TFT Data Science - What is the most overpowered comp of all time (in NA)?

Dec 29, 2021

Introduction

In early December, there was a bracket conducted by Riot Mortdog asking TFT players what, in their opinions, was the most overpowered (OP) team comp of all time. Players voted in the bracket and the results can be found here: https://twitter.com/Mortdog/status/1468361897426632708/photo/1.

There are many factors influencing the poll, such as recency bias, different definitions of OP, etc. Influenced by this, my goal in this study is to perform a **data-driven analysis** using some data science techniques to give a more data driven answer to the question: **what is the most OP comp of all time?**

Methods

The general idea is as follows:

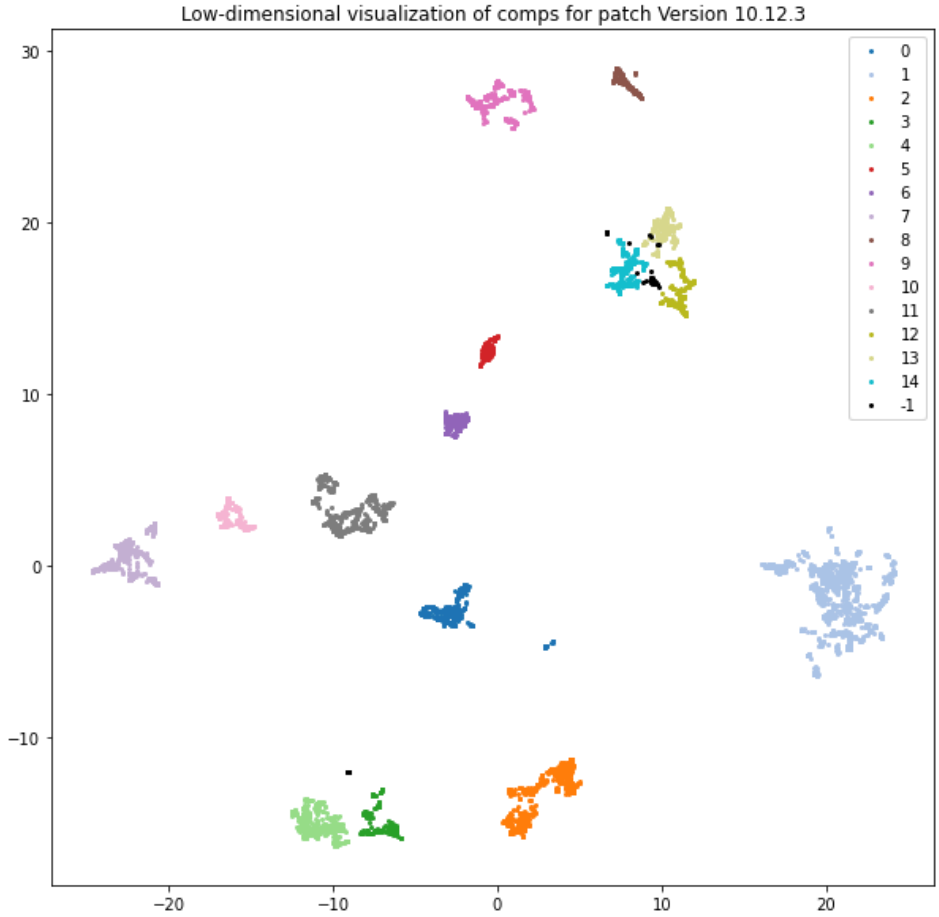
1. Pull about **~1500 games from each patch of TFT for Sets 2-6**. These games were played by players who were in **Masters/GM/Challenger in the NA server** at the end of the season. I did not include Set 1 because of some technical issues.
2. For each patch, NOT INCLUDING b patches (because of technical issues), **find the most played team comps in that specific meta through some data science techniques** (i.e. clustering).
3. For each comp, compute the frequency played, the average placement and analyze the data. I present a metric which I call the **OP-score** which takes into account **both frequency of play and average placement.**

The specific techniques used are very similar to the document I wrote here: https://docs.google.com/document/d/1VK6LSpgaHRR-pNm3XnOMWB4sEKJjXanPJ8kvxo4LzOw/edit?usp=sharing and have posted on Reddit previously.

