Examine Marketing Performance via Web Log Analysis

Szuhui Wu

2/13/2018

Questions

- In 2017 a marketing firm was hired to drive up traffic to the company website.
    - Was there increased traffic in 2017?
    - Did the Ad/Campaign bring in *Quality* visitors in 2017?

- How did visitors navigate the site?

- What other interesting patterns exist?

Definition

- Visitor
  - Each unique IP / user_agent combination

- Session
  - For each visitor, a new session is constructed if
    - idle time >=30mins, OR
    - has "gclid" (google ad click) or "utm_" campaign tag

- Page
  - A hit of an actual web page url or a pdf download
  - Excludes non pages (images, styles etc.)
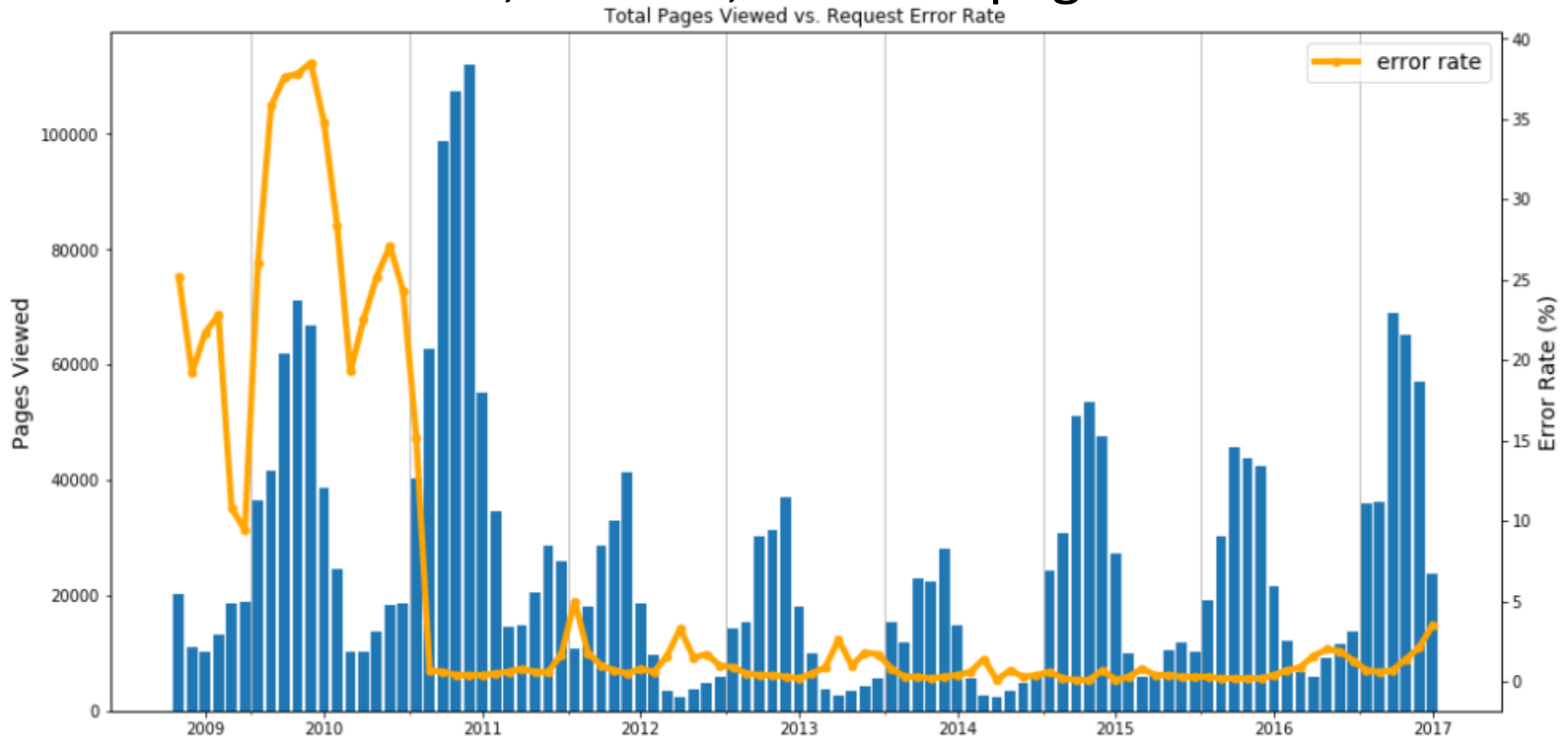
Referer(sic) Types
(where the traffic originated)

- Organic Search
  - Search Engine: Google, Bing, Yahoo, Baidu…
- Direct/Referer Hidden
  - No referer / Hidden (some visitors hide their info)
  - From company homepage
- Referral/Ad/Campaign
  - gclid tag (google ad click) / campaign tag (utm)
  - other local sites
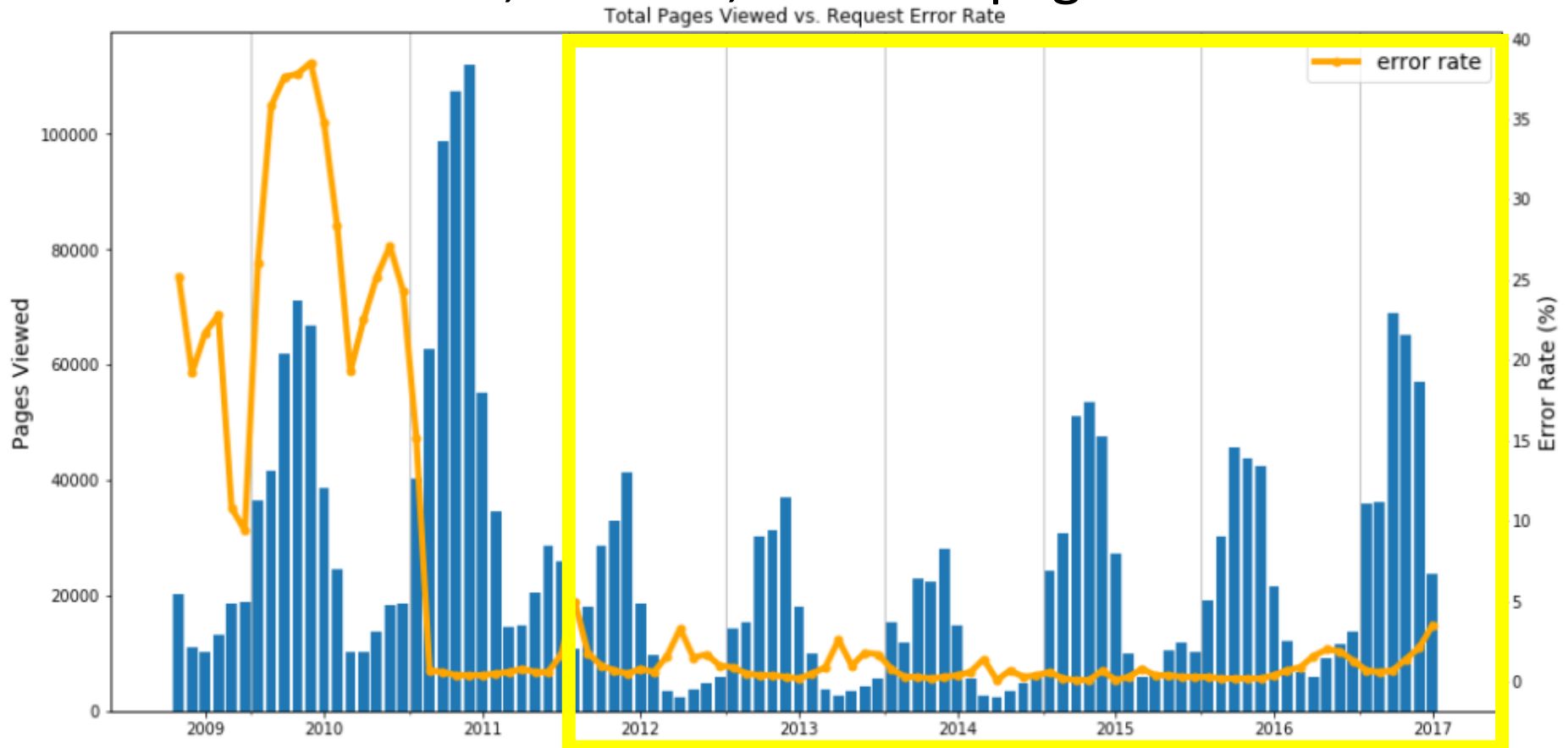 - other referral sites (Facebook, email…)

