Project 1 - Proposal and Data Assembly

Andrew Kerr

2023-10-15

# Research Question

One of my favorite games to play is Team Fight Tactics (TFT), a real-time strategy game centered around managing gold, augments (game modifiers different for each player), units, and traits (each unit has 1-3 traits), where 8 players fight it out to be the last one standing.

One method of measuring the skill of a player is by looking at their tier. The tiers in TFT go as follows (from lowest to highest):

* Iron
* Bronze
* Silver
* Gold
* Platinum
* Diamond
* Masters
* Grandmasters
* Challenger

For example, a Silver player is ranked lower than a Gold player.

I am interested in whether you can determine a players tier based on the results of a match of TFT. In other words,

* Are there distinct play styles or mannerism associated with different tiers in TFT?

I have collected a random sample of matches from different tiers of ranked TFT. The variables I have decided to include are:

* placement: a measure of how well a player performed in a match compared to other players in the match
  + I may create subgroups for each placement
* level: a measure of how well strong a player was in a match
* last\_round: what round a player is eliminated from a match
* gold\_left: amount of gold a player had when the were eliminated from a match
* augment\_1-3: game modifier selected by a player during the match
* traits: active traits a player had when they were eliminated
* units: units a player was using when they were eliminated
* game\_length: how long (in seconds) a match lasted
* tier: what tier the player currently is

# Data

Over the summer I worked on using the Riot API (company that owns TFT) to collect my match history. I have re-purposed the code to now collect match history for any tier of player. I provide the tier I desire, and my program acquires 205 randomly selected players from the tier and adds their 5 most recently played matches to my database. For each match I collect information on all 8 players and how they performed in the match. Below is an observation from a data set created from a combination of my SQL tables (traits and units will need to be re-structured for this project):

head(df\_TFT, 1)

## tier match\_id name game\_length tft\_game\_type placement level  
## 1 CHALLENGER NA1\_4801065569 Spethom 2108 RANKED 4 9  
## last\_round gold\_left augment\_1  
## 1 34 3 TFT9\_Augment\_CyberneticBulk1  
## augment\_2 augment\_3  
## 1 TFT9\_Augment\_CapriciousForge <NA>  
## traits  
## 1 [{"name": "Set9\_Armorclad", "style": 1, "num\_units": 2, "tier\_total": 3, "tier\_current": 1}, {"name": "Set9\_Bastion", "style": 0, "num\_units": 1, "tier\_total": 4, "tier\_current": 0}, {"name": "Set9\_Challenger", "style": 0, "num\_units": 1, "tier\_total": 4, "tier\_current": 0}, {"name": "Set9\_Demacia", "style": 0, "num\_units": 1, "tier\_total": 4, "tier\_current": 0}, {"name": "Set9\_Multicaster", "style": 0, "num\_units": 1, "tier\_total": 3, "tier\_current": 0}, {"name": "Set9\_Noxus", "style": 0, "num\_units": 2, "tier\_total": 4, "tier\_current": 0}, {"name": "Set9\_Preserver", "style": 0, "num\_units": 1, "tier\_total": 4, "tier\_current": 0}, {"name": "Set9\_Shurima", "style": 3, "num\_units": 6, "tier\_total": 4, "tier\_current": 3}, {"name": "Set9\_Slayer", "style": 1, "num\_units": 2, "tier\_total": 3, "tier\_current": 1}, {"name": "Set9\_Strategist", "style": 2, "num\_units": 3, "tier\_total": 4, "tier\_current": 2}, {"name": "Set9b\_Darkin", "style": 3, "num\_units": 2, "tier\_total": 2, "tier\_current": 2}]  
## units  
## 1 [{"name": "", "tier": 2, "rarity": 0, "itemNames": [], "character\_id": "TFT9\_Cassiopeia"}, {"name": "", "tier": 2, "rarity": 1, "itemNames": ["TFT9\_Item\_StrategistEmblem"], "character\_id": "TFT9\_Naafiri"}, {"name": "", "tier": 1, "rarity": 1, "itemNames": [], "character\_id": "TFT9\_Taliyah"}, {"name": "", "tier": 2, "rarity": 4, "itemNames": ["TFT\_Item\_GiantsBelt"], "character\_id": "TFT9\_JarvanIV"}, {"name": "", "tier": 2, "rarity": 4, "itemNames": ["TFT\_Item\_JeweledGauntlet", "TFT\_Item\_PowerGauntlet"], "character\_id": "TFT9\_Mordekaiser"}, {"name": "", "tier": 2, "rarity": 4, "itemNames": ["TFT\_Item\_Leviathan", "TFT\_Item\_HextechGunblade", "TFT\_Item\_StatikkShiv"], "character\_id": "TFT9\_Azir"}, {"name": "", "tier": 2, "rarity": 4, "itemNames": ["TFT\_Item\_RedBuff", "TFT\_Item\_DragonsClaw"], "character\_id": "TFT9\_Nasus"}, {"name": "", "tier": 2, "rarity": 6, "itemNames": ["TFT9\_Item\_OrnnPrototypeForge", "TFT4\_Item\_OrnnTheCollector", "TFT9\_Item\_OrnnDeathfireGrasp"], "character\_id": "TFT9b\_Aatrox"}, {"name": "", "tier": 1, "rarity": 6, "itemNames": ["TFT\_Item\_WarmogsArmor"], "character\_id": "TFT9\_KSante"}]

# Multileveled Structure

I plan to talk to you more about this, but here is what I am considering:

* Observational units: individual player outcomes in a match
* Level 1: player (8 players from each match)
* Level 2: match (10k+ matches, I am currently collecting them overnight, there was a bug in my code :( )

# Variable Chart

kable(tbl, format = "simple")

| Name | VariableRole | Type | Values |
| --- | --- | --- | --- |
| Tier | Response | Categorical | Iron, Bronze, Silver, Gold, Platinum, Diamond, Masters, Grandmasters, Challenger |
| placement | Level 1 Predictor | Categorical | 1 to 8 |
| level | Level 1 Predictor | Categorical | 1 to 10 |
| last round | Level 1 Predictor | Categorical | 1 to 50 |
| gold left | Level 1 Predictor | Quantitative | 0 to 200 |
| legend augment count | Level 1 Predictor | Categorical | how many legend augments were chosen => 0 to 3 |
| game length | Level 2 Predicotr | Quantitative | 0 to 3600 (seconds) |
| avg unit level | Level 1 Predictor | Quantitative | each unit can be between level 1 and 3 => 1 to 3 |
| traits | Level 1 Predictor | Categorical | list of active traits (will dummify; total of 28) |
| avg level | Level 2 Predictor | Quantitative | mean level of a match |