**Task 3**: Detect unstandardized data and provide examples of these cases.

Neke od osnovnih mogućih nestandardizovanih podataka su:

* Drugačiji zapis naziva biznisa (viška razmak, bez apostrofa, razlika u velikim i malim slovima, isl)
* Greška u zapisu naziva grada

Za neke biznise sa najvećim frekvencijama ćemo uraditi pretragu nestandardizovanih naziva.

copy(

select distinct name, count(name)

from ybusiness

where name ilike '%Donald%' and name ilike '%mc%' and categories ilike '%food%'

group by name)

to 'D:\Filmovi\McDonalds.csv'

with csv delimiter '|' header

|  |  |
| --- | --- |
| **name** | **count** |
| McDonald | 1 |
| Mcdonald's | 3 |
| McDonald's | 805 |
| McDonald's - MGM Grand The District Food Court | 1 |
| Mcdonalds | 5 |
| McDonalds | 22 |

where name ilike '%starbuck%'

|  |  |
| --- | --- |
| **name** | **count** |
| Starbucks | 1066 |
| Starbucks Coffee | 6 |
| Starbucks Reserve | 1 |

Where name ilike '%subway%' and name not ilike '%station%'

|  |  |
| --- | --- |
| **name** | **count** |
| Subway | 768 |
| Subway Mahogany | 1 |
| Subway sandwich shop | 1 |

where name ilike '%horton%' and name ilike '%tim%'

|  |  |
| --- | --- |
| **name** | **count** |
| Tim Horton | 7 |
| Tim Horton's | 61 |
| Tim Horton's - University Of Calgary | 1 |
| Tim Horton Donuts | 4 |
| Tim hortons | 1 |
| Tim Hortons | 333 |
| Tim Hortons Cafe & Bake Shop | 1 |

where name ilike '%pizza%' and name ilike '%hut%'

|  |  |
| --- | --- |
| **name** | **count** |
| Pizza Hut | 320 |
| Pizza Hut Express | 2 |
| Pizza Hut Inside Target | 1 |

where name ilike '%taco%' and name ilike '%bel%'

|  |  |
| --- | --- |
| **name** | **count** |
| Kentucky Fried Chicken & Taco Bell | 1 |
| Taco bell | 1 |
| Taco Bell | 313 |
| Taco Bell Cantina | 3 |
| Taco Bell of Canada | 2 |

where name ilike '%burger%' and name ilike '%king%' and name not ilike '%hero%'

|  |  |
| --- | --- |
| **name** | **count** |
| Burger King | 302 |
| BURGER KING | 1 |
| Burger King | 1 |
| Burger King Canada | 2 |
| Burger King Restaurant | 7 |
| Burger King Restaurants | 4 |
| Burger King Unit 4232 | 1 |

where name ilike '%walgreen%' or (name ilike 'wal' and name ilike 'green')

|  |  |
| --- | --- |
| **name** | **count** |
| Health Care Clinic At Walgreens | 1 |
| Walgreen's | 1 |
| Walgreens | 301 |
| Walgreens Mail Service Pharmacy | 1 |
| Walgreens Pharmacy | 1 |

where name ilike '%wendy%' and (categories ilike '%burger%' or categories ilike '%food%')

|  |  |
| --- | --- |
| **name** | **count** |
| Wendy's | 294 |
| Wendy's Noodle Cafe | 1 |
| Wendy's Old Fashioned Hamburgers | 2 |
| Wendy's Restaurant | 1 |
| Wendy's Richmond Hill | 1 |
| Wendys | 4 |

where name ilike '%UPS%' and name ilike '%store%'

|  |  |
| --- | --- |
| **name** | **count** |
| The UPS Store | 279 |
| UPS UPS Store | 1 |

where name ilike '%cvs%'