Final Data

- 1/2012 – 6/2017, focus on 2017
- Excludes bots, outliers, and non-pages

Final Data

- 1/2012 – 6/2017, focus on 2017
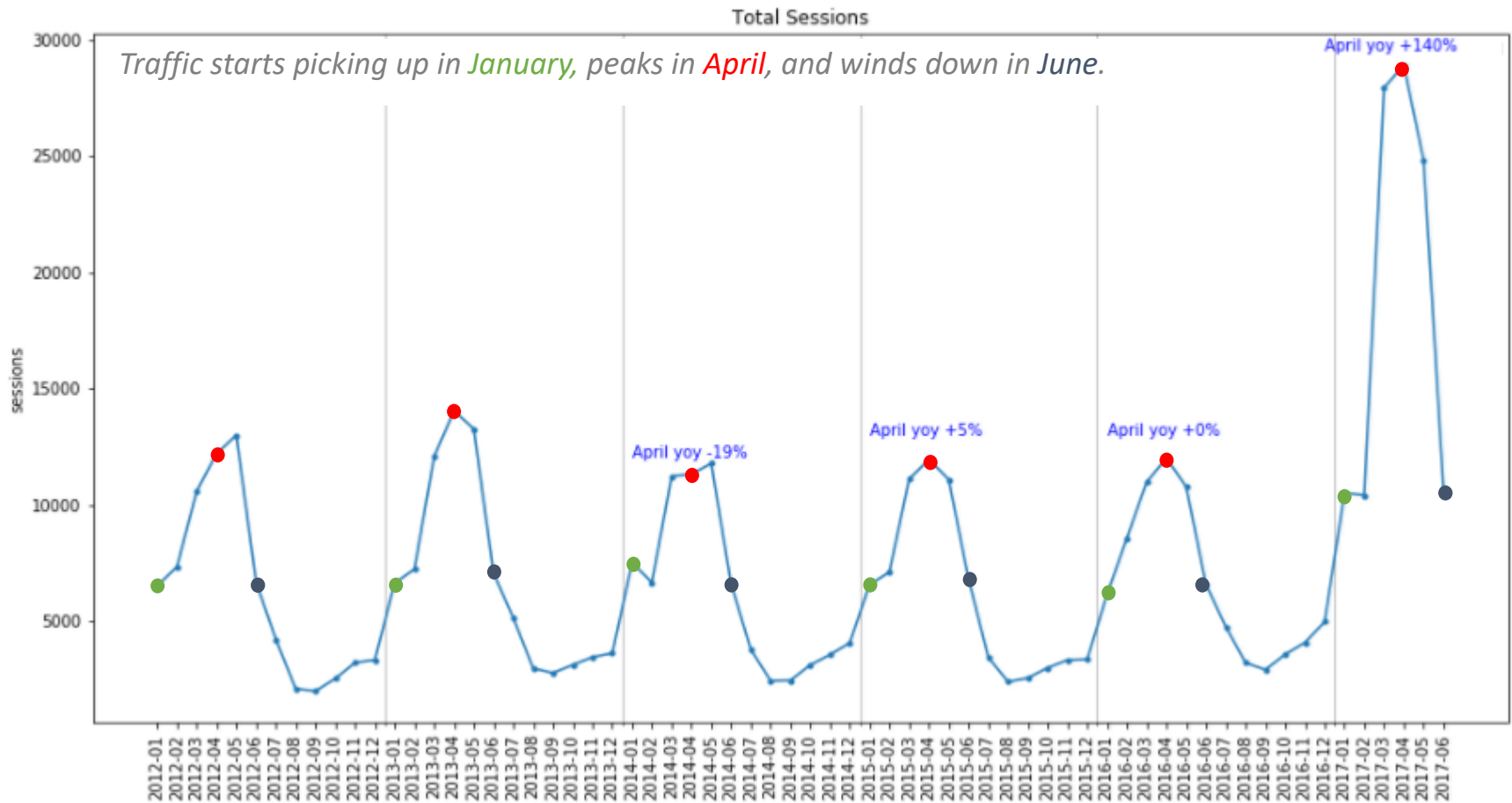- Excludes bots, outliers, and non-pages

Total Pages Viewed vs. Request Error Rate

Final Data

- 1/2012 – 6/2017, focus on 2017
- Excludes bots, outliers, and non-pages

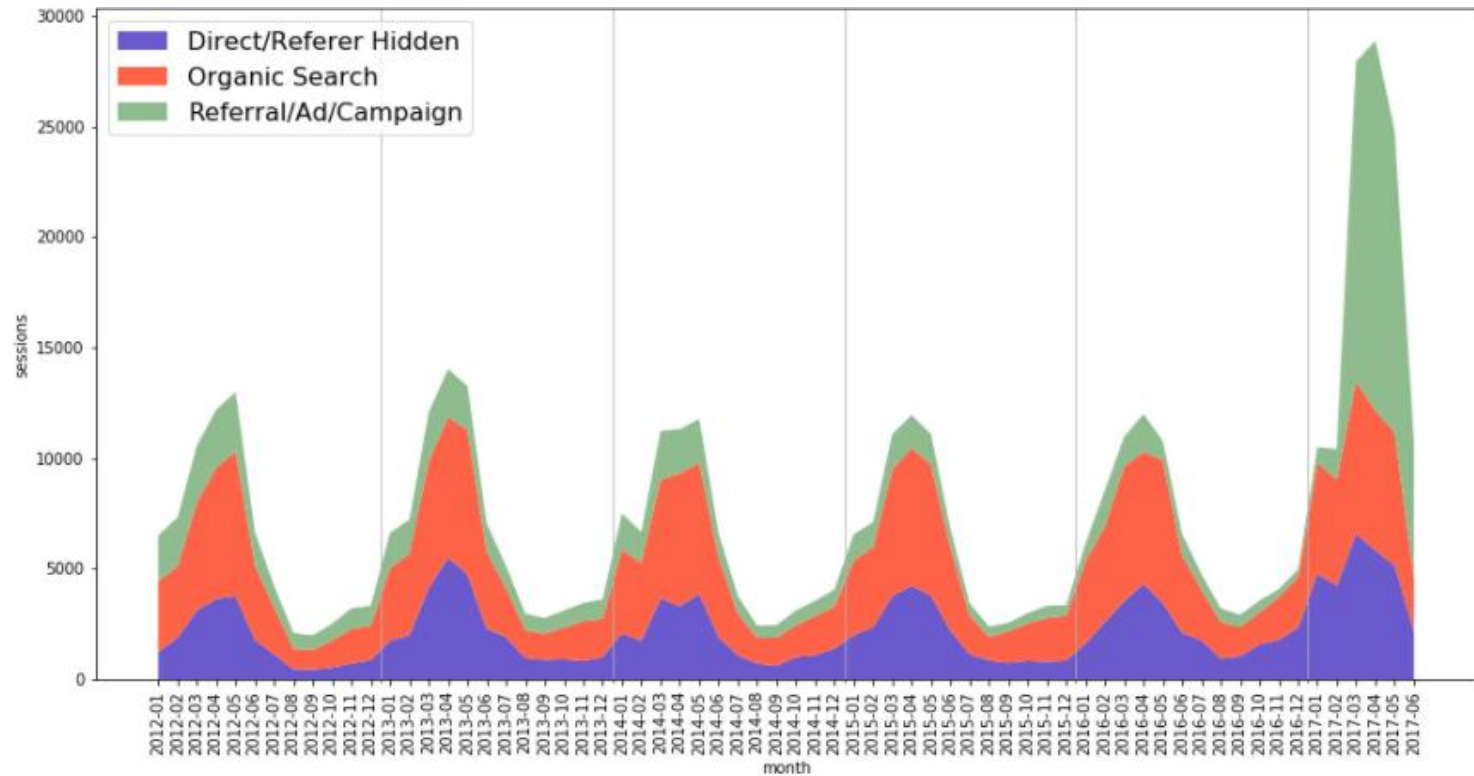Total Pages Viewed vs. Request Error Rate

2017 Marketing Efforts

- Was there increased traffic in 2017?

- Did the Ad/Campaign bring in *Quality* visitors in 2017?

As expected, traffic increased in 2017
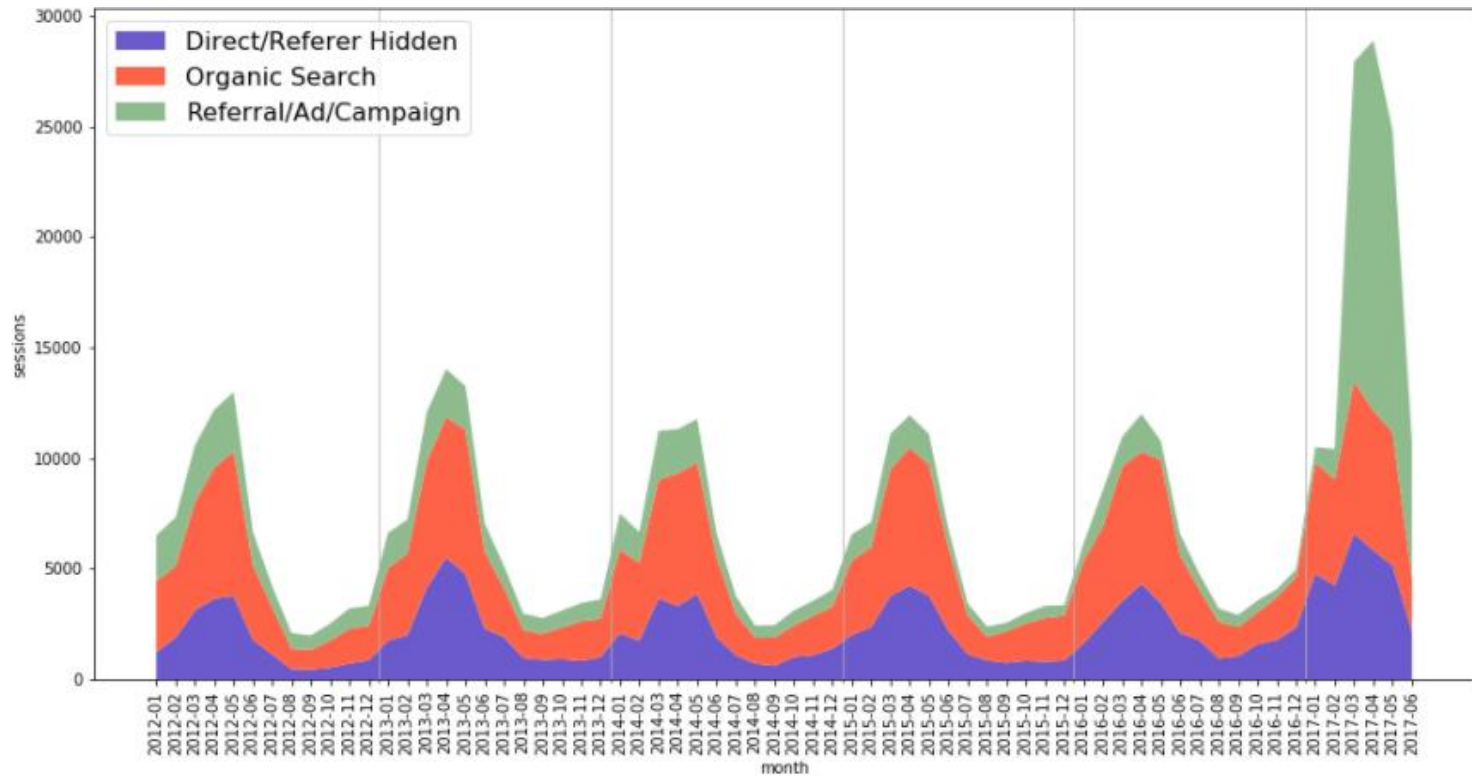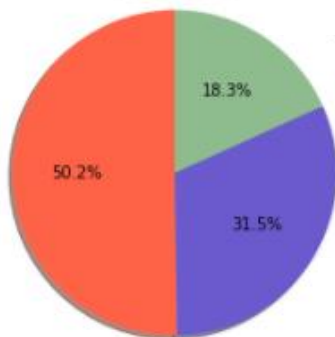

Total Sessions

Traffic starts picking up in January, peaks in April, and winds down in June.

