Wikipedia Dataset Analysis

Revature Project 1

Christopher Chee

Data Used:

All Pageviews for October 20, 2020

```
class PageMapper extends Mapper[LongWritable, Text, Text, IntWritable] {
   override def map(key: LongWritable, value: Text, context: Mapper[LongWritable, Text, Text, IntWritable]#Context): Unit = {
     val line = value.toString
     val words = line.split("\\s+").filter(_.length > 0)

     try
     if ((words(0).equalsIgnoreCase("en")) || (words(0).equalsIgnoreCase("en.m")))
        context.write(new Text(words(1)), new IntWritable(words(2).toInt))
     }
}
```

```
class PageReducer extends Reducer[Text, IntWritable, Text, IntWritable] {
  override def reduce(key: Text, values: lang.Iterable[IntWritable], context: Reducer[Text, IntWritable, Text, IntWritable]#Context): Unit = {
    var count = 0
    values.forEach(count += _.get())
    context.write(key, new IntWritable(count))
  }
}
```

```
! 72
!! 25
!!! 148
!!!!!!! 19
!!!Factor You!!! 23
!!!Factor You!!!_And_Then_Some 19
!!!Factor You!!!_and_Then_Some 14
!!!_(!!!_album) 16
!!!_(American_band) 14
!!!_(Chk_Chk_Chk) 14
```

Format of MapReduce output

```
job.setInputFormatClass(classOf[KeyValueTextInputFormat])
FileInputFormat.setInputPaths(job, new Path(args(0)))
FileOutputFormat.setOutputPath(job, new Path(args(1)))

job.setMapperClass(classOf[InverseMapper[Text, Text]])
job.setReducerClass(classOf[SortReducer])

class TextToLongComparator extends WritableComparator(classOf[Text], true) {
   override def compare(a: WritableComparable[_], b: WritableComparable[_]): Int = {
        -a.toString.toLong.compareTo(b.toString.toLong)
   }
}
```

```
Main_Page 5961008

Special:Search 1476831
- 544714

Jeffrey_Toobin 321459

C._Rajagopalachari 210558

The_Haunting_of_Bly_Manor 185139

Robert_Redford 178779

Jeff_Bridges 159163

Bible 151484

Chicago_Seven 149966
```

Top 10 viewed pages

Data Used:

Clickstream for the month of Sept.

All Pageviews for the month of Sept.

```
class InternalMapper extends Mapper[LongWritable, Text, Text, LongWritable] {
   override def map(key: LongWritable, value: Text, context: Mapper[LongWritable, Text, Text, LongWritable]#Context): Unit = {
     val fields = value.toString.split("\\s+").filter(_.length > 0)

     if (fields(2).equalsIgnoreCase("link"))
        context.write(new Text(fields(0)), new LongWritable(fields(3).toLong))
   }
}
```

```
class InternalReducer extends Reducer[Text, LongWritable, Text, LongWritable] {
   override def reduce(key: Text, values: lang.Iterable[LongWritable], context: Reducer[Text, LongWritable, Text, LongWritable]#Context): Unit = {
     var count = 0
     values.forEach(count += _.get().toInt)
     context.write(key, new LongWritable(count))
   }
}
```

```
internal views.views
      internal_views.page
111
                                     1938
!!! (album)
                                     121
!!! (disambiguation)
                                     19
!Hero
                                     24
!Oka Tokat
                                     101
!T.O.O.H.!
                                     53
!Women Art Revolution
                                     18
! (The Dismemberment Plan album)
                                    124
 (The Song Formerly Known As)
                                     20
```

Format of MapReduce output

internal_views.page	internal_views.views	
!!	133	
111	1938	
!!!_(album)	121	
!!!_(disambiguation)	19	
!Hero	24	
!Oka_Tokat	101	
!T.O.O.H.!	53	
!Women_Art_Revolution	18	
!_(The_Dismemberment_Plan_album)	124	
! (The Song Formerly Known As)	20	

