CAPSTONE PROJECT

LEAGUE OF LEGENDS

PREPARED BY VIC VOSKOVSKY

version: 1.0
github repository: here
6.8.20

PROJECT OVERVIEW:

- Introduction & Problem Statement
- Data Collection Process
- EDA
- Modeling Results
- Key Observations
- Challenges & Limitations
- Conclusion & Recommendation

INTRODUCTION

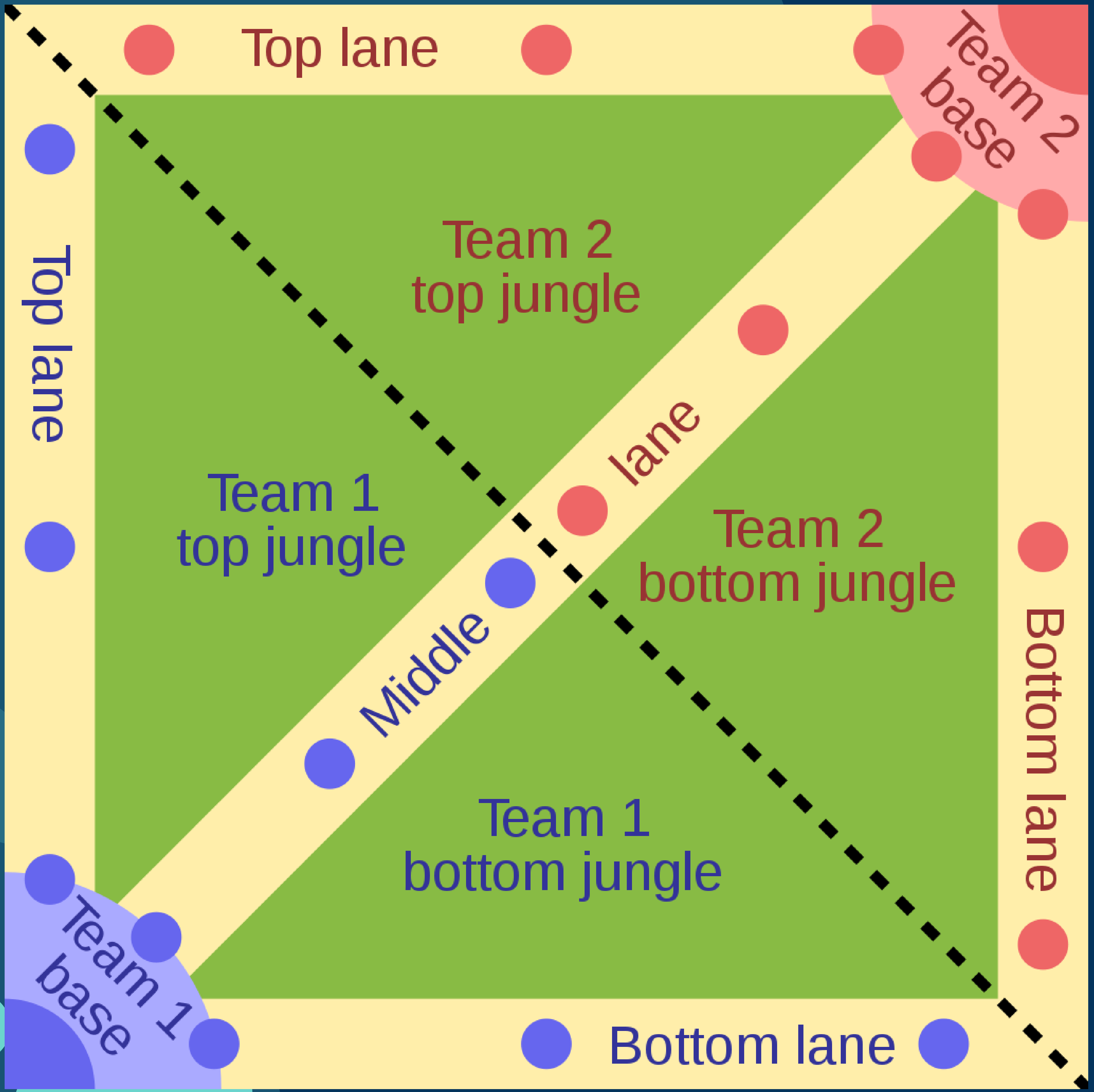
"SETTLING FOR THE DATA WE HAVE."

