

Name Collision Analysis Data Analysis Part 2

Agenda

“Can we predict what strings are going to make the Internet go boom and is there a way to mitigate any of these if we do discover them so?”

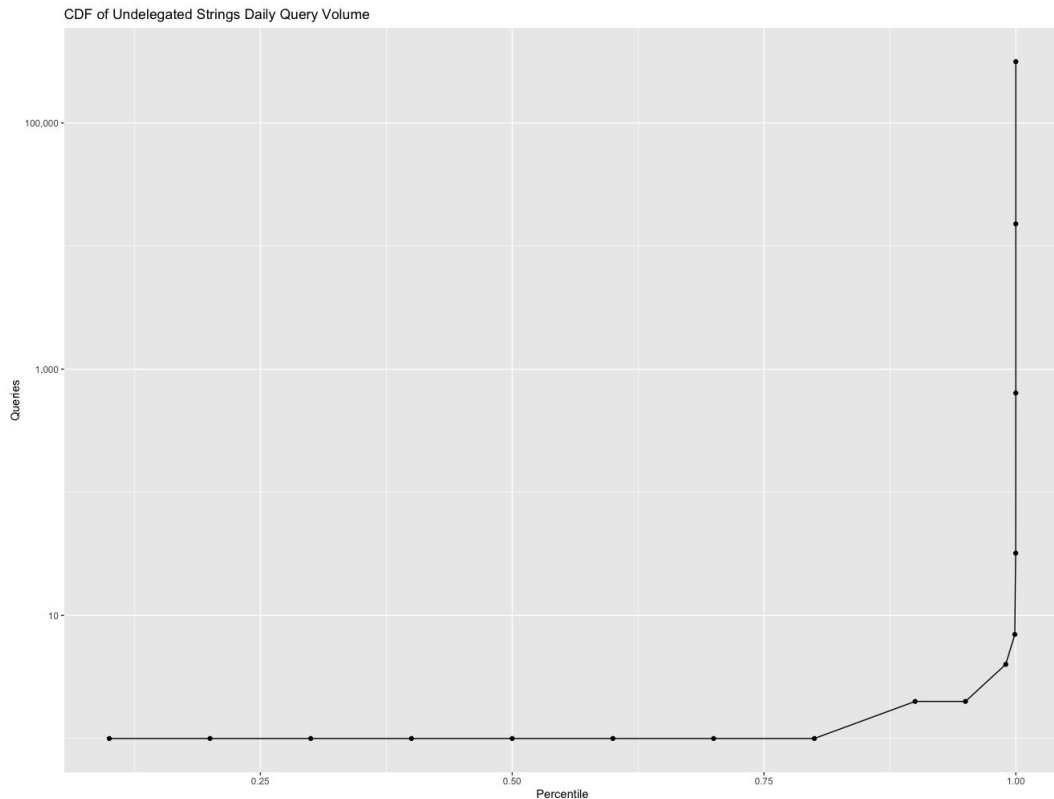
1. How big is the ocean? How many fish are in it? How many sharks?
2. Continue deliberating how we incorporate these data exploration case studies back into guidance that we must provide for the Board’s 10 questions.

String Query Volume

Proposed TLD	2012 rank	2013 rank	Rank change	2012 count	2013 count	Change 2012-2013
home	1	1	0	595,024	952,944	60%
corp	2	2	0	122,794	144,507	18%
ice	21	3	18	1,837	19,789	977%
global	4	4	0	10,838	12,352	14%
med	29	5	24	1,262	10,801	756%
site	3	6	-3	13,013	10,716	-18%
ads	5	7	-2	7,799	10,563	35%