April yoy +140%

April yoy -19%

April yoy +5%

April yoy +0%

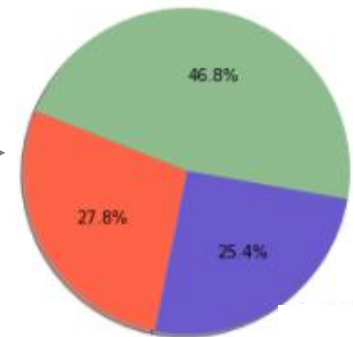# Sessions by Referer Types
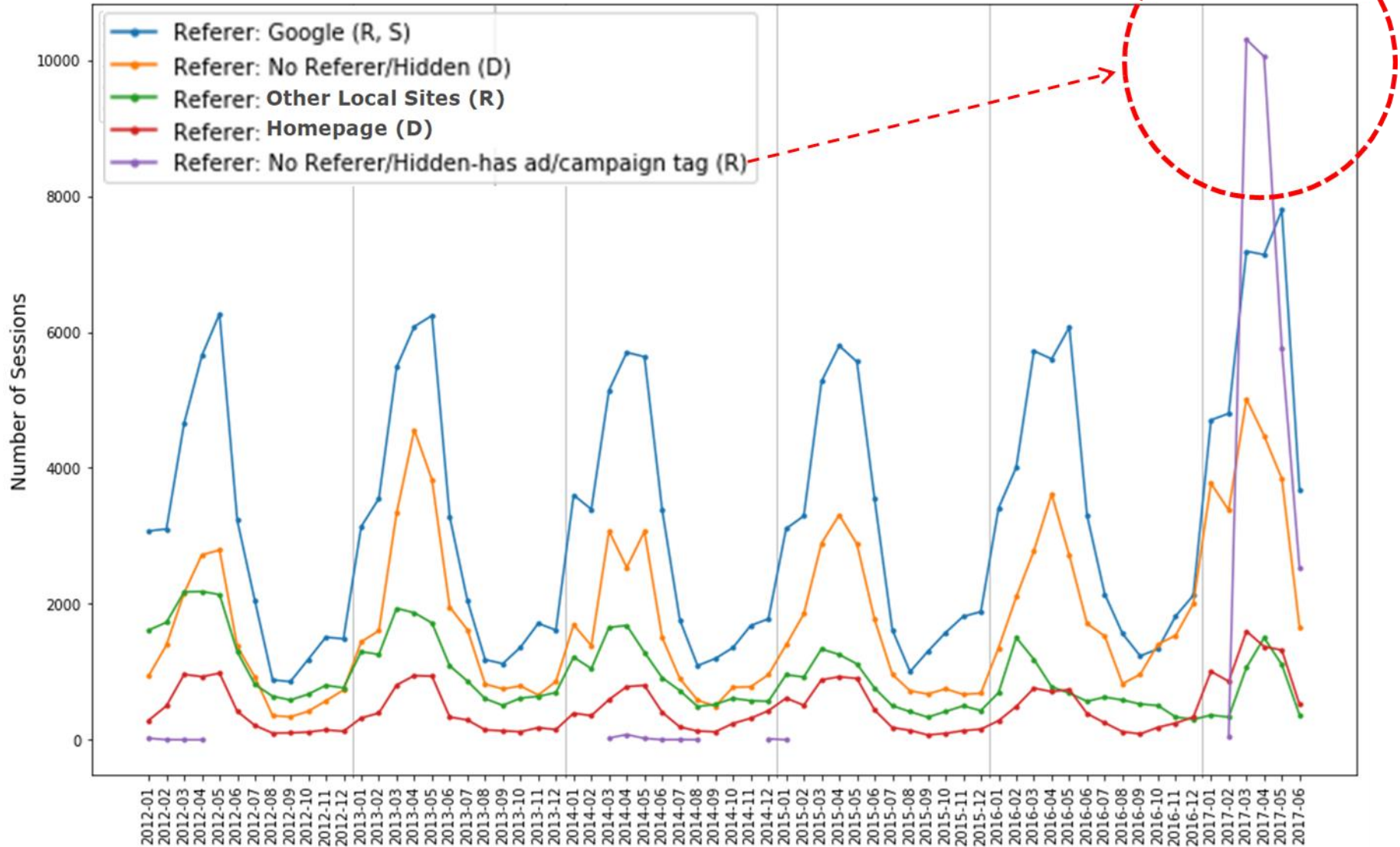
Sessions by Referer Types

2012 - 2016

2017 Jan-Jun

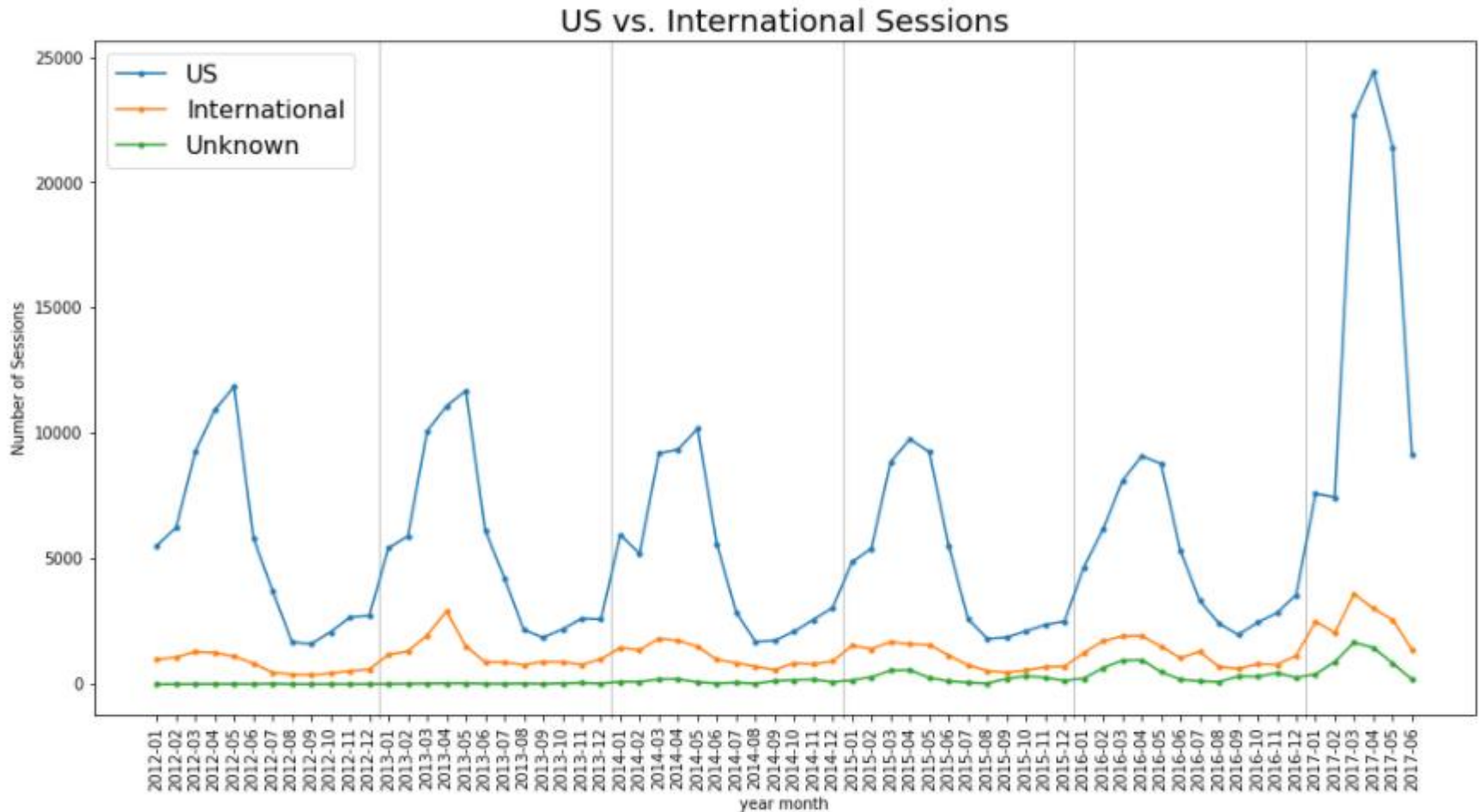Big percentage shift to Referral/Ad/Campaign in 2017

Top Referer URLs

US vs. International Traffic

- The increase was mostly in the United States.

