



## 1. Introduction

The GAISE report provides many useful recommendations for educators. The 2016 report added 2 new emphases for the “Teach statistical thinking” recommendation:

- teach statistics as an investigative process of problem-solving and decision-making
- give students experience with multivariable thinking

For educators, there is always demand for interesting data sets and examples that can promote learning.

**The GOAL of this work is to share 1) resources and 2) experiences for using data from the Call of Duty franchise for developing data science skills.**

## 2. Data Collection

The data were collected by the presenter while playing Call of Duty: Black Ops Cold War.

**The primary data describe a player’s performance in an online multiplayer match.** Each observation represents a unique match and contains pre and post game data as shown in Fig 1. All observations correspond to the same player.

There are **2 gameplay datasets**:

- one is “easy” to use
- one is more challenging (typos, increased NA’s) to illustrate realities of data analysis

Secondary data include **3 additional datasets** containing information about Maps, Weapons, and Game Modes.

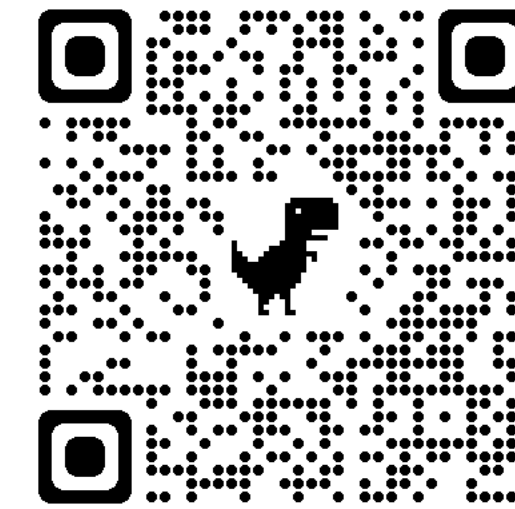
Fig 1: Preview of gameplay datasets. 12 of 25 variables shown.

	Map1	Map2	Choice	MapVote	Result	Eliminations	Deaths	Score	Damage	TotalXP	XPType	GameType
1	Moscow	Miami Strike	Miami Strike	5 to 0	100-97	22	17	4070	634	11002	10% Boost	HC - TDM
2	Moscow	WMD	Moscow	2 to 0	76-89	20	15	5305	560	9451	10% Boost	HC - TDM
3	NA	NA	Yamantau	NA	100-92	18	11	3335	483	12948	10% Boost	HC - TDM
4	Drive-In	Jungle	Drive-In	2 to 0	80-100	10	19	2170	280	11502	Double XP + 10%	HC - TDM

## 3. Resources

You may obtain the following resources:

- Data dictionary
- 5 datasets
- Examples with code



by visiting:

<https://github.com/matthewdslifko/CallOfDutyProject>

## 4. Experiences

**Where have I used this data?**

- University-level courses with 7 to 40 students
- Introductory R/Data Science courses (No prerequisites)
- Introductory Statistical Learning (R and Intro Stat prerequisites)

**How have I used this data?**

- For practice applying data wrangling, data visualization, and modeling concepts
- To motivate challenges of real datasets
- Types of activities:
  - mini-projects and final projects
  - take-home exam
  - class demonstrations

**Suggestions:**

- Limit use to a college classroom because of mature content of game
- Have students work in groups initially to help those unfamiliar with such games
- Provide ample background information so that lack of familiarity with game is not a problem
- Leverage student knowledge of the game

## 5. EDA Example

**Background:** “Damage” represents the amount of damage issued by the player on the opposing team’s players, weapons, and vehicles.