JOIN

```
total_views.page
                                total_views.views
                                1016
                                398
                                4488
                                411
                                372
      You!!! And Then Some
                                217
!!!F You!!! and Then Some
                                153
!!! (!!! album)
                                68
!!! (American band)
                                104
!!! (Chk Chk Chk)
                                77
```

```
SELECT PAGE, INTERNAL_VIEWS, TOTAL_VIEWS, ROUND(INTERNAL_FRACTION * 100, 2) AS INTERNAL_PERCENTAGE FROM INTERNAL_FRACTION
LIMIT 10;
```

page	internal_views	total_views	internal_percentage	
/r/	64	1	6400.0	
/\	56	2	2800.0	
Health//Disco	209	8	2612.5	
Strange_haircuts_//_cardboard_guitars_//	and computer samples 26	1	2600.0	
List of listed buildings in Musselburgh,	East_Lothian 662	28	2364.29	1
Flourish_//_Perish	19	1	1900.0	
Lost Forever // Lost Together	463	29	1596.55	
2006 Chicago Rush season	185	12	1541.67	
Baeolidia_gracilis	121	8	1512.5	
Finally // Beautiful Stranger	282	19	1484.21	

Invalid results

```
SELECT PAGE, INTERNAL_VIEWS, TOTAL_VIEWS, ROUND(INTERNAL_FRACTION * 100, 2) AS INTERNAL_PERCENTAGE FROM INTERNAL_FRACTION
WHERE INTERNAL_FRACTION < 1 AND TOTAL_VIEWS > 500000
LIMIT 10;
```

page	internal_views	total_views	internal_percentage
Dune (2020 film)	1201459	1278838	93.95
Cobra Kai	2241751	2459988	91.13
COVID-19_pandemic_by_country_and_territory	1093321	1207880	90.52
Christopher Nolan	612734	680233	90.08
Schitt's_Creek	1339942	1493588	89.71
Bill & Ted Face the Music	458119	518671	88.33
Elizabeth_II	922145	1065045	86.58
The_Babysitter:_Killer_Queen	663863	767589	86.49
72nd_Primetime_Emmy_Awards	465626	539024	86.38
2020 US Open (tennis)	436071	536585	81.27

Filtered results

Data Used:

Clickstream for the month of Sept.

All Pageviews for the month of Sept.

referrals.referrer	referrals.referred	referrals.clicks
Bathtubs_Over_Broadway	Industrial_musical	97
Industrial_music	Industrial_musical	67
Skittles_Commercial:_The_Broadway_Musical	Industrial_musical	14
Yvonne_Craig	Industrial_musical	23
Kander_and_Ebb	Industrial_musical	15
Descendants_of_Ibn_Saud	Saud_bin_Abdulaziz_bin_Nasser_Al_Saud	20
Nasser_bin_Abdulaziz_Al_Saud	Saud_bin_Abdulaziz_bin_Nasser_Al_Saud	70
Abdullah_of_Saudi_Arabia	Saud_bin_Abdulaziz_bin_Nasser_Al_Saud	59
LGBT_rights_in_Saudi_Arabia	Saud_bin_Abdulaziz_bin_Nasser_Al_Saud	142
List_of_gay,_lesbian_or_bisexual_people:_A	Saud_bin_Abdulaziz_bin_Nasser_Al_Saud	66

JOIN

```
total_views.views
      total views.page
                                1016
                                398
                                4488
                                411
                                372
      You!!! And Then Some
                                217
     You!!! and Then Some
                                153
!!! (!!! album)
                                68
!!! (American band)
                                104
!!! (Chk Chk Chk)
```

referral_fractions.referrer	referral_fractions.referred	referral_fractions.internal_views	referral_fractions.total_views	referral_fractions.internal_fractio
11	Double-negation_translation	14	291	0.048109965635738834
11	Retroflex_click	76	335	0.22686567164179106
11	!_(disambiguation)	145	781	0.1856594110115237
11	Chess_annotation_symbols	1262	4108	0.30720545277507305
H	Double_factorial	404	7530	0.053652058432934926
11	!!!_(disambiguation)	19	181	0.10497237569060773
111	Strange Weather, Isn't It?	94	242	0.3884297520661157
111	!!!_(disambiguation)	19	181	0.10497237569060773
111	Tyler Pope	121	765	0.15816993464052287
111	As If (album)	94	248	0.3790322580645161