Specifically, for every player in every game, we can treat their team composition instance as a data point. The goal is to group together these data points (i.e. team comps instances) into clusters. By detecting “clusters” of data points, I can discern popularly played team comps.

For example, if a team comp is 3 socialite kaisa + taum, this specific instance should be “grouped together” with 3 socialite kaisa + fiora, even if the exact champions differ slightly. I should then end up with a cluster of data points representing the 3 socialite kaisa “composition”.

Example of clustering



In the above figure, for patch 10.12, each data point represents a single instance of a player in a match and their team composition. The colours indicate clusters of points i.e. points that should be within the same overarching team comp.

For example, in the **middle-right blue cluster, also labelled as 1,** the aggregate statistics of team comp instances within the cluster are:

Average placement: 4.427675772503359,
Frequency played: 0.2233 (22.33% of team comp instances lie within cluster)
Most played champions:
Irelia 95.61%
Vi 95.52%

Vayne 93.51%
Leona 90.73%
Fiora 87.42%
Ekko 77.3%
Thresh 67.31%
WuKong 50.43%

From this, we can see that this blue cluster represents the Cybernetic comps in set 3.5 because Irelia, Vi, Vayne, Leona, Fiora, Ekko are all played at a high rate within this cluster. Therefore, about 22% of players use a Cybernetic comp in each lobby in this patch, and they place slightly better than average (average is 4.5). I would say this comp is actually quite OP if it has 22% play rate and still performs better than average.

Technical Details - reader may skip if uninterested

Step 1 - Pulling data
I pulled match data using the Riot API for TFT, Riot's own API for pulling game data. Once you have a match id, you can access all of the units, traits, and more for each player in each game. However, there is a technical issue in that they don't have an interface for retrieving old matches, see https://github.com/RiotGames/developer-relations/issues/278. Riot please fix this if you can!! :).

So in order to get games for old sets, I had to scrape lolchess's html for match ids. I took about 20 challenger NA players who have achieved > Masters ratings in each of the 6 sets and obtained all of their matches for each set.

Step 2 - Clustering and team comp classification
Basically, we

1. Treat each composition in a game as a vector of traits, e.g for Set 6, the datapoint [0 5 3 1 2 …] may indicate 0 arcanists, 5 bruisers, 3 chemtechs, 1 twinshot, 2 snipers, etc. Each composition is a datapoint in high-dimensional space (dimension = # of distinct traits). Usually this is about ~20 dimensional.
2. To cluster this high-dimensional data, we perform non-linear dimensionality reduction. I used UMAP in this study; see https://umap-learn.readthedocs.io/en/latest/.
3. To actually cluster the low-dimensional projection, we use HDBSCAN, see https://hdbscan.readthedocs.io/en/latest/.

Why choose this clustering strategy in particular? No particular reason, honestly, other than that it seems to work decently and has been empirically tested in different fields. I found that UMAP seems a bit more reasonable than t-SNE (a sentiment echoed in the biology field, see https://www.nature.com/articles/nbt.4314) and used HDBSCAN because it's known to perform decently and has easy parameter choices. I'm sure someone who has free time can come up with a better method of classification!

Step 3 - Analyzing team compositions.
Given a cluster of data points, representing one team composition, we can obtain the following information

1. Number of team comp instances in the cluster, i.e **frequency of comp play**
2. Average placement (top 1-8) of team comp instances in the cluster, i.e. **average placement of team comp**
3. The most frequently played units in each cluster, i.e. **representative units for each team comp.**

How do we measure how OP a comp is?
To understand how OP a comp is, we need both the frequency of play and the average placement. If a comp has average placement 3 but is played only 3% of the time, is this as OP as a comp which is played 20% of the time and has avg placement 3.2?

I would, in fact, argue the latter comp is much stronger than the former statistically (there is a larger sample size). However, due to TFT champion pool mechanics, the latter comp is even more OP than the statistics show, because as % play goes up, a comp's placement should go down because of champion pool depletion.

The OP-score -- warning: statistics/math below

Basic definition: The OP-score is a measure of how OP the comp is by taking into account both the frequency of play (how often the comp is played) and how good the average placement is. It is a measure of how unlikely it is for a dice that rolls 1-8 with equal probability to have average result < the comps average placement. So if a comp is played 100 times and has average placement 2.5, what is the probability that rolling 1-8 100 times gives an average score of 2.5? How unlikely this is is the OP-score. See the document for full details.