**Task:** Explore the distribution of Damage.

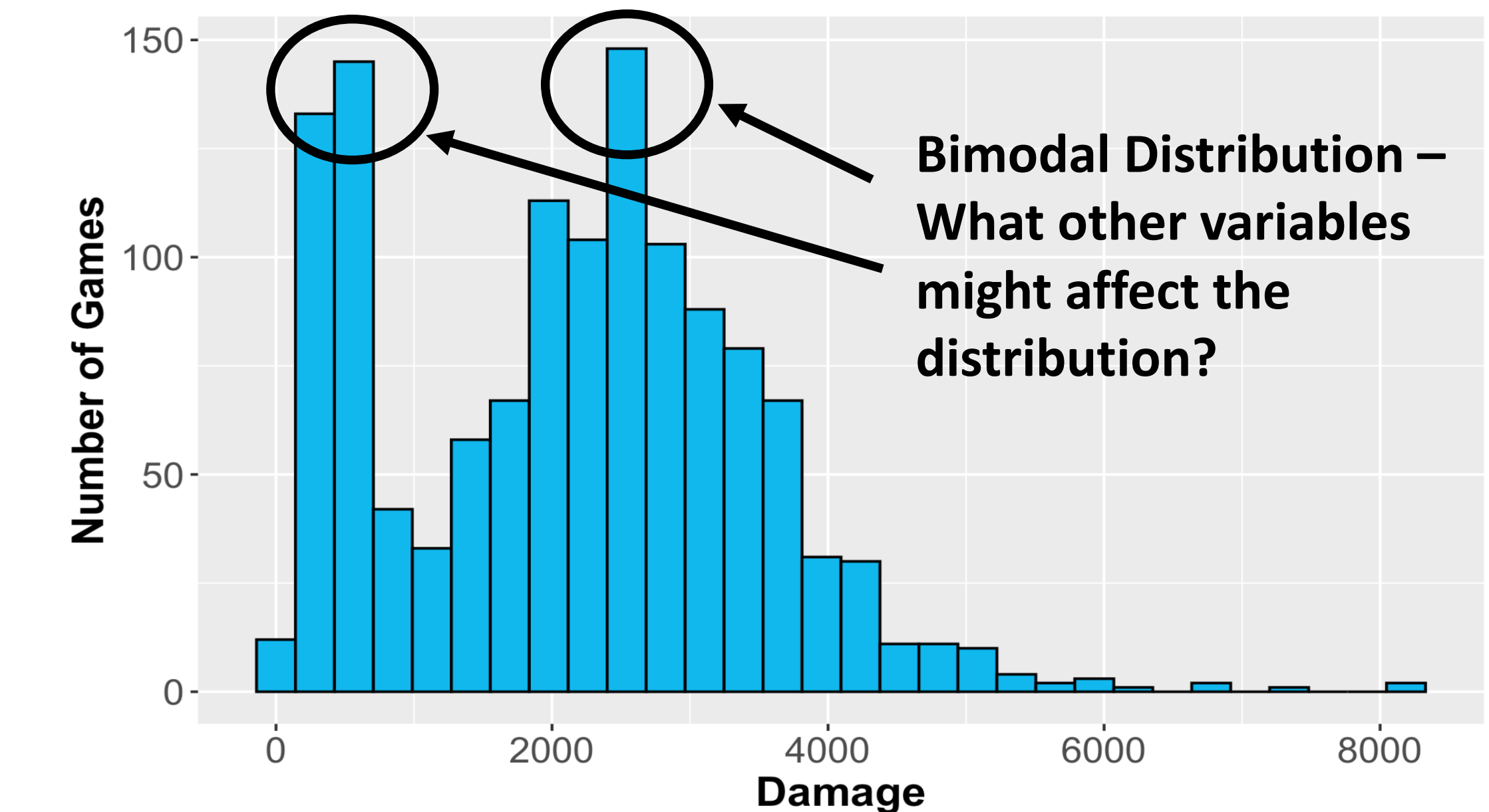


Fig 2: Distribution of damage variable.

**Additional background:** Some values in the “GameType” variable include an “HC” designation. Unlike Core games, players begin with less health and health does not regenerate in Hardcore (HC) games.

