:23.000	NO	15380556	6640790	Y

Content of logs

timestamp	countrycode	systemuserid	trackid	tracktype
0028-02-12 18:31:19.000	NO	13356188		
0028-02-12 20:25:54.000	NO			
0028-02-12 22:51:47.000				
0028-02-12 23:59:59.000				Y
0028-02-13 00:00:00.000			6640791	Y
0028-02-13 00:00:00.000		15380556	6640784	Y
0028-02-13 00:00:00.000	NO	13356188	20659854	Y
0028-02-13 00:00:00.000	NO	13356188	20659855	Y
0028-02-13 00:37:23.000	NO	15380556	6640790	Y

1 590 422 377

Analysis environment

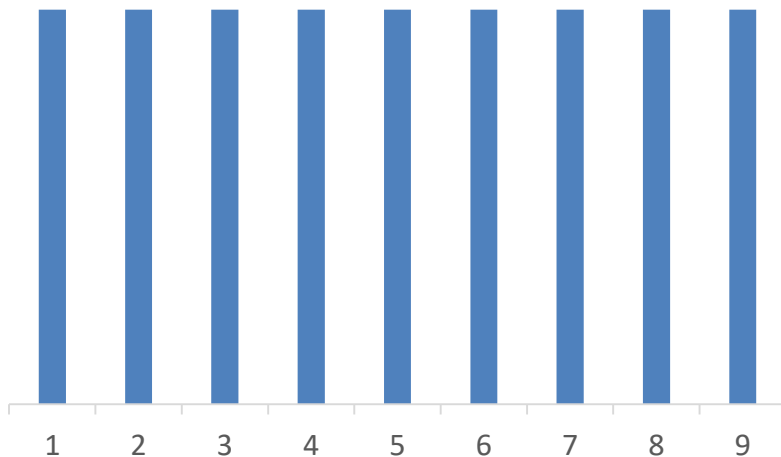
- Hardware
 - Desktop computer
 - Windows 10
 - 32 GB memory
 - Intel i7-6700 processor 3.40 GHz
- Software
 - Python
 - Pandas
 - Subprocess
 - Excel

Pandas



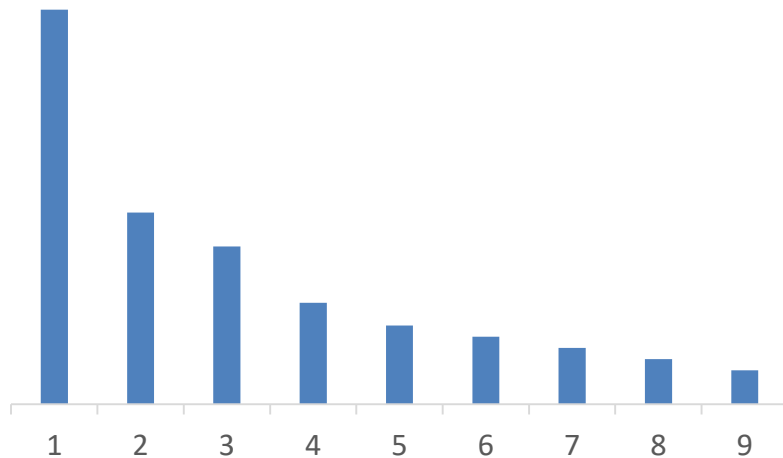
Bendford's law

- Are all numbers from 1 to 9 equally likely?



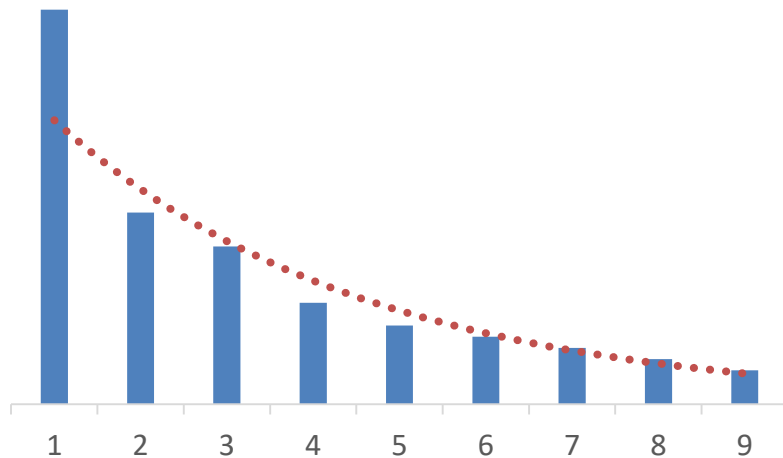
Bendford's law

- Are all numbers from 1 to 9 equally likely?



Bendford's law

- Are all numbers from 1 to 9 equally likely?