Statistical derivation: As a measure of OP-ness, I propose using the log p-value of the t-test for the mean statistic under a normality assumption. I will call this the **OP-score.**

Assume, under the null hypothesis, a composition has an equal chance of placing 1-8. After sampling (i.e. playing) this composition multiple times, the sample mean, by the central limit theorem and appropriate assumptions, is roughly normal.

Rigorously, let X be the sample mean of the random variable taking on 1,2,...,8 with equal probability, Y be the actual average placement, N be the number of times a comp is played and S be the standard deviation of this random variable. S can be easily calculated by the standard formulas for variance and is ~ 2.291.

Mathematically, $P(X < Y) = P([X - 4.5] * \sqrt{N} / S < [Y - 4.5] * \sqrt{N}/ \text{std}) \sim Pr(N(0,1) < [Y-4.5] * \sqrt{N}/S)$ where N(0,1) is the standard normal distribution.

Calling **Pr(N(0,1)< [Y - 4.5] * sqrt(N) / S) the p-value,** then my **OP-score is -log(p-value)**.

OP-score interpretation
The OP-score has the interpretation as follows: Let's say we have an OP team comp which has N players and average placement Y. Assume that you play the average team comp which attains 1-8 at equal chance. The chance that this average comp has average placement less than or equal to the OP team comp after N plays is about 1/(2.713)^(OP-score). This is about 0.00004629 when OP-score = 10. So the OP-score measures how unlikely it is that the average comp, after N plays, performs as well as the our OP team comp.

Results

Important notes about results

- **B-patches are lumped together with A patches**. So 11.24 and 11.24b are lumped together. This means that some OP-comps that were patched in a B-patch may not be well represented (i.e. **Warweek was nerfed in 10.21b**).
- I only took comps with > 2.5% play frequency.
- Items are not accounted for. I do not differentiate between, say, a 6 rebel comp with protector spat Aurelion Sol vs a 6 rebel comp with no protector spat Aurelion Sol.

Interpretations:

Consistency is king: for a comp to have low average placement and high frequency of play, it must be *consistently OP*. Yordle-lords are thus not an OP comp in this formulation since it's relatively hard to hit Veigar, but if you do, it's strong.

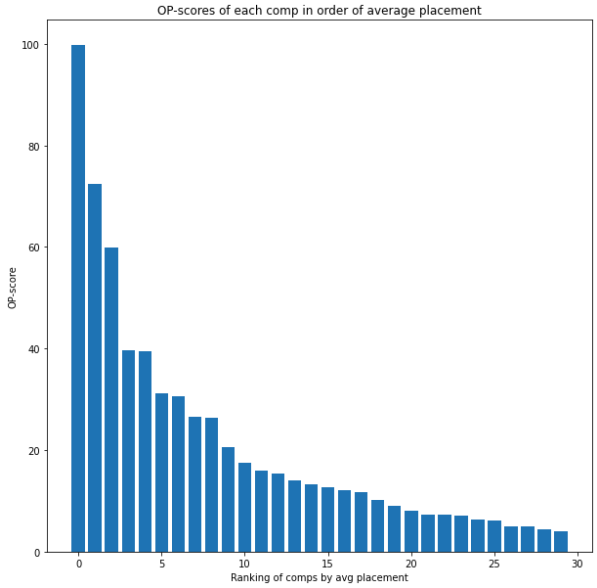
But too frequent is not OP: however, if a comp is played too often, it is impossible to have a low average placement; if we have 8 people in the lobby all playing the comp, the average placement is at best 4.5. The statistics of the OP-score does not take into account this sort of placement mechanic with multiple comps at once and thus **undervalues extremely frequent comps with > 10% play rate. I believe this is the main problem with my study and any further study should correct for this effect.**

OP comps are high-performing versions of the comp: Frequently you'll see two different clusters for the same comp, say, chemtech bruisers. You may see a "high-performing version" with legendaries at a high usage rate, and a "low-performing" version without many legendaries. This is inevitable in the clustering and represents whether a player "hits" with their comp.