**Updated Task:** Explore the distribution of Damage using this information.

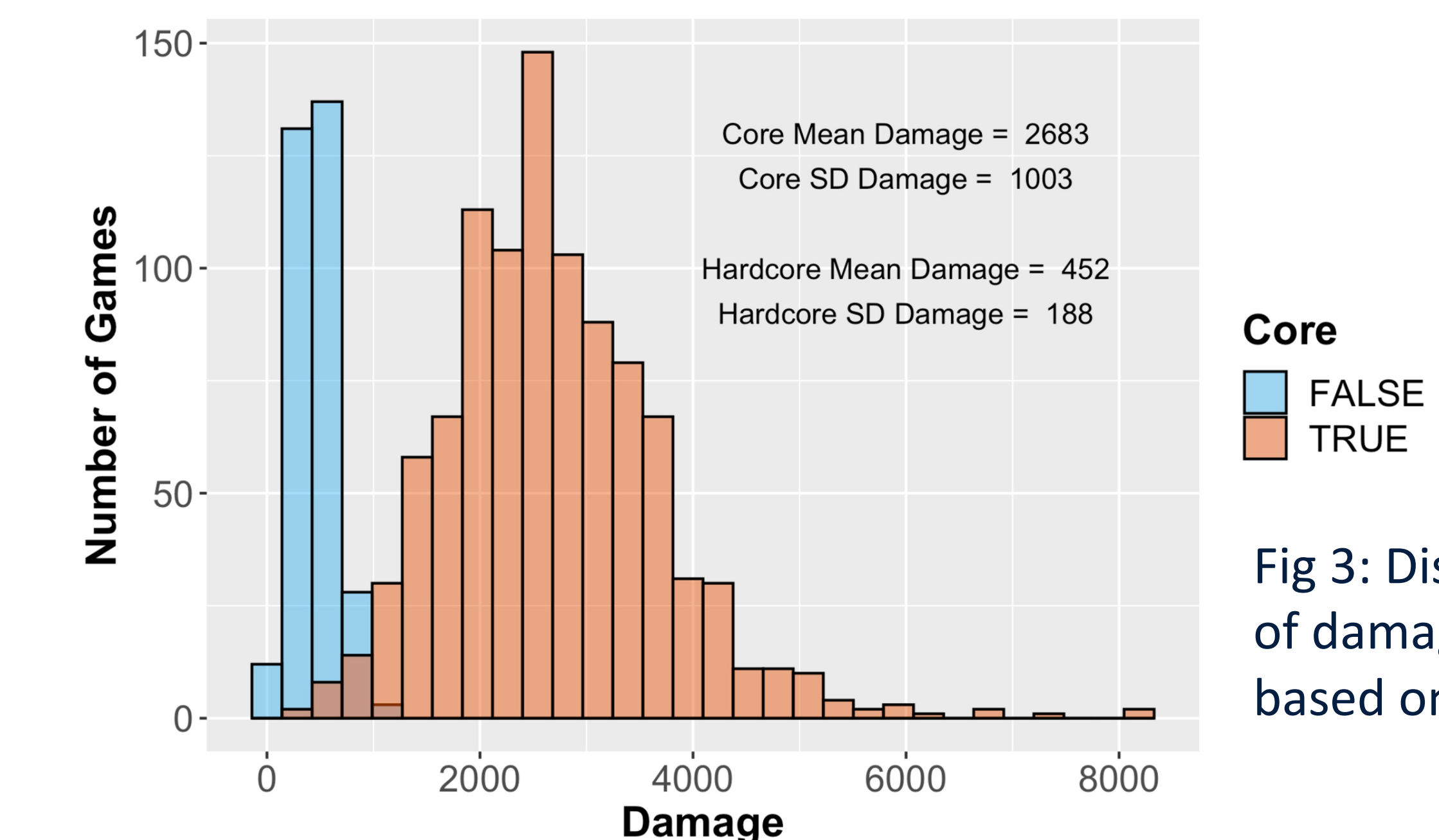


Fig 3: Distribution of damage variable based on Core

This example provides an opportunity to discuss:

- Data visualization and process of exploration
- Complementary roles of visualization and summary statistics
- Annotating plot with text
- String processing and new variable creation (to create Core variable)
- Multivariable thinking

Take a handout or visit the GitHub page for more examples