Resulting table from join

```
SELECT *
FROM REFERRAL_FRACTIONS
WHERE REFERRER='Hotel_California'
ORDER BY INTERNAL FRACTION DESC;
```

referral_fractions.referrer	referral_fractions.referred	referral_fractions.internal_views	referral_fractions.total_views	referral_fractions.internal_fraction
Hotel California	Eagles (band)	132994	139366	0.954278661940502
Hotel California	Jethro Tull (band)	51257	61744	0.8301535371857994
Hotel California	Steely Dan	66279	85960	0.7710446719404375
Hotel California	The Twilight Zone (1959 TV series)	31414	42170	0.7449371591178563
Hotel California	American Horror Story: Hotel	53807	73110	0.7359731910819314
Hotel California	Desperado	1255	1718	0.7305005820721769
Hotel California	Cameron Crowe	38391	58203	0.6596051749909798
Hotel California	Anton LaVey	36586	56460	0.6479985830676586
Hotel California	John Fowles	4350	7491	0.5806968362034441
Hotel_California	Hotel_California_(disambiguation)	144	248	0.5806451612903226

Hotel_California -> Eagles_(band) (95.4%)

```
FROM REFERRAL_FRACTIONS
WHERE REFERRER='Eagles_(band)'
ORDER BY INTERNAL_FRACTION DESC
LIMIT 10;
```

referral_fractions.referrer	referral_fractions.referred	referral_fractions.internal_views	referral_fractions.total_views	referral_fractions.internal_fraction
Eagles_(band)	Eagles_discography	16815	13799	1.2185665627944053
Eagles_(band)	2008 Universal Studios fire	26158	22063	1.185604858813398
Eagles_(band)	Deep_Purple	106701	98931	1.0785395881978348
Eagles_(band)	Earth, Wind & Fire	103052	99649	1.0341498660297646
Eagles_(band)	Led_Zeppelin	242143	239290	1.0119227715324501
Eagles_(band)	Yes_(band)	81239	80624	1.0076280015876167
Eagles_(band)	Emerson, Lake & Palmer	34792	34941	0.9957356686986635
Eagles_(band)	Fleetwood_Mac	262893	274440	0.9579252295583734
Eagles_(band)	Grammy Award for Album of the Year	38177	40811	0.9354585773443435
Eagles (band)	Crosby, Stills, Nash & Young	55053	58977	0.933465588280177

Hotel_California -> Eagles_(band) (95.4%) -> Emerson,_Lake_&_Palmer (99.6%)

```
SELECT *
FROM REFERRAL_FRACTIONS
WHERE REFERRER='Emerson,_Lake_&_Palmer' AND INTERNAL_FRACTION < 1
ORDER BY INTERNAL_FRACTION DESC
LIMIT 10;
```

referral_fractions.referrer	referral_fractions.referred	referral_fractions.internal_views	referral_fractions.total_views	referral_fractions.internal_fraction
Emerson,_Lake_&_Palmer	Atomic_Rooster	7230	7542	0.958631662688942
Emerson, Lake & Palmer	King_Crimson	54767	63808	0.8583093029087262
Emerson, Lake & Palmer	Asia_(band)	22870	28514	0.8020621449112717
Emerson, Lake & Palmer	The_Nice	3897	5743	0.6785652098206513
Emerson, Lake & Palmer	Moog_Music	1634	2617	0.6243790599923577
Emerson, Lake & Palmer	Emerson, Lake & Powell	1418	2305	0.6151843817787419
Emerson, Lake & Palmer	HRGiger	20205	36074	0.5600986860342629
Emerson, Lake & Palmer	The Crazy World of Arthur Brown	3025	5456	0.5544354838709677
Emerson, Lake & Palmer	Welcome_Back_My_Friends_to_the_Show_That_Never	Ends - Ladies and Gentlemen 1366	2765	0.49403254972875227
Emerson, Lake & Palmer	Progressive rock	20568	42038	0.4892716113992102

Hotel_California -> Eagles_(band) (95.4%) -> Emerson,_Lake_&_Palmer (99.6%) -> Atomic_Rooster (95.9%)

```
SELECT *
FROM REFERRAL_FRACTIONS
WHERE REFERRER='Atomic_Rooster' AND INTERNAL_FRACTION < 1
ORDER BY INTERNAL_FRACTION DESC
LIMIT 10;
```

referral_fractions.referrer	referral_fractions.referred	referral_fractions.internal_views	referral_fractions.total_views	referral_fractions.internal_fraction
Atomic_Rooster	Emerson, Lake & Palmer	34792	34941	0.9957356686986635
Atomic_Rooster	Colosseum (band)	4822	5122	0.9414291292463881
Atomic Rooster	Cactus (American band)	3699	5200	0.7113461538461539
Atomic Rooster	The Crazy World of Arthur Brown	3025	5456	0.5544354838709677
Atomic Rooster	Hard Stuff	331	599	0.5525876460767947
Atomic Rooster	Dexys Midnight Runners	13884	25479	0.5449193453432238
Atomic Rooster	Homework (Atomic Rooster album)	119	221	0.5384615384615384
Atomic Rooster	List of Atomic Rooster members	306	609	0.5024630541871922
Atomic Rooster	Made in England (Atomic Rooster album)	514	1031	0.498545101842871
Atomic Rooster	Death Walks Behind You	1011	2126	0.4755409219190969