Proposed TLD	2012 rank	2013 rank	Rank change	2012 count	2013 count	Change 2012-2013
network	12	8	4	4,417	8,711	97%
group	7	9	-2	6,505	8,580	32%
cisco	9	10	-1	5,231	8,284	58%
box	8	11	-3	6,152	7,694	25%
prod	14	12	2	4,107	7,004	71%
iinet	6	13	-7	7,668	5,427	-29%
hsbc	10	14	-4	4,924	5,249	7%
inc	11	15	-4	4,622	5,208	13%
win	18	16	2	2,511	5,199	107%
dev	13	17	-4	4,344	5,058	16%
office	15	18	-3	3,833	4,006	5%
business	20	19	1	1,920	3,279	71%
host	16	20	-4	2,965	3,127	5%
star	31	21	10	1,040	2,435	134%
mail	25	22	3	1,505	2,383	58%
ltd	19	23	-4	1,962	1,990	1%
google	23	24	-1	1,644	1,859	13%
sap	169	25	144	91	1,735	1808%
app	17	26	-9	2,573	1,720	-33%
world	27	27	0	1,441	1,650	14%
mnet	30	28	2	1,132	1,568	38%
smart	26	29	-3	1,475	1,331	-10%
web	33	30	3	815	1,126	38%
orange	32	31	1	924	1,072	16%
red	24	32	-8	1,603	1,043	-35%
msd	43	33	10	534	956	79%
school	37	34	3	696	872	25%
bank	39	35	4	622	780	25%
casa	28	36	-8	1,283	771	-40%
telefonica	45	37	8	519	768	48%
zone	51	38	13	435	701	61%
movistar	118	39	79	156	660	324%

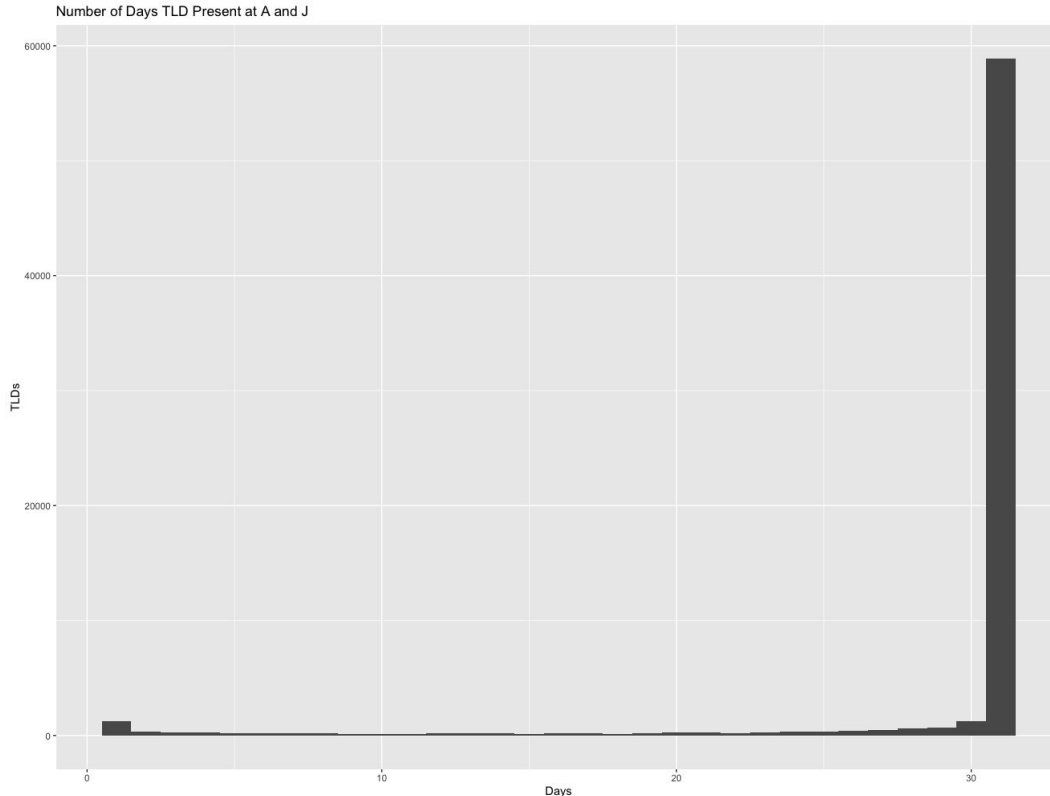
String Query Volume



- One day of A and J observed 3,430,602,835 strings for the pattern `^[a-z]+\.\$'`

Percentile	Queries	TLDs
0.1000000	1	3087542551.5000
0.2000000	1	2744482268.0000
0.3000000	1	2401421984.5000
0.4000000	1	2058361701.0000
0.5000000	1	1715301417.5000
0.6000000	1	1372241134.0000
0.7000000	1	1029180850.5000
0.8000000	1	686120567.0000
0.9000000	2	343060283.5000
0.9500000	2	171530141.7500
0.9900000	4	34306028.3500
0.9990000	7	3430602.8350
0.9999000	32	343060.2835
0.9999900	640	34306.0283
0.9999990	15145	3430.6028
0.9999999	314716	343.0603

