DATA ANALYSIS OF PUBG

Naiqing Cai, Yuchen Zeng, Hao Pan, Qintao Ying

Department of Statistics, University of Wisconsin - Madison

Data Description

We are going to analyze 19 GB PUBG Data collected from kaggle created by KP.

- **Aggregate**
  Each match's meta information and player statistics are summarized (as provided by pubg). It includes various aggregate statistics such as player kills, damage, distance walked, as well as metadata on the match itself such as queue size, fpp/tpp, date, etc.
- **Deaths**
  Death that occurred within the 720k matches. Each row documents an event where a player has died in the match.
- **Data Resource**
  https://www.kaggle.com/skihikingkevin/pubg-match-deaths
  https://www.kaggle.com/c/pubg-finish-placement-pr data
- **Read Data**
  agg = read.csv('pubg-match-deaths/aggregate/agg_match_stats_0.csv')

Statistical Method and Computation

- **Statistical Method**
  Linear Regression
  Hypothesis Testing (one-way ANOVA)
- **Computation**
  HPC
  R
  Python

Maps

Fig. 1: Erangle Map.

Fig. 2: Miramar Map

Goal

- **Aggregate**
  1. Distance
  Ride Distance vs Rank
  Walk Distance vs Rank
  2. Team Member
  Team Size vs Rank
  3. Kills
  Kills and Rank
- **Deaths**
  1. Maps (Erangel and Miramar)
  Death vs Map
  2. Weapon Comparison
  Powerful Weapon
  3. Duration
  Average Player Damage vs. Average Survival Time
- **Prediction**
  Predict Top10 based on player statistics.
- **Strategy Analysis**
  Give some actionable suggestions for achieving a better rank.

Fig. 3: PUBG Game Image