Practice Homework 3.2 : Restaurant Data Processing

This homework deals with the following topics:

- Joining & filtering data

The Assignment

This assignment is entirely optional and focuses on practicing how to load a dataset and the basic use of joining and filtering the data.

To begin this assignment, we have provided you with a .ipynb file in Codio, which contains the instructions as inline comments. We have also provided a dataset for the exercise called "Restaurant Data with Consumer Ratings". Please continue to use this database to complete practice assignments 3.1 through 3.3. Details about the dataset are described in the following sections.

For each question, there are clear instructions in each cell. Follow those instructions and write the code after each block of:

```
# YOUR CODE HERE
raise NotImplementedError()
```

Make sure to delete the line raising an error! Please use the exact variable name if it is specified in the comment.

We'll run a Python test script against your program to test whether the code in each cell is correct.

Dataset Overview:

userID	latitude	longitude	smoker	drink_level	dress_preference	ambience	transport	marital_status
U1001	22.139997	-100.978803	FALSE	abstemious	informal	family	on foot	single
U1002	22.150087	-100.983325	FALSE	abstemious	informal	family	public	single
U1003	22.119847	-100.946527	FALSE	social drinker	formal	family	public	single

We will use "Restaurant Data with Consumer Ratings" as the dataset for all practice assignments. The link to this dataset is shown here:
https://www.kaggle.com/datasets/uciml/restaurant-data-with-consumer-ratings

This dataset contains many different tables, each representing different pieces of information about each restaurant or customer. For this assignment we are going to use "geoplaces2.csv", "chefmozaccepts.csv", "chefmozcuisine.csv", "rating_final.csv" and "userprofile.csv".

The file “geoplaces2.csv” presents a comprehensive overview of various restaurants, each uniquely identified by a ‘placeID’, the geographical coordinates ‘latitude’ and ‘longitude’, as well as a specific metric representation (‘the_geom_meter’), to pinpoint each restaurant's exact location. Essential details like the restaurant's ‘name’, ‘address’, ‘city’, and ‘state’ provide context about its locale.

placeID	latitude	longitude	the_geom_meter	name	address	city	state
134999	18.915421	-99.184871	0101000020957F00	Kiku Cuernavaca	Revolucion	Cuernavaca	Morelos
132825	22.1473922	-100.983092	0101000020957F00	puesto de tacos	esquina santos dego	s.l.p.	s.l.p.
135106	22.1497088	-100.9760928	0101000020957F00	El Rinc◆n de San Fr	Universidad 169	San Luis Potosi	San Luis Potosi

The table titled "chefmozcuisine.csv" shows a list of cuisines of restaurants, each represented by a unique identifier ‘placeID’, along with the type of cuisine they specialize in, denoted by the column ‘RCuisine’.

placeID	Rcuisine
135110	Spanish
135109	Italian
135107	Latin_American

The file titled "chefmozaccepts.csv" relates restaurants, identified by the unique ‘placeID’, to their accepted modes of payment, denoted in the ‘Rpayment’ column.

placeID	Rpayment
135110	cash
135110	VISA

The table "userprofile.csv" captures detailed profiles of users. Each user, identified by the ‘userID’, has associated geolocation data including ‘latitude’ and ‘longitude’. The table also provides personal preferences and characteristics, such as smoking habits (denoted by ‘smoker’), level of alcohol consumption (‘drink_level’), style preferences (‘dress_preference’), favorite ambience (‘ambience’), mode of transport (‘transport’), and marital status (‘marital_status’).

userID	latitude	longitude	smoker	drink_level	dress_preference	ambience	transport	marital_status
U1001	22.139997	-100.978803	FALSE	abstemious	informal	family	on foot	single
U1002	22.150087	-100.983325	FALSE	abstemious	informal	family	public	single
U1003	22.119847	-100.946527	FALSE	social drinker	formal	family	public	single

The table "rating_final.csv" provides user ratings for various places, indexed by ‘placeID’. Each record captures a user's unique ID (denoted by ‘userID’), their overall satisfaction (‘rating’), their opinion on the quality of food (‘food_rating’), and their assessment of the service (‘service_rating’).

userID	placeID	rating	food_rating	service_rating
U1077	135085	2	2	2
U1077	135038	2	2	1
U1077	132825	2	2	2

What to submit

Before submitting, click “Validate” to check if your answers are right. When you are satisfied with your answers, in Codio go to the “Education” menu and select “Mark as Completed.”

Grading:

- This practice assignment consists of three parts: Parts 1 and 2 are worth 35 points per question. Part 3 is worth 30 points. The total score is 100.