Hotel_California -> Eagles_(band) (95.4%) -> Emerson,_Lake_&_Palmer (99.6%) -> Atomic_Rooster (95.9%) -> Emerson,_Lake_&_Palmer

Data Used:

Pageviews during peak internet traffic hours

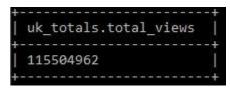
Used 1 week of pageview data from Sept. 21 - 25 (Mon – Fri)

Assumptions:

Traffic during peak hours is representative of traffic all the time Wikipedia traffic during peak hours is representative as well

```
| uk_views.page | uk_views.views |
| ! | 18 |
| !! | 7 |
| !!! | 92 |
| !!!!!!! | 4 |
| !!!Fact You!!! | 5 |
| !!!Fact You!!!_And_Then_Some | 2 |
| !!!_(!!!_album) | 1 |
| !!!_(American_band) | 1 |
| !!!_(Chk_Chk_Chk) | 9 |
| !!!_(album) | 12 |
```

```
INSERT OVERWRITE DIRECTORY 'user/hive/uk_total' SELECT SUM(VIEWS) FROM UK_VIEWS;
```



Sum of views during UK peak hours

```
INSERT OVERWRITE DIRECTORY 'user/hive/uk_percentage'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t'
SELECT PAGE, VIEWS, VIEWS / TOTAL_VIEWS * 100
FROM UK_VIEWS, UK_TOTALS
ORDER BY VIEWS / TOTAL_VIEWS DESC;
```

uk_percentage.page	uk_percentage.views	uk_percentage.view_percentage
Shooting_of_Breonna_Taylor	251305	0.2175707395150695
Amy Coney Barrett	173273	0.15001346868544055
Ruth_Bader_Ginsburg	154539	0.13379425205992448
Ratched_(TV_series)	124418	0.10771658450482846
Dennis_Nilsen	119860	0.10377043368924704
Enola_Holmes_(film)	115567	0.10005371024666455
Rosemary_West	107911	0.09342542357617502
Fred_West	104913	0.0908298640884363
Millie_Bobby_Brown	78851	0.06826633127674636
S. P. Balasubrahmanyam	66924	0.05794036796445161

Percentage of total views per article

uk_us.page	uk_us.uk_view_percent	uk_us.us_view_percent	uk_us.difference
Dennis_Nilsen	0.10377043368924704	0.006216127892937832	0.09755430579630921
Rosemary West	0.09342542357617502	0.001891453646464705	0.09153396992971032
Fred West	0.0908298640884363	0.0030371035823942836	0.08779276050604201
The_7.39	0.043744441039684515	5.866056763502978E-4	0.043157835363334215
Moors_murders	0.03501321441065017	0.0017240283236915133	0.033289186086958654
Janet Leach (appropriate adult)	0.031143250798177834	7.45392346105708E-4	0.030397858452072126
2020-21 UEFA Europa League	0.033051393930591484	0.0056718927528176056	0.02737950117777388
Appropriate_Adult	0.02712870465253259	4.619622542365949E-4	0.026666742398295997
David_Morrissey	0.021270081886179053	7.820037803239243E-4	0.02048807810585513
Rory Delap	0.02311242697954396	0.002925212457592544	0.020187214521951418

Dennis_Nilsen comparison $\frac{0.10377}{0.006216} = 16.694$

```
INSERT OVERWRITE DIRECTORY 'user/hive/aus_us'

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

SELECT AUS_PERCENTAGE.PAGE,

AUS_PERCENTAGE.VIEW_PERCENTAGE,

US_PERCENTAGE.VIEW_PERCENTAGE,

(AUS_PERCENTAGE.VIEW_PERCENTAGE - US_PERCENTAGE.VIEW_PERCENTAGE)

FROM AUS_PERCENTAGE JOIN US_PERCENTAGE

ON AUS_PERCENTAGE.PAGE = US_PERCENTAGE.PAGE

ORDER BY (AUS_PERCENTAGE.VIEW_PERCENTAGE - US_PERCENTAGE.VIEW_PERCENTAGE) DESC;
```

