The purpose of the project is to analyze the benefits and risks of stealing a base in baseball. By examining relevant data sources, we aim to gain insights into the success rates, contextual factors, and strategic implications of stolen base attempts. The specific data sources that can be considered for this project include:

1. Game Logs: Retrieve data from game logs that provide detailed information about individual games, including play-by-play data, base running events, and stolen base attempts. These logs capture the context in which stolen base attempts occur and can be used to analyze the outcomes.
2. Player Statistics: Access player statistics datasets that include information on stolen base attempts, successful steals, caught stealing, base running speed, and other relevant metrics. These statistics can help evaluate individual player performance and identify trends related to stolen base success rates.
3. Team Performance Data: Analyze team performance data to understand the overall success rates of stolen bases for different teams. This data can provide insights into how teams strategize and execute stolen base attempts collectively.
4. Pitcher-Catcher Data: Explore data specific to pitchers and catchers, such as their effectiveness in controlling the running game. This information can help assess the defensive capabilities of pitchers and catchers in deterring stolen base attempts.
5. Game Situation Data: Consider contextual data, including inning, score differentials, base-out states, and pitch counts. These variables can be used to analyze the impact of different game situations on the success rates and risks associated with stealing bases.
6. Historical Data: Access historical data to examine trends and changes in stolen base success rates over time. This data can provide insights into how strategies and approaches to stolen bases have evolved in the game of baseball.

Recap of required info :

* Play by play data of individual games
* Base running events (context of each stolen base attempt)
* Stolen base attempts (in a game)

The goal of these three pieces of data is to identify if there’s specific situations where attempting to steal is particularly beneficial or risky.

* Player stats (nb of stolen base attempts, success rates, caught stealing, base running speed)

Are there specific players that should attempt more base stealing? Are there ways to identify if a rookie is gonna be a good base stealer?

* Team success rates for stealing bases and success rates against other teams

Is that strategy gonna change depending on which team they play against? Are certain teams better at stealing bases? If you are a player wanting to develop that part of your skills, are there certain teams you might be interested in joining?

* Pitcher – catcher : Stolen base prevention, caught stealing, pickoff attempts, pop time, runners caught stealing
* Throwing arm of pitcher