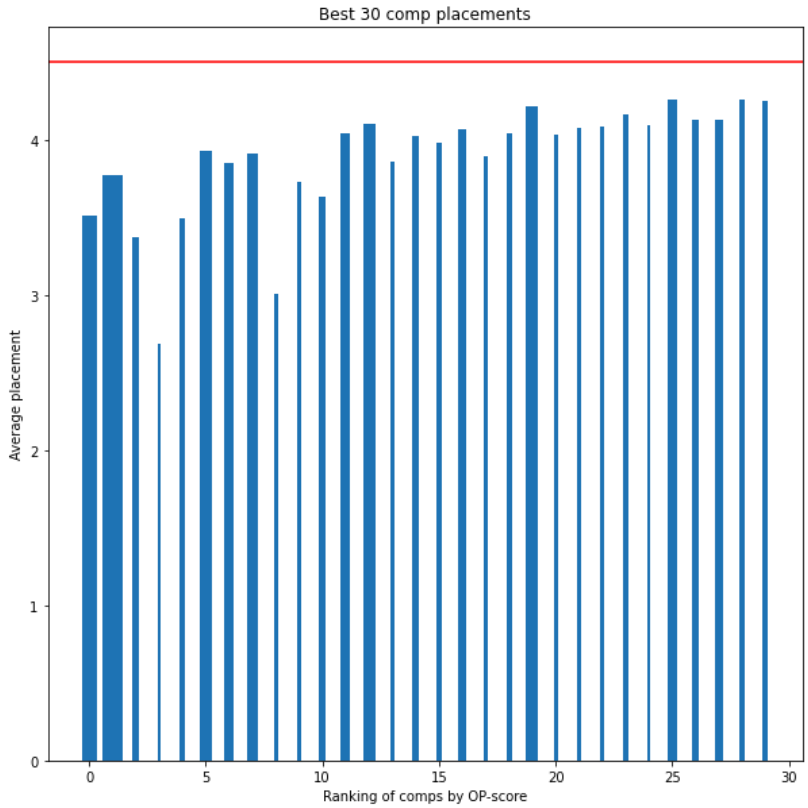
Meta shifts throughout the patch make play frequency lower: Because it may only be apparent halfway through the patch that, say, 3 Socialite Kaisa is OP, games at the start of the patch won't be playing 3 socialite kaisa. Therefore the frequency of play is lower than what you might expect.

Example OP-score and average placement plots:

Below is the plot of OP-scores for compositions in set 2. Each bar represents a single cluster i.e. team comp.



For the same ordering, I can also plot the average placement of each comp. The red line is 4.5 (average placement for an average comp) and the width of the bar indicates the frequency of play for the comp.



Results - Set 2

Most OP Comp - Blender

OP-score	99.75	Sivir 99.61% Yasuo 95.91% Nocturne 94.75% MasterYi 93.29% Khazix 92.9% RekSai 88.81% Janna 72.76% QiyanaWind 26.85% QiyanaInferno 21.11% QiyanaOcean 20.23% QiyanaWoodland 19.84%
Average placement	3.50	
Play frequency	0.1028	
Game version	9.24	

Comments:
The comp with the highest OP score in Set 2 was the infamous blender, with Sivir, Yasuo, Nocturne, Master Yi, Khazix, RekSai, Janna, and Qiyana. While the average placement is higher than some other comps, the frequency of play was a staggering 10%, which, for a comp with average placement << 4.5, is quite staggering. Notice the patch version 9.24, the peak of blender.

2nd place - 6 Shadow 10.x

OP-score	72.35	Sion 99.05% Kindred 98.03% MasterYi 94.01% Malzahar 90.95% Veigar 89.27% Senna 86.57% Janna 45.77% Yasuo 34.6% Karma 32.55% LuxShadow 18.03%
Average placement	3.77	
Play frequency	0.1395	
Game version	10.4	

Comments: In some ways, 6 shadow was even more OP than blender because it was viable for multiple patches. In my analysis, 6 shadow 10.3 and 10.5 was still the 3rd and 4th most OP comps.