Sessions Originating Internationally

- Most international traffic originated from China.
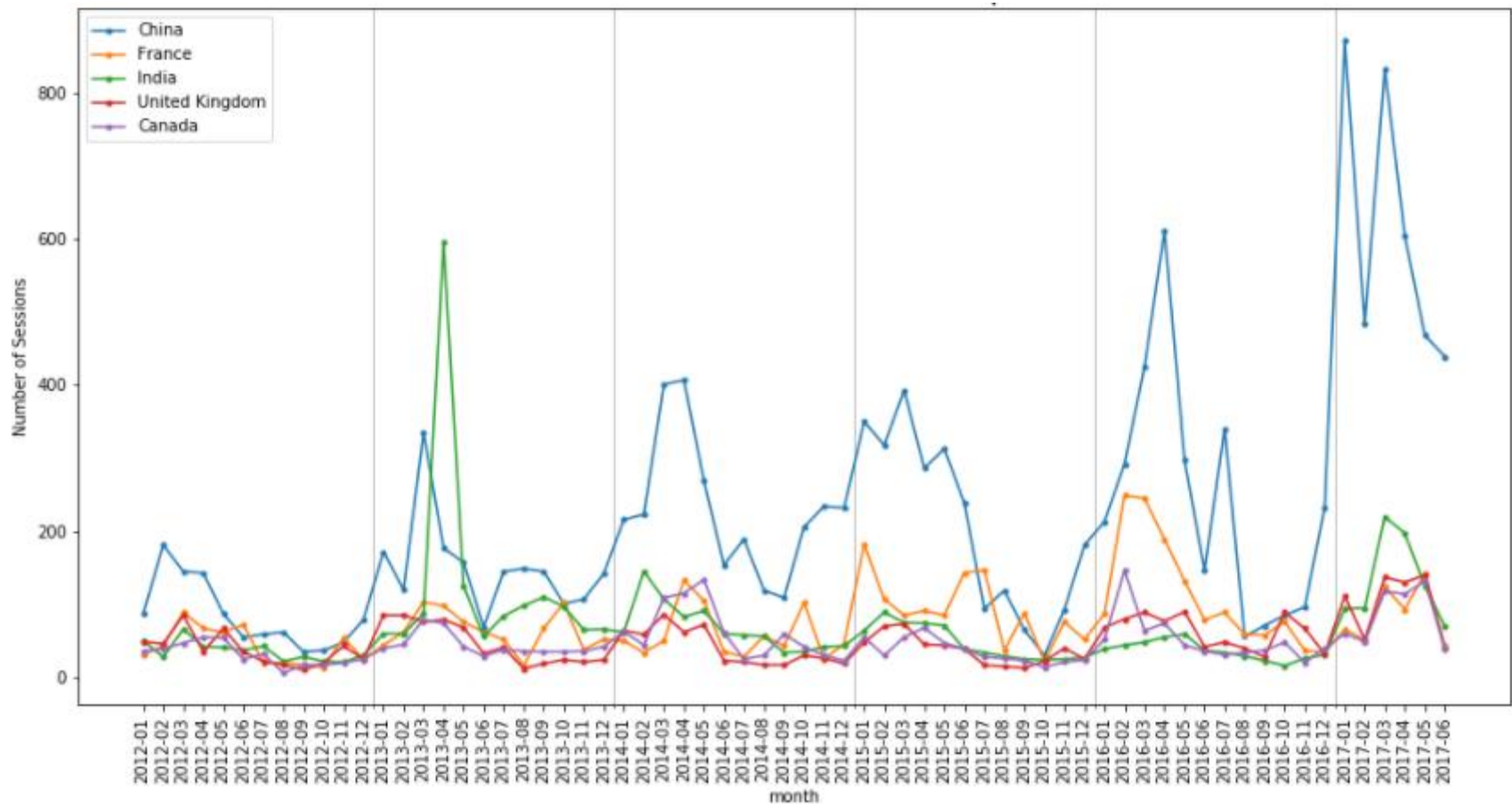
- The Ad/Campaign did drive up traffic in 2017.
- Did the Ad/Campaign bring in *Quality* visitors?

Visitor Return Rate

In 2017, visitor return rate improved slightly.
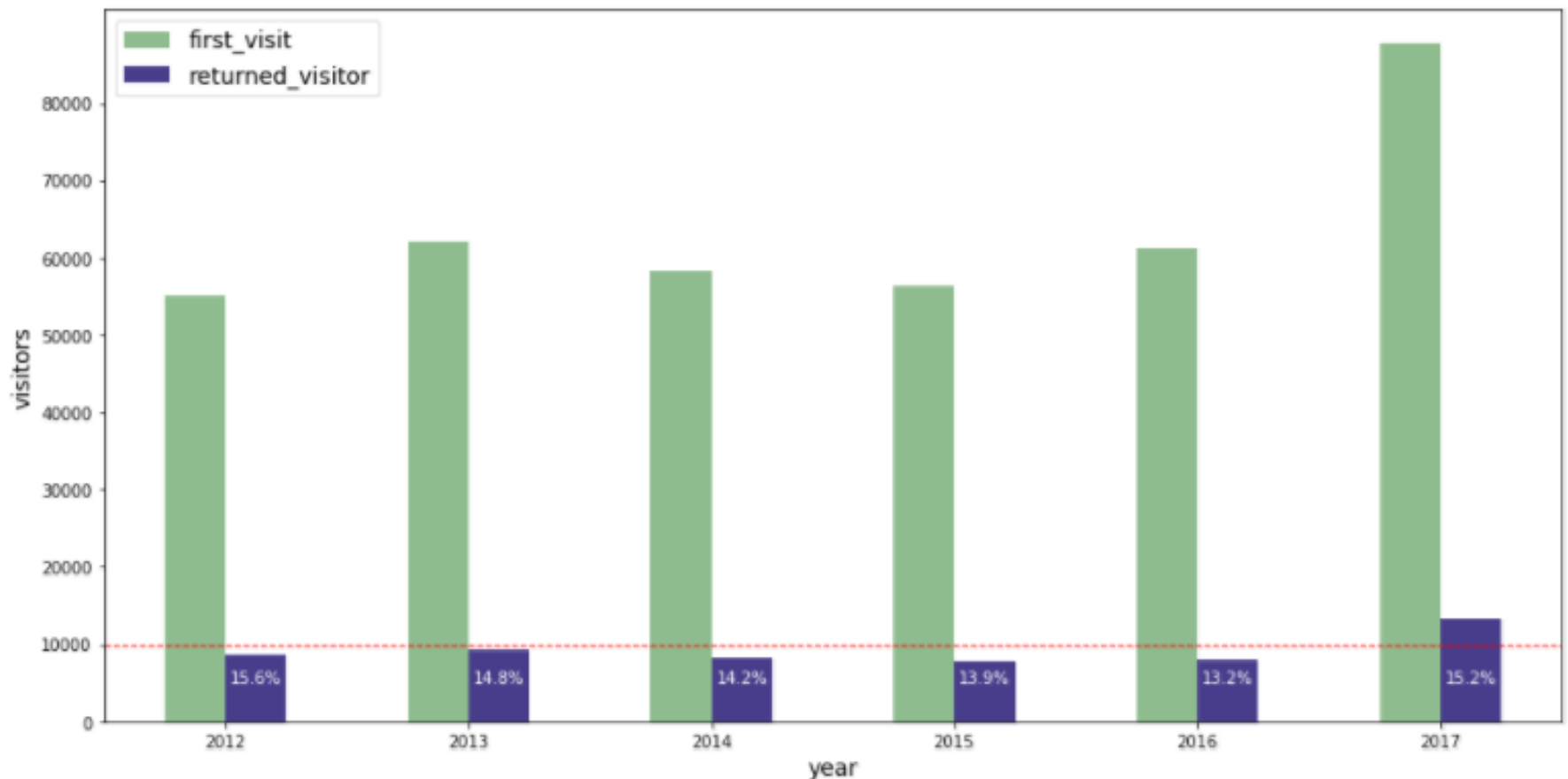
Visitor Return Rate

Consistently, Referral/Ad/Campaign traffic is <u>least</u> likely to generate returned visitors
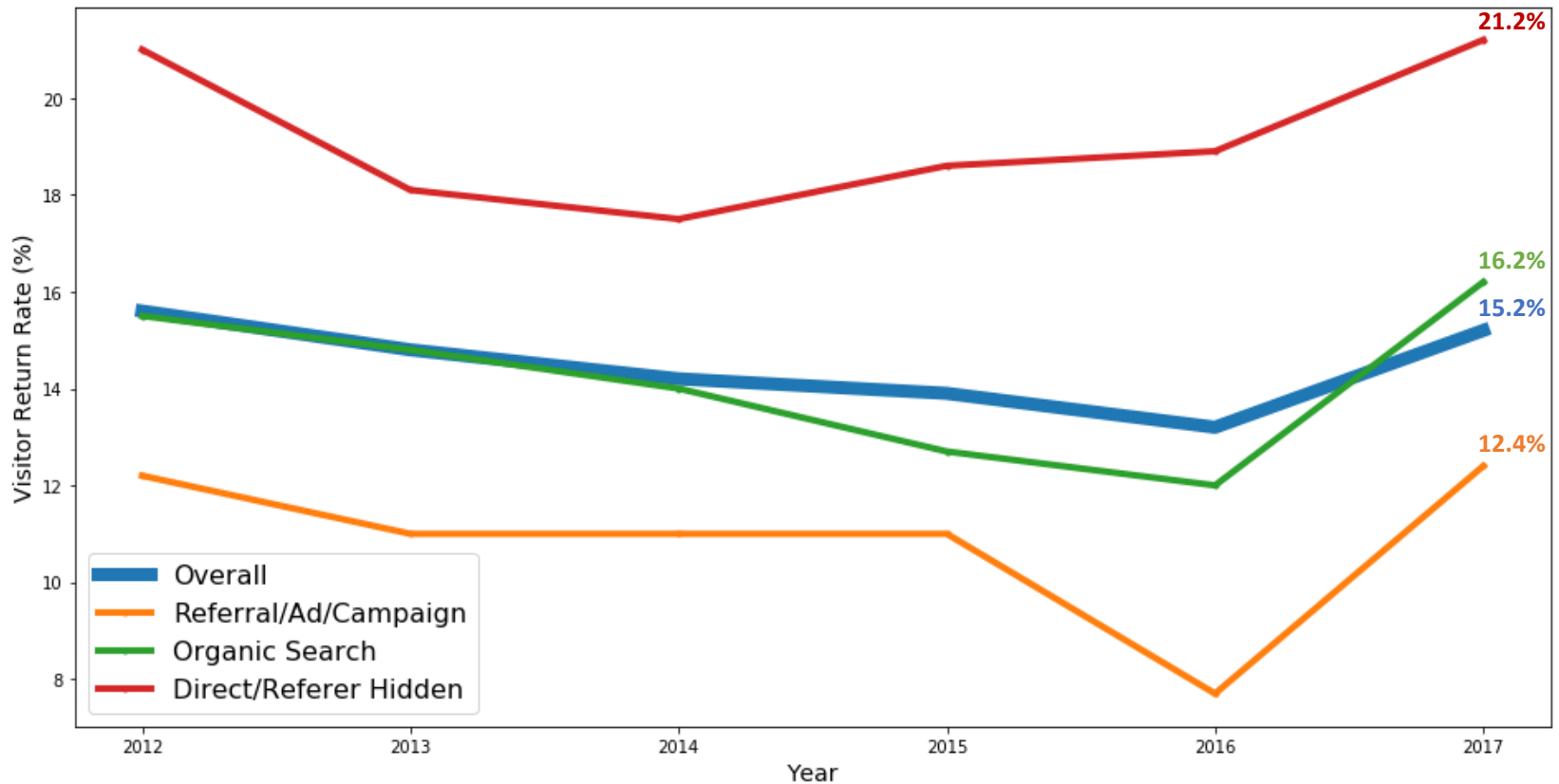
Bounce Rate (% of single-page sessions)

