April Trinh (BID: 013075686)
Dr. Markus Eger
CS 4990.02
19 Feb. 2022

Lab 1 Report

Data Set

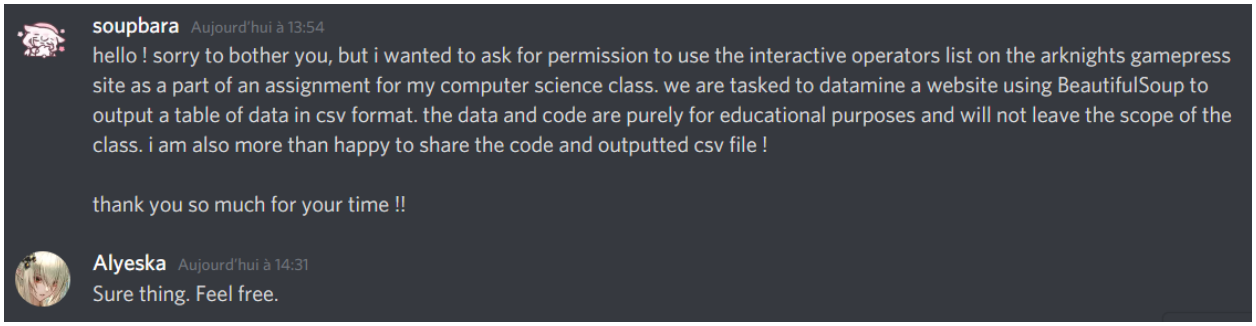
The data I collected is taken from the GamePress website for the mobile tower defense game *Arknights*. The values represent the different characters, or "operators," that are currently playable in the game, alongside the categories they fall under and their individual statistics.

The data was extracted from the Interactive Operators List using Python 3 and BeautifulSoup. The extracted data was then outputted into CSV format via Pandas.

While Arknights and its relevant data is originally owned by Hypergryph Network Technology Co. Ltd. (the creator), the data and content compiled on GamePress is protected under copyright. As quoted from their Terms of Use under "Intellectual Property Rights,"

> "the Site [GamePress] is [GamePress Media Inc] proprietary property and all source code, databases, functionality, software, website designs, audio, video, text, photographics, and graphics o the Site (collectively, the 'Content')... are owned or controlled by [GamePress Media Inc] or licensed to [GamePress Media Inc], and are protected by copyright and trademark laws... of the United States, international copyright laws, and international conventions."

Fortunately, users are able to copy any portion of the content for personal and non-commercial use, which is what this lab falls under. However, it is important to note that "Prohibited Activities" includes "Systematically retrieve data or other content from the Site to create or compile, directly or indirectly, a collection compilation, database, or directory without written permission from us." As I needed to compile the data into a CSV file, I reached out to the site lead of the Arknights GamePress team, Alyeska, on Discord. The following is a screenshot of their permission for me to use the data collected for the lab.



Data Format

The following table provides the data's label, type, valid range of values, and a brief description of what it represents in the game.

In order to understand what the values represent, it is important to understand the mechanics of a tower defense game and *Arknights* itself. Enemies travel on a fixed path from a spawn point to an end point. The goal of each map is to prevent this by depleting the enemies health before they can reach the end points by placing operators (the "towers") on certain tiles so they may target the enemies. In order to place the operators on the map, they first must be available to deploy; this means that the player must collect enough "deployment points," and each operator "costs" a certain amount. Operators can be removed from the field at any point in time, whether it be by proxy of the player or if the operator's health is fully depleted. In order for the operator to be "redeployed," the player must possess enough "deployment points" and wait a cool down period called "redeployment time".

Some discrepancies to note are:

- *"Justice Knight"*'s name includes the quotation marks.
- Operators who were announced but not yet implemented into the game will have certain values listed as 0 when it is otherwise not possible.
    - At the time of this report, these operators are *Goldenglow* and *Quercus*.
- Regarding Target Number, "Multi (Block #)" means that an operator's attack only strikes enemies that they are currently blocking. This is not to be confused with "AOE," which means that an operator's attack strikes *all* enemies in range regardless of if they are currently blocked by themself or not.