Honourable Mentions - **Ocean/Mage 9.23, Light 10.2, and Electric Zed 10.4. See Notebook for more statistics.** Sion 99.05%

Results - Set 3

Most OP Comp - 6 Rebels 10.6

OP-score	201.15	Gangplank 99.78% MissFortune 99.33% Sona 97.66% Jinx 90.77% AurelionSol 89.1% Ziggs 87.65% Lulu 87.1% Yasuo 75.64% MasterYi 73.53% Malphite 44.27% Kayle 11.57%
Average placement	2.98	
Play frequency	0.09	
Game version	10.6	

Comments: This is the most OP comp in Set 3 by far. The average placement was extremely low, but more importantly, the play frequency was **9%.** While this may not seem that high, if you take a look at the clustering, this is extremely high compared to many other high performing clusters. This means that hitting all 4 legendaries, MF, GP, and Asol, Lulu was extremely possible -- this suggests to me that Rebels not only had a strong end game, it had a strong mid-game too.

If you look at the patch notes for 10.7, you can see just how badly rebels got nerfed. Anyone who played this patch (fondly?) remembers a 1-star GP 1-shotting an entire team with just jeweled gauntlet and GA.

2nd place - Demo Mech 10.8

OP-score	76.20	KaiSa 100.0% Annie 100.0% Shaco 100.0% Fizz 99.72% Rumble 99.72% Kayle 99.17% Lux 88.95% KhaZix 70.44% Ekko 19.06%
Average placement	3.05	
Play frequency	0.036	
Game version	10.8	

Comments: Mech is an example of an “if you hit” comp. When you hit 3 star mechs, it was extremely strong, as shown by the low placement, but the frequency was low. This doesn’t mean that people didn’t play mech, but rather that this specific iteration of mech with all 8 champions > 70% frequency in the list was not hit often.

However, my analysis shows that Mech was viable for multiple patch versions of Set 3 and thus was relatively consistent throughout. The fifth most OP comp was Mech in 10.7.

Honourable mentions - Rebels 10.10, Mech 10.7, 3-Star Xayah/Jarvan 10.10, Cyber 10.7, Stargirls 10.8, Void. See Notebook for more statistics.

Results - Set 3.5

Most OP Comp - Mystic Vanguard Cassiopeia 10.12

OP-score	104.18	Cassiopeia 99.4% Karma 99.0% Soraka 98.6% Lulu 98.2% WuKong 97.8% Thresh 96.61% Jayce 92.81% Nautilus 88.02% Mordekaiser 84.63% Urgot 12.57%
Average placement	3.05	
Play frequency	0.050	
Game version	10.12	

Comments: Surprisingly, this is the most OP comp in this set. In 10.13 Cassiopeia was nerfed a decent amount. Notice that this version relied on legendaries Thresh/Lulu at a high rate. Nevertheless, this comp was able to hit legendaries with relatively good consistency.

2nd place - Versatile Mech 10.16

OP-score	60.37	Rumble 99.39% Gangplank 99.39% Fizz 99.08% Annie 98.47% AurelionSol 93.25% Viktor 83.44% Soraka 79.45% Karma 74.54% Ekko 44.17% Shaco 42.02% Lulu 26.99%
Average placement	3.14	
Play frequency	0.033	
Game version	10.16	

Comments: Mech was also strong in set 3.5. This version of mech was more versatile than the previous iteration, which relied on Shaco and Kha'zix.

Honourable mentions - Riven casters 10.13, 6-rebel Jinx 10.12. Nothing else too OP in this set.

Results - Set 4

Most OP Comp - Spirit hunter Aphelios 10.20 "Moonman"

OP-score	71.19	Yuumi 100.0% Aphelios 100.0% Kindred 100.0% Ahri 97.76% Teemo 93.03% Shen 75.12% Lissandra 56.97% Diana 52.74% Azir 39.8%
Average placement	3.17	
Play frequency	0.04	
Game version	10.20	

Comments: This is the dreaded Moonman comp with Aphelios played at 100%. I did not incorporate chosen statistics in this set, but from the playrate, it seems like it was rare that 3 different moonlight champs were taken, and probably only one of Lissandra/Diana taken to complement a Chosen Aphelios as Lis/Diana were picked ~ 50%. This comp was hard nerfed in 10.20b but surprisingly it still shows up when the 10.20a/10.20b versions are clumped together. I bet if 10.20a was it's own game version, the OP-score would be much much higher.