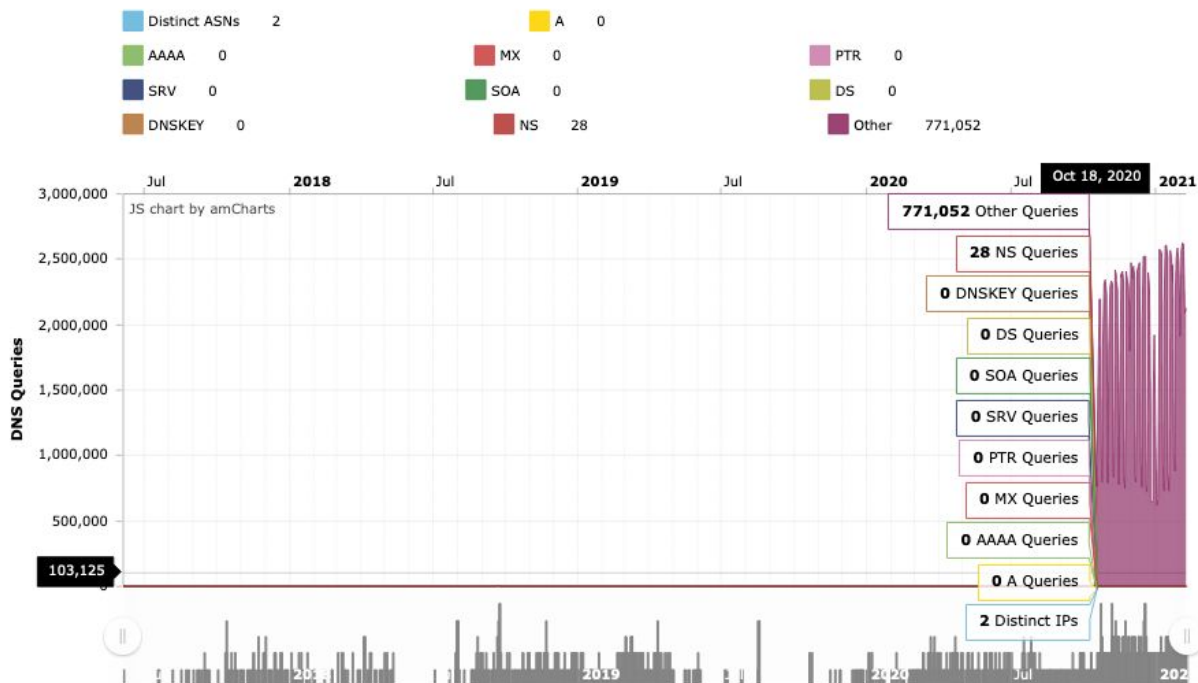
String Query Persistence



- One month of A and J root data matching previous REGEX, in which a TLD receive ≥ 500 queries for 1 out of the 31 days in January 21
- Vast majority (~86%) of strings are present every day during month
- 9,735 out of 68,615 strings were seen less than 31 days.