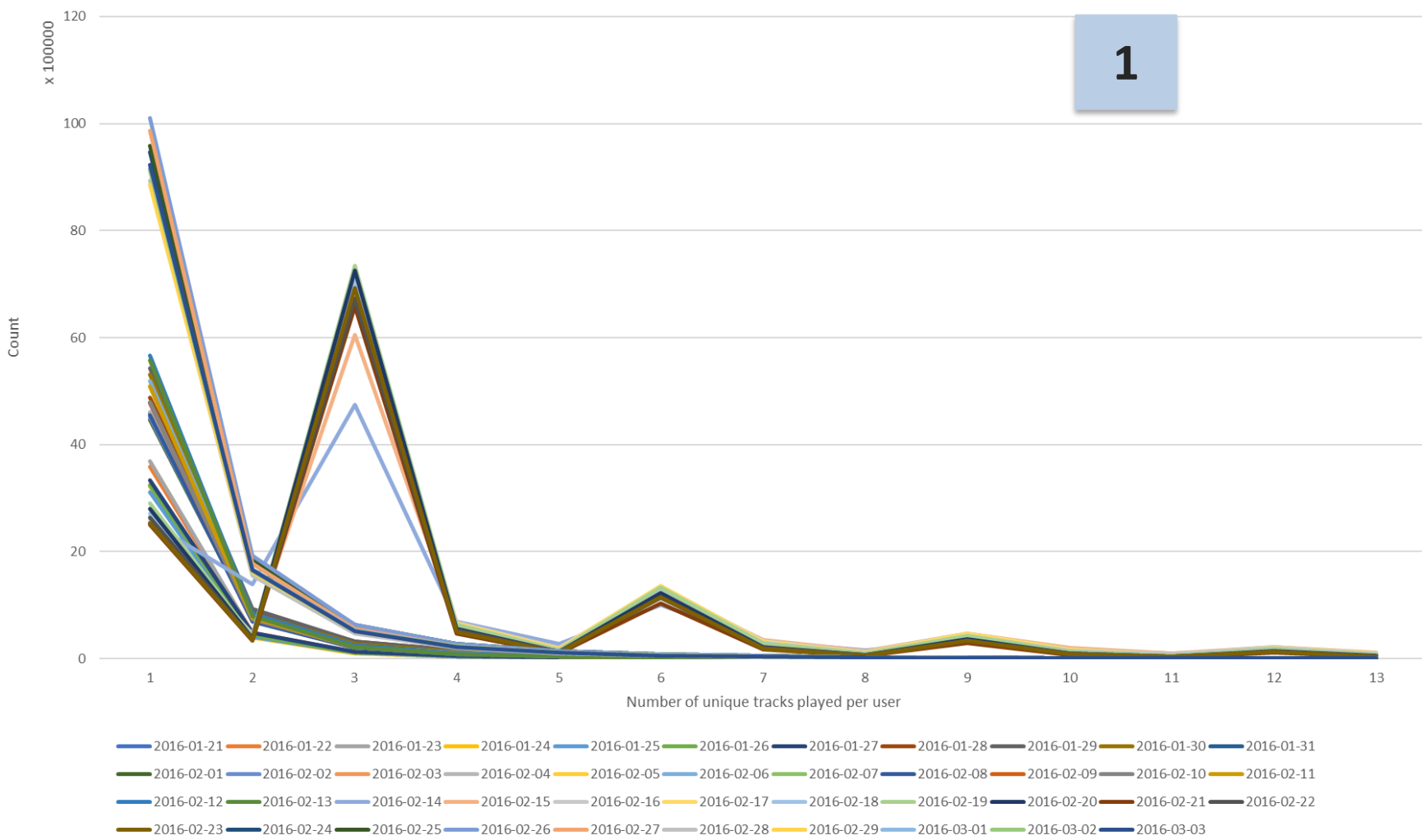


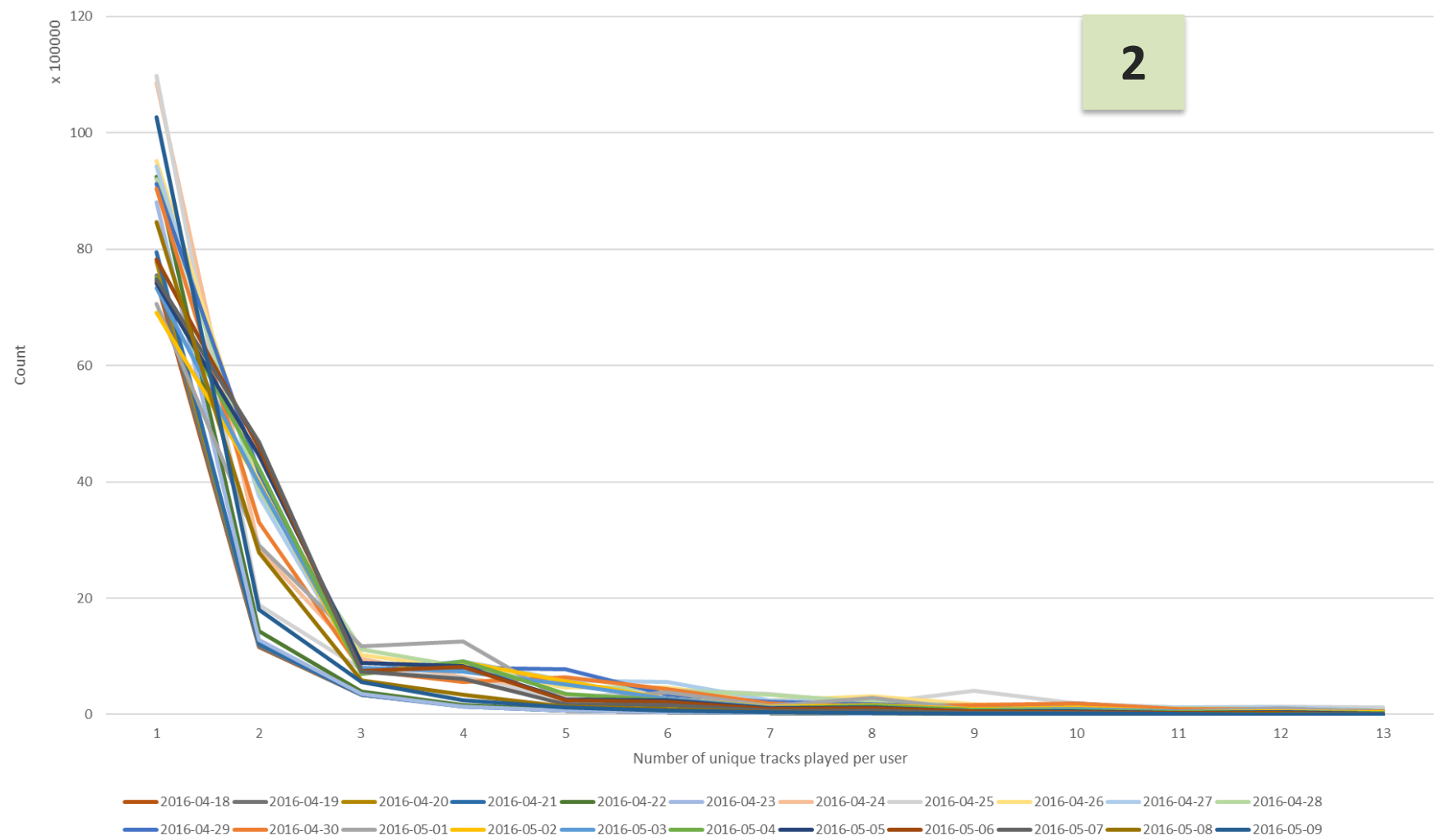
Methods

1. Descriptive statistical analysis
2. **Logical impossibilities**
3. **Unique tracks per user**
4. **Tracks per user**
5. Popular tracks
6. **Number of unique tracks**
7. System user frequency
8. Binning
9. **Modulo six**

