CSGO Pro Matches Data Structure

General Purpose:

This dataset is a collection of each CS:GO professional match played from 11/2015 to 03/2020. CS:GO is a competitive video game and in this game, matches are played between two teams of 5 players on maps that the teams draft pick. Matches are normally conducted as best of 3, where the two teams play three games, each on a different map, but there are also best of 1 and best of 5 matches. In this dataset, each match has an id which is common amongst the files, and details about these matches are located throughout. These details include:

- **match_id**: id of match
- **team_1 / team_2**: the two teams playing (this is two fields/columns)
- **_map**: the map played during the game
- **result_1 / result_2**: the scores of the respective teams (this is two fields)
- **map_winner**: winner of the map
- player_name: name of the player on the team
- rating: their rating after the match (out of 2.0)
- date: date the match was played
- t1_picked_1 / t2_picked_1 / left_over: the maps that each team picked and the map selected from process of elimination (this is three fields/columns)

The dataset was created by Mateus Dauernheimer Machado, so many thanks to him. It is important to note that, since there are not only BO1 matches, the info in "results.csv" is for individual maps and "players.csv" and "picks.csv" refer to the overall match which may have more than 1 map played. My data structure allows users to see a team's past record of their performances on specific maps and sometimes even versus specific opponents. The data can then be used to draw possible conclusions on how future matches will turn out based on a team's past history with the map and enemy team.

Data Structure Description

In my dictionary data structure, I chose to have each team's sub-dictionary to be split into the maps they've played, and in those maps (which are also dictionaries) you can find each match in which the team played a game on that map. The fields I decided to display are listed in the "general purpose" section. I ignored "economy.csv" and other fields in the csv's I used because they were too round-specific and irrelevant for my purpose.

General Example:
{Team: {**Map**: {**match id**: [**win/loss**, **score**, **opponents**, whose map pick, [mvp name, mvp rating], date], **match id 2**: [...], ...}, **Map 2**: {**match id**: [...], …}, …}, **Team 2**: { … }, …}, ...}

Data-based Example:
{ 'Liquid' : { **'Dust2'** : { **2336445** : [**'Won'** , **'16:9'** , **'Astralis'** , "Liquid's Pick" , ['dupreeh' , 1.15] , '2019-09-28'] , … }, **'Inferno'** : { … } , …. }, **'fnatic'** : { … } , … }**

Spaces between punctuation added for clarity

Bold: Data taken from "results.csv"
Red: Data taken from "picks.csv"
Gold: Data taken from "players.csv"