LEAGUE OF LEGENDS

- Released in 2009, League of Legends has amassed a following of approx. 100 million players world-wide.
- Developer Riot Games has been pioneering a start-up like environment since its creation.
- Over 11 years, Riot has released 137 unique champions, averaging 12 champions per year.
- Coinciding with its 10th anniversary, Riot has brought together 100 million viewers online for its 2019 Worlds Tournament.

In 2020, 86 teams around the world will face off for a chance to take the LoL Worlds Trophy home, but first they have to perform in two other preliminary tournaments, the Spring Split and Summer Split.

THE MAP



- Two teams face off against each other.
- First team to destroy the enemy base wins.
- Teams split between 3 lanes & the jungle.

Each team consists of 5 players:

- Bottom Lane = ADC(Bot P) + SUPPORT(P)
- Middle Lane = Middle(P)
- Top Lane = Top(P)
- The Jungler = Jungler(P)

- Bottom Lane consists of the Attack Damage Carry player & her support player.
  - ADC is usually a pure ranged damage dealing champion.
  - Support is usually a champion that provides the ADC with sustainability.
- Middle Lane is usually a magic dealing champion.
- Top Lane is usually a champion that can take more damage.
- The Jungler is usually a self sustaining champion that splits his time between all lanes.

CHANGES IN 2020

In early 2020, the game has seen some changes. According to my understanding, the battle to make the game more balanced continues. Riot expands a map wide buff according to the first dragon element killed, they add new elemental dragon specific buffs, elder dragon buffs & quality of life map improvements.

Elemental Rifts:
Every game, Summoner's Rift presents a variety of distinct and memorable environments for players to master.

Elemental Buffs & Dragon Soul:
The elemental buffs are more equal in both satisfaction and strength.

Elder Dragon:
Elder buff is more satisfying; Elder buff doesn't favor the team with more dragons.

Side Lane Alcoves & Brush:
The side lanes offer greater outplay opportunities by allowing players to turn fights through clever environmental play.

THE RULES STAY THE SAME

The first split of the season is the Spring Split that took place on January, 2020. After countless bug fixes, character and map balancing endeavors the overall core goals for winning have stayed relatively the same throughout the years.

Gold Spend:

The more gold you earn from killing minions, monsters and opponents the more you can buy. As you buy more items your base statistics exponentially increase, and often some items give unique buffs*. To gain gold you simply last hit an enemy minion, slay neutral monsters in the jungle, kill enemy champions, take down an enemy tower or inhibitor!

Note: Gold can only be gained from last-hitting champions or NPC's* & you are also given gold in increments of around 1g per second.

* **Buffs:** "Term used in some video games to describe increases in the power of a game element." source

* **NPC:** Non-Player Character.

Environmental bosses:

Slay environmental monsters on your own or with your team to gain valuable, limited-time buffs, unique abilities and base-stat* increases.

* **Base-Stat:** Term used to refer to the numerical attributes unique to a Champion. Take player numerical attributes at a given level (1-18) & subtract all points gained from items and limited-time buffs.

Level-up!:

As always, there is only one way to get better in a game, level up! Starting at level 1 you will receive experience points for contributing in fights and slaying NPCs.

PROBLEM STATEMENT

Assume you are a Data Scientist working for a Professional League of Legends team.
You are tasked with analyzing, generating inferences and ranking players based on a proprietary algorithm that the team can use to train for the 2020 Worlds Tournament.
You are provided with access to the Riot Games API and you have reading permissions for all matches completed in the 2020 Spring Split, including your own team.
You are interested in using this data to generate a player ranking score, uncover hidden clues on what the best team combination might be & in general be able to offer your recommendation on who will go to win the Summer Split and later the Worlds Tournament.

Lets see if we can save our players some valuable grinding time by looking at the performance of the collective & generating some inferences for management in the process.

DATA COLLECTION PROCESS

SETTLING FOR THE DATA WE HAVE

The data that the Riot API provides is vast and simple to use after you set up an authentication account. Riot is actually one of the only developers that allows for this kind of functionality, making this process a breeze.

But I think we can do one better by using a dataset generously provided by Tim Sevenhuysen from Oracleselixir.com, complete will all the data we might need for analysis including a few extra columns created by Tim.