2nd place - Dusk 10.19

OP-score	32.86	Riven 90.88% Cassiopeia 88.82% Azir 85.99% Thresh 64.17% Zilean 57.0% Aatrox 55.48% Jhin 55.48% Kayn 51.68% Sett 41.48% Vayne 40.72% Lillia 37.89% Sejuani 27.25%
Average placement	3.91	
Play frequency	0.0921	
Game version	10.19	

Comments: Straight dusk comp, probably with Riven carry. Riven 3-star was hard nerfed in version 10.20, probably as a response to this comp. Dusk was nerfed too. Note the extremely high play frequency. Probably more OP than the score lets on.

Honourable mentions - Spirit Sharpshooters 10.21, "Legendary Comp" 10.23 with just a bunch of legendaries? 10.24 Ninjas, Vanguard/Mystic/Ahri 11.1.

Results - Set 4.5

Most OP Comp - Slayers 11.4

OP-score	22.55	Pyke 97.36% Tryndamere 94.99% Olaf 94.2% Darius 91.03% Samira 77.57% Aatrox 48.02% Sejuani 47.76% _Zed 46.17% LeeSin 29.29% Braum 28.76% Tristana 27.18%
Average placement	3.76	
Play frequency	0.038	
Game version	11.4	

Comments: This is a slayer comp. I didn't play 4.5 too much so I can't make any more comments. It seems like bugs with olaf and tryndamere were fixed in 11.4b and 11.5 with a Samira rework.

2nd place - Warlords 11.3

OP-score	17.47	JarvanIV 90.93% Nidalee 90.19% Katarina 83.68% Tryndamere 79.14% Vi 67.35% Azir 67.19% Pyke 65.54% Garen 65.29% Samira 57.13% Sejuani 21.43% Sivir 17.97% Sett 17.97%
Average placement	4.14	
Play frequency	0.121	
Game version	11.3	

Comments: A straight up warlord comp. The play frequency is quite high and the average placement is strong. Warlords were pretty hard nerfed in the subsequent patch 11.4.

Honourable mentions - Nothing of note/I am not qualified to say much about these comps.

Results - Set 5

Most OP Comp - Skirmisher Jax 11.10

OP-score	86.52	Rell 99.84% Jax 99.68% Trundle 99.03% Pantheon 98.06% Kennen 84.93% LeeSin 83.14% Viego 82.98% Diana 64.83% Mordekaiser 43.6% Darius 34.04% Nidalee 21.56%
Average placement	3.31	
Play frequency	0.062	
Game version	11.10	

Comments: I didn't play much Set 5/5.5 so I can't comment much. This comp actually is quite OP from the OP-score. It seems that Jax/Pantheon/Skirmisher all got changed in the subsequent patch 11.11.

2nd place - Forgotten 11.12

OP-score	67.85	Rell 98.85% Kindred 98.85% Hecarim 98.15% Ryze 98.15% Thresh 89.61% Draven 87.53% Katarina 71.82% Viego 66.28% Viktor 27.02%
Average placement	3.25	
Play frequency	0.043	
Game version	11.12	

Comments: From what I hear, something about Ryze was broken during this patch and fixed on 11.13. I can't comment on much else besides that.

Honourable mentions - Ryze 11.13, Dragonslayer 11.12? Dawnbringer Karma Carry 11.12. I did not play Set 5 much -- interpretations may be fuzzy.

Results - Set 5.5

Most OP Comp - Revenant/Invoker

OP-score	57.28	Fiddlesticks 99.36% Nunu 99.36% Teemo 99.04% Heimerdinger 98.73% Ivern 98.09% Brand 96.18% Lulu 95.22% Volibear 90.13% Gwen 44.59% Lux 8.28%
Average placement	3.16	
Play frequency	0.0314	
Game version	11.16	

Comments: Again, I didn't play set 5.5 much. In patch 11.17, almost all of the champs in this comp either got a small or a big nerf.

2nd place - Skirmishers? 11.16

OP-score	13.99	Tristana 100.0% Kennen 100.0% Lulu 99.2% Poppy 97.07% MissFortune 89.33% Thresh 81.33% Ziggs 49.07% Teemo 35.47% Kled 25.33% Lux 14.13%
Average placement	3.93	
Play frequency	0.0375	
Game version	11.16	

Comments: Don't know much about this comp. Note the OP-score is not very high compared to any of the other comps listed.

