VNIVERSITAT DE VALÈNCIA

STRUCTURES Laboratory

DATA AND ALGORITHMS

Degree in Data Science (1st)

Academic year 2023-24

Practice No. 6: Dictionaries
Completion period: Week of**06 to 05/10/2024**

Work methodology

Phases	Guy	Dedication
1. Solve the tasks set out in this scriptbefore to start the face-to-face session in the laboratory.	No presential	Max. 4.5 hrs.
2. Solve a problem in the laboratorynew exercise based on the resolution of previous tasks.	In person	Max. 3 hrs.

Introduction

In this practice the structure will be applied**dictionary**to store information on the temporal evolution of the population census in Spanish municipalities, managed by the National Institute of Statistics (INE). Again, we use this data set as an example for an information search problem. In the Virtual Classroom you can download the file "EvolMunicipios2022.csv" that contains the data of that record. Each line of the file reflects the identifying data of a municipality and the number of inhabitants registered in the municipality in each year from 1998 to 2022. As is known, the file includes data from the 8,137 Spanish municipalities.

Exercises (Phase 1, prior to the laboratory)

Homework 1

The program must be written**"pr6_v1.py"**to read and load the population census information into memory according to the INE format. The data reading process must result in a**dictionary**, which will associate the INE code of a municipality (key) with an object of type Municipality (value), which will contain all its associated information: code, name and population list by year.

To verify the correctness of the process, run the program reading the data from the file **"EvolMunicipios_small.csv"**. The dictionary of municipalities must have a size**10**. If you display each element of the dictionary you should get something like this (the format depends on how you do it):

28079: 28079 Madrid ', ' 3265038', '3273049', '3255944', '3213271', '3132463', '3128600', '3155359', '3099834', '3092759', '3016788', '2957058', '2882860', '28 79052' , '2881506']

08019: 08019 Barcelona ', ' 1615448', '1619337', '1621537', '1615908', '1595110', '1605602', '1593075', '1578546', '1582738', '1527190', '1505325', '1496266', '15 03451' , '1505581']

27028 : 27028 Lugo ['97211', '97613', '98519', '98276', '98025', '97995', '98268', '98134', '98560', '98761', '98457', ' 98007', '97635', '96678', '95416', '93853', '93450', '92271', '91426', '91158', '89509', '88901', '88235', '87480' , '86620']

42173 : 42173 Soria ['39450', '39695', '39821', '39398', '39112', '38881', '39171', '39168', '39516', '39753', '40147', ' 39987', '39838', '39528', '39078', '38205', '38004', '37200', '35769', '35178', '35112', '34640', '34088', '34045' , '33882']

38038 : 38038 Santa Cruz de Tenerife ['208688', '208563', '209194', '207312', '204856', '203692', '203585', '203811', '205279', '206593', '206965 ', '222271', '222643', '222417', '221956', '220902', '223148', '221567', '219446', '220022', '217415', '214153', '215132', '213050', '211930']

46078 : 46078 Burjassot ['38880', '38712', '38632', '38024', '37584', '37575', '37324', '37546', '37641', '38148', '38175', ' 38205', '38170', '38433', '37667', '37402', '37756', '37394', '37330', '37213', '36168', '35682', '35171', '34869' , '34426']

46126 : 46126 Foios ['7496', '7502', '7452', '7367', '7342', '7234', '7208', '7199', '7094', '7078', '7034', ' 6962', '6781', '6671', '6477', '6358', '6287', '6193', '6035', '5906', '5695', '5434', '5372', '5304' , '5265']

46248 : 46248 Turís ['7019', '6910', '6722', '6646', '6623', '6627', '6609', '6528', '6588', '6678', '6637', ' 6529', '6546', '6499', '6317', '6053', '5788', '5556', '5305', '5100', '4954', '4804', '4733', '4655' , '4560']

