Qatar 2022 Football World Cup has arrived, and we can’t be more excited.

In this project I’m going to analyze the statistic behaviour of Qatar 2022 Football World Cup Panini collectible album completion.

The album contains 638 slots, that will be filled with each of all football players in Qatar 2022 World Cup. Once we get all of them, we have completed the album.

How are we going to get a football player? They can be bought in a pack of 6 randomized players. Players will be numbered from 1 to 638 (player ID)

Having said that, we can build our model:

The number of different players collected in a certain moment is a random variable, that depends on the number of packs I have bought. For instance, if I buy 2 packs and I get: player ID: 128, 135, 456, 12, 220, 603, 94, 557, 447, 309, 135, 298; then my random variable takes the value of 11, because I got 11 different Player ID and one is repeated (Player ID 135)

Problem to solve: What’s the probability of album completion having bought a fixed number of packs (X)? Special scenario: what’s the probability of album completion having bought 107 packs?\*

To solve this problem, this project aims to create a code in RStudio that simulates thousands of people buying X number of packs and see how many of them have completed the album.

\*Note: 107 packs is the minimum number of packs needed to complete the album