|  |  |
| --- | --- |
| **name** | **count** |
| Cvs | 1 |
| CVS | 1 |
| Cvs 1742 | 1 |
| Cvs 2772 | 1 |
| CVS Caremark | 1 |
| Cvs Front Store | 1 |
| CVS Minute Clinic | 1 |
| Cvs Pharmacy | 9 |
| CVS Pharmacy | 250 |
| Cvs Store | 1 |
| Cvs Store 2877 | 1 |
| Cvs/Pharmacy | 1 |

where name ilike '%great%' and name ilike '%clip%'

|  |  |
| --- | --- |
| **name** | **count** |
| Great Clip For Hair | 1 |
| Great Clips | 240 |
| Great Clips - Anderson Springs | 1 |
| Great Clips - Camelback Colonnade | 1 |
| Great Clips - Fry's Sundome Plaza | 1 |
| Great Clips - Odana | 1 |
| Great Clips Hair Salon | 1 |

where name ilike '%circle%' and (name ilike '% k' or name ilike '%k %')

|  |  |
| --- | --- |
| **name** | **count** |
| circle k | 1 |
| Circle K | 238 |
| Circle K Food | 1 |
| Circle K Stores | 9 |
| Shell - Circle K #6347 | 1 |

where name ilike '%enterprise%' and name ilike '%rent%'

|  |  |
| --- | --- |
| **name** | **count** |
| Enterprise Motorcycle Rental | 1 |
| Enterprise Rent-A-Car | 227 |
| Enterprise Rent-A-Truck | 1 |
| Enterprise Rent a Car | 1 |
| Enterprise Truck Rental | 1 |

where name ilike '%domino%'

|  |  |
| --- | --- |
| **name** | **count** |
| Domino's | 35 |
| Domino's Pizza | 210 |
| Domino Foods | 1 |
| Dominos | 3 |
| Dominos' Pizza | 1 |
| Dominos Pizza | 3 |

where name ilike '%dunkin%' and name ilike '%donut%'

|  |  |
| --- | --- |
| **name** | **count** |
| Baskin Robbins & Dunkin' Donuts | 1 |
| Dunkin' Donuts | 187 |
| Dunkin' Donuts & Baskin-Robbins | 1 |
| Dunkin' Donuts (Canada) Ltee | 1 |
| Dunkin' Donuts/Baskin-Robbins | 1 |
| Dunkin Donut | 2 |
| Dunkin Donuts | 25 |
| Dunkin Donuts Baskin Robbins | 1 |
| Dunkin Donuts & Baskin Robbins | 5 |
| Dunkin Donuts Baskin-Robbins | 1 |
| Dunkin Donuts Baskin Robbins | 3 |

where name ilike '%7%' and name ilike '%eleven%'

|  |  |
| --- | --- |
| **name** | **count** |
| 7-Eleven | 178 |
| 7-Eleven Bar | 1 |
| 7 Eleven | 1 |

Sljedeće ćemo prikazati neke gradove čija su imena nestandardizovana.

|  |  |
| --- | --- |
| **city** | **count** |
| 110 Las Vegas | 1 |
| C Las Vegas | 1 |
| Henderson and Las vegas | 1 |
| La Vegas | 2 |
| Lake Las Vegas | 1 |
| Las Vegas | 5 |
| las vegas | 8 |
| las Vegas | 1 |
| Las vegas | 14 |
| Las Vegas | 29370 |
| LAS VEGAS | 19 |
| Las Vegas | 4 |
| Las Vegas & Henderson | 1 |
| Las Vegas East | 1 |
| Las Vegas Nevada | 1 |
| Las Vegas Nv | 1 |
| Las Vegas NV | 1 |
| Las Vegas Valley | 1 |
| Las Vegas, | 1 |
| Las Vegas, NV | 2 |
| Las Vegass | 1 |
| LasVegas | 2 |
| Los Vegas | 1 |
| N E Las Vegas | 1 |
| N Las Vegas | 26 |
| N W Las Vegas | 1 |
| N. Las Vegas | 12 |
| North Las Vegas | 1548 |
| NORTH LAS VEGAS | 1 |
| North Las Vegas | 1 |
| South Las Vegas | 4 |
| Vegas | 1 |

