a.

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

.appName('readfile')\

.getOrCreate()

country = spark.read.json('country.json')

country[country.Continent == 'North America'][['Name']].show()



**Rdd**

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

.appName('readfile')\

.getOrCreate()

country = spark.read.json('country.json')

ans = country.rdd.filter(lambda r: r['Continent'] == 'North America').map(lambda r:r['Name']).collect()

for item in ans:

print(item)

output:

Aruba

Anguilla

Netherlands Antilles

Antigua and Barbuda

Bahamas

Belize

Bermuda

Barbados

Canada

Costa Rica

Cuba

Cayman Islands

Dominica

Dominican Republic

Guadeloupe

Grenada

Greenland

Guatemala

Honduras

Haiti

Jamaica

Saint Kitts and Nevis

Saint Lucia

Mexico

Montserrat

Martinique

Nicaragua

Panama

Puerto Rico

El Salvador

Saint Pierre and Miquelon

Turks and Caicos Islands

Trinidad and Tobago

United States

Saint Vincent and the Grenadines

Virgin Islands, British

Virgin Islands, U.S.

b.

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

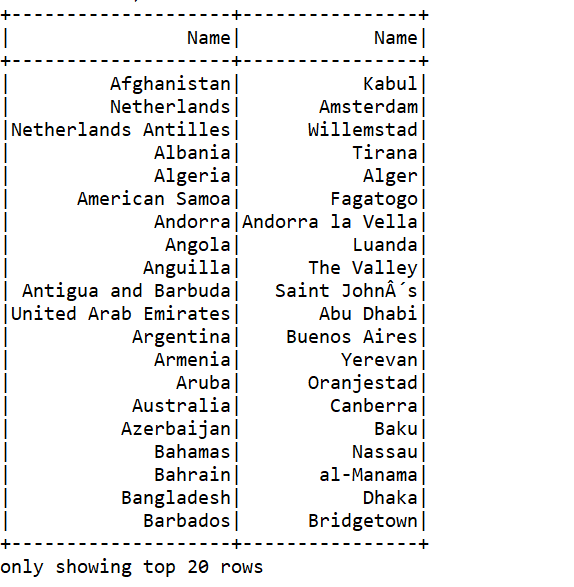
.appName('readfile')\

.getOrCreate()

country = spark.read.json('country.json')

city = spark.read.json('city.json')

country.join(city, country.Capital== city.ID).select(country['Name'],city['Name']).show()



**Rdd:**

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

.appName('readfile')\

.getOrCreate()

country = spark.read.json('country.json')

city = spark.read.json('city.json')

country\_rdd = country.rdd.map(lambda r:(r['Capital'],r['Name']))

city\_rdd = city.rdd.map(lambda r:(r['ID'],r['Name']))

ans = country\_rdd.join(city\_rdd).map(lambda r:r[1]).collect()

for item in ans:

print(item[0]+' '+item[1])

output:

Angola Luanda

Anguilla The Valley

Albania Tirana

Armenia Yerevan

American Samoa Fagatogo

Azerbaijan Baku

Burundi Bujumbura

Bangladesh Dhaka

Bahamas Nassau

Belarus Minsk

Bolivia La Paz

Barbados Bridgetown

Brunei Bandar Seri Begawan

Bhutan Thimphu

Botswana Gaborone

Canada Ottawa

Switzerland Bern

Chile Santiago de Chile

CÃ´te dÂ’Ivoire Yamoussoukro

Cameroon YaoundÃ©

Congo, The Democratic Republic of the Kinshasa

Congo Brazzaville

Costa Rica San JosÃ©

Cyprus Nicosia

Germany Berlin

Dominica Roseau

Ecuador Quito

Egypt Cairo

Eritrea Asmara

Ethiopia Addis Abeba

Finland Helsinki [Helsingfors]

Fiji Islands Suva

France Paris

Gabon Libreville

United Kingdom London

Ghana Accra

Guinea Conakry

Gambia Banjul

Equatorial Guinea Malabo

Grenada Saint GeorgeÂ´s

Guatemala Ciudad de Guatemala

French Guiana Cayenne

Guyana Georgetown

Iran Teheran

Israel Jerusalem

Italy Roma

Jamaica Kingston

Jordan Amman

Japan Tokyo

Kazakstan Astana

Cambodia Phnom Penh

Kiribati Bairiki

Saint Kitts and Nevis Basseterre

Laos Vientiane

Lebanon Beirut

Liberia Monrovia

Liechtenstein Vaduz

Luxembourg Luxembourg [Luxemburg/LÃ«tzebuerg]

Latvia Riga

Macao Macao

Morocco Rabat

Moldova Chisinau

Macedonia Skopje

Mali Bamako

Malta Valletta

Myanmar Rangoon (Yangon)

