Beating CodeWars

Data analysis as a coding weapon

Beating CodeWars

Data analysis as a coding weapon

A Kevin Björk Production

Outline

Outline

- What is codewars? How do you measure "winning"?

Outline

- What is codewars? How do you measure "winning"?
- The scraping journey

Outline

- What is codewars? How do you measure "winning"?
- The scraping journey
- Basic information about the DF

Outline

- What is codewars? How do you measure "winning"?
- The scraping journey
- Basic information about the DF
- EDA: Fun Facts
- EDA: Beating CodeWars

Outline

- What is codewars? How do you measure "winning"?
- The scraping journey
- Basic information about the DF
- EDA: Fun Facts
- EDA: Beating CodeWars
- Improvements and moving forward

CodeWars

C++	JavaScript
Java/C#	PHP(Without MySQL)
Ruby	Pascal
Perl	Lisp
Visual Basic	Haskell
Python	C
Assembly	Cobra
	Delphi

CodeWars

- A website where you solve community made coding problems (katas) and gain "levels" by doing so
- The problems vary in difficulty from 8 kyu (easy) to 1 kyu (hard)

CodeWars

- A website where you solve community made coding problems (katas) and gain "levels" by doing so
- The problems vary in difficulty from 8 kyu (easy) to 1 kyu (hard)
- The more katas you solve, the faster you do it and the harder they are, the more you level up
- The more levels you gain, the more you WIN

C++	JavaScript
Java/C#	PHP(Without MySQL)
Ruby	Pascal
Perl	Lisp
Visual Basic	Haskell
Python	C
Assembly	Cobra
	Delphi

Scraping katas

Scraping katas

- Problem: codewars isn't built with pages but instead infinite scrolling
- Solution: use selenium to scroll down to the bottom before scraping
- Learning: this takes a long time

Scraping katas

- Problem: codewars isn't built with pages but instead infinite scrolling
- Solution: use selenium to scroll down to the bottom before scraping
- Learning: this takes a long time
- Problem: every kata has a unique URL that needs to be accessed
- Solution: loop through every URL when scraping
- Learning: this takes a very long time

Scraping katas

- Problem: codewars isn't built with pages but instead infinite scrolling
- Solution: use selenium to scroll down to the bottom before scraping
- Learning: this takes a long time
- Problem: every kata has a unique URL that needs to be accessed
- Solution: loop through every URL when scraping
- Learning: this takes a very long time
- Problem: weird quirks in the webpages result in weird quirks when scraping
- Solution: ???
- Learning: if you don't want your webpage to be scraped, build it in an unintuitive way

Basic information about the DataFrame

Basic information about the DataFrame

- Only use data from python compatible katas not in Beta up to the 2021-04-30

Basic information about the DataFrame

- Only use data from python compatible katas not in Beta up to the 2021-04-30
- First katas were created in 2013

Basic information about the DataFrame

- Only use data from python compatible katas not in Beta up to the 2021-04-30
- First katas were created in 2013
- 4061 rows, 21 columns

5 kyu 307

Beta

No Return #1: Multiply

3 1 67% of 120 305 danilkamyshov

Status: Testing & feedback needed

Estimated Rank: 7 kyu

7 kyu **Kaprekar Split**

4 1 86% of 39 112 nimaij ⚠ 1 Issue Reported

Stats:

Created	Oct 12, 2020	% of votes with a positive feedback rating	13% of 4
Published	Oct 12, 2020	Total "Very Satisfied" Votes	0
Warriors Trained	30	Total "Somewhat Satisfied" Votes	1
Total Skips	0	Total "Not Satisfied" Votes	3
Total Code Submissions	26	Total Rank Assessments	3
Total Times Completed	8	Average Assessed Rank	7 kyu
Python Completions	8	Highest Assessed Rank	7 kyu
Total Stars	1	Lowest Assessed Rank	8 kyu