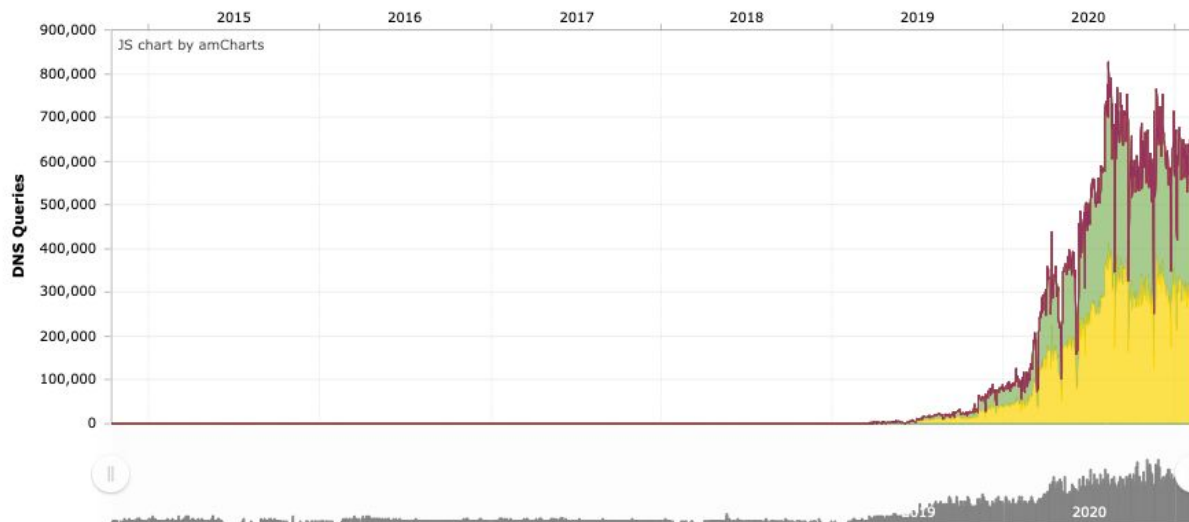
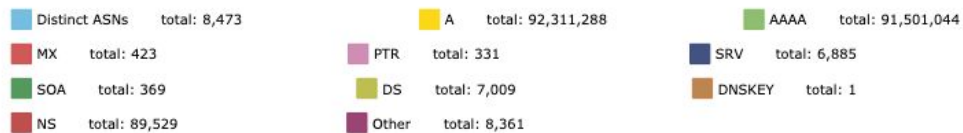
String Query Rapid Volume Increase

A and J Root Traffic for J051M946



IDN Strings

A and J Root Traffic for XN--C6H



IDN String Query Volume

TLD	Return Type	Total	A	AAAA	MX	PTR	SRV	SOA	DS	DNSKEY	NS	Other	RD	Countries	ASNs
xn--c6h	NXD	467	258	229	0	0	0	0	0	0	0	0	0	2	1
xn--9hb	NXD	23871	21528	2221	0	3	0	0	83	0	23	0	0	31	72
xn--6ii	NXD	68830	34541	34231	0	0	0	0	0	0	25	0	0	1	4
xn--mgbqly7cvaftr	NXD	231275	107	0	0	0	0	0	0	0	230920	0	0	5	4
xn--mgb2ddes	NXD	233369	137	1	0	0	0	0	0	0	232690	0	0	5	4
xn--mgbtf8fl	NXD	234777	101	0	0	0	0	0	0	0	233889	0	0	5	4
xn--mix082f	NXD	246012	11	0	0	0	0	3	0	0	244875	19	0	4	4
xn--nnx388a	NXD	260231	4	0	0	0	0	0	0	0	260022	27	0	5	4
xn--mgba3a4fra	NXD	260869	3	0	0	0	0	0	0	0	260532	0	0	4	2
xn--mgbqly7c0a67fbc	NXD	265647	0	0	0	0	0	0	0	0	265536	0	0	4	2
xn--mgbal9a5eva00b	NXD	265667	0	0	0	0	0	0	0	0	265444	0	0	4	2
xn--mgberp4a5d4a87g	NXD	265673	0	0	0	0	0	0	0	0	265549	0	0	4	2
xn--0qah2a36aa4215c	NXD	289503	145765	143891	0	0	0	0	0	0	0	3	0	4	12