aus_us.page	aus_us.aus_view_percent	aus_us.us_view_percent	aus_us.difference
S. P. Balasubrahmanyam	0.32380010645457113	0.075400220361343	0.24839988609322813
Dean_Jones_(cricketer)	0.07238597143434429	0.014293679828992843	0.05809229160535145
F5 Networks	0.06548221326297168	0.007738999010643865	0.05774321425232782
List of Bollywood actresses	0.03801161897189277	0.0018223444223224537	0.03618927454957032
Liu Chuyu	0.03252593809548109	9.021221937140273E-4	0.03162381590176706
List of Indian film actresses	0.030716735526306237	0.001958917412889283	0.028757818113416954
Dennis Nilsen	0.03471286443835748	0.006216127892937832	0.028496736545419645
S. P. Charan	0.03450588570822143	0.010560136267593613	0.02394574944062782
Agha Mohammad Khan Qajar	0.021654343068694546	0.001404398162033602	0.020249944906660943
Anurag Kashyap	0.025761158447940854	0.007182422937641094	0.01857873551029976

List of Bollywood actresses $\frac{0.03801}{0.001022} = 20.86$

us_uk.page	us_uk.us_view_percent	us_uk.uk_view_percent	us_uk.difference
Tyler_Herro	0.12242205419484475	0.0230630784502574	0.09935897574458735
Amy Coney Barrett	0.2351535965258859	0.15001346868544055	0.08514012784044536
Shooting of Breonna Taylor	0.2894347784453299	0.2175707395150695	0.07186403893026039
Ryan Fitzpatrick	0.04377534575653362	0.006620494797444286	0.03715485095908933
Jamal Murray	0.041187452131063965	0.006094976248726007	0.03509247588233796
Chrishell Stause	0.041239284049170655	0.007493184578511873	0.033746099470658784
Anne Heche	0.03684673467184353	0.004292456284259026	0.032554278387584505
Darren Waller	0.03639793607934831	0.0040786126573506	0.03231932342199771
Allegiant Stadium	0.03312059567017382	0.0053755266375482644	0.02774506903262556
The Killers	0.02973424368720352	0.0021375705054125726	0.02759667318179095

vs UK vs Australia (+ India)

Tyler Herro

$$\frac{0.1224}{0.02306} = 5.307 \qquad \frac{0.1224}{0.03383} = 3.618$$

US Page Views vs UK Page Views

us_aus.page	us_aus.us_view_percent	us_aus.aus_view_percent	us_aus.difference
Amy Coney Barrett	0.2351535965258859	0.060953995145390814	0.1741996013804951
Shooting of Breonna Taylor	0.2894347784453299	0.14866383788786555	0.14077094055746434
Tyler Herro	0.12242205419484475	0.03383829244195526	0.08858376175288948
Ratched (TV series)	0.10891737134147594	0.06547178987368425	0.04344558146779169
Ryan Fitzpatrick	0.04377534575653362	0.0070129555829551715	0.03676239017357845
Chrishell Stause	0.041239284049170655	0.004769941478675008	0.03646934257049564
Anne Heche	0.03684673467184353	0.0028078625332845497	0.03403887213855898
Darren_Waller	0.03639793607934831	0.004595721972013726	0.03180221410733459
Jamal Murray	0.041187452131063965	0.010188118500653822	0.03099933363041014
Allegiant Stadium	0.03312059567017382	0.004713357365400404	0.028407238304773410

Amy Coney Barrett

$$\frac{0.2351}{0.1500} = 1.567 \qquad \frac{0.2351}{0.06095} = 3.857$$

Shooting of Breonna Taylor

$$\frac{0.2894}{0.2175} = 1.33 \qquad \frac{0.2894}{0.1486} = 1.947$$

Data used:

MediaWiki enwiki history up to October 2020 Pageviews for October 20, 2020

Assumptions:

All revisions that were reverted were vandalizations

Total views over a day are representative of normal traffic

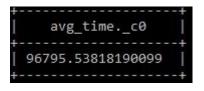
```
CREATE TABLE REVISIONS_SECONDS AS
SELECT PAGE_TITLE, REVISION_SECONDS_TO_IDENTITY_REVERT
WHERE REVISION_SECONDS_TO_IDENTITY_REVERT > 0
FROM WIKHISTORY;
```

```
revisions seconds.page title
                                    revisions seconds.revision seconds to identity revert
Tityus serrulatus
                                    591
                                    1168
Sandbox
Michael Matricciani
                                    2552
Werner Fischer (sailor)
                                    35326
The Little Match Girl
                                    97
Eurodog/sandbox266
                                    51
                                    194
DannyS712 bot III/Redirects.json
                                    874
Main Page
                                    17
Maxie_the_Fox/sandbox
```