Bounce rate of a website = total number of bounces across all the pages on the website (in a given time period) / total number of entrances across all the pages on the website (in the same time period)

(the lower the better)

Single-page sessions account for most of the growth
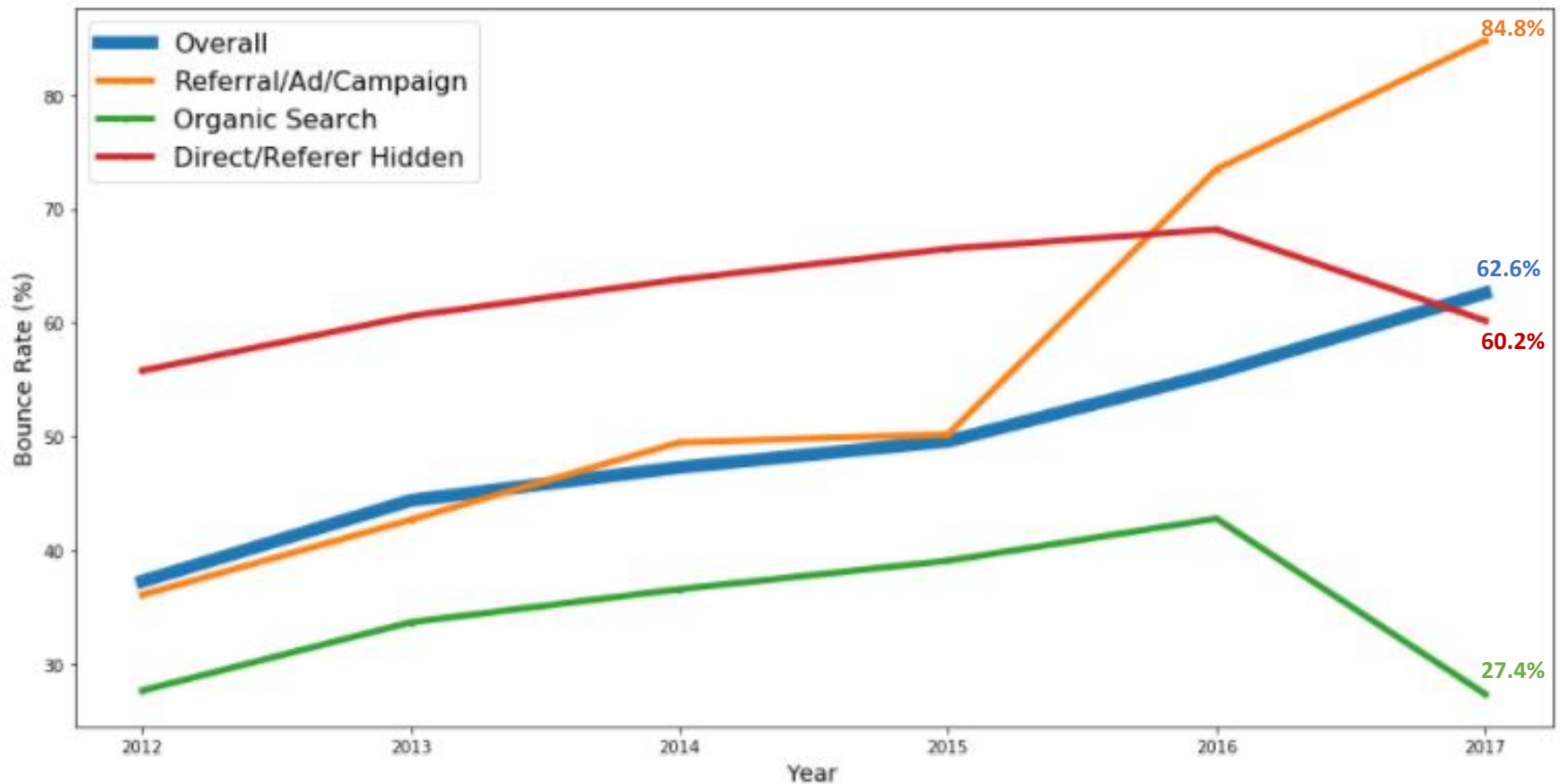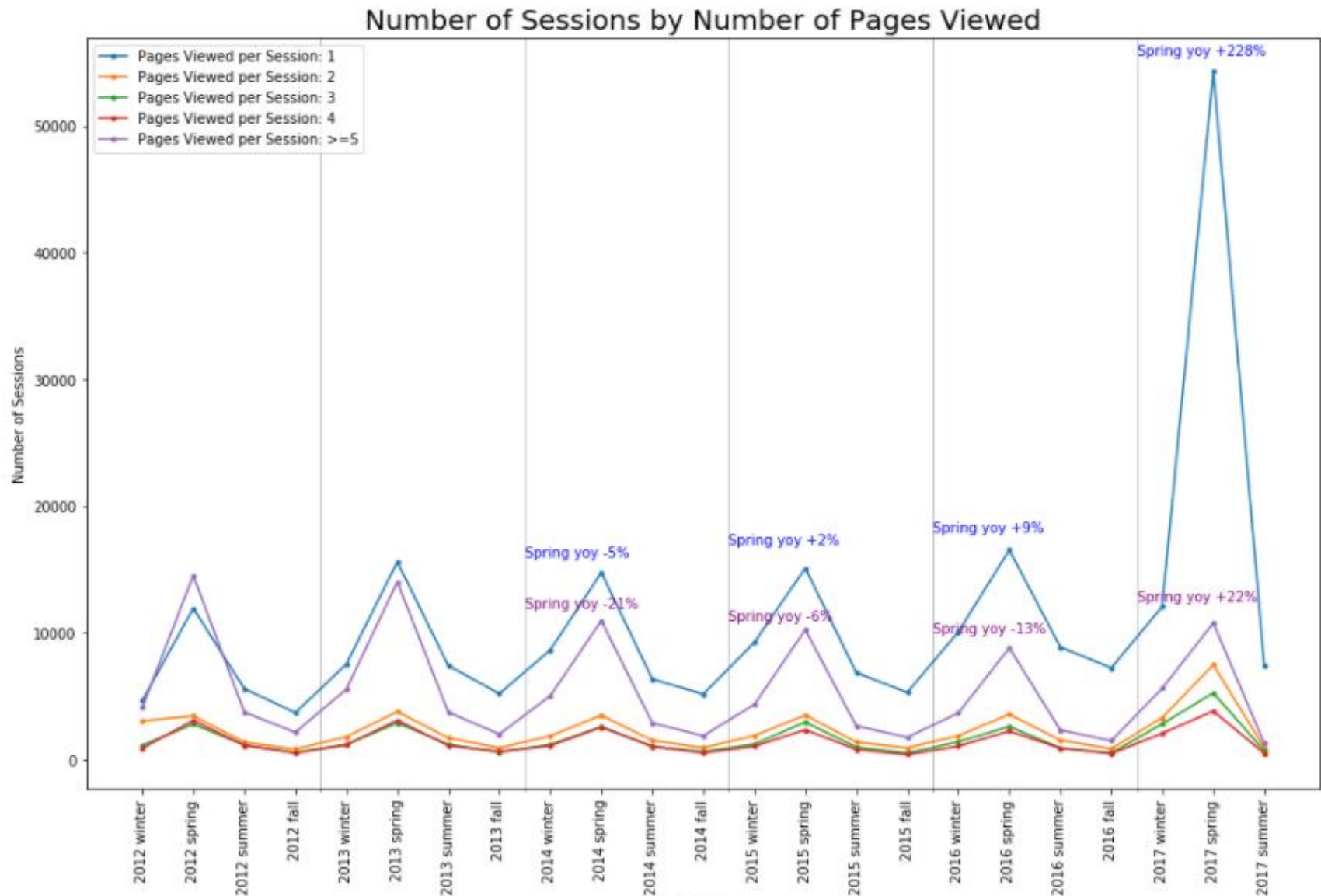