Results for

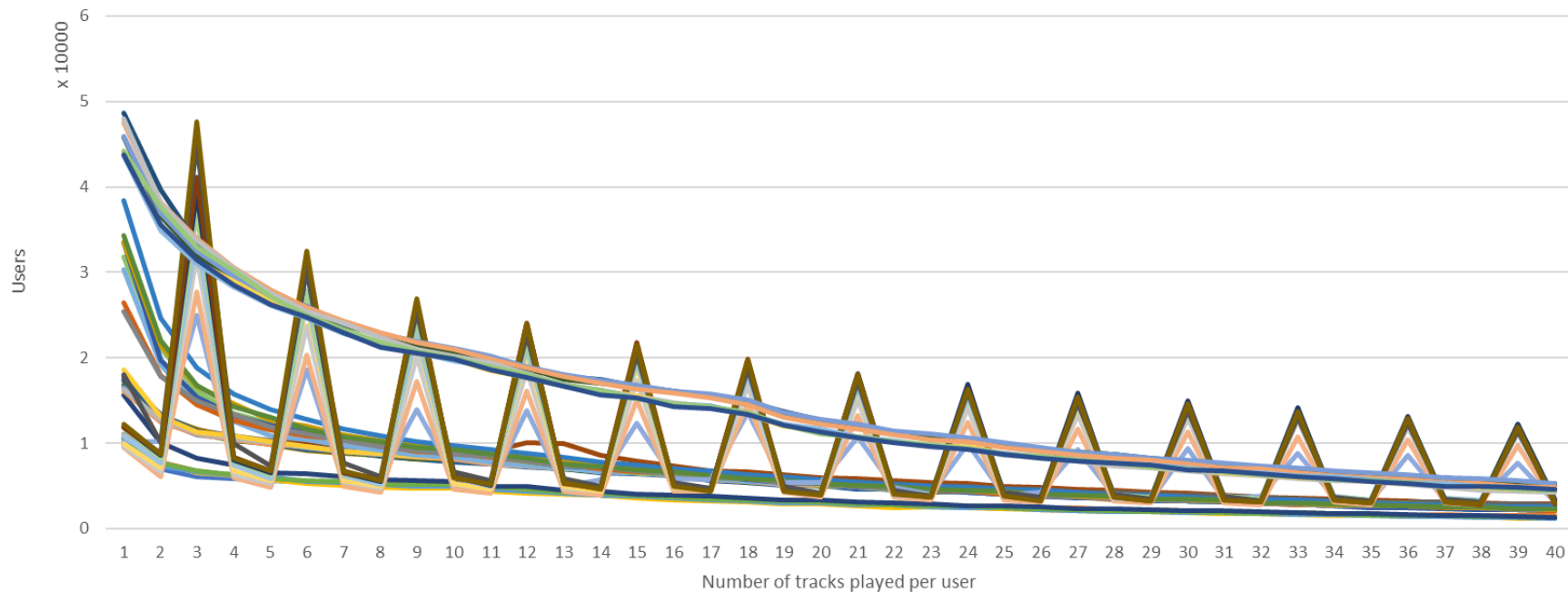
3. UNIQUE TRACKS PER USER



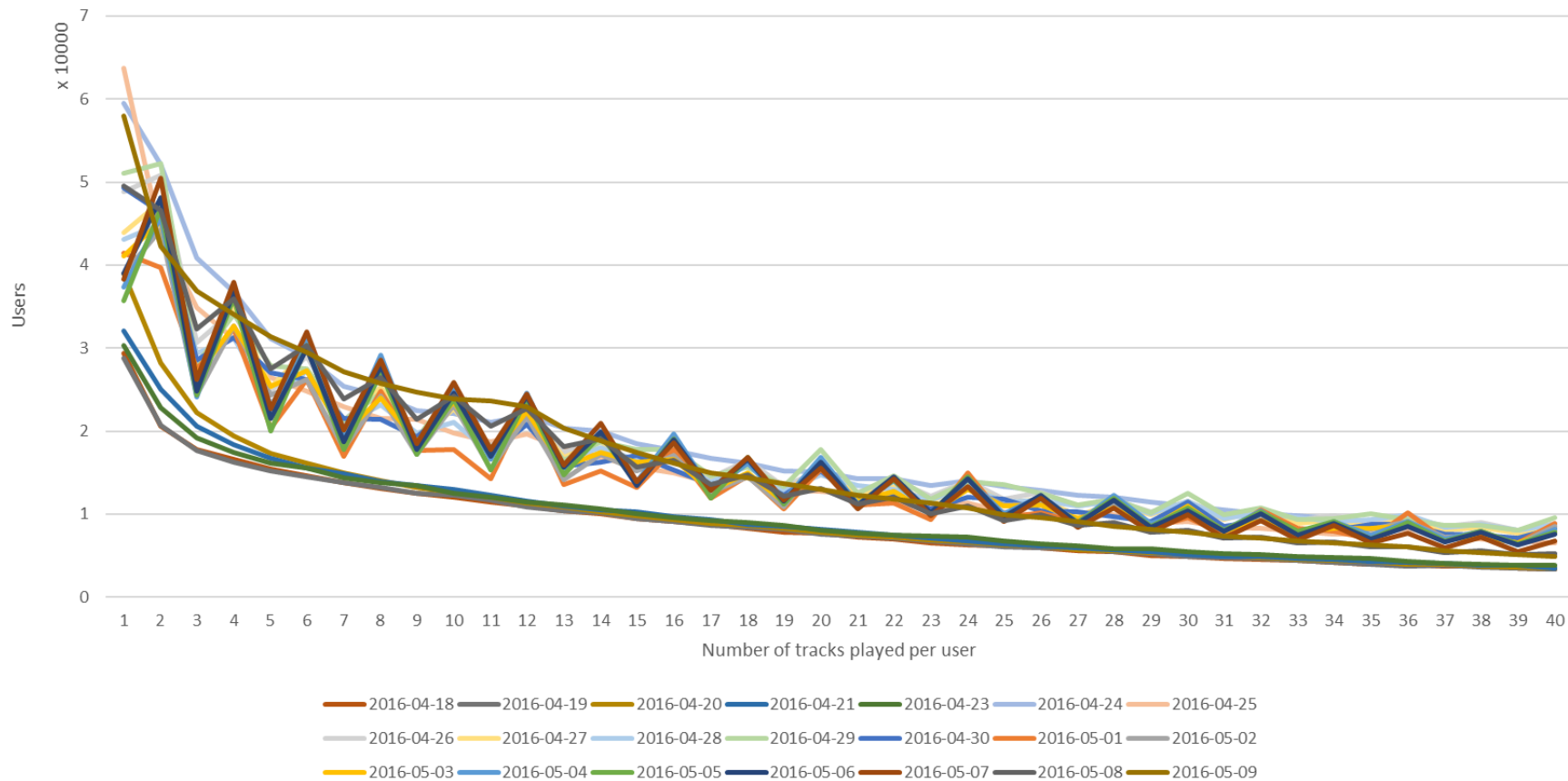


Results for

4. TRACKS PER USER

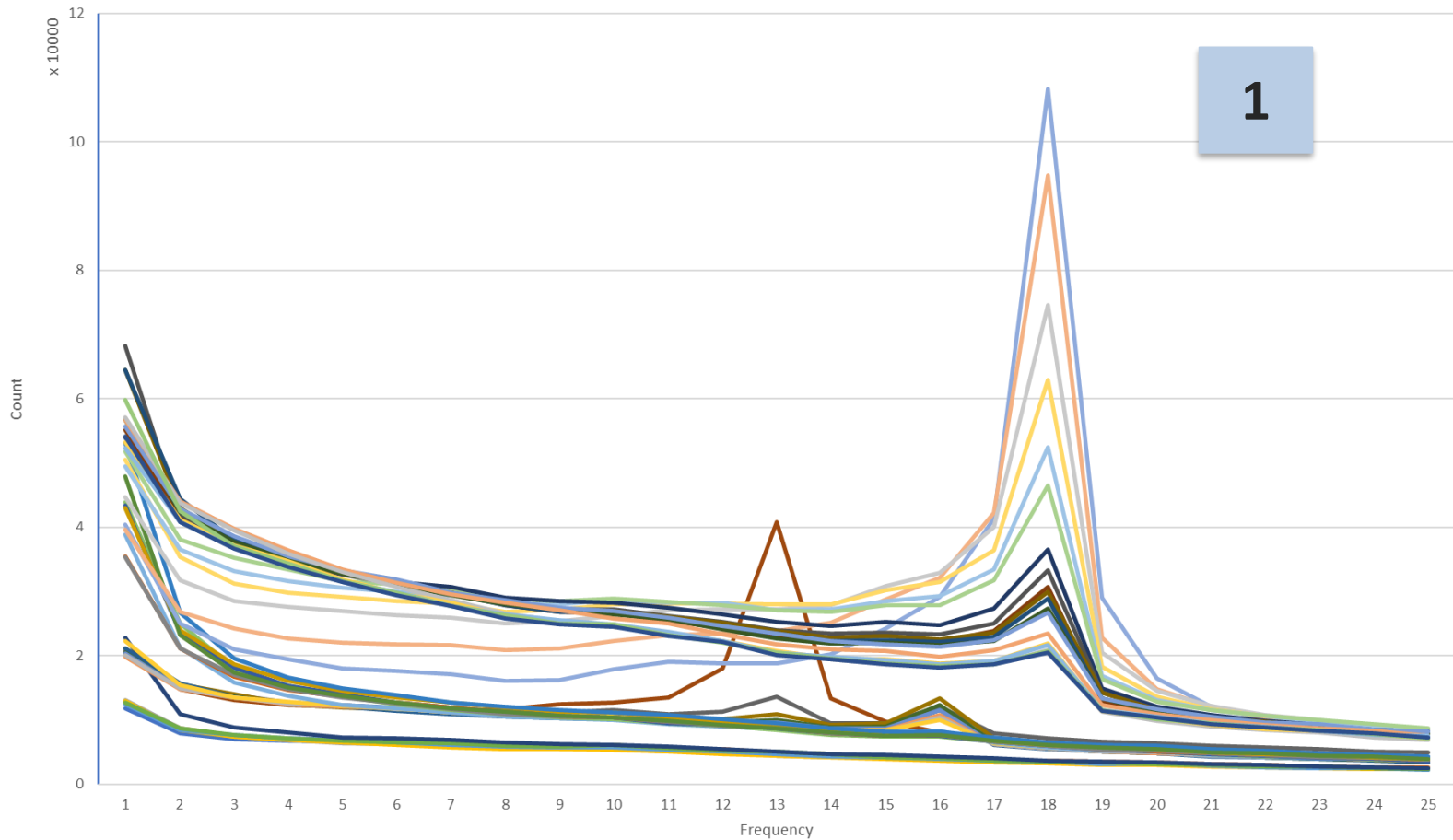


2016-01-21 2016-01-22 2016-01-23 2016-01-24 2016-01-25 2016-01-26 2016-01-27 2016-01-28 2016-01-29
 2016-01-30 2016-01-31 2016-02-01 2016-02-02 2016-02-03 2016-02-04 2016-02-05 2016-02-06 2016-02-07
 2016-02-08 2016-02-09 2016-02-10 2016-02-11 2016-02-12 2016-02-13 2016-02-14 2016-02-15 2016-02-16
 2016-02-17 2016-02-18 2016-02-19 2016-02-20 2016-02-21 2016-02-22 2016-02-23 2016-02-24 2016-02-25
 2016-02-26 2016-02-27 2016-02-28 2016-02-29 2016-03-01 2016-03-02 2016-03-03

