Assignment-1

Data Analytics and Visualization (CS/IT314)

12 January, 2022

Instructions

1. Create"Assignment1_Branch_YourId.py" file. *YourId* will be the student ID and *Branch* will be "CS" or "IT". For example, if student ID is 202018001 and Branch is CS, then file name will be "Assignment1_CS_202018001.py".
2. A student has to write the code for following tasks.
3. Upload your "Assignment4_Branch_YourId.py" file into the Google form shared on Google class room. Fill the Google form.
4. Please be present during lab session. The deadline of the submission is 11:59 pm, 21 January, 2022.

Get help from:
Book: chapter 3: Python-Data-Science-Handbook, Jake VanderPlas
Data Source: https://data.gov.in/

Tasks

The assignment is uploaded on google Class. The assignment consists of two data files (a csv and a json). Both the file consists of **District Level *Manufacturing* (json file) and *Service* (csv) MSME Registered Enterprises under UDYAM Registration till last date**. MSME stands for Micro, Small and Medium Enterprise.

Json file: Number of Manufacturing MSME
CSV file: Number of Service MSME

1. Create two data frames by reading above two files.
2. Find out total "Small" Manufacturing MSME in India.
3. Create a dataframe having state wise total number of "Micro","Small" and "Medium" Services MSE (as shown below) and save the results as a CSV file.

STATE_NAME	MICRO	SMALL	MEDIUM
ANDAMAN AND NICOBAR ISLANDS	2884	111	6
ANDHRA PRADESH	115320	5799	422
ARUNACHAL PRADESH	1201	53	6
ASSAM	41051	2182	200
BIHAR	167835	3908	245
CHANDIGARH	9010	687	76
CHHATTISGARH	58905	2726	257
DADAR AND NAGAR HAVELI	2511	98	3

4. Join the both the data frame based on common STATE_NAME, DISTRICT_NAME, Lg_Dist_Code and Last_Updated. The result should look like below. "x" and "y" in below image represent the manufacturing MSME and service MSME respectively.

STATE_NAME	Lg_Dist_Code	DISTRICT_NAME	MICRO_x	SMALL_x	MEDIUM_x	Total_x	Last_Updated	MICRO_y	SMALL_y	MEDIUM_y	Total_y
ANDHRA PRADESH	502	ANANTHAPUR	11370	483	33	11886	2022-01-11	7797	238	10	8045
ANDHRA PRADESH	503	CHITOOR	13037	592	51	13680	2022-01-11	9623	351	19	9993
ANDHRA PRADESH	505	EAST GODAVARI	16711	1044	123	17878	2022-01-11	12347	615	49	13011
ANDHRA PRADESH	506	GUNTUR	15955	1678	169	17802	2022-01-11	11386	630	53	12069
ANDHRA PRADESH	510	KRISHNA	20460	1673	165	22298	2022-01-11	15189	1007	79	16275
ANDHRA PRADESH	511	KURNOOL	11229	558	37	11824	2022-01-11	8972	348	16	9336
ANDHRA PRADESH	517	PRAKASAM	9079	524	44	9647	2022-01-11	6527	231	7	6765
ANDHRA PRADESH	515	SPSR NELLORE	9782	670	52	10514	2022-01-11	7515	251	19	7885

5. Create a Pivot Table having rows STATE_NAME and columns Service and Manufacturing "MSME" as show in below. Use "Sum" to add up district wise number.

STATE_NAME	MEDIUM_x	MEDIUM_y	MICRO_x	MICRO_y	SMALL_x	SMALL_y
ANDAMAN AND NICOBAR ISLANDS	7	6	3905	2884	140	111
ANDHRA PRADESH	1030	422	153814	115320	10599	5799
ARUNACHAL PRADESH	11	6	2172	1201	99	53
ASSAM	309	200	65544	41051	3345	2182
BIHAR	413	245	247394	167835	5226	3908