Number of Sessions by Number of Pages Viewed

Legend:
- Pages Viewed per Session: 1
- Pages Viewed per Session: 2
- Pages Viewed per Session: 3
- Pages Viewed per Session: 4
- Pages Viewed per Session: >=5

Spring yoy +228%
Spring yoy -5%
Spring yoy -21%
Spring yoy +2%
Spring yoy -6%
Spring yoy +9%
Spring yoy -13%
Spring yoy +22%

Y-axis: Number of Sessions

X-axis labels: 2012 winter, 2012 spring, 2012 summer, 2012 fall, 2013 winter, 2013 spring, 2013 summer, 2013 fall, 2014 winter, 2014 spring, 2014 summer, 2014 fall, 2015 winter, 2015 spring, 2015 summer, 2015 fall, 2016 winter, 2016 spring, 2016 summer, 2016 fall, 2017 winter, 2017 spring, 2017 summer

* Spring = March, April, May

Visitors Path from Referral/Ad/Campaign

referer type	%	→	entry page	%	→	2nd page	%	→	3rd page	%
Referral/Ad/Campaign	46.8		**1** Ad/Campaign Lead (achieve)	26.3		exit	92.7			
						Home page	5.7		exit	29.7
						Courses Query Main Page	0.6		Courses Query Main Page	42.0
						ce	0.3			4.0
						ne	0.1			3.9
						Tuits students	0.1		others	20.4
						others	0.5			
			2 Ad/Campaign Lead (undergrad-grad)	16.0		exit	95.4			
						Home page	3.5		exit	38.5
						Courses Query Main Page	0.4		Courses Query Main Page	35.5
						ne	0.1			3.0
						HS	0.1			2.7
						Visiting students	0.1		others	20.3
						others	0.4			
			3 Home page	15.9		exit	81.1			
			4 Ad/Campaign Lead (nutrition)	7.5		Courses Query Main Page	9.5		exit	47.3
			5 A...)	6.1	Appl... ...ns			1.0		Courses Query On-line	11.3
			6 Ad/C... ...)	5.6		...ts	0.8			5.6
			7 ...e	4.6		...ce	0.7			5.5
			8 ...)	3.6		...e	0.6		others	30.3
			9 ...)	3.4		others	6.3			
			10 others	11.0						
Organic Search	27.8									
Direct/Referer Hidden	25.4									

Did Ad/Campaign Help?

- Traffic did increase.

- Visitor return rate improved slightly.

- Website bounce rate worsened.

- Ad/Campaign lead pages had very high bounce %

→ These are mixed results (What does 'good' looks like?)

→ Other possibility exist where weblog data is unclear
  - Visitors filled out the form and left?

→ Form submission data would be useful

→ Better understanding of the campaign would help

How did Organic/Direct Visitors Navigate the Site?

- Ad/campaign traffic visitors did not continue to other pages (84.8 bounce %)

- How about Organic Search & Direct visitors?
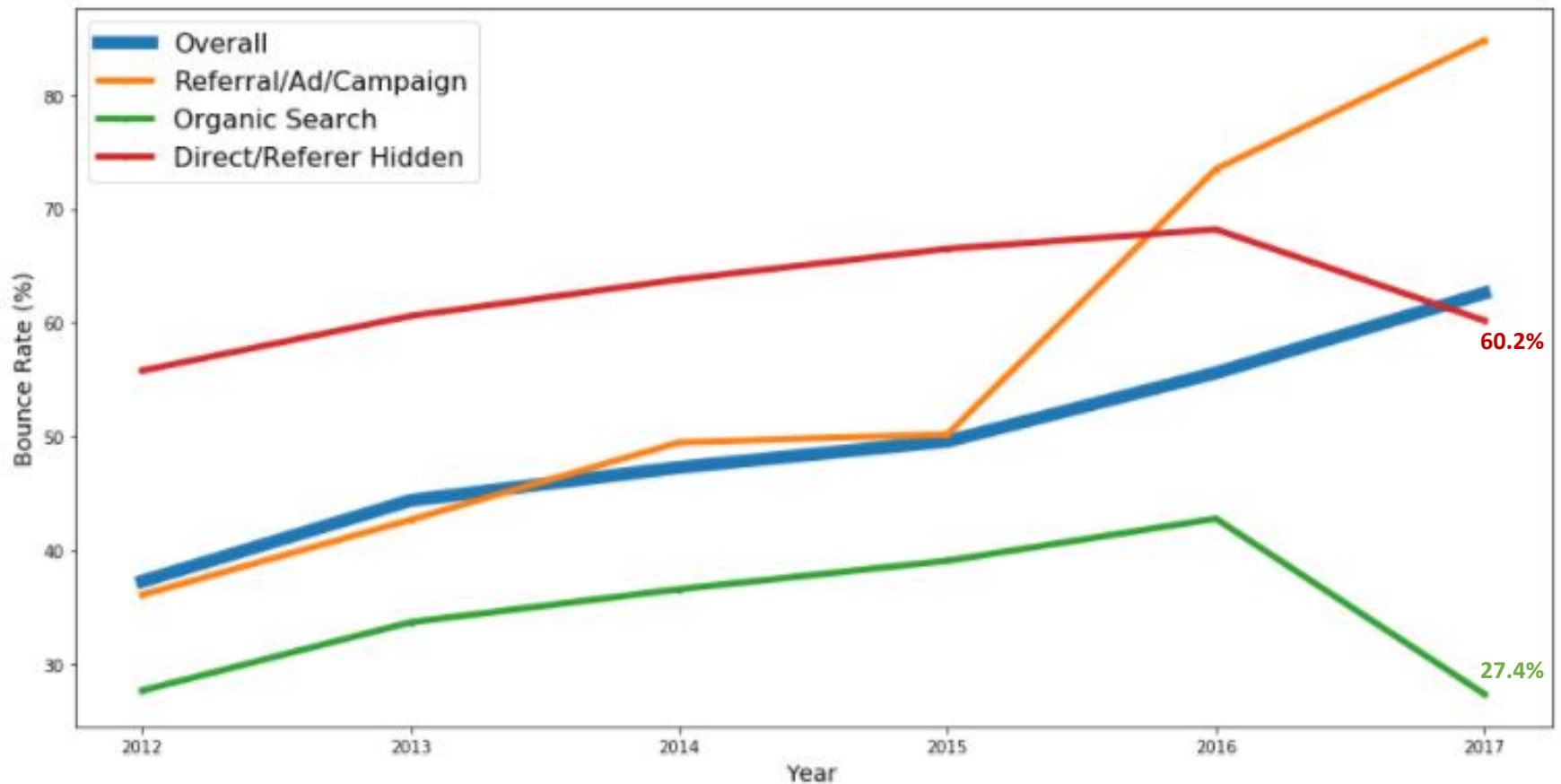
Visitors from Organic Search / Direct

- 43.1% Single-page sessions
- 14.3% Two-page sessions
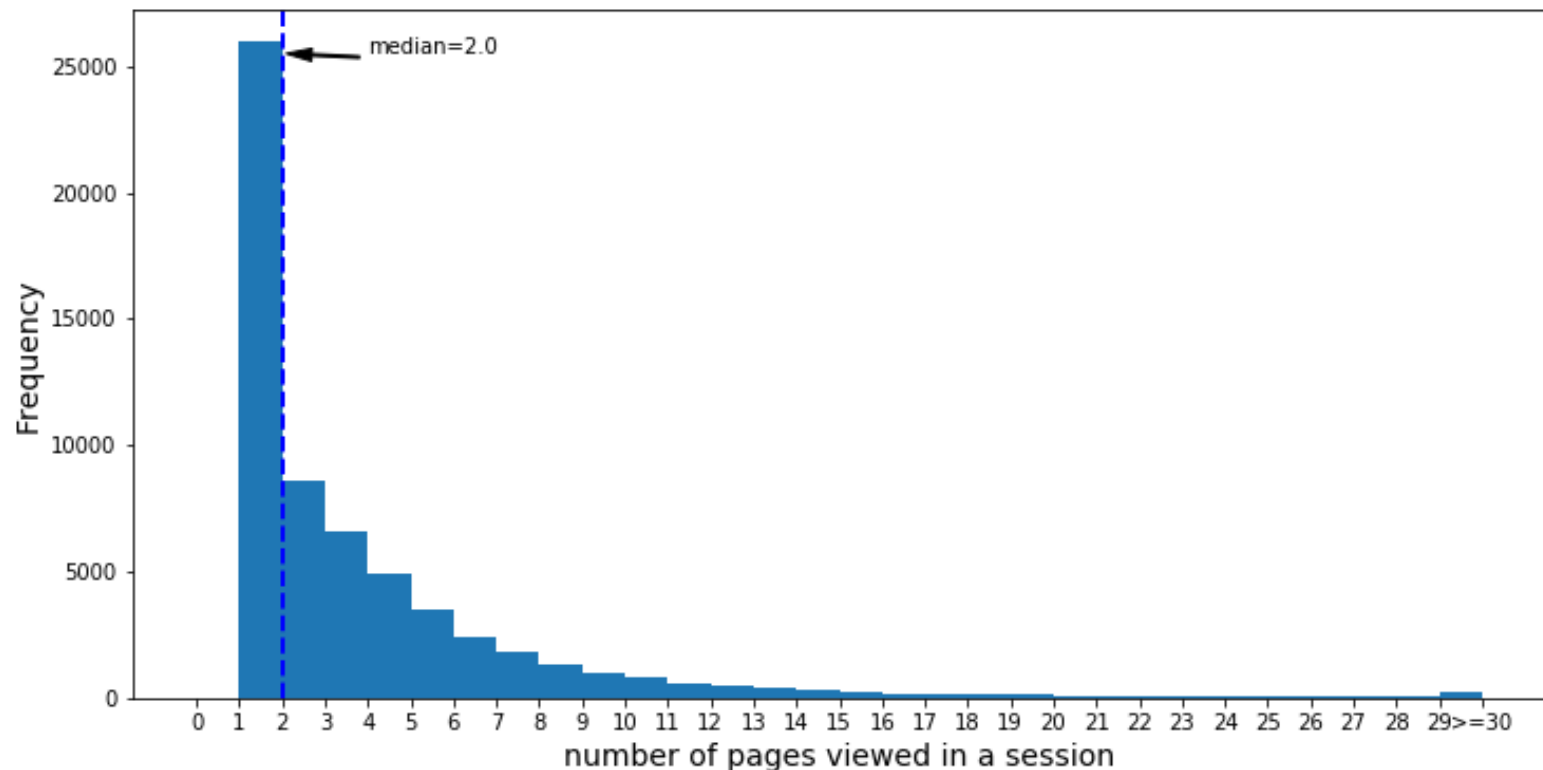
Bounce Rate (% of single-page sessions)

