

Evil Geniuses SWE OA README

This is Shanrong Zhang(zhangshanrong@outlook.com). I am so happy to finish the OA. I will answer the question in this readme file.

Requirements

This script requires Python 3.6+ and the following libraries:

pandas seaborn matplotlib scipy numpy You can install these libraries using pip:

```
pip install pandas seaborn matplotlib scipy numpy
```

Question 1

a) Utilizing Pandas, I processed a Parquet file detailing the status of two teams and 10 players in a game. And I specified the engine as pyarrow, which can help to read the json type data.

```
Q1 a
   round_num  tick side  ...  map_name  utility_used  player
0          1   2511   T  ...  de_overpass         0.0  Player0
1          1   2527   T  ...  de_overpass         0.0  Player0
2          1   2543   T  ...  de_overpass         0.0  Player0
3          1   2559   T  ...  de_overpass         0.0  Player0
4          1   2575   T  ...  de_overpass         0.0  Player0
```

b) Considering chokepoints as triangles and quadrilaterals based on their coordinates, I applied the Barycentric method to verify a point's presence within these areas. The 'is_within_boundary' function with O(n) complexity was implemented using standard Python libraries. The solution involves several key functions.

```
def is_point_in_triangle(pt, v1, v2, v3):

def is_point_in_quadrilateral(pt, v1, v2, v3, v4):

def is_point_on_line(pt, v1, v2):

def is_within_boundary(self, z_min, z_max):
```

c) Under most circumstances, if the player is alive, the inventory typically is not none.

```
def extract_weapon_classes(self):

Q1 c
0  [Pistols, Grenade]
1  [Pistols, Grenade]
2  [Pistols, Grenade]
3  [Pistols, Grenade]
4  [Pistols, Grenade]
Name: weapon_classes, dtype: object
```

Question 2

a) No, it was not.I initially plotted Team2's T-side movements over 15 rounds and noticed the light blue entry area wasn't frequently used. Using a function to count players entering this boundary per round, I found that only two players utilized this path in the 16th round, indicating it wasn't a common strategy for the team.

```
Q2 a
round_num
16      2
Name: within_boundary, dtype: int64
```

```
def plot_player_movements(self, round_num, team):

def count_players_within_boundary_each_round(self, team, side):
```

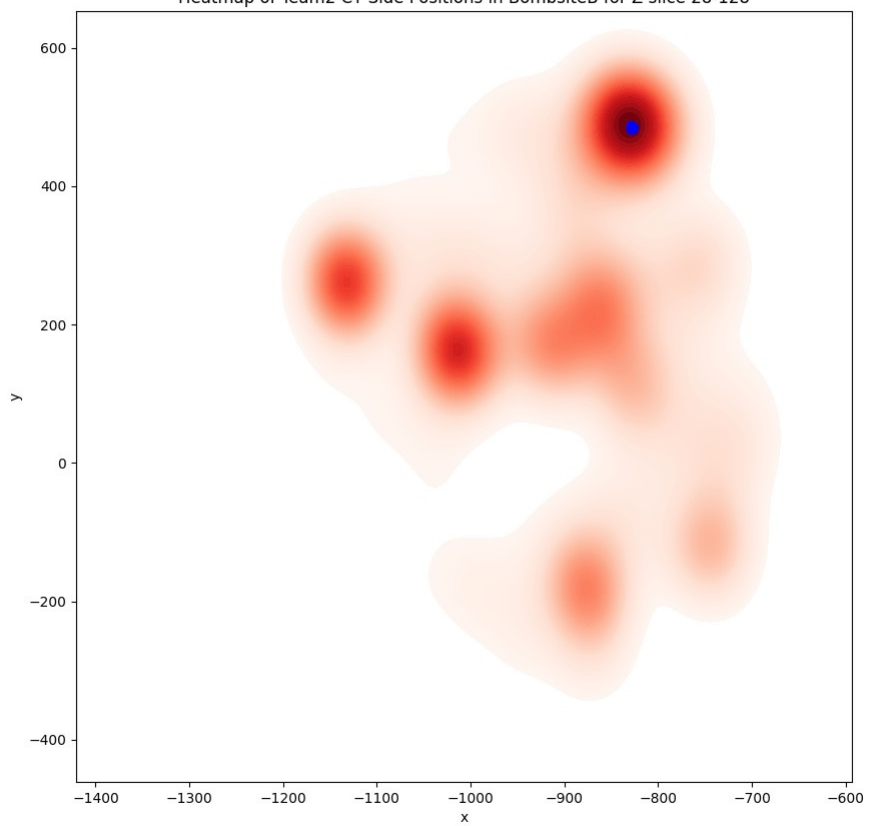
b) The average timer stands at 28.56 seconds. I first filtered the top two players entering the area with a rifle or SMG per round. Then, using the round number, I identified each round's start time. For a broader approach to average timer calculation, I utilized ticks, given their persistence even upon bomb planting.

```
def get_first_two_entry_rows(self):

def get_average_timer(self):
```

c) Given the game's 3D nature, creating a 3D heatmap could be complex. So, I split the z-axis into two, identified the top five frequented player locations, and found this to be at the suspected "BombsiteB" waiting point. The coordinates is (-826, 485, 96).

Heatmap of Team2 CT Side Positions in BombsiteB for Z slice 28-128





(You might not want to face a situation like that!)

Question 3

1. Use Command Line for Customized Data Analytics

I intend to create frequently-used classes and inquire with stakeholders about the methods they most commonly employ. They can then conduct queries via the command line interface, with predefined commands for tasks like calculating average timers, generating heatmaps, querying player stats, etc. The arguments for these commands will serve as analysis parameters, like team name, side (T or CT), and player name.

To ensure user-friendliness, I will offer detailed documentation elucidating each command's function, required arguments, expected results, and usage. This will feature command examples for the coaching staff to adapt for their needs.

Moreover, to stay aligned with stakeholder requirements, this approach will undergo regular updates.

2. Build a web dashboard

If time allows, I can build a simple web-based data dashboard where I can implement some common data retrieval logic on the backend. On the frontend, users would only need to interact with the mouse or drag column name tags to perform common group-by operations and generate charts based on their selections.

During my internship at Kwai, a similar system was built internally. In that system, specific data tables could be selected based on the requirements of the product manager, and common filtering, sorting, and statistical operations could be performed on the web page. I believe this will greatly facilitate stakeholders.