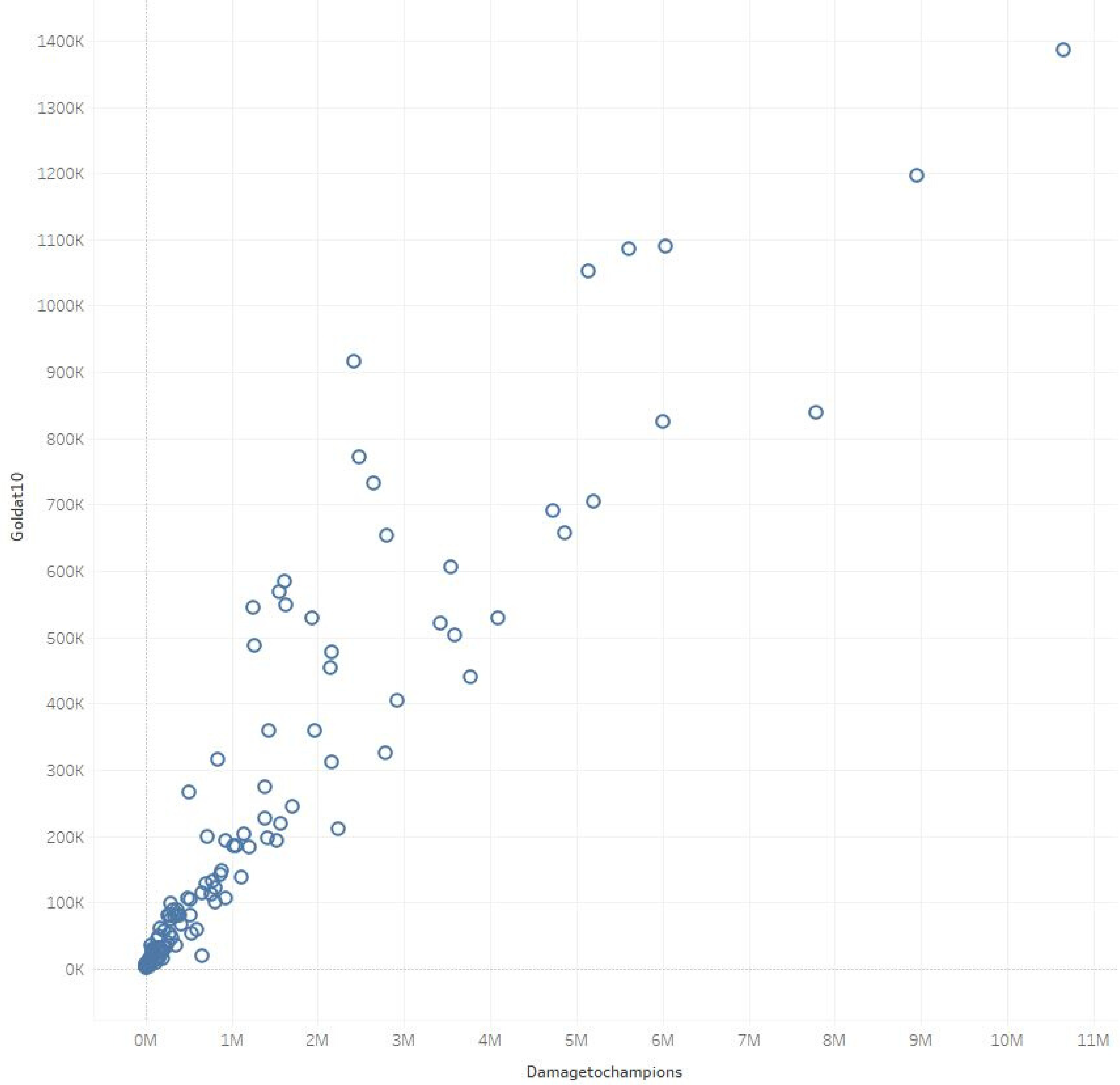
Credit to Tim for doing most of the heavy lifting with this dataset.

Please check him out below!

EDA & VISUALIZATIONS

- My EDA process began by analyzing the 2020 spring data.
- The dataset contained 99 columns and 15,312 rows.
- I sorted my columns into 5 main groups and noted:
  - _l for Record Labels like Game ID, URL, names.
  - _c for Categories like champion name, map position.
  - _t for Time Stamp values.
  - _n for Numerical columns.
  - _b for binary columns like win/loss.

- I started by testing some assumptions:
  a. I began by examining each field individually to confirm some initial assumptions I had about the data. For example, an Attack Damage Carry would output more damage then a support player.
  b. In the next slide you can see that the more gold you earn the likelier you are to output more damage, like the graph below which only shows gold at 10 minutes, we can still see a high correlation between damage output at the end of the match and gold at 10 minutes.
  c. A support player will have a higher vision score, etc.