Honourable mentions - Nothing that OP/don't know enough.

Results - Set 6 (so far, patch 11.24b)

Most OP Comp - 7 Innovators 11.23

OP-score	43.05	Seraphine 98.98% Zilean 97.95% Heimerdinger 97.95% Jayce 95.9% Ezreal 88.74% Singed 87.03% Janna 59.39% Orianna 40.96% Yuumi 35.15% Taric 33.45% Jhin 28.67%
Average placement	3.30	
Play frequency	0.0293	
Game version	11.23	

Comments: 7 innovators were pretty strong in patch 11.23 because of Innovator augments, and the dragon was nerfed in 11.24. If you could hit dragon, it was very strong.

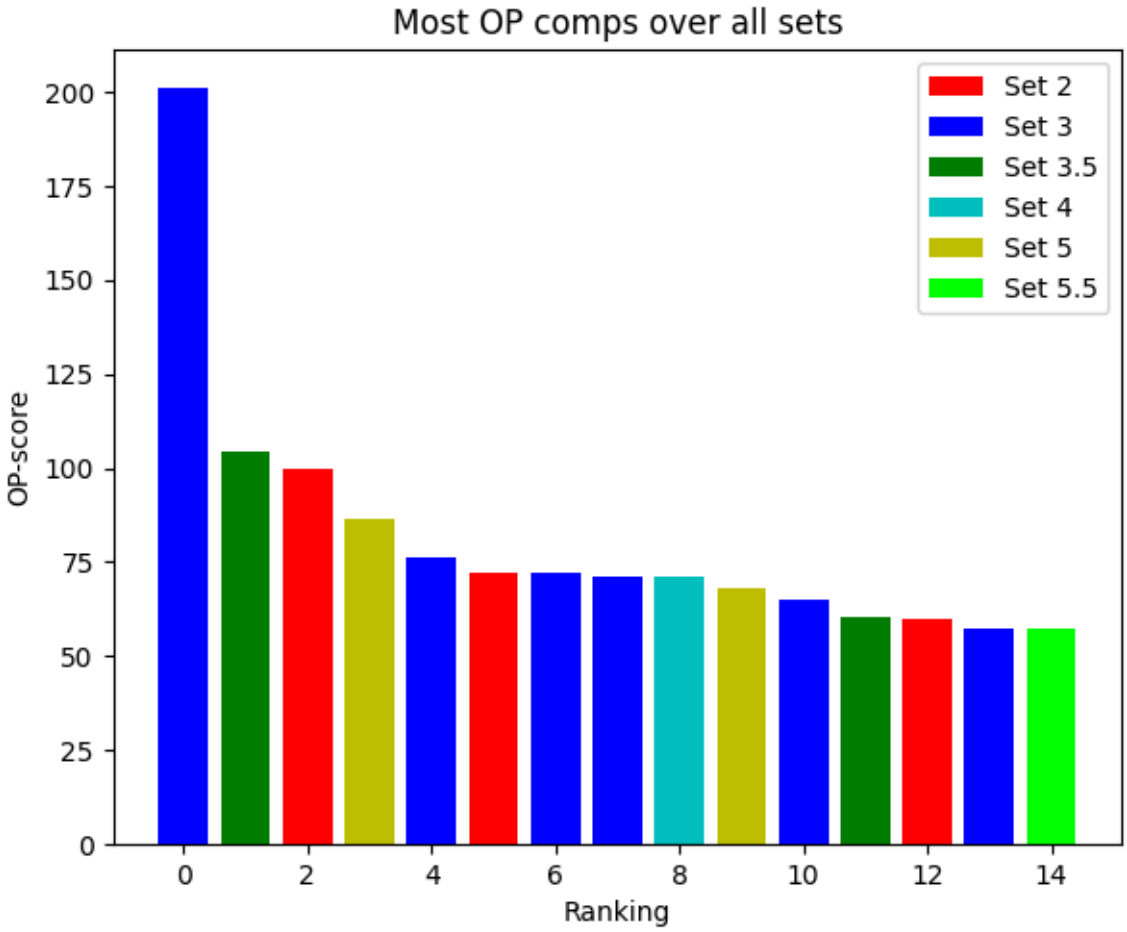
2nd place - Katarina Assassins 11.23

OP-score	23.05	Ekko 99.39% Katarina 98.77% Talon 97.34% Shaco 94.07% Leona 91.82% Blitzcrank 88.96% Taric 43.15% Akali 42.54% Twitch 36.81% Braum 14.72% Seraphine 10.43%
Average placement	3.84	
Play frequency	0.049	
Game version	11.23	

Comments: Katarina sins was a pretty strong comp in 11.23. Kat 3 was haard nerfed in 11.24 as a result.