	kata name	developer	kata rank	publish date	total stars	no. collections kata part of	times attempted	times skipped	total code submissions	total times completed	...	percentage of votes with positive feedback [%]	total votes	total very satisfied	total somewhat satisfied	total not satisfied
0	(Ready for) Prime Time	boatmeme	5 kyu	Aug 29, 2013	48	10	5428	1166	20600	3338	...	88	198	156	35	7
1	Impress your friends with Brown's Criterion!	Raffaele Fiorillo	6 kyu	Mar 8, 2021	1	0	90	8	133	51	...	87	27	21	5	1
2	Convert Lambda To Def	The-Kingfisher	6 kyu	Mar 26, 2021	3	3	203	2	1216	119	...	83	45	34	7	4
3	Fibo akin	g964	5 kyu	Jun 28, 2016	81	14	6640	3899	2498	778	...	88	193	156	28	9
4	Map over a list of lists	Paul Robertson	7 kyu	Apr 5, 2021	17	3	1520	16	1568	666	...	91	132	112	16	4
5	Getting along with Bernoulli's numbers	g964	5 kyu	Nov 8, 2017	40	9	1356	308	742	85	...	86	25	20	3	2
6	Balanced parentheses string	ecolban	4 kyu	Apr 16, 2021	5	2	75	2	91	6	...	100	4	4	0	0
7	Merged String Checker	Abbe	5 kyu	May 25, 2015	956	149	26962	7358	181612	5557	...	83	817	599	158	60

