Dr. Sandro Andreotti

Programming for Data Science

Winter term 2023/2024

Assignment 3
Issued: 2023-10-19

1 Comprehensions I

Read the *phonebook.csv* file and store content in a proper data structure (e.g. list of lists or dictionary of lists). Use the proper comprehension (list, dict, set,..) to generate to the following collections:

1. all unique forenames separately for females and males (if gender can be inferred from title entry).
2. all entries (complete rows) for persons where forename and surname are longer than 8 characters.
3. all surnames (unique) containing more vowels than consonants
4. all tuples (forename, surname, title) sorted by surname
5. (optional) a rank list containing the ranks of the rows when ordered by surname (first) and forename (second). Example: Assuming a list `["a","c","z","b","y"]` the associated rank list reads `[0,2,4,1,3]`.

Remark: Each of the collections should be produced comprehension expressions.

2 Comprehensions II

Use list comprehension to generate from two lists of integers, A and B a list containing for each element a_i of A a list with values $[a_i^{b_0}, a_i^{b_1}, ...]$. Example: $A = [1, 2, 3, 4]$ and $B = [1, 3, 5]$ produces $[[1, 1, 1], [2, 8, 32], [3, 27, 243], [4, 64, 1024]]$

Hint: Nested comprehension

3 Sets

Apply proper set methods to identify:

- All names that appear as surname and forename
- All names that appear as surname or forename
- All names that are either only a surname or only a forename
- All surnames that are not also a forename