Bounce rate of a website = total number of bounces across all the pages on the website (in a given time period) / total number of entrances across all the pages on the website (in the same time period)

(the lower the better)

Honorable Mention Topic:
Bot Traffic (aka Spiders, Crawlers)

- How I identified bots:
  - Python 'user_agents' package captures most of the known bots
  - Keywords (8LEGS, python, email address…)
  - Outliers Exclusion
    - IP's with 'more than 30 sessions per season (3 months)'
    - IP's with 'more than 50 pages per session'
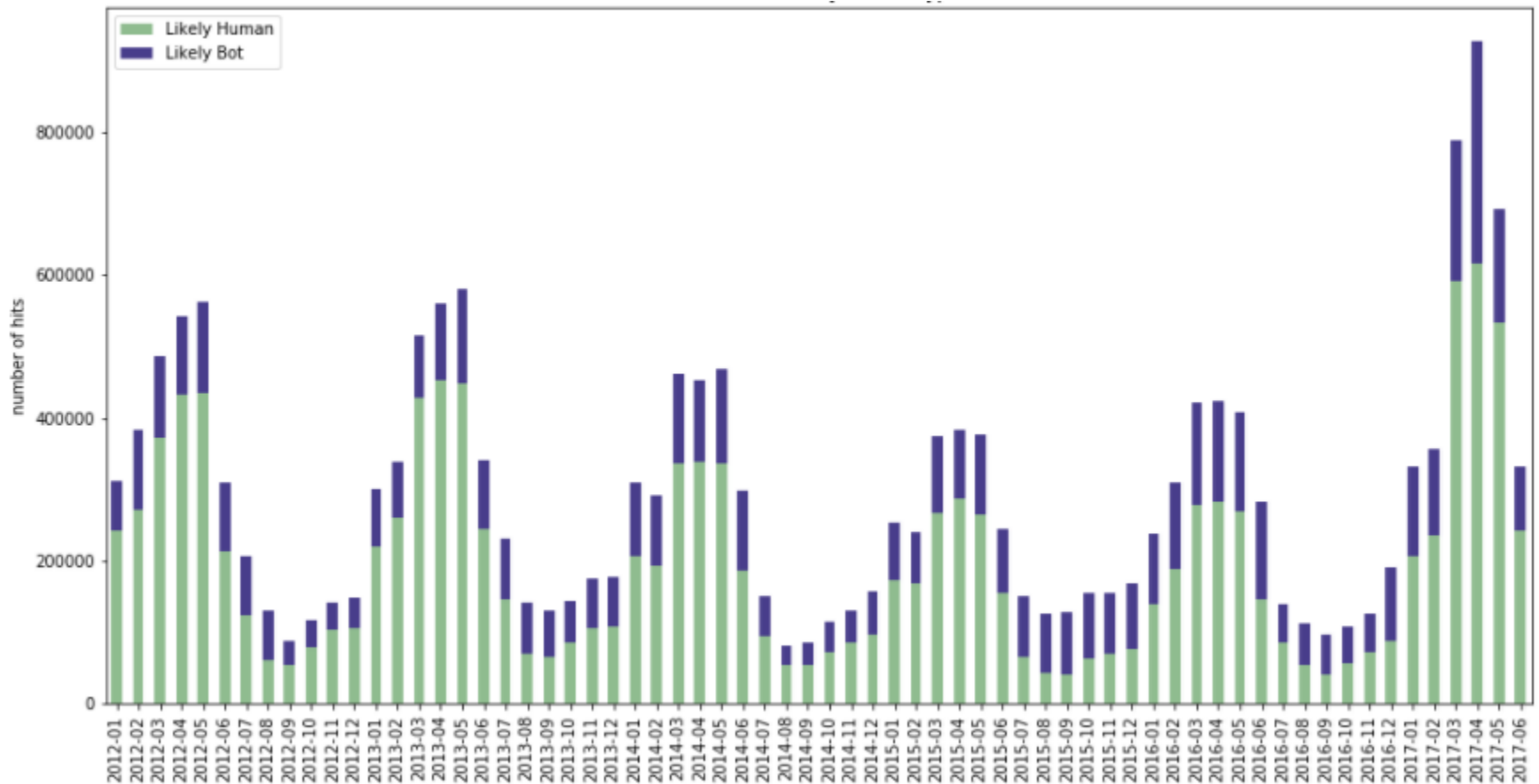   - IP's with 'session(s) longer than 2hrs'

Bots – by Number of Unique IPs
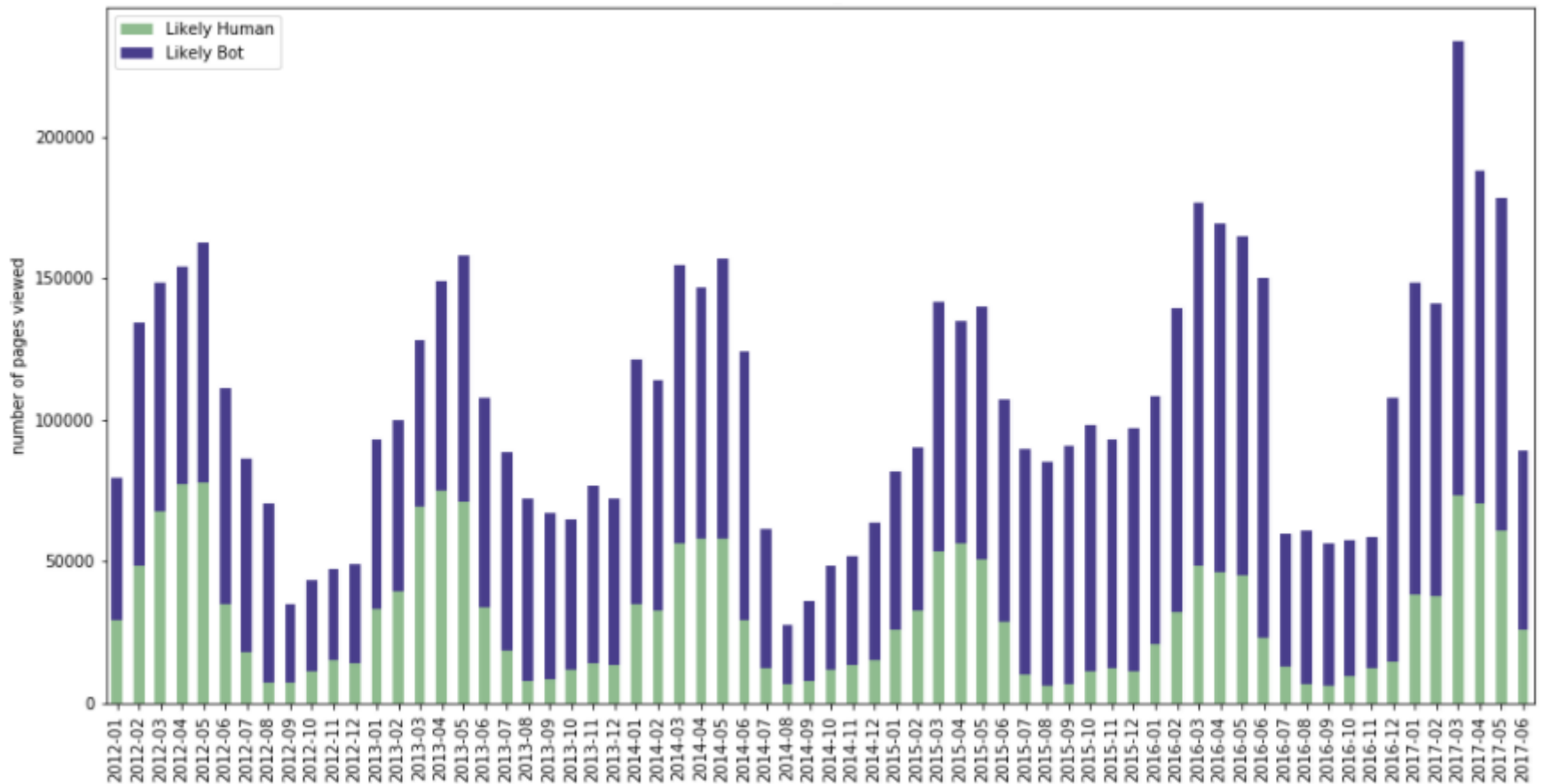
- 7.2 % of unique IPs are bots in 2017
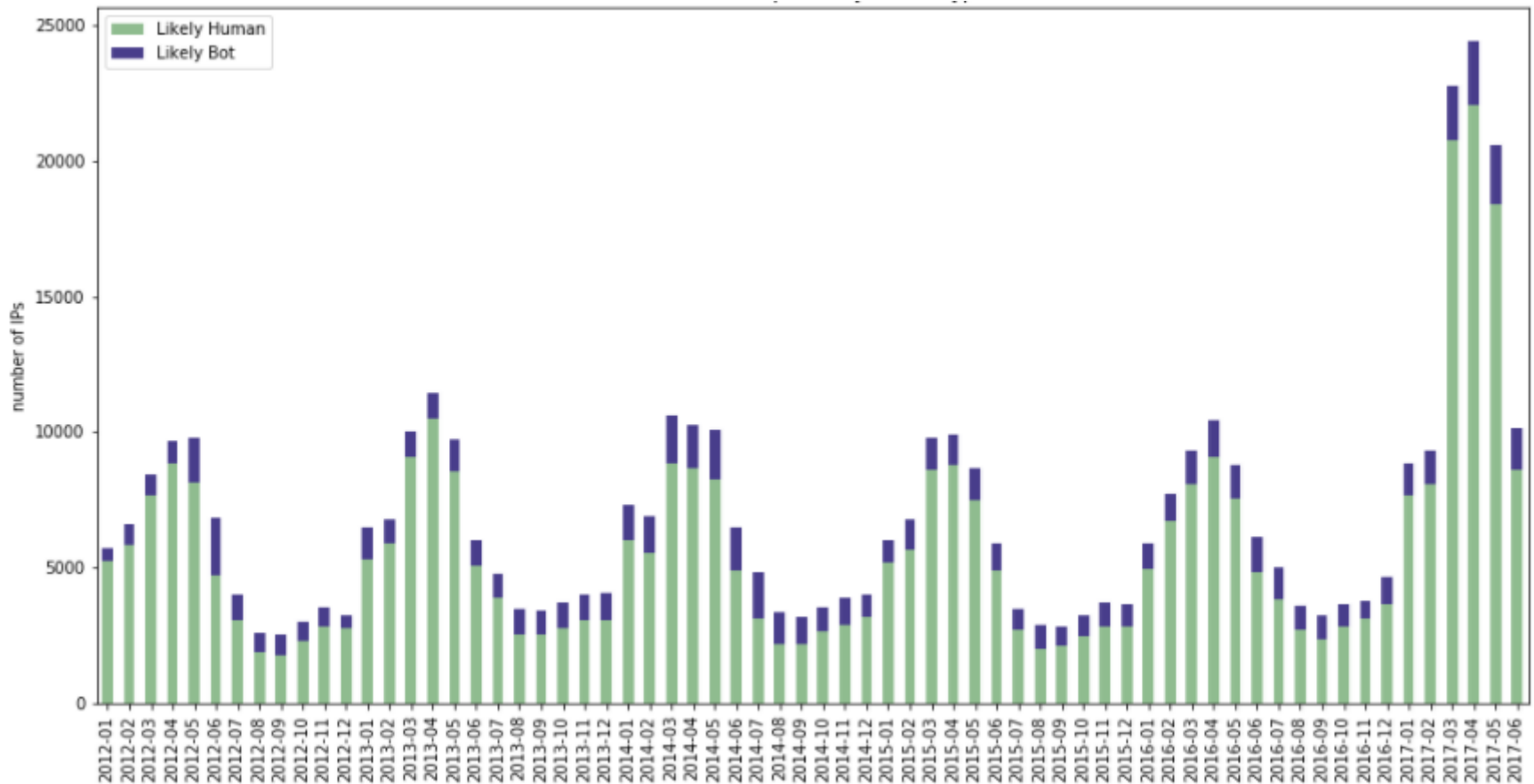
Bots – by Number of Hits