EDA: Fun Facts

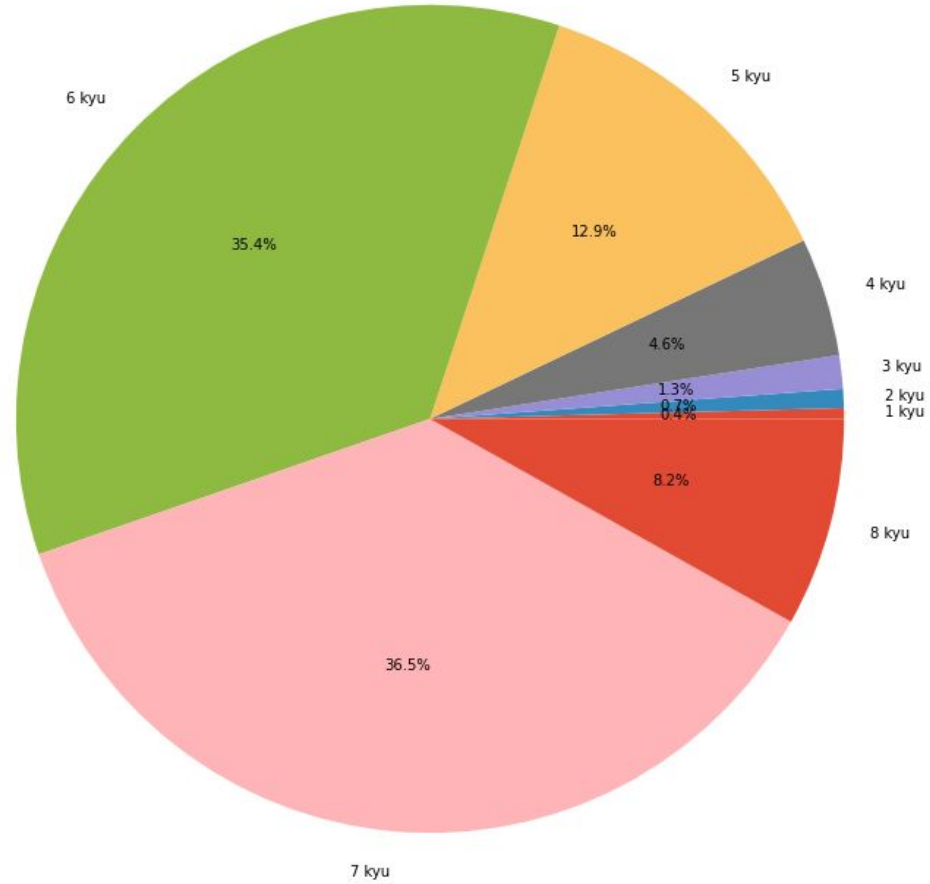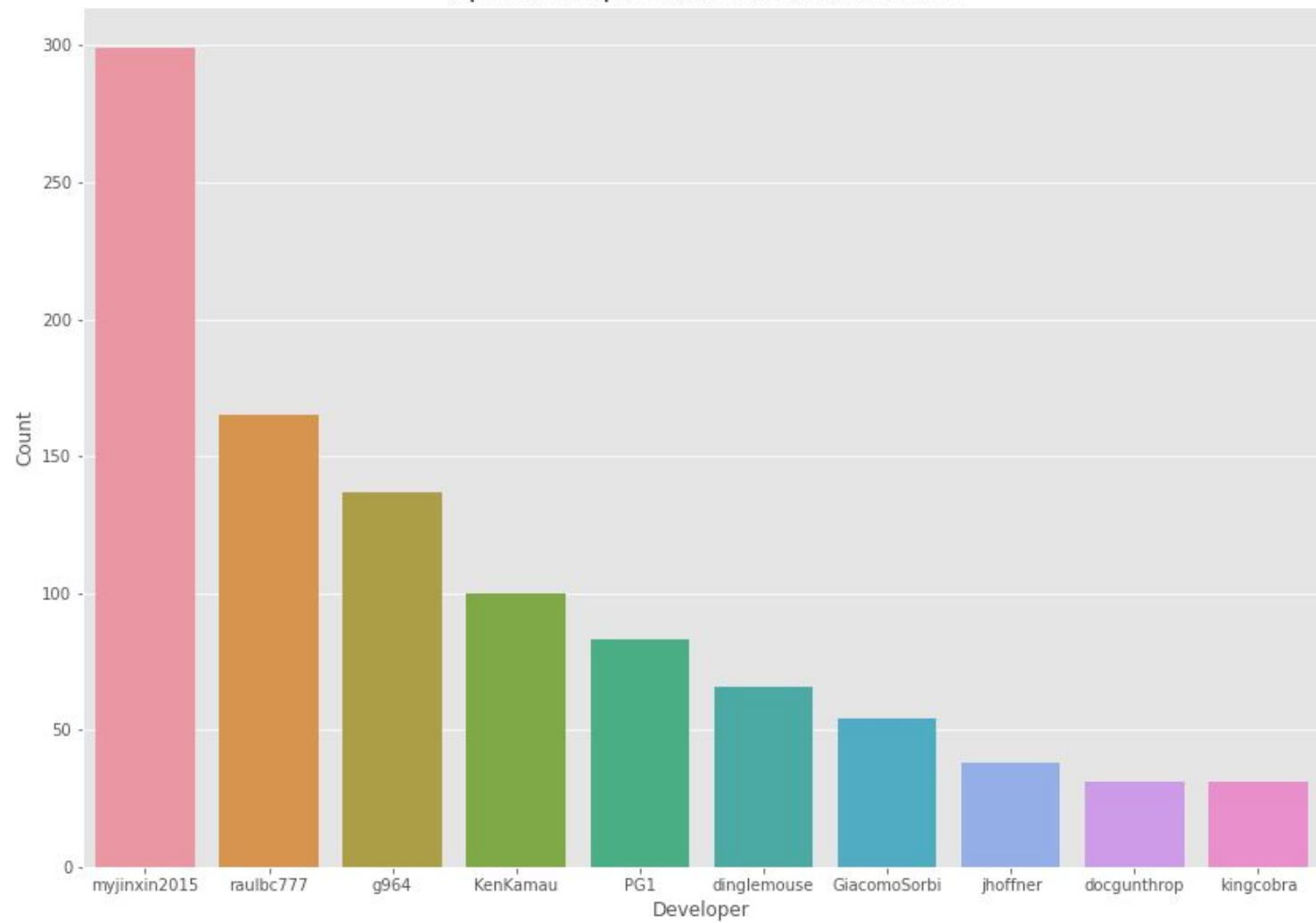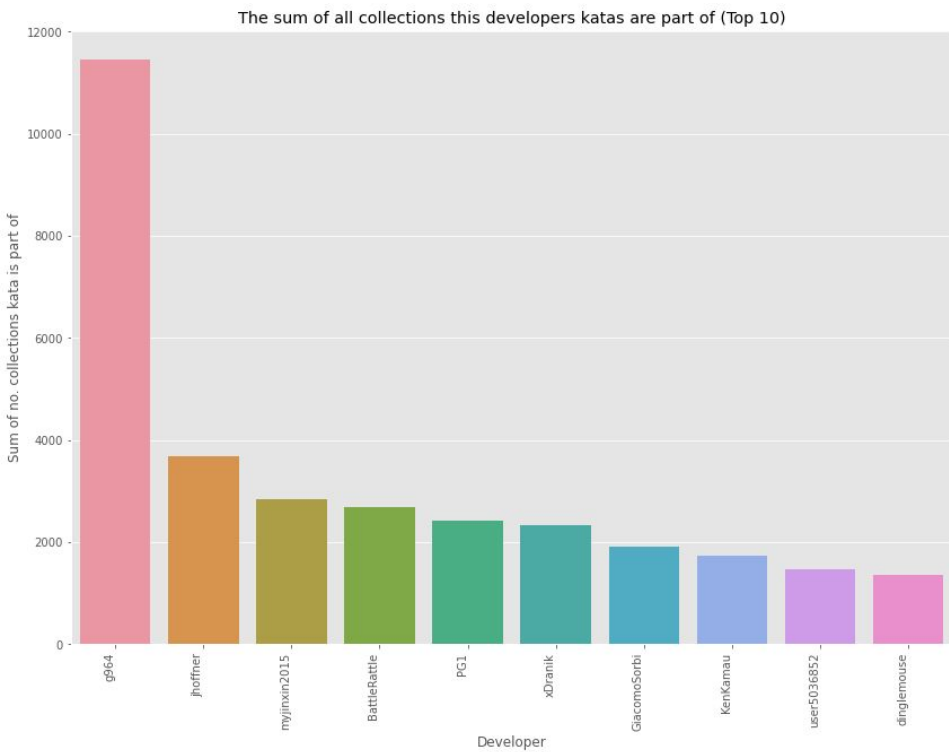|	total stars	no. collections kata part of	times attempted	times skipped	total code submissions	total times completed	percentage of votes with positive feedback [%]	total votes	total very satisfied	total somewhat satisfied	total not satisfied	total rank assessment
total stars	1	0.97	0.32	0.9	0.58	0.22	-0.014	0.7	0.7	0.66	0.51	0.065
no. collections kata part of	0.97	1	0.3	0.85	0.55	0.21	0.0062	0.72	0.73	0.68	0.51	0.08
times attempted	0.32	0.3	1	0.39	0.92	0.99	-0.048	0.46	0.44	0.5	0.59	0.24
times skipped	0.9	0.85	0.39	1	0.63	0.28	-0.061	0.76	0.76	0.76	0.58	0.085
total code submissions	0.58	0.55	0.92	0.63	1	0.88	-0.073	0.68	0.66	0.71	0.73	0.2
total times completed	0.22	0.21	0.99	0.28	0.88	1	-0.043	0.38	0.36	0.42	0.55	0.29
percentage of votes with positive feedback [%]	-0.014	0.0062	-0.048	-0.061	-0.073	-0.043	1	-0.062	-0.043	-0.099	-0.28	-0.025
total votes	0.7	0.72	0.46	0.76	0.68	0.38	-0.062	1	1	0.99	0.76	0.31
total very satisfied	0.7	0.73	0.44	0.76	0.66	0.36	-0.043	1	1	0.98	0.73	0.32
total somewhat satisfied	0.66	0.68	0.5	0.76	0.71	0.42	-0.099	0.99	0.98	1	0.8	0.28
total not satisfied	0.51	0.51	0.59	0.58	0.73	0.55	-0.28	0.76	0.73	0.8	1	0.24
total rank assessment	0.065	0.08	0.24	0.085	0.2	0.29	-0.025	0.31	0.32	0.28	0.24	1