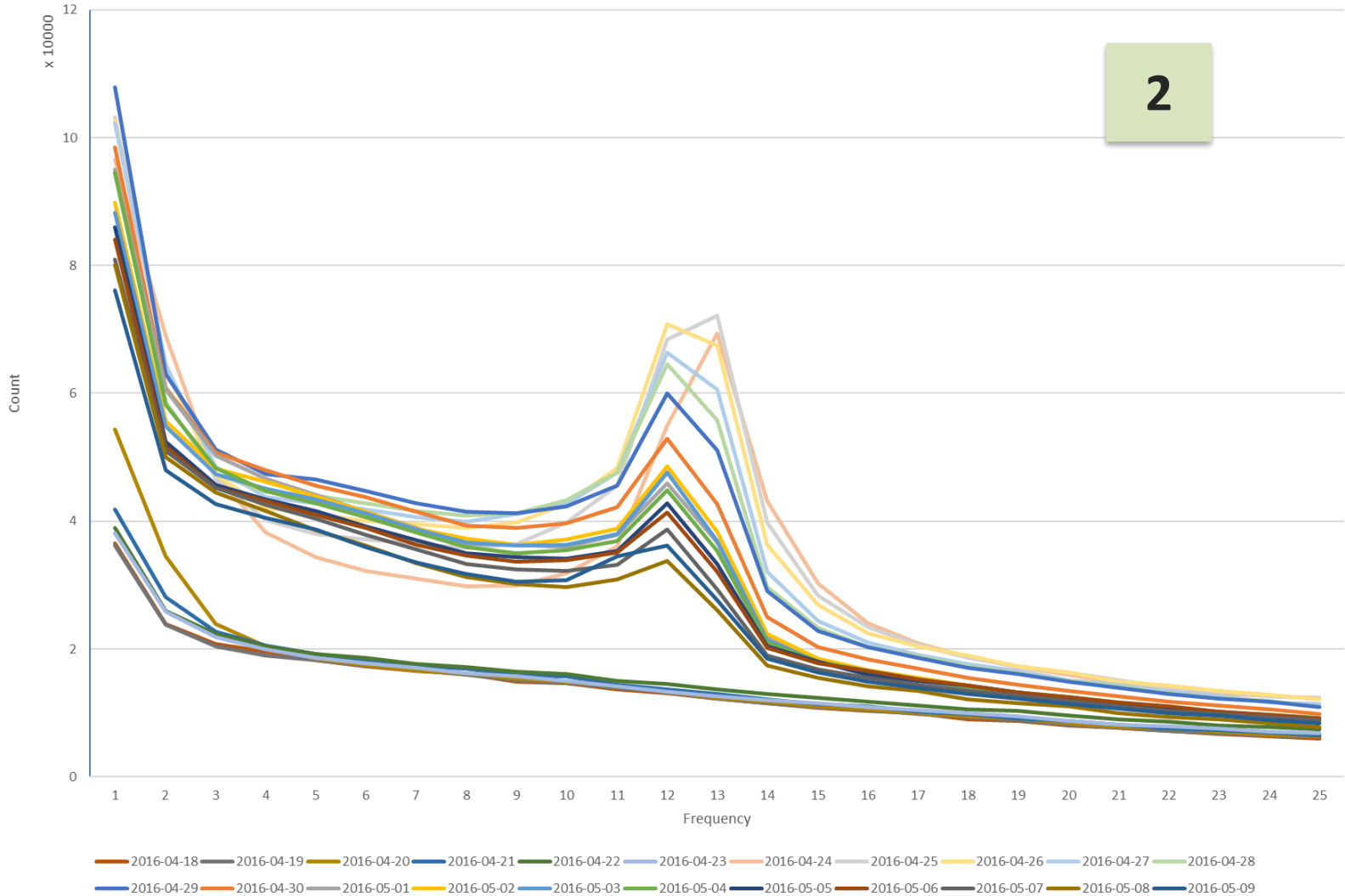


Results for

6. NUMBER OF UNIQUE TRACKS



1



Results for

2. LOGICAL IMPOSSIBILITIES

2. Logical impossibilities

Impossible scenarios	Timestamp	Systemuserid	Trackid
Equal duplicates	✓	✓	✓
Unequal duplicates	✓	✓	✗

