

# Proposal dataset & plan

## League dataset

- The data is about League of Legends games, thus placing this dataset in the 'Gaming' domain.
- The dataset contains 1000 records of real ranked games.
- It's formatted in 10 files of JSON, 100 records in each file.

## Source

The [Developer Portal \(https://developer.riotgames.com/getting-started.html\)](https://developer.riotgames.com/getting-started.html) links to *ten* Amazon-hosted json-files. Direct links to these files:

- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches1.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches1.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches2.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches2.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches3.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches3.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches4.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches4.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches5.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches5.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches6.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches6.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches7.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches7.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches8.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches8.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches9.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches9.json>)
- <https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches10.json> (<https://s3-us-west-1.amazonaws.com/riot-developer-portal/seed-data/matches10.json>)

One sample record:

```
{
  "gameId":2585563902,
  "platformId":"NA1",
  "gameCreation":1504029097863,
  "gameDuration":3509,
  "queueId":420,
  "mapId":11,
  "seasonId":9,
  "gameVersion":"7.17.200.3955",
  "gameMode":"CLASSIC",
  "gameType":"MATCHED_GAME",
  "teams":[
    {
      "teamId":100,
      "win":"Win",
      "firstBlood":true,
      "firstTower":false,
      "firstInhibitor":false,
      "firstBaron":false,
      "firstDragon":false,
      "firstRiftHerald":false,
      "towerKills":9,
      "inhibitorKills":1,
      "baronKills":0,
      "dragonKills":2,
      "vilemawKills":0,
      "riftHeraldKills":0,
      "dominionVictoryScore":0,
      "bans":[ ]
    },
    {
      "teamId":200,
      "win":"Fail",
      "firstBlood":false,
      "firstTower":true,
      "firstInhibitor":true,
      "firstBaron":true,
      "firstDragon":true,
      "firstRiftHerald":false,
      "towerKills":9,
      "inhibitorKills":3,
      "baronKills":2,
      "dragonKills":3,
      "vilemawKills":0,
      "riftHeraldKills":0,
      "dominionVictoryScore":0,
      "bans":[ ]
    }
  ],
  "participants":[ ],
  "participantIdentities":[ ]
}
```

## Research Question

**Does a 'first-something' really have a big impact on whether or not a team eventually wins the game?**

*Clarification: does a team have a much higher chance to win a game, when they claimed first blood, reached for the first tower, or when they killed the first dragon?*

## Plan

I will look for correlations between 'firsts' (firstBlood, firstTower, etc...) and whether or not a team has won that game. Additionally, I could check if a combination of either one of them strongly impacts the win percentage of said team. For example: maybe firstBaron and firstDragon together strongly impact the statistical chance of winning a game up to 80%.