Number of katas published over the years
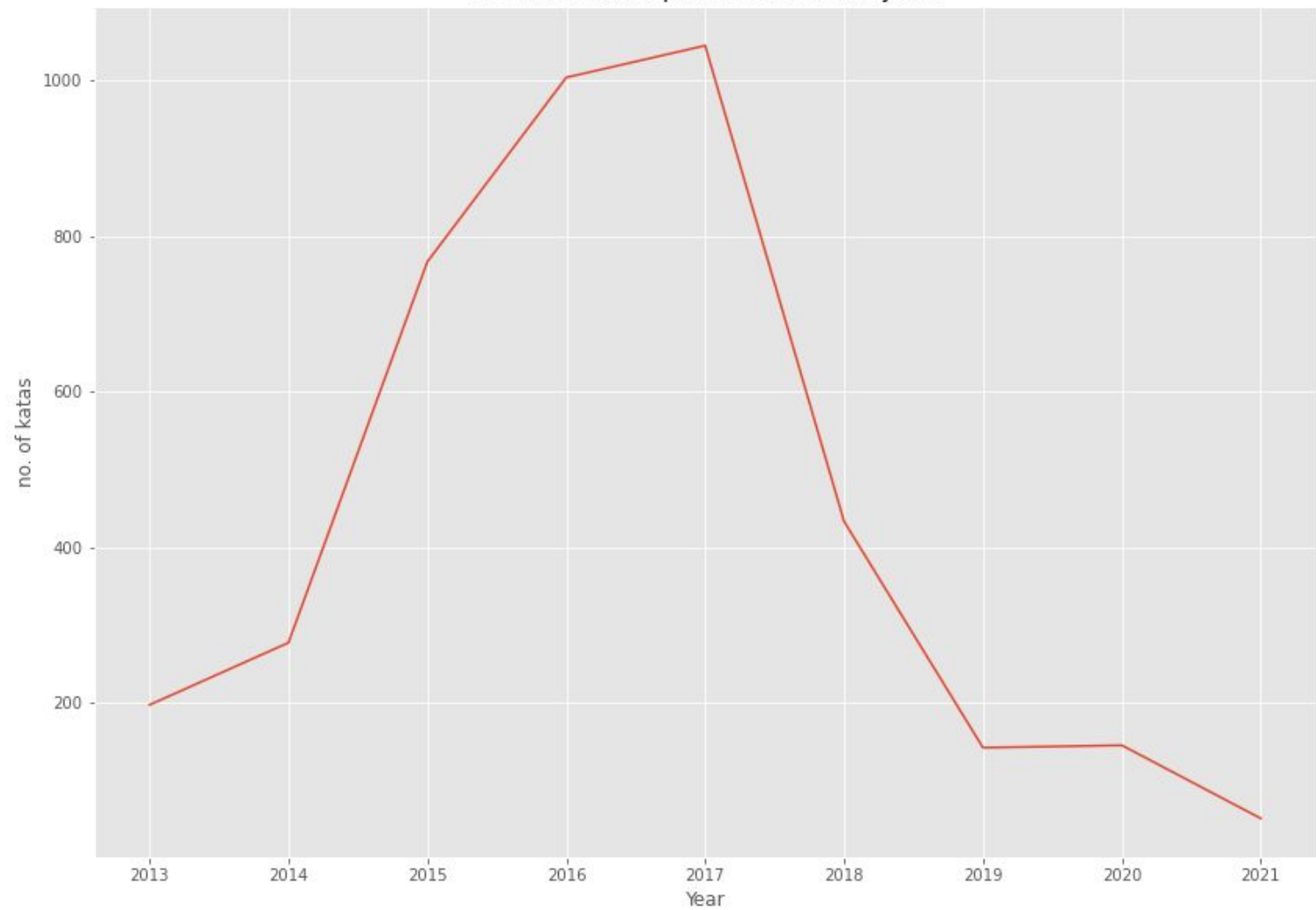
Rank distribution among katas