Impossible scenarios	= 1	= 2	>= 3
Equal duplicates			
Unequal duplicates			

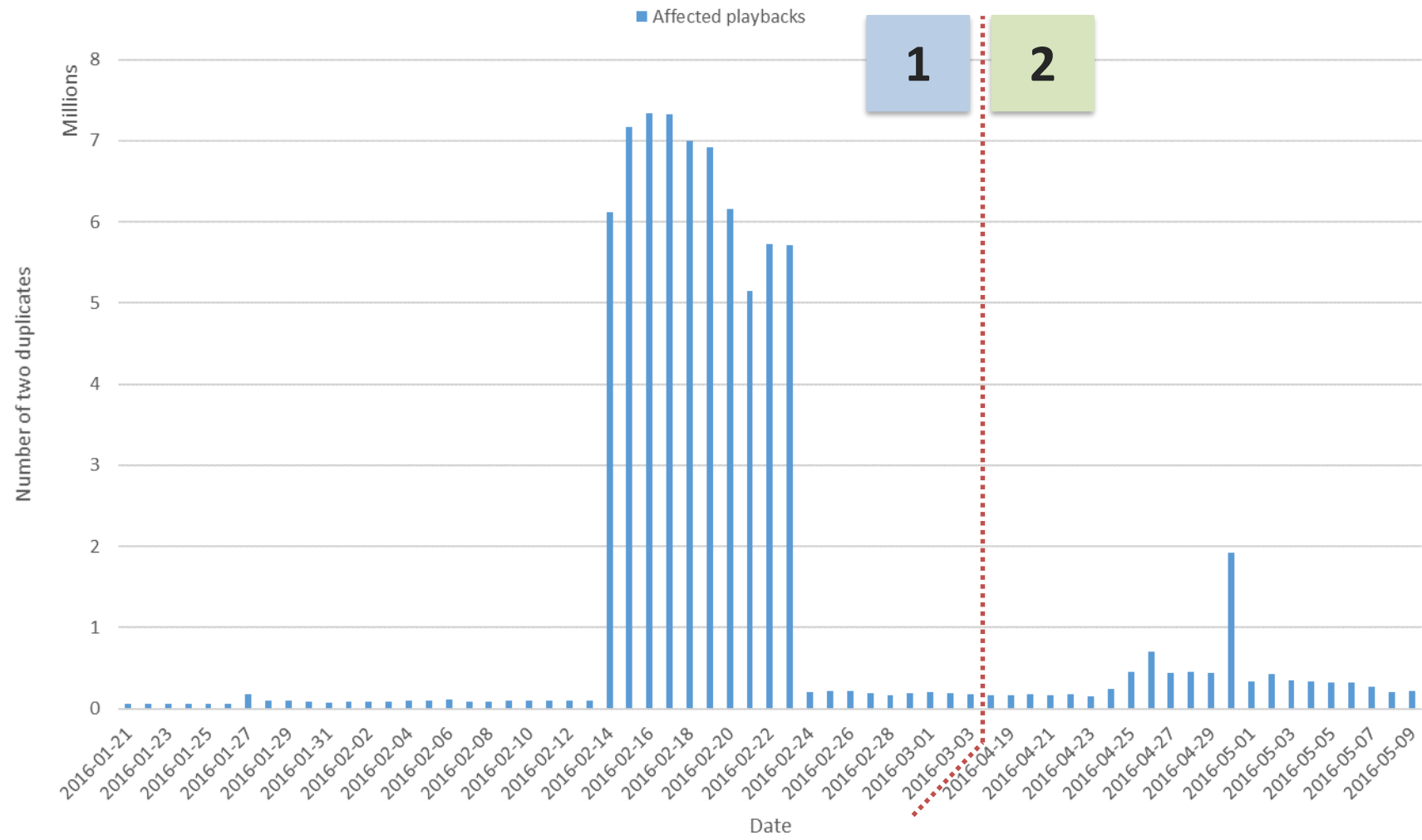
2. Logical impossibilities

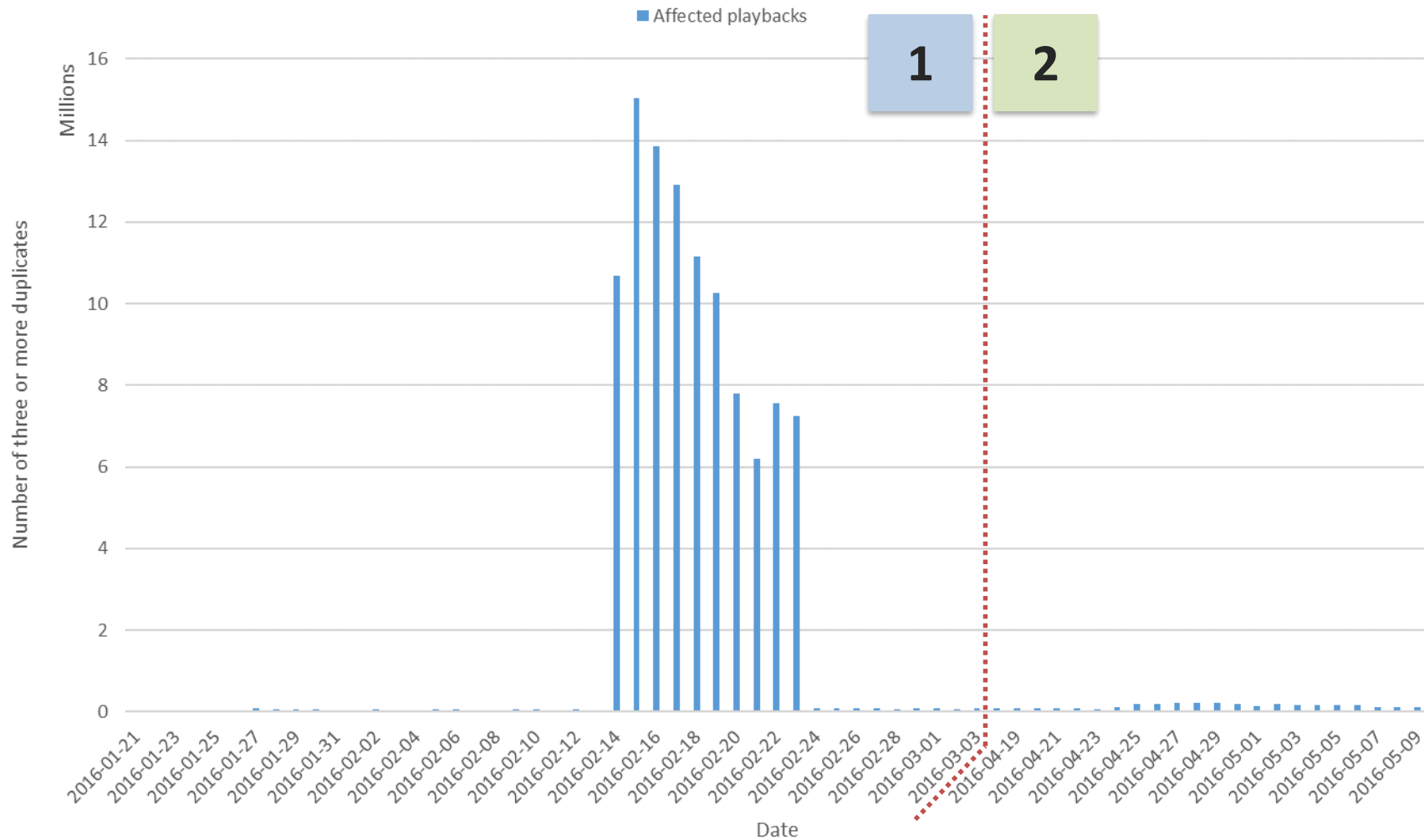
- Equal duplicates example

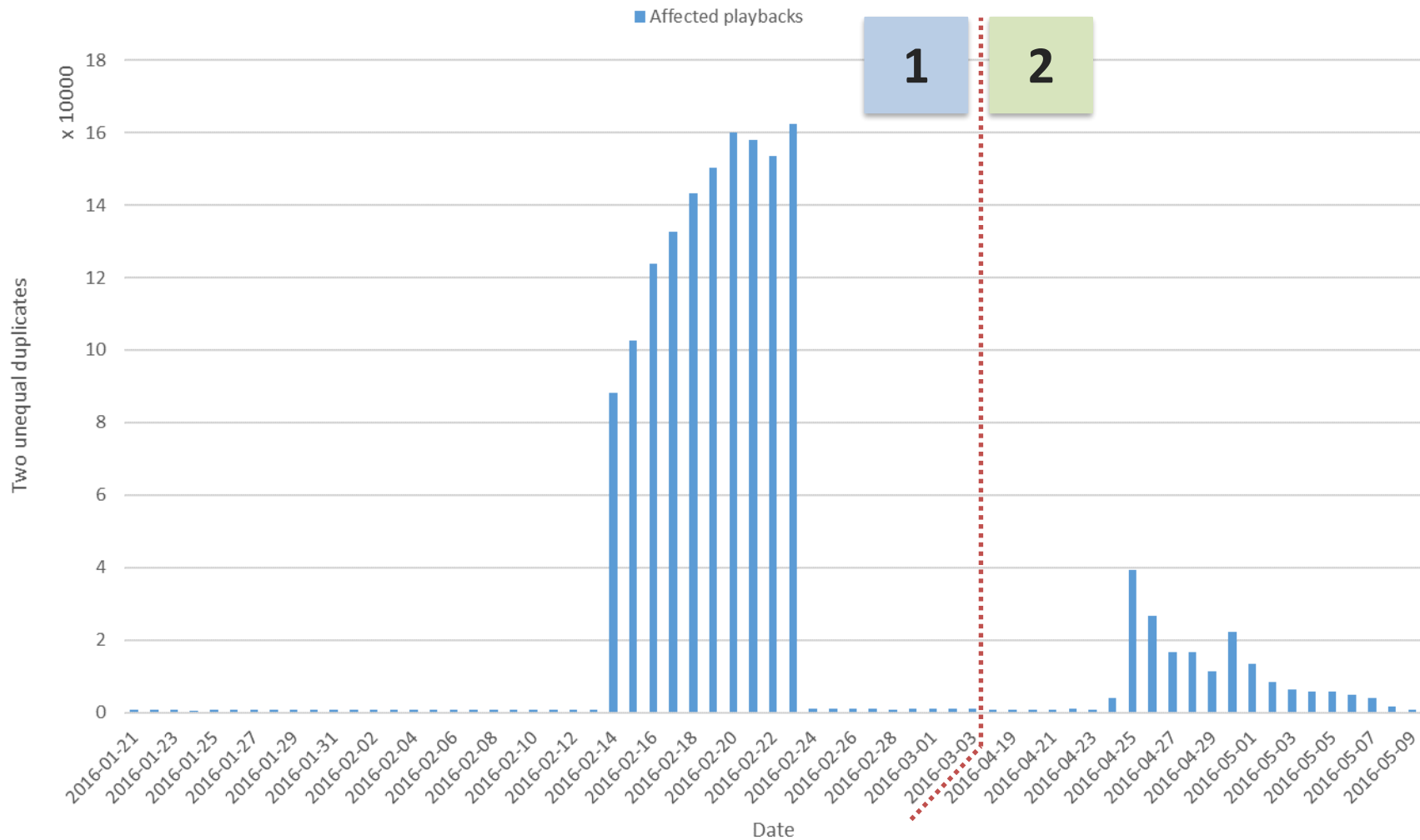
Timestamp	Systemuserid	Trackid
0028-02-12 18:31:19.000	13866374	20659857
0028-02-12 18:31:19.000	13866374	20659857