Mongolia Ulan Bator

Mozambique Maputo

Martinique Fort-de-France

Malawi Lilongwe

Malaysia Kuala Lumpur

Mayotte Mamoutzou

Namibia Windhoek

Niger Niamey

Norfolk Island Kingston

Nigeria Abuja

Nicaragua Managua

Nauru Yaren

Panama Ciudad de PanamÃ¡

Pitcairn Adamstown

Peru Lima

Philippines Manila

Papua New Guinea Port Moresby

Poland Warszawa

North Korea Pyongyang

Portugal Lisboa

Palestine Gaza

French Polynesia Papeete

Romania Bucuresti

Russian Federation Moscow

Senegal Dakar

Singapore Singapore

Svalbard and Jan Mayen Longyearbyen

Somalia Mogadishu

Sao Tome and Principe SÃ£o TomÃ©

Slovenia Ljubljana

Sweden Stockholm

Swaziland Mbabane

Seychelles Victoria

Syria Damascus

Togo LomÃ©

Thailand Bangkok

East Timor Dili

Tonga NukuÂ´alofa

Trinidad and Tobago Port-of-Spain

Turkey Ankara

Tuvalu Funafuti

Tanzania Dodoma

Ukraine Kyiv

Uruguay Montevideo

Holy See (Vatican City State) CittÃ del Vaticano

Saint Vincent and the Grenadines Kingstown

Vietnam Hanoi

Wallis and Futuna Mata-Utu

Yemen Sanaa

Yugoslavia Beograd

South Africa Pretoria

Zambia Lusaka

Zimbabwe Harare

Aruba Oranjestad

Afghanistan Kabul

Andorra Andorra la Vella

Netherlands Antilles Willemstad

United Arab Emirates Abu Dhabi

Argentina Buenos Aires

Antigua and Barbuda Saint JohnÂ´s

Australia Canberra

Austria Wien

Belgium Bruxelles [Brussel]

Benin Porto-Novo

Burkina Faso Ouagadougou

Bulgaria Sofija

Bahrain al-Manama

Bosnia and Herzegovina Sarajevo

Belize Belmopan

Bermuda Hamilton

Brazil BrasÃ­lia

Central African Republic Bangui

Cocos (Keeling) Islands West Island

China Peking

Cook Islands Avarua

Colombia SantafÃ© de BogotÃ¡

Comoros Moroni

Cape Verde Praia

Cuba La Habana

Christmas Island Flying Fish Cove

Cayman Islands George Town

Czech Republic Praha

Djibouti Djibouti

Denmark KÃ¸benhavn

Dominican Republic Santo Domingo de GuzmÃ¡n

Algeria Alger

Western Sahara El-AaiÃºn

Spain Madrid

Estonia Tallinn

Falkland Islands Stanley

Faroe Islands TÃ³rshavn

Micronesia, Federated States of Palikir

Georgia Tbilisi

Gibraltar Gibraltar

Guadeloupe Basse-Terre

Guinea-Bissau Bissau

Greece Athenai

Greenland Nuuk

Guam AgaÃ±a

Hong Kong Victoria

Honduras Tegucigalpa

Croatia Zagreb

Haiti Port-au-Prince

Hungary Budapest

Indonesia Jakarta

India New Delhi

Ireland Dublin

Iraq Baghdad

Iceland ReykjavÃ­k

Kenya Nairobi

Kyrgyzstan Bishkek

South Korea Seoul

Kuwait Kuwait

Libyan Arab Jamahiriya Tripoli

Saint Lucia Castries

Sri Lanka Colombo

Lesotho Maseru

Lithuania Vilnius

Monaco Monaco-Ville

Madagascar Antananarivo

Maldives Male

Mexico Ciudad de MÃ©xico

Marshall Islands Dalap-Uliga-Darrit

Northern Mariana Islands Garapan

Mauritania Nouakchott

Montserrat Plymouth

Mauritius Port-Louis

New Caledonia NoumÃ©a

Niue Alofi

Netherlands Amsterdam

Norway Oslo

Nepal Kathmandu

New Zealand Wellington

Oman Masqat

Pakistan Islamabad

Palau Koror

Puerto Rico San Juan

Paraguay AsunciÃ³n

Qatar Doha

RÃ©union Saint-Denis

Rwanda Kigali

Saudi Arabia Riyadh

Sudan Khartum

Saint Helena Jamestown

Solomon Islands Honiara

Sierra Leone Freetown

El Salvador San Salvador

San Marino San Marino

Saint Pierre and Miquelon Saint-Pierre

Suriname Paramaribo

Slovakia Bratislava

Turks and Caicos Islands Cockburn Town

Chad NÂ´DjamÃ©na

Tajikistan Dushanbe

Tokelau Fakaofo

Turkmenistan Ashgabat

Tunisia Tunis

Taiwan Taipei

Uganda Kampala

United States Washington

Uzbekistan Toskent

Venezuela Caracas

Virgin Islands, British Road Town

Virgin Islands, U.S. Charlotte Amalie

Vanuatu Port-Vila

Samoa Apia

c.

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

.appName('readfile')\

.getOrCreate()

country = spark.read.json('country.json')

country[['Continent']].distinct().show()



d.

from pyspark.sql import SparkSession

spark = SparkSession\

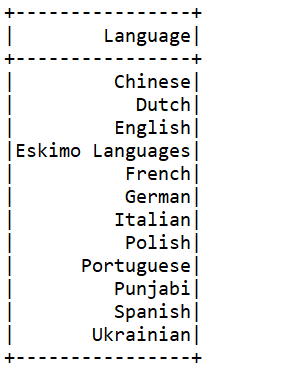
.builder\

.appName('readfile')\

.getOrCreate()

countrylanguage = spark.read.json('countrylanguage.json')

countrylanguage[countrylanguage.CountryCode == 'CAN'][['Language']].show()



e.

from pyspark.sql import SparkSession

spark = SparkSession\

.builder\

.appName('readfile')\

.getOrCreate()

import pyspark.sql.functions as fc

country = spark.read.json('country.json')

country.groupBy('Continent').agg(fc.avg("LifeExpectancy").alias("avg\_le"),fc.count("\*").alias("cnt")).filter('cnt>=20').orderBy(fc.desc('cnt')).limit(1).select('Continent','avg\_le').show()