Top 10 developers who made the most katas

The sum of all collections this developers katas are part of (Top 10)

The sum of all collections this developers katas are part of (Top 10)

The sum of all collections this developers katas are part of (Bottom 5)
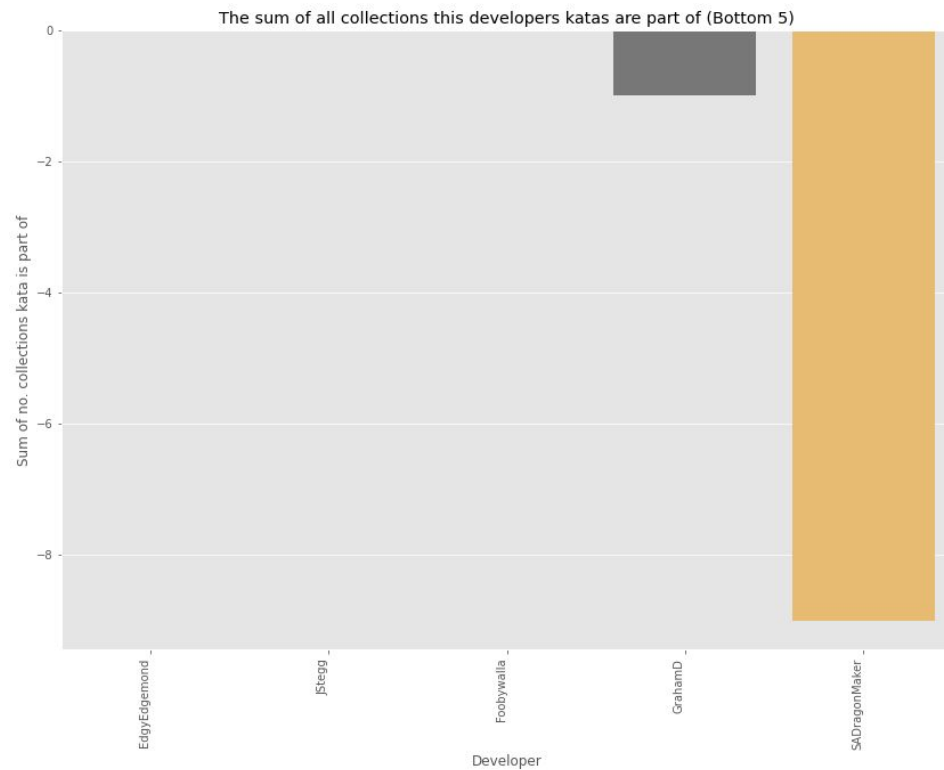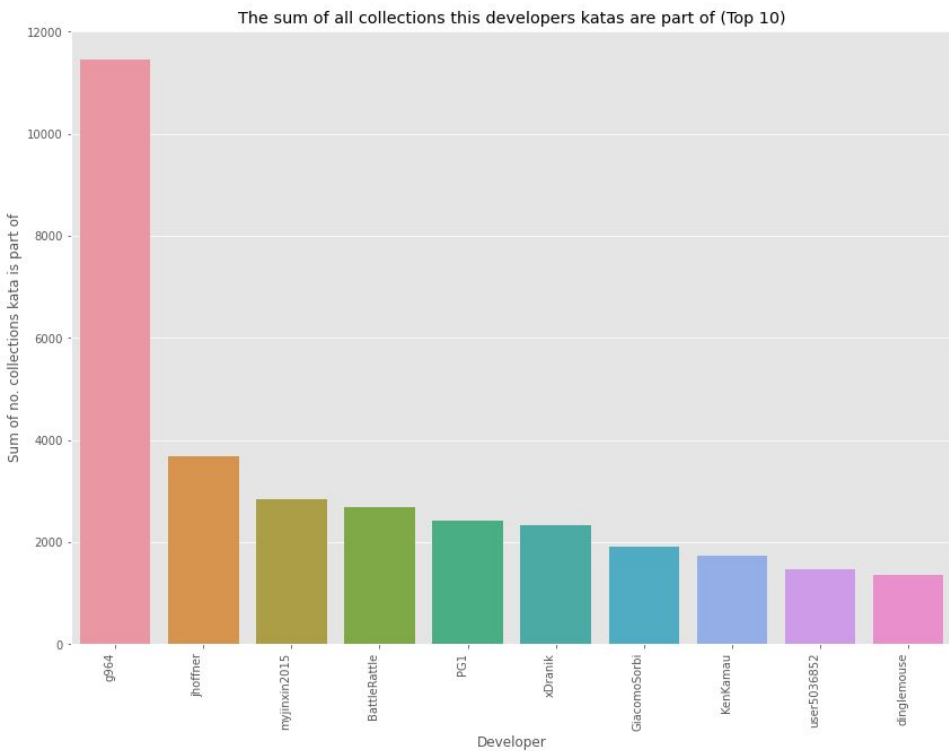
EDA: Max Experience Gain

	kata name	kata rank	developer	completions/attempts	skips/attempts	completions/submissions	average of comp/att and comp/sub
0	(Ready for) Prime Time	5 kyu	boatmeme	0.614959	0.214812	0.162039	0.388499
1	Impress your friends with Brown's Criterion!	6 kyu	Raffaele Fiorillo	0.566667	0.088889	0.383459	0.475063
2	Convert Lambda To Def	6 kyu	The-Kingfisher	0.586207	0.009852	0.097862	0.342034
3	Fibo akin	5 kyu	g964	0.117169	0.587199	0.311449	0.214309
4	Map over a list of lists	7 kyu	Paul Robertson	0.438158	0.010526	0.424745	0.431451
...	...	...	...	...	...	...	...
4056	Javascript filter - 1	7 kyu	ineiti	0.531301	0.191791	0.291026	0.411164
4057	Extract the domain name from a URL	5 kyu	xianpants	0.470869	0.193174	0.122074	0.296472
4058	Return to Sanity	8 kyu	begriffs	0.854822	0.017165	0.363961	0.609391
4059	Basic Training: Add item to an Array	8 kyu	jhoffner	0.833210	0.039148	0.405402	0.619306
4060	Ninja vs Samurai: Attack + Block	5 kyu	jhoffner	0.516905	0.217149	0.171823	0.344364

