**Pandas - Monster Task**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Name | Strength | Speed | Type |
| 0 | Mike | 1 | 3 | Monster |
| 1 | Sully | 99 | 65 | Monster |
| 2 | Randall | 35 | 99 | EvilMonster |
| 3 | Boo | 5 | 5 | Human |

Figure : Monster Table #1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Name | Strength | Speed | Type | Color | MCODE |
| 90 | Fungus |  | 3 | Monster | Orange | 223423 |
| 91 | Celia | 40 |  | Monster | Purple | 223422 |
| 92 | Waternoose | 51 | 35 | EvilMonster | Gray | 223234 |

Figure : Monster Table #2

|  |  |  |
| --- | --- | --- |
|  | Species | FavouriteFood |
| 0 | Monster | Screams |
| 1 | Human | Candy |
| 2 | EvilMonster | <NULL> |
| 3 | Donut | Flour |

Figure : Monster Food

***Note: This exercise is intended to be done in order. Later tasks may depend on the results of previous tasks.***

**DataFrame Construction**

Task 1: Construct the monster table #1 using the following:

* A list of lists
* A dictionary
* A list of dictionaries

Assign the result to variable df\_monster1.

Task 2: Save the data in monster table #2 to a csv file. Now use the pd.read\_csv(…) function to construct the dataframe. Assign the result to df\_monster2.

Task 3: Construct the monster food table using the method of your choice. Assign the result to df\_monster2.

**Indexes**

Task 1: Change the index of monster table 2 to [4,5,6].

Task 2: Use the reset\_index() method on monster table 2. What does it do?

**Indexing**

Task 1: Using the .loc accessor, how would you get the following:

|  |  |
| --- | --- |
|  | Using .loc |
| Mike’s speed |  |
| Boo’s strength |  |
| Celia’s type |  |
| The favourite food for humans |  |

Task 2: Using the .iloc accessor, how would you get the following:

|  |  |
| --- | --- |
|  | Using .iloc |
| Mike’s speed |  |
| Boo’s strength |  |
| Celia’s type |  |
| The favourite food for humans |  |

Task 3: How would you use slicing to select the first three rows of monster table #1?

Task 4: How would you select the strength and speed columns in monster table #1?

Task 5: How would you use Boolean indexing to select rows where the type is ‘Monster’?

**Concatenation and Joining**

Task 1: Concatenate monster table 1 and monster table 2 (row-wise). Assign the result to variable df\_monster.

Task 2: Merge df\_monster with the Monster Food table. Assign the result to variable df\_monster. What kind of join is most suitable?

Task 3: Make sure you can replicate the two tasks above using .append and .join.

**Filling in Missing Values**

Task 1: Assign the FavouriteFood of all EvilMonsters to ‘ExtremeScreams’.

Task 2: Assign the speed of Celia to the average speed of all monsters.

Task 3: Assign the color green to Mike, and blue to Sully. If there are any others missing colors, fill in their values with an appropriate value.

Task 4: Use a forward fill to fill in the rest of the blanks.

**Dropping / Sorting**

Task 1: Drop the MCODE column from df\_monster.

Task 2: Sort df\_monster by Type, then Name.

**Element Wise Operations**

Task 1: Create a new column in df\_monster called ‘AverageStat’. The value should be the average of the strength and speed for the row.

Task 2: Repeat the above task. Instead of using an average, use min and max.

**Aggregation**

Task 1: How would you calculate the maximum AverageStat across all rows?

Task 2: Write code to find the row with the maximum AverageStat. This is the strongest monster.

**Grouping and Aggregation**

Task 1: Find the average, maximum and maximum speed by Type.

Task 2: Join the results of task 1 back to df\_monster.

**Pivot Tables**

Task 1: Create a new column ‘IsStrong’. The column should be assigned to True if strength is above 50, and False otherwise.

Task 2: Create a pivot table that shows the average speed for groupings of IsStrong and Type.