

# **Mini Project 2 - Apache Spark**

**CIE 427**

**Mazen Hassan**

**201801897**

**Ibrahim Hamada**

**201800739**

## I. Champion win, pick, and ban rates

We can calculate these requirements as follows:

$$\text{pick rate} = \frac{\text{champion occurrence}}{\text{total number of champions}}$$

$$\text{win rate} = \frac{\text{wins number per champion}}{\text{picks number per champion}}$$

$$\text{ban rate} = \frac{\text{bans number per champion}}{\text{total number of games}}$$

## II. Champion Synergies or duos

Synergy between two champions can be defined as the probability of the champion pair to belong to the same winning team over the probability of the champion pair to occur in any team.

It can be simplified in terms of counts as follows:

$$\text{Synergy} = \frac{\text{count of occurrences of a pair in the same winning team}}{\text{total number of occurrences of the pair in any team}}$$

## III. Item win, pick rates

We can group the items' columns along with the win status and aggregate the count of items and sum of win columns.

We can find the win rate by dividing the total wins per item over the count of that item.

We can find the pick rate by dividing the item count by the total items.

## **IV. Item Synergies**

We can find item-champion synergy by dividing the number of times a champion won with an item by the number of times a champion bought that item.

We can find the item-class synergy in a similar way by dividing the number of times a champion won with an item by the number of times a champion belonged to a class.

## **V. Item suggestion**

This can be defined as items with the highest win rate for a specific champion. It can be found by aggregating the number of items and the number of wins, then dividing number of wins by the total number of choosing that item by the champion.

## **VI. More requirements to get more insights**

- Distribution of champion win rate over different lanes.
- The best lane for each champion.
- Distribution of players activity per time.