copy(

select distinct city, count(city)

from ybusiness

where city ilike '%vegas%'

group by city)

to 'D:\Filmovi\WW\LasVegas.csv'

with csv delimiter '|' header

where city ilike '%charlotte%'

|  |  |
| --- | --- |
| **city** | **count** |
| charlotte | 3 |
| Charlotte | 9509 |
| CHARLOTTE | 7 |
| Charlotte | 2 |
| North Charlotte | 1 |
| South Charlotte | 1 |

where city ilike '%henderson%' and city not ilike '%ville%'

|  |  |
| --- | --- |
| **city** | **count** |
| Central Henderson | 1 |
| henderson | 3 |
| Henderson | 4892 |
| HENDERSON | 8 |
| Henderson (Green Valley) | 1 |
| Henderson and Las vegas | 1 |
| Henderson Nevada | 1 |
| Las Vegas & Henderson | 1 |
| W Henderson | 1 |

where city ilike '%Calgary%'

|  |  |
| --- | --- |
| **city** | **count** |
| calgary | 1 |
| Calgary | 7736 |
| CALGARY | 1 |
| East Calgary | 1 |
| Northeast Calgary | 1 |
| Northwest Calgary | 3 |
| Southeast Calgary | 3 |
| SW Calgary | 1 |

where city ilike '%mesa%'

|  |  |
| --- | --- |
| **city** | **count** |
| East Mesa | 1 |
| mesa | 3 |
| Mesa | 6080 |
| MESA | 6 |
| Mesa Arizona | 1 |
| Mesa AZ | 1 |

where city ilike '%mont%' and (city ilike '%real%' or city like '%réal%')

|  |  |
| --- | --- |
| **city** | **count** |
| Communauté-Urbaine-de-Montréal | 2 |
| montreal | 1 |
| Montreal | 153 |
| Montréal | 6449 |
| Montreal-Est | 1 |
| Montréal-Est | 1 |
| Montreal-Nord | 7 |
| Montréal-Nord | 3 |
| Montreal-Ouest | 4 |
| Montréal-Ouest | 4 |
| Montreal-Quest | 1 |
| Montreal-West | 3 |
| Montréal-West | 1 |
| Montréal (Québec) | 1 |
| Old Port of Montreal | 1 |
| Vieux-Montréal | 1 |
| West Montréal | 1 |

where city ilike '%phoenix%'

|  |  |
| --- | --- |
| **city** | **count** |
| Metro Phoenix | 1 |
| North Phoenix | 1 |
| phoenix | 9 |
| Phoenix | 18766 |
| PHOENIX | 6 |
| Phoenix | 1 |
| Phoenix AZ | 1 |
| Phoenix Metro Area | 2 |
| Phoenix Valley | 1 |
| Phoenix, | 1 |
| Phoenix, AZ | 2 |

where city ilike '%pittsbur%'

|  |  |
| --- | --- |
| **city** | **count** |
| East Pittsburgh | 6 |
| Pittsburch | 1 |
| Pittsburg | 8 |
| Pittsburgh | 7017 |
| Pittsburgh | 2 |
| St.Pittsburgh | 1 |

where city ilike '%Scottsdale%'

|  |  |
| --- | --- |
| **city** | **count** |
| North Scottsdale | 1 |
| Old Scottsdale | 1 |
| Old Town Scottsdale | 2 |
| scottsdale | 5 |
| Scottsdale | 8837 |
| SCOTTSDALE | 3 |
| Scottsdale, AZ | 1 |
| Westworld Scottsdale | 1 |

where city ilike '%toronto%'

|  |  |
| --- | --- |
| **city** | **count** |
| Downtown Toronto | 4 |
| North Toronto | 1 |
| toronto | 3 |
| Toronto | 18906 |
| TORONTO | 2 |
| Toronto-Etobicoke | 1 |
| Toronto-North York | 1 |
| Toronto-West | 1 |
| Toronto Division | 1 |
| Toronto Scarborough | 1 |
| West Toronto | 1 |