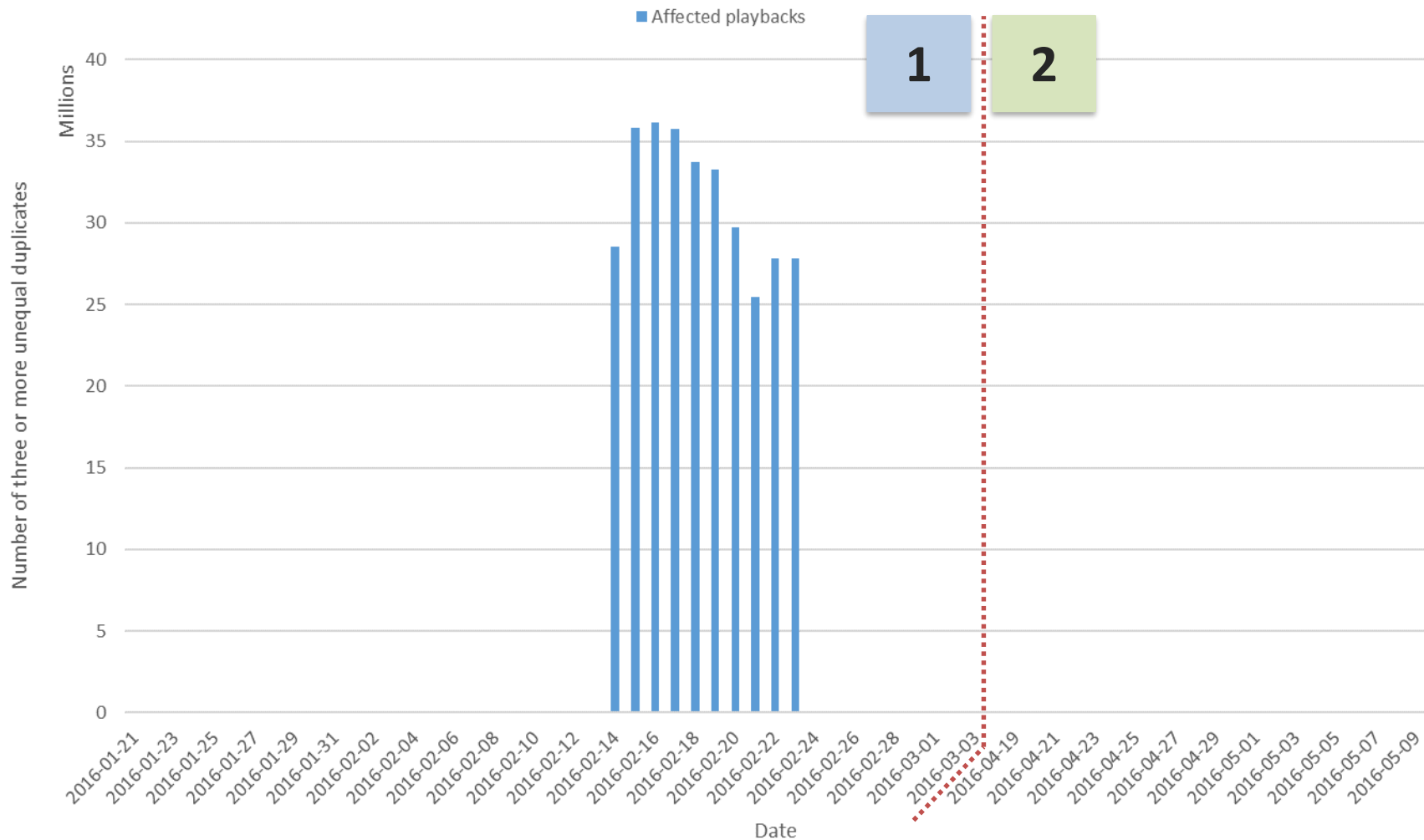
- Unequal duplicates example

Timestamp	Systemuserid	Trackid
0028-02-12 18:31:19.000	13866374	20659857
0028-02-12 18:31:19.000	13866374	6640790









Results for

9. MODULO SIX

9. Modulo six

- Modulo-operasjon finner *resten* av et opprinnelig heltall etter en divisjon med et annet tall
- Formel: $(a \bmod n) = a - \left\lfloor \frac{a}{n} \right\rfloor * n$
- Eksempel: $7 \bmod 3 = 1$

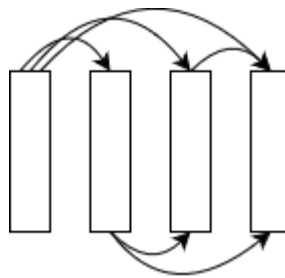
$$7 - \left\lfloor \frac{7}{3} \right\rfloor * 3 = 7 - [2,33] * 3 = 7 - (2 * 3) = 1$$

9. Modulo six

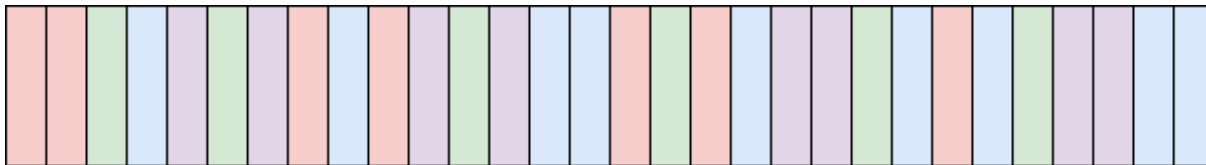
14:16:00.000 - 17:16:00.000	1540390560 - 1540401360
03:00:00.000	349200
03:00:00.000 mod 00:06:00.000	349200 mod 360
00:00:00.000	0

9. Modulo six

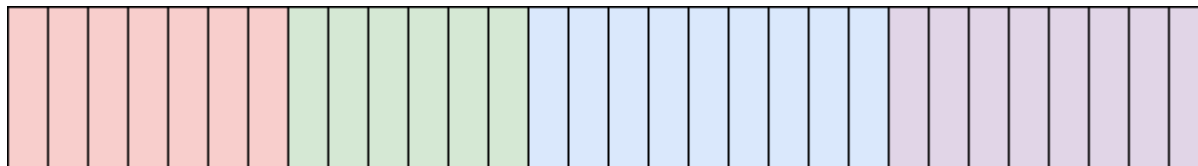
- Number of comparisons per user: $(n - 1)!$
- Example: one user had listened to a track 4 times
 - $n = 4$
 - $(4 - 1)! = 3! = 3 * 2 * 1 = 6$ comparisons
- Parallelise work!