Abbreviations:

- Operator: OP
- Health Points: HP
- Attack: ATK

- Defense: DEF
- Potential: Pot
    - Refers to duplicate copies of operators obtained, and each duplicate copy increases or decreases certain stats
- Area of Attack: AOE
- Single Target: ST
- Multi (Block #): Multi

Label	Data Type	Valid Range	Description
Name	String	Unique Values	Unique ID, name of the operator
Class	String	1 of 8 types	General category the OP is classified as
Archetype 1	String	1 from list	Subset category, based off of an OP's class and utility
Archetype 2	String	N/A or 1 from list	Subset category, based off of an OP's class and utility
Rarity	Integer	1 – 6	Dictates rate of appearance through banner pulls, number of skills, number of talents, number of rank up tiers, etc.
Base HP	Integer	Positive non-zero	HP of a max level OP at 0/6 Pot and 0% Trust
Base ATK	Integer	Positive non-zero	ATK of a max level OP at 0/6 Pot and 0% Trust
Base Def	Integer	Positive non-zero	DEF of a max level OP at 0/6 Pot and 0% Trust
HP (Max Potential)	Integer	Positive non-zero	HP of a max level OP with 6/6 Pot and 0% Trust
ATK (Max Potential)	Integer	Positive non-zero	ATK of a max level OP with 6/6 Pot and 0% Trust
DEF (Max Potential)	Integer	Positive non-zero	DEF of a max level OP with 6/6 Pot and 0% Trust
HP (Max Trust)	Integer	Positive non-zero	HP of a max level OP with 0/6 Pot and 100% Trust
ATK (Max Trust)	Integer	Positive non-zero	ATK of a max level OP with 0/6 Pot and 100% Trust
DEF (Max Trust)	Integer	Positive non-zero	DEF of a max level OP with 0/6 Pot and 100% Trust
Full HP	Integer	Positive non-zero	HP of a max level OP with 6/6 and 100% Trust
Full ATK	Integer	Positive non-zero	ATK of a max level OP with 6/6 and 100% Trust
Full DEF	Integer	Positive non-zero	DEF of a max level OP with 6/6 and 100% Trust
Resistance	Integer	Positive non-zero	"Defense" against enemies dealing arts damage
Base Cost	Integer	1 – 99	Cost to deploy a max level OP with 0/6 Pot onto the field
Max Cost	Integer	1 – 99	Cost to deploy a max level OP with 6/6 Pot onto the field

Block Count	Integer	0 – 3	Number of enemies an OP can block
Base Redeployment Time	Float	Positive non-zero	Time in seconds it takes to redeploy an OP with 0/6 Pot
Max Redeployment Time	Float	Positive non-zero	Time in seconds it takes to redeploy an OP with 6/6 Pot
Attack Interval	Float	Positive non-zero	Time between an OP striking an enemy in seconds
Target Number	String	1, "Multi," or "AOE"	Number of targets an operator can hit per strike
Damage Type	String	"Arts" or "Physical"	Type of damage dealt

Data Mining

Of the three tasks, the operator data collected is best suited for clustering. We can determine similarities in stats between operators within a single class. Because of the availability of different archetypes, the similarities can be further broken down into smaller categories. It would also be interesting to examine the potential differences between operators that fall into the archetype "AOE" and those with a Target Number of "AOE," likewise for archetype "ST" and Target Number of "ST."

I would hope to see patterns between the classes and archetypes to improve my strategy in the game, which includes team composition and to know which operators I should look out for regarding specific stats. It would also be interesting to see which operators fare the best on certain maps, as *Arknights* carries a wide variety of map mechanics.

This data would be representative for those who wish to design a similar tower defense game. It is important to design characters that both feel unique and also blend well with existing groups. It would also be significant to make sure that the game is balanced for both casual players and those who are looking for more unconventional ways of play.