In this short report, the whole scheme is subject to the loss-chasing pattern in gambling as the irresponsible behaviour of interest.

The data and features required for this task include the following:
(Statistics of each game)

- **Player_ID**: the unique identifier of each player.
- **Game_ID**: the unique identifier of each game.
- **Date**: the date and time at which this game took place.
- **Odds**: the odds of this game, which helps us decide if this player tended to make irrational decisions.
- **Bet**: the amount of money bet by the player in this game.
- **Bonus**: the bonus of this game; 0 if there wasn't any.
- **Outcome**: the outcome of this game – win/loss.
- **Profit**: the profit of this round, calculated as (Bet * Odds + Bonus - Bet) if the outcome was a win; 0 otherwise.
- **Balance**: the amount of money left in the player's account at the time of play.
- **BetRatio**: (Bet/Balance), which could help understand how cautious this player was.

Since we are more interested in the long-term behaviour per player, we could group the data above and derive some information about each player. Potential new features are:

- **GamesWon**: the total number of games this player had won.
- **GamesLost**: the total number of games this player had lost.
- **Avg_Odds**: the average odds value.
- **Wins_in_Last5**: the number of wins in the last 5 games, which could be an essential indication of the loss-chasing pattern.
- **TotalProfit**: the total profit this player has made over time.
- **TotalLoss:** the total loss of this player over time.
- **NetProfit:** (TotalProfit – TotalLoss).
- **Avg_BetRatio**: the average of bet ratio over time.
- **Max_BetRatio/Min_BetRatio**: loss-chasing pattern sometimes accompanies sudden increases in wager amounts. For someone loses a substantial amount in previous rounds, a reckless player is quite likely to bet more for the chance of winning the money back.

Before getting into the algorithm, feature scaling should also be implemented as the last part of pre-processing, since "TotalProfit" and "TotalLoss" are at a much greater scale than other features.

The next step is to implement algorithms to detect irresponsible behaviours. The best choice for this situation would be clustering algorithms such as K-means, by which we get to define the number of clusters we are looking for. The clusters represent different player groups. There are generally 3 types of players:

- Cautious players: risk averse players. Players in this group tend to make conservative decisions and stop betting when they do not see a win. Some characteristics of this group include small "Avg_Odds", "NetProfit" being close

to 0, small "GamesWon" and "GamesLost", small "TotalLoss", stable "Avg_BetRatio" and "Max_BetRatio/Min_BetRatio" being slightly above 1.

- Strategic players: these players are some of the most rational and collected ones. They usually play a lot of games and are well-versed of the tricks, which helps them keep everything under their control. Their "NetProfit" tend to be positive, "Avg_Odds" are relatively low, and other statistics of theirs are more flexible.

- Irrational players: these players tend to make audacious moves in exchange of higher wins. Their game stats may show a negative "NetProfit" and a constantly high "Avg_BetRatio". For our interest group, a classic pattern would be that "Avg_BetRatio" stays high while "Wins_in_Last5" stays low.