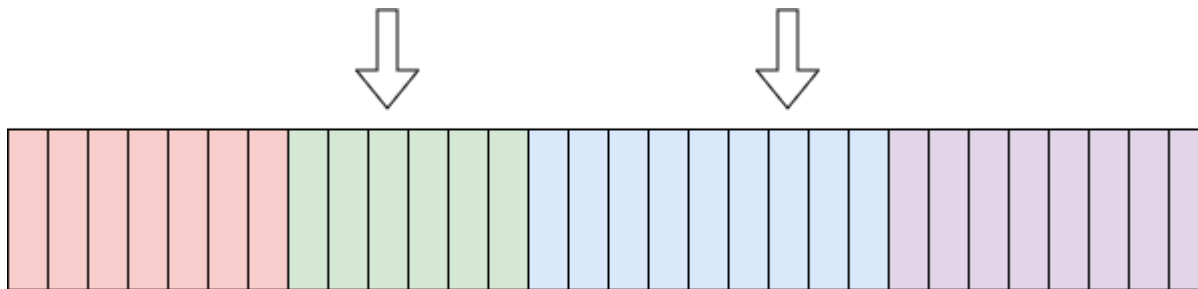
9. Modulo six



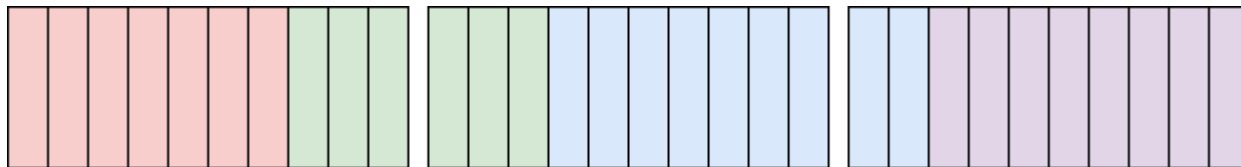
9. Modulo six



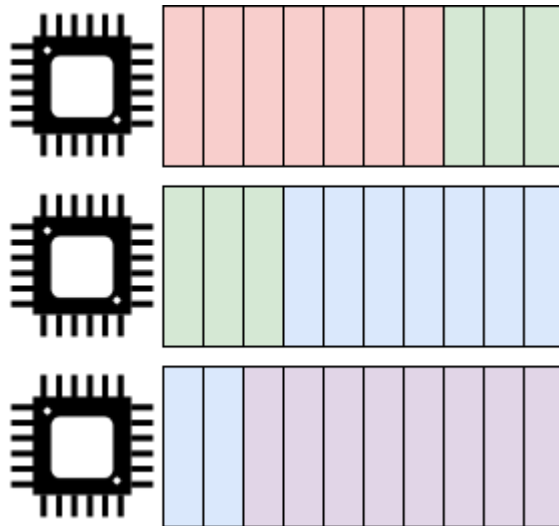
9. Modulo six



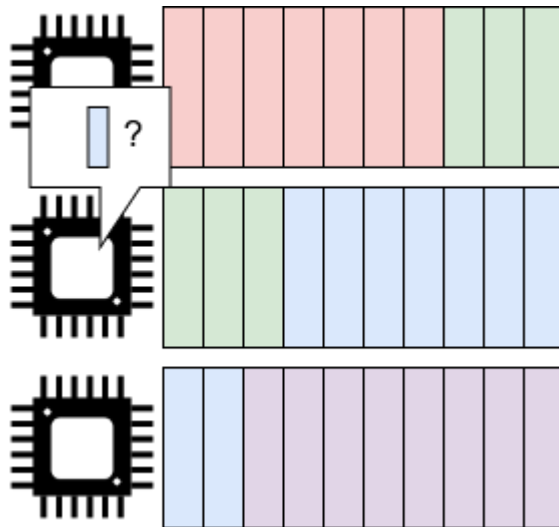
9. Modulo six



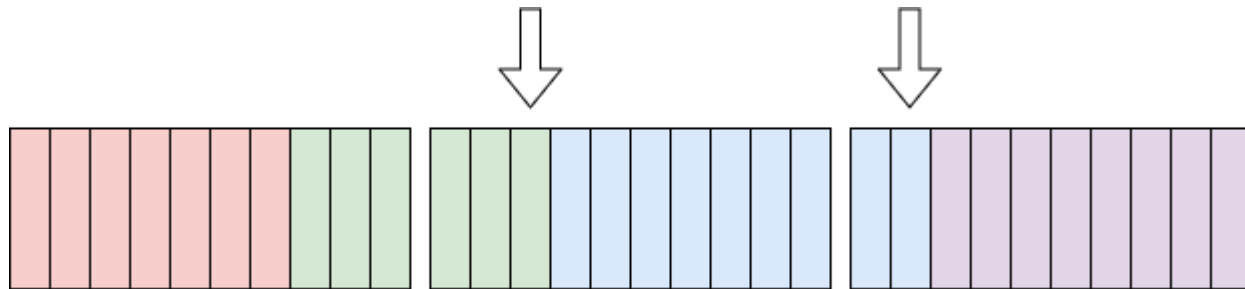
9. Modulo six



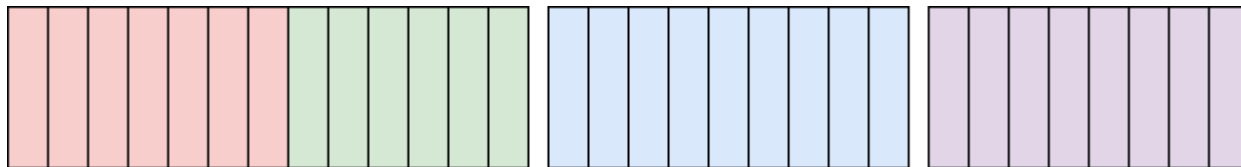
9. Modulo six



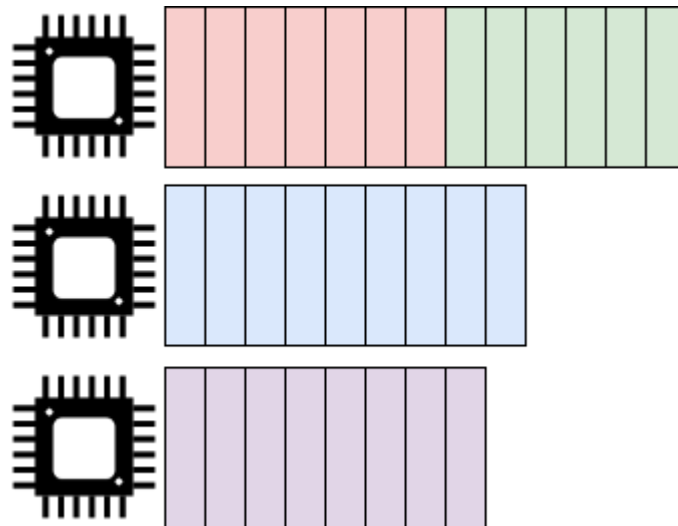
9. Modulo six

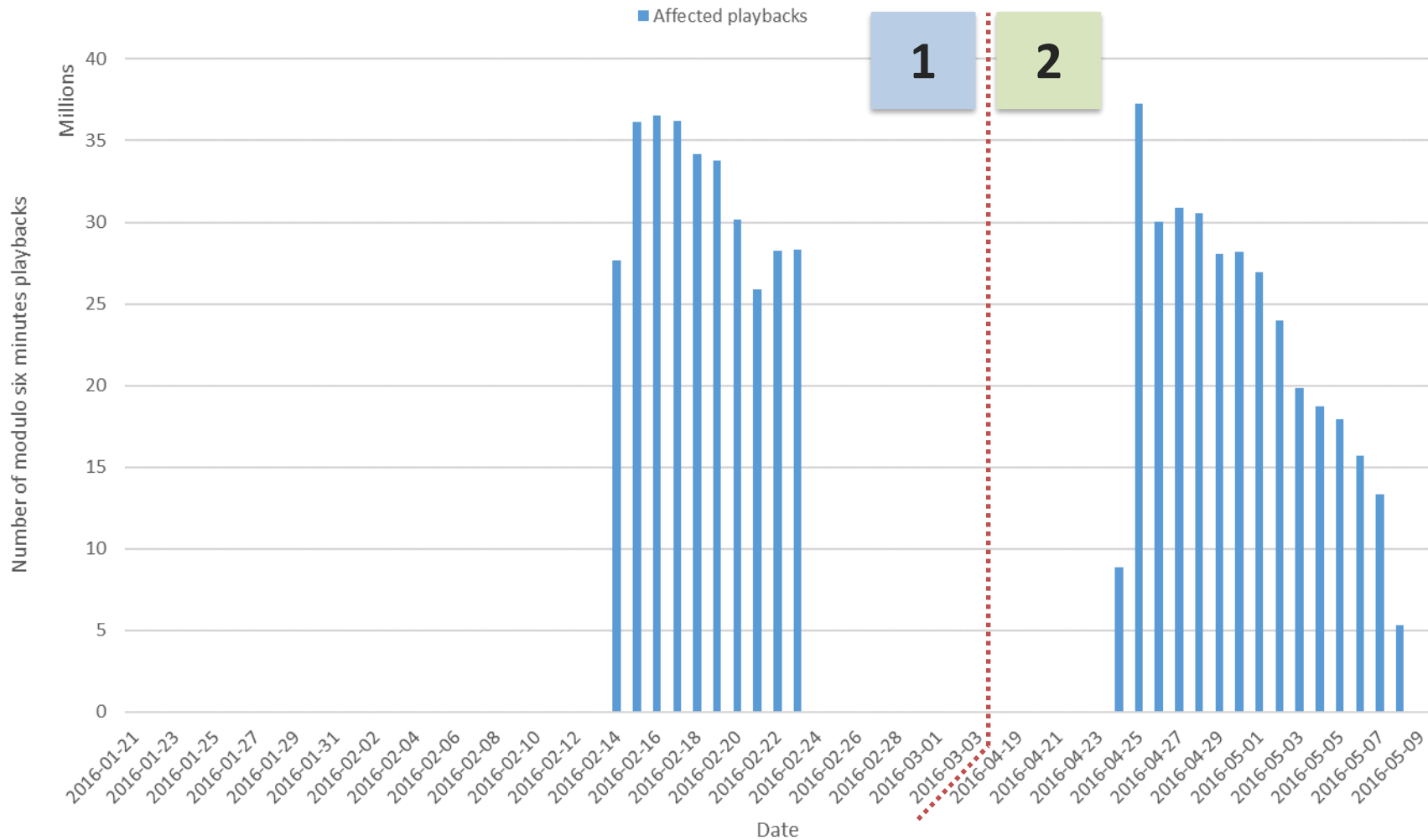


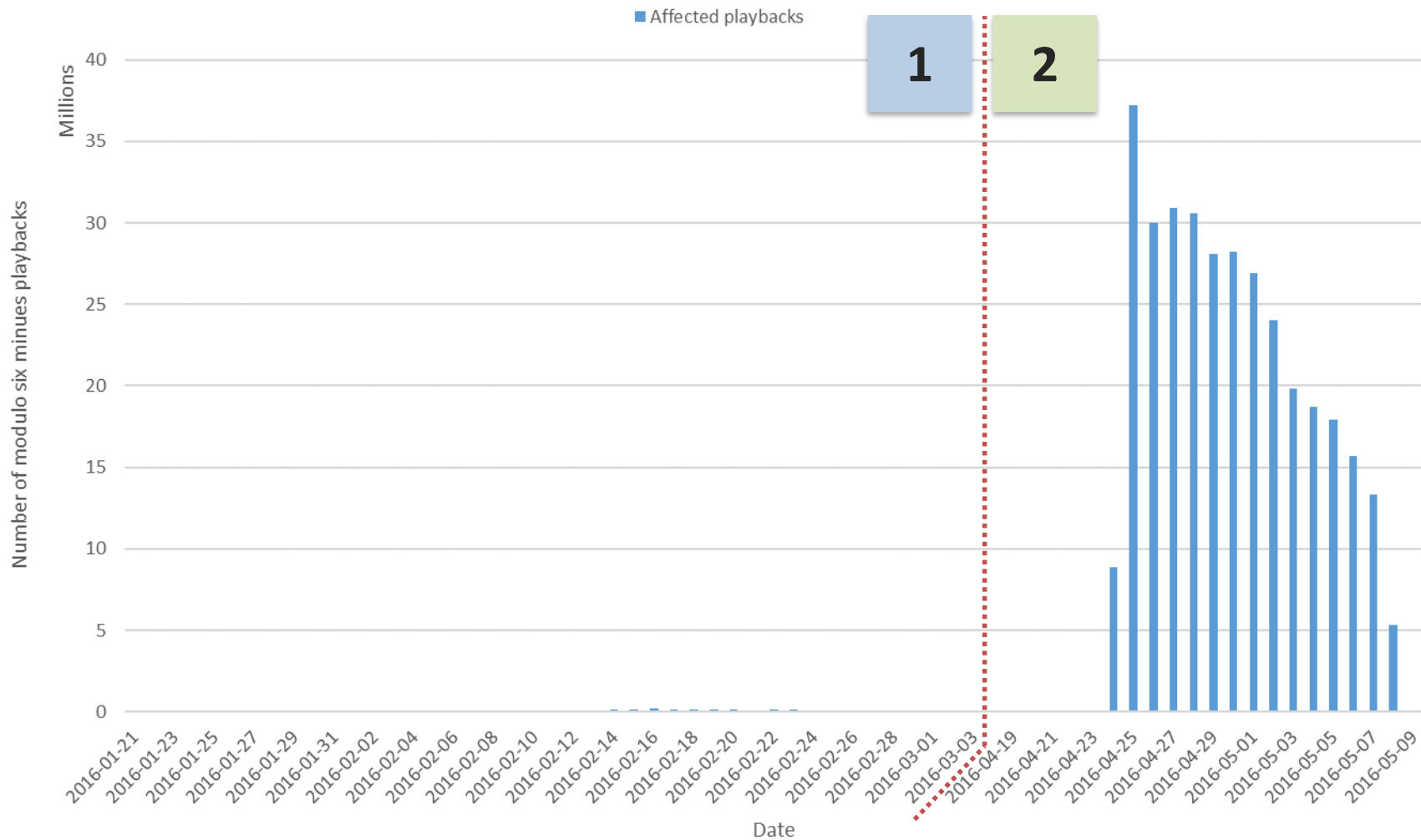
9. Modulo six



9. Modulo six