46250 : 46250 Valencia ['792492', '789744', '800215', '794288', '791413', '787808', '790201', '786189', '786424', '792303', '797028', ' 798033', '809267', '814208', '807200', '797654', '805304', '796549', '785732', '780653', '761871', '746612', '739014', '739412' , '739412']

48020 : 48020 Bilbao ['344127', '346405', '350184', '346843', '345821', '345110', '345122', '345141', '346574', '349356', '351629', ' 352700', '353187', '354860', '353340', '353168', '354145', '353173', '352317', '353567', '353950', '353943', '354271', '357589' , '358467']

Task 2

In the File "**provinces.csv**" the identification codes (INE) and the names of the provinces of Spain are available. The program must also read this file and store it in a second dictionary, which associates the name of the province (value) with each code (key).

The dictionary obtained must contain**52**keys and the following content:

{'04': 'Almeria', '11': 'Cadiz', '14': 'Cordoba', '18': 'Granada', '21': 'Huelva', '23': 'Jaen', ' 29': 'Malaga', '41': 'Seville', '22': 'Huesca', '44': 'Teruel', '50': 'Zaragoza', '33': 'Asturias', '07' : 'Balearics (Illes)', '35': 'Palmas (Las)', '38': 'Santa Cruz de Tenerife', '39': 'Cantabria', '05': 'Avila', '09': 'Burgos', '24': 'Leon', '34': 'Palencia', '37': 'Salamanca', '40': 'Segovia', '42': 'Soria', '47': 'Valladolid' ', '49': 'Zamora', '02': 'Albacete', '13': 'Ciudad Real', '16': 'Cuenca', '19': 'Guadalajara', '45': 'Toledo' , '08': 'Barcelona', '17': 'Girona', '25': 'Lleida', '43': 'Tarragona', '03': 'Alicante/Alacant', '12': 'Castellon/ Castello', '46': 'Valencia/Valencia', '06': 'Badajoz', '10': 'Caceres', '15': 'Coruna (A)', '27': 'Lugo', '32 ': 'Ourense', '36': 'Pontevedra', '28': 'Madrid', '30': 'Murcia', '31': 'Navarra', '01': 'Araba/Alava', '20 ': 'Gipuzkoa', '48': 'Bizkaia', '26': 'Rioja (La)', '51': 'Ceuta', '52': 'Melilla'}

Task 3

Using the dictionaries created in the previous tasks, you must make the program ask the user for a province code and display on the screen: (1) the name of the province and (2) the population of the year 2022 ofall the municipalities of that province. To do this, you should know that all the municipalities of a province have INE codes whose first two characters match the code of the province. For example, the codes of all the municipalities in the province of Valencia (code 46) begin with "46".

If you run the program with the file**"EvolMunicipios_small.csv"**and you search for the province code 46, the following result should be obtained:

Province code: 46
Population in 2022 of the municipalities of the Valencia/
Valencia province

46078 Burjassot 38880
46126 Foios 7496
46248 Turis 7019
46250 Valencia 792492

If the expected results have been obtained, you can perform the same process with the file " **EvolMunicipios2022.csv**". In this case, for the province**46**you should get (first and last lines only):

Province code: 46
Population in 2022 of the municipalities of the Valencia/
Valencia province

46001 Ademuz 1000
46002 Ador 1706
46003 Atzeneta d'Albaida 1146
46004 Agullent 2361
46005 Alaquàs 29537
46006 Albaida 5989
46007 Albal 16845
46008 Albalat de la Ribera 3444 46009
Albalat dels Sorells 4084 46010 Albalat
dels Tarongers 1368 ...

46256 Vilamarxant 10348
46257 Castelló 6932
46258 Villar del Arzobispo 3581 46259
Villargordo del Cabriel 602 46260
Vinalesa 3525
46261 Yatova 2169
46262 Yesa, La 227
46263 Zarra 370
46902 Gátova 412
46903 San Antonio de Benagéber 9874
46904 Benicull de Xúquer 1102