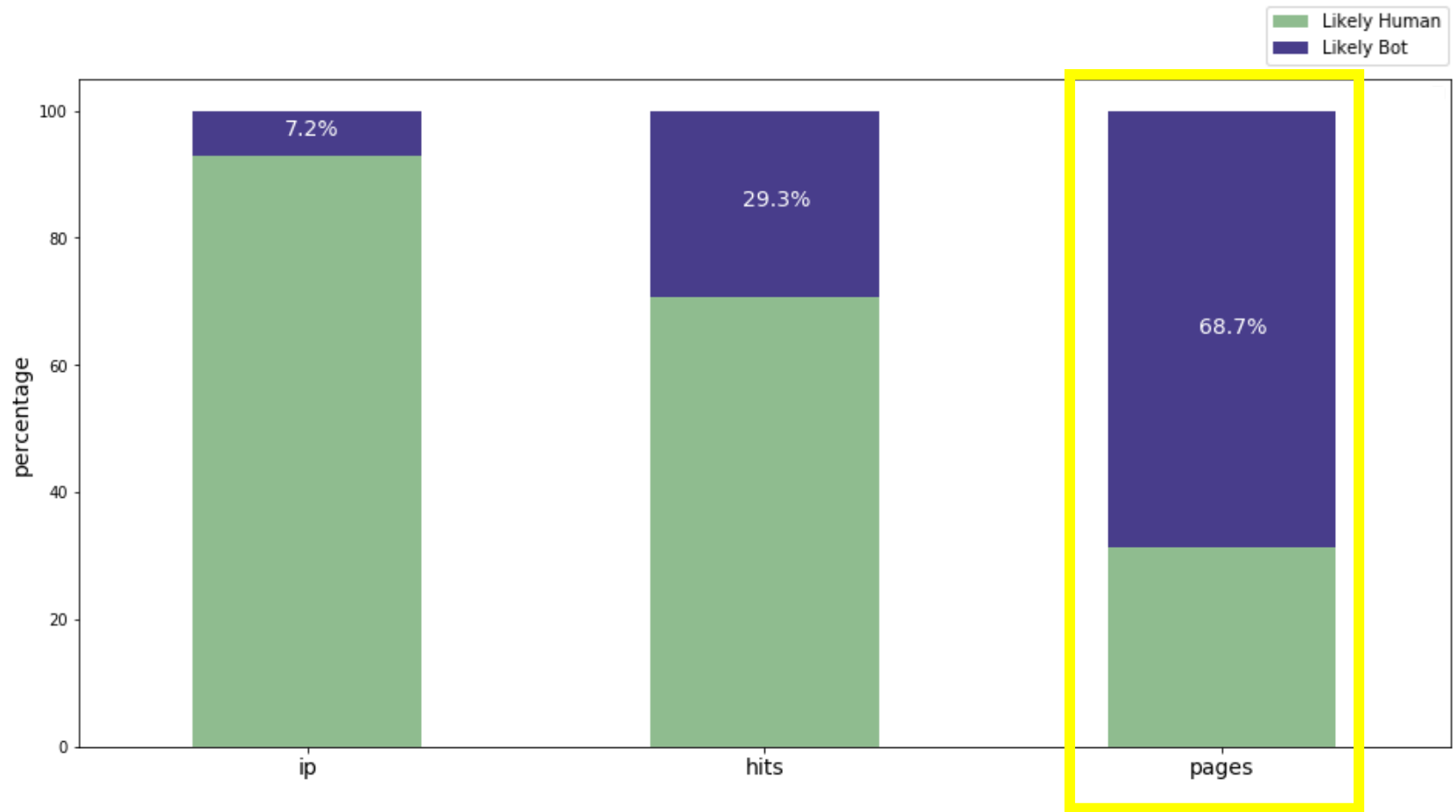- 29.3 % of hits are bots in 2017

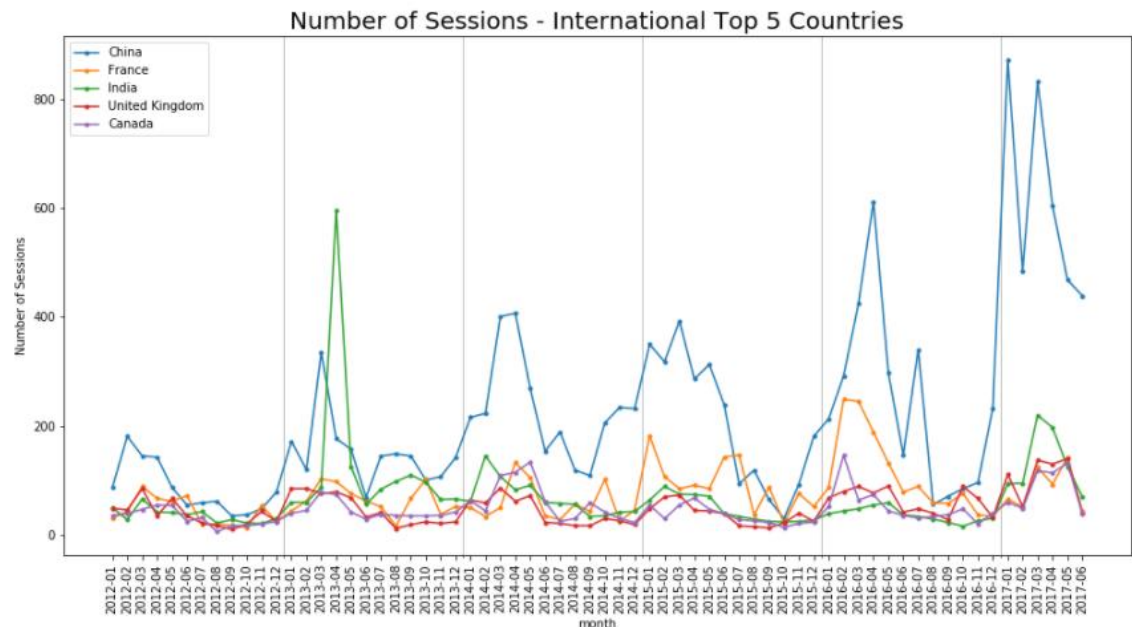# Bots – by Number of Pages Viewed

- 68.7 % of pages are bots in 2017

2017 Bot Percentages

- Do we know if this is normal?
- Integrity of all the data shown today depends on classification of bots.
- Classifying Bots is not an easy task.

Number of Sessions - International Top 5 Countries

Looking Forward

- How can these analyses be helpful?

- Form submission data would be useful

- Better understanding of web campaign would help

- Learn from bot behavior experts

- Machine Learning bot classification may be superior

- Finance data
  - Do they tell the same story?

- Other insights, suggestions…

Questions?