Time (in seconds) before reverting change on page

CREATE TABLE AVG_TIME AS

SELECT AVG (REVISION_SECONDS_TO_IDENTITY_REVERT)
FROM REVISIONS_SECONDS
WHERE REVISION_SECONDS_TO_IDENTITY_REVERT > 0;

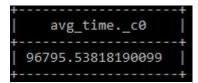


Average time before vandalization is reverted

CREATE TABLE AVG_TIME AS

SELECT AVG(REVISION_SECONDS_TO_IDENTITY_REVERT)
FROM REVISIONS_SECONDS
WHERE REVISION_SECONDS_TO_IDENTITY_REVERT > 0;

SELECT AVG(VIEWS) AS AVG VIEWS FROM PAGEVIEWS_20_10_20;



Average time before vandalization is reverted

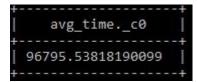


Average views an article receives per day

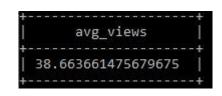
CREATE TABLE AVG_TIME AS

SELECT AVG(REVISION_SECONDS_TO_IDENTITY_REVERT)
FROM REVISIONS_SECONDS
WHERE REVISION_SECONDS_TO_IDENTITY_REVERT > 0;

SELECT AVG(VIEWS) AS AVG VIEWS FROM PAGEVIEWS 20 10 20;



Average time before vandalization is reverted



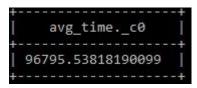
Average views an article receives per day

$$\frac{96795.538 \ seconds}{revert} \ x \ \frac{1 \ minute}{60 \ seconds} \ x \ \frac{1 \ hour}{60 \ minutes} \ x \ \frac{1 \ day}{24 \ hours} = \frac{1.12 \ days}{revert}$$

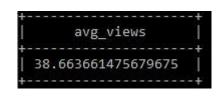
CREATE TABLE AVG_TIME AS

SELECT AVG(REVISION_SECONDS_TO_IDENTITY_REVERT)
FROM REVISIONS_SECONDS
WHERE REVISION_SECONDS_TO_IDENTITY_REVERT > 0;

SELECT AVG(VIEWS) AS AVG VIEWS FROM PAGEVIEWS_20_10_20;



Average time before vandalization is reverted



Average views an article receives per day

$$\frac{96795.538 \ seconds}{revert} \ x \ \frac{1 \ minute}{60 \ seconds} \ x \ \frac{1 \ hour}{60 \ minutes} \ x \ \frac{1 \ day}{24 \ hours} = \frac{1.12 \ days}{revert}$$

$$\frac{1.12 \ days}{revert} \ x \ \frac{38.663 \ views}{day} = \frac{43.303 \ views}{revert}$$

6. Which popular English Wikipedia pages have the lowest percentage of people click an internal link?

6. Which popular English Wikipedia pages have the lowest percentage of people click an internal link?

Data Used:

Clickstream for the month of Sept.

All Pageviews for the month of Sept.

6. Which popular English Wikipedia pages have the lowest percentage of people click an internal link?

```
SELECT PAGE, INTERNAL_VIEWS, TOTAL_VIEWS, ROUND(INTERNAL_FRACTION * 100, 2) AS INTERNAL_PERCENTAGE FROM INTERNAL_FRACTION
WHERE INTERNAL_FRACTION < 1 AND TOTAL_VIEWS > 500000
ORDER BY INTERNAL_PERCENTAGE ASC
LIMIT 10;
```

page	internal_views	total_views	internal_percentage
F5_Networks	513	1487955	0.03
Flag_of_Scotland	3815	527267	0.72
Bible	32110	3170711	1.01
Microsoft_Office	23379	2136261	1.09
Gmail	6667	508260	1.31
Google Classroom	11300	790056	1.43
123Movies	8201	573309	1.43
Main Page	2379287	165044119	1.44
XXXX	49387	2056847	2.4
The Pirate Bay	14763	570331	2.59

10 popular pages with lowest percentage

Questions?