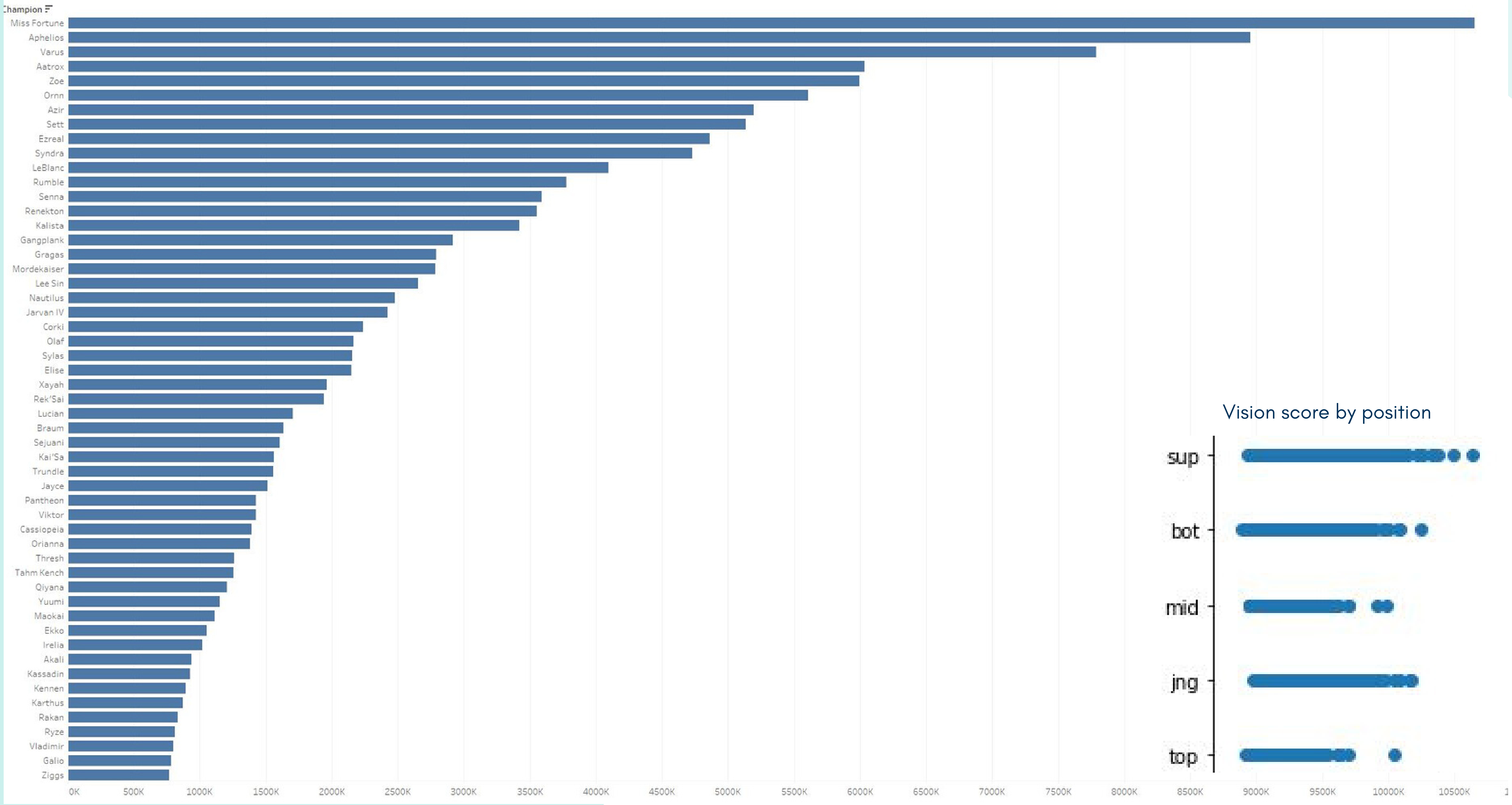


2020 Spring Split Gold@10 vs. Total Damage Output

* **PVP**: Player vs. Player.

2020 Spring Split Champion list sorted by PVP damage ascending
&
Vision score by position

* **PVP**: Player vs. Player.



MODELING & FEATURE ENGINEERING

- After cleaning this dataset, I began by first running a Logistic Regression on my numerical data to make sure it does a decent job predicting the result.