4061 rows × 7 columns

Top 10 rank 5 katas with the highest average of comp/att and comp/sub
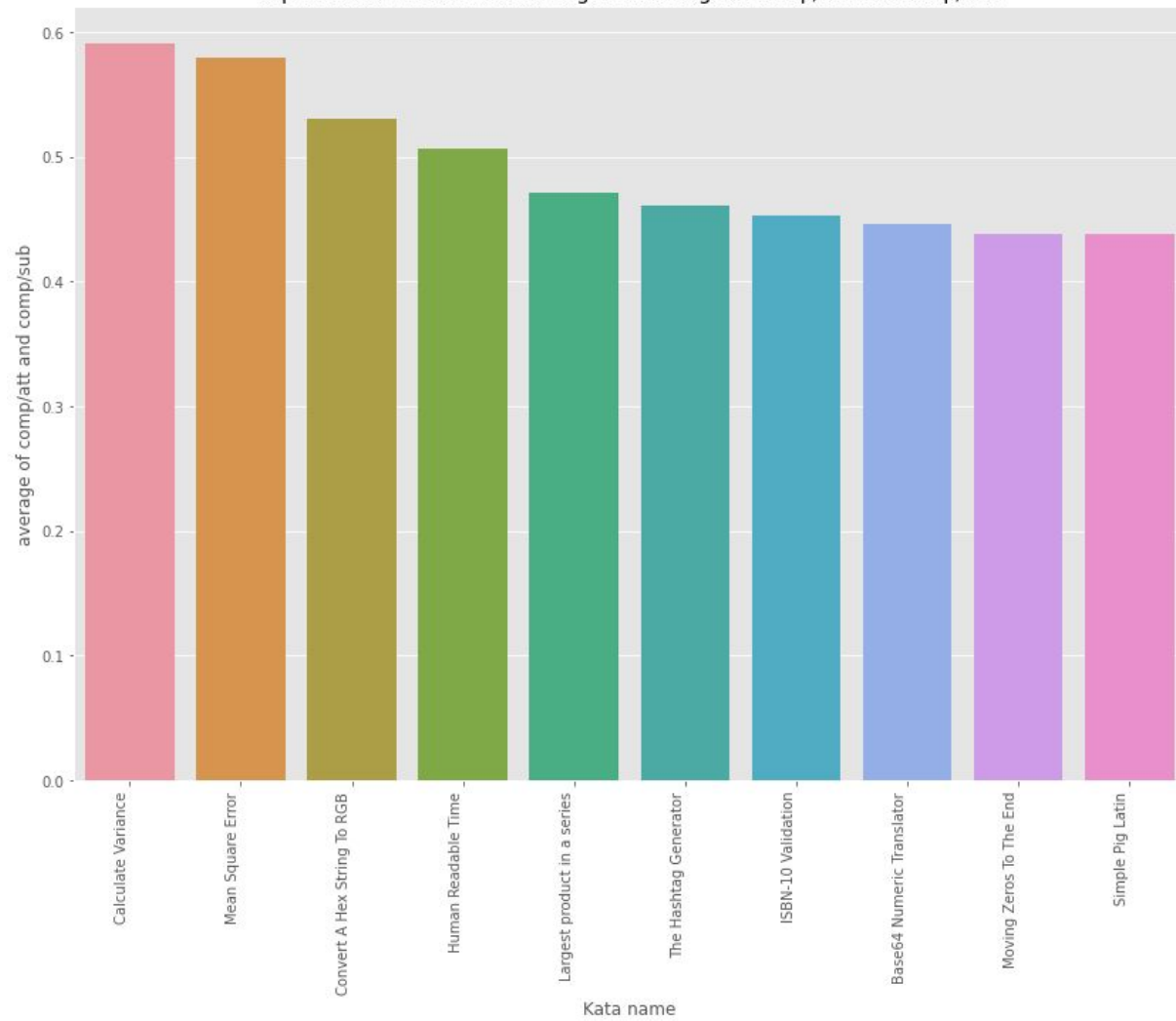
Improvements and moving forward

Improvements and moving forward

- Make sure you only need to scrape the URL's once

Improvements and moving forward

- Make sure you only need to scrape the URL's once
- Use the max experience gain analysis to make an application

The End