١

السعوديه

اليمن

سوريا

澳門

臺灣

ايران

السعوديه

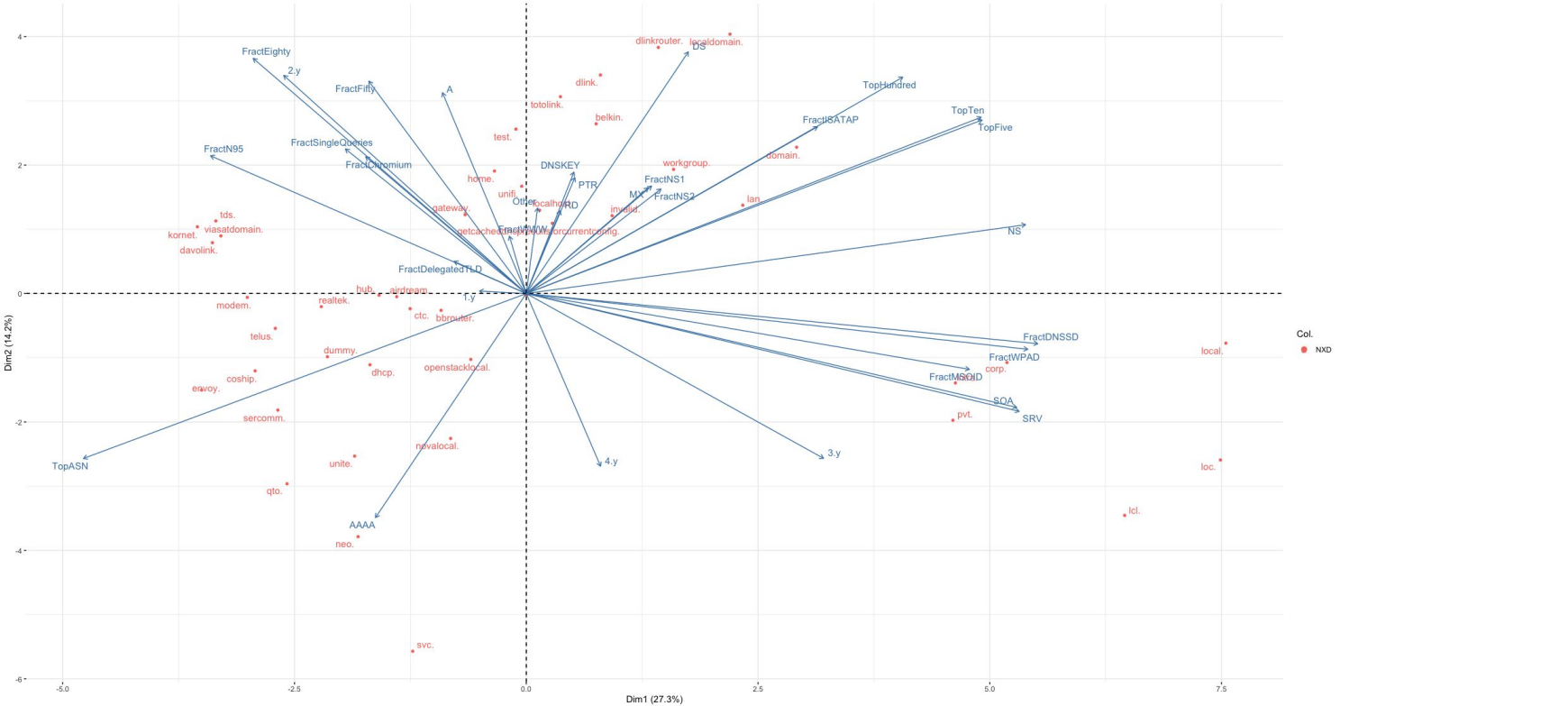
پاكستان

السعوديه

١ -

PCA Analysis of Top Strings

PCA - Biplot



String Query Volume & Source Affinity/Concentration

tld	request_count	sld_count	qname_count	dstip_count	dstip24_count	asn_count	date
dhcp	73,079,967	830,634	5,611,411	34,629	8,909	2,900	2021-02-08
svc	47,726,588	101,497	12,498,196	30,224	6,832	2,605	2021-02-08
bbrouter	30,787,187	5,094,605	5,919,697	22,241	4,779	2,033	2021-02-08
novalocal	11,386,555	76,386	596,335	24,298	6,961	2,626	2021-02-08
openstacklocal	10,680,435	201,293	2,090,396	31,036	6,416	1,746	2021-02-08
sercomm	10,395,878	476	8,555,996	4,945	1,498	510	2021-02-08
telus	9,178,900	2,933,467	2,991,455	11,134	3,005	922	2021-02-08
realtek	9,150,083	4,868,688	5,085,186	28,208	6,342	2,959	2021-02-08
coship	9,021,391	4,227,154	4,287,714	3,753	739	267	2021-02-08
ctc	7,201,080	900,629	1,468,549	11,942	2,746	711	2021-02-08
unite	5,412,302	59	414	361	172	95	2021-02-08
dummy	2,499,527	23,393	24,677	39,522	15,992	6,761	2021-02-08
neo	2,388,550	2,704	587,488	3,168	915	409	2021-02-08
envoy	1,529,341	5,021	5,205	3,860	274	84	2021-02-08
qto	27,265	170	334	315	76	27	2021-02-08
modern	884	191	254	449	208	113	2021-02-08