- The initial model produced a promising score but was slightly under-fit:
  - Train - 98.3%.
  - Test - 97.6%.
  - Cross Val Score - 97.7%
- At this point I was comfortable knowing that the numerical values in this dataset do a good job at predicting the result of a game.

- Some of my rows also had missing categorical data, and a few players were registered as unknowns, possibly due to the Riot API not having this data, but otherwise the numerical information was intact. I filled categorical values with "missing" and numerical's with 0's where appropriate. Missing player names can always be filled in later.

I engineered several columns to create my own ranking score per player, those consisted of:

- **player_dmg_score_s** => Ranked based on damage to opposing champions.
- **champ_success_score**_s => Ranked from least to most 'winning' champion regardless of team.
- **player_deaths_score_s** => Ranked based on player deaths as compared to average.
  - Its important that every player focuses on not dying.
- **combo_kill_score_s** => Combination score consisting of double, triple, quadra and penta kills.
  - Applying more weight to players with out standing team-play capabilities, like participating in multiple pentakills.
  - Otherwise, 1 point of participation given even if the combo was not theirs.
- **position_weight_s** => -5 for bot(ADC), -4 for mid, -3 for top, -2 for jng, +1 for sup.
- **assists_weight_s** => Ranked based on assists, hoping to provide counter weight for support players.
- **player_vision_score_s** => Ranked based on players warding performance.
- **player_gold_score_s** => Ranked based on players gold earnings per minute.
- **blue_side_advantage** => .4% subtracted from blue side team.

KEY OBSERVATIONS

- My model actually did surprisingly well, considering it was the first split of the year.
- I think I would be comfortable using this dataset to predict the Summer Split and it would be great to combine Spring and Summer to predict Worlds in the future.
- My ranking sometimes tells a different story depending on the metrics that I generated, which tells me there is a ton more analysis I can do here to improve my rankings.
- I can use the final score per play as weights for the Summer Split which should be happening this month.
- I need to get a more granular score for the top 3 players/teams which would mean incorporating summer data or creating more features.

CHALLENGES & LIMITATIONS

- The biggest challenge when it comes to game data is the granularity. Its amazing to be able to even have this data, but its exciting to think of all the other metrics we could use that we're not currently measuring like:
  - Ability accuracy
  - Player itemization score
  - Reaction time
  - How many "clutch" plays each player made & so on.

CONCLUSIONS

- **Top 10 players:** Average Final Score
  - Teddy = 25.5*
  - Zven = 25.3*
  - Fastlegged = 25.2
  - K1ng = 25.2
  - Vulcan = 25.1
  - Feiz = 25.1
  - Upset = 25.1
  - Licorice = 25.1
  - Ruler = 25.0
  - HiRit = 24.9

- **Top 10 players:** Max Final Score
  - Deft = 1347.1
  - Chovy = 1333.7
  - Keria = 1319.8
  - Teddy = 1301.6
  - Pyosik = 1232.6
  - Effort = 1228.1
  - Nuguri = 1221.0
  - Faker = 1219.9
  - Doran = 1218.9
  - Aiming = 1209.5

- **Top 10 players:** Top Winners
  - Effort = 36
  - Faker = 36
  - Teddy = 36
  - Pyosik = 35
  - Chovy = 35
  - Deft = 35
  - Keria = 35
  - Canna = 34
  - Doinb = 33

- **Top 10 teams:** Average Final Score
  - Cloud9 = 24.56
  - T1 = 24.14
  - Cloud9 Acad = 24.12*
  - Team Dynamics = 24.06
  - Origen = 24.05
  - Dignitas Acad = 23.98
  - Gen.G = 23.87
  - Talon Esports = 23.79
  - ahq eSports = 23.78
  - Machi Esports = 23.74

- **Top 10 players:** Max Final Score
  - DragonX = 6521.71
  - T1 = 6154.64
  - DAMWON G = 5911.26
  - SANDBOX G = 5856.74
  - Griffin = 5671.62
  - KT Rolster = 5639.55
  - Gen.G = 5608.63
  - Afreeca Freecs = 5240.47
  - Hanwha Life = 5060.78
  - Top Esports = 5058.91

- **Top 10 players:** Top Winners
  - T1 = 216
  - DragonX = 210
  - FunPlus Phoenix = 198
  - JD Gaming = 192
  - Gen.G = 186
  - Top Esports = 186
  - Invictus Gaming = 174
  - Cloud9 = 156
  - G2 Esports = 156
  - DAMWON Gaming = 156