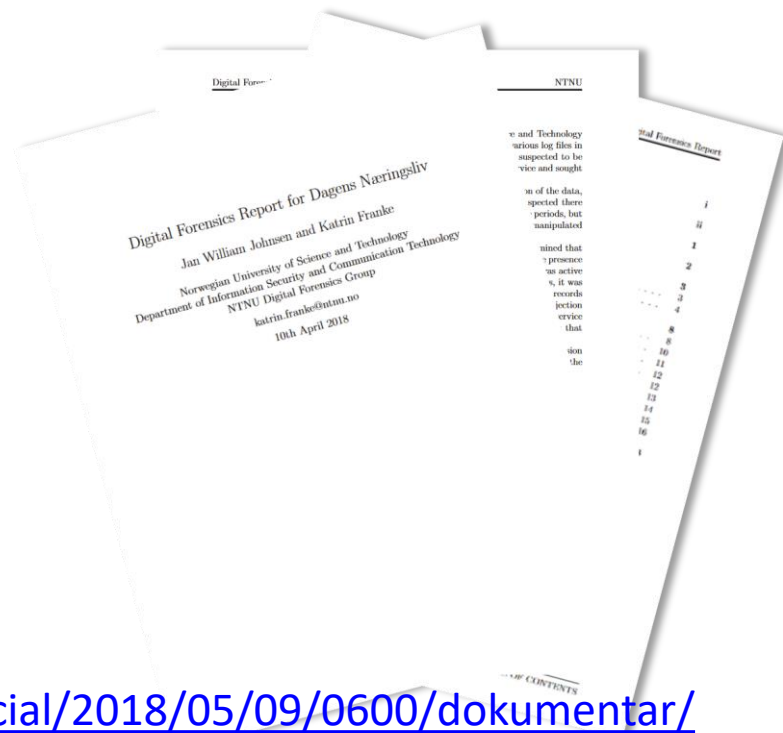






Report

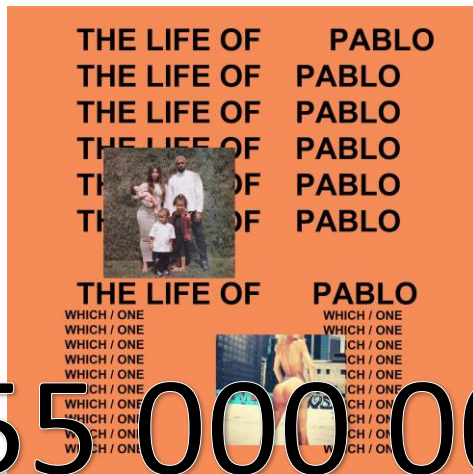
- More details, including:
 - Source code
- Internet search query:
*“Dagens Næringsliv NTNU
Digital Forensics Report”*



https://www.dn.no/staticprojects/special/2018/05/09/0600/dokumentar/strommekuppet/data/documentation/NTNU-rapport_til_publicisering.pdf

Conclusion

- Likely that suspicious playbacks was not a result of a cyber attack
 - Anomalies would target artists more broadly



155 000 000



165 000 000