Honourable mentions - Academy 11.23, Lux 11.23 are the next OP comps but their OP-scores are still relatively low < 10.

So what's the most OP comp of all time?



The most OP comps are (note the graph starts from 0 but the list below starts from 1):

1. 6 Rebels + Legendaries - 10.6 -- **SET 3**
2. Mystic Vanguard Cass - 10.12 -- **SET 3.5**
3. Nocturne Blender - 9.24 -- **SET 2**
4. Skirmisher Jax - 11.10 -- **SET 5**
5. Shaco Mech - 10.8 -- **SET 3**
6. 6 Shadow - 10.4 -- **SET 2**
7. 6 Rebels + Legendaries - 10.10 -- **SET 3**
8. Xayah/Jarvan 3-star Celestials - 10.10 -- **SET 3**
9. Moonman Aphelios w/ Spirits - 10.20 -- **SET 4**
10. Forgotten (Shadow Blue Ryze??) - 11.12 -- **SET 5**
11. Shaco Mech - 10.7 -- **SET 3**
12. Versatile Mech - 10.16 -- **SET 3.5**
13. 6 Shadow - 10.3 -- **SET 2**
14. 6 Cybernetic - 10.7 -- **SET 3**

15. Revenant/Invoker - 11.16 -- **SET 5.5**

Conclusion: 6 Rebel 10.6 was by far the most busted comp of all time according to the OP-score. It is the Wayne Gretzky of busted comps. **Gangplank 1's ultimate in patch 10.6 did more damage than Gangplank 2 in patch 10.7.** Apparently there was also a bug where Rebel's shields scaled with AP.

Final thoughts: I think the results are pretty neat. However, I am not satisfied with the OP-score's statistical foundations yet because 1. it does not take into account champion pool depletion and 2. the phenomenon where two copies of the same comp can't both get 1st in the same game. Therefore, comps with high frequency have lower OP-score than they should have. I truly believe that Blender >> Mystic/Vanguard Cass in terms of OP-ness and that Shadow is probably the 3rd most OP comp because these comps have play rates > 10%.

FAQ:

Where's warweek?

Warwick was patched in 10.21b. Unfortunately, A patches and B patches were collapsed together. Secondly, Warwick was played a lot in patch 10.21, and very frequent comps are undervalued. The cluster which seemed to correspond to Warwick is the following:

Partition: 8, Version: 10.21
mean placement: 4.381443298969073,
OP-score: 5.4
percentage played: 0.2522
Warwick 90.21%
Irelia 78.91%
Shen 75.1%
Lux 69.15%
Jax 55.31%
LeeSin 52.06%
Yone 48.49%
Ashe 46.0%
Morgana 36.99%

Notice how at a staggering 25% play rate, the mean placement was 4.38 which was a decent amount lower than 4.5 given the massive play rate. The OP-score undervalues frequently played comps. I still think this comp is very impressive.

Where's set 1?

I had problems with scraping matches for set 1.

Where's xxx comp?

Look at my notebooks if you want to find out how OP a specific comp was.

Where can I find code to reproduce your work/look at other comps?

I did my work in a set of jupyter notebooks. See
Set 2 Notebook
Set 3 Notebook
Set 3.5 Notebook
Set 4 Notebook
Set 4.5 Notebook
Set 5 Notebook
Set 5.5 Notebook
Set 6 Notebook

Simply go to one of the links, and scroll through the boxes where I output results. Under the second code block, all comps over all patches are shown for that set. The formatting should be somewhat obvious (hopefully).

How can you be confident in your clusterings?

I'm not! :) OP-ness of comps shift if you mess around with the algorithms a bit. Nothing is true for sure when it comes to clustering, but the big picture shouldn't change that much.