**API**

World Bank API was used in order to get Population, GDP and Unemployment rates among countries.

Please refer below to see the breakdown of the World Bank API

for single indicator:

domain country *(all countries)* date

<http://api.worldbank.org/v2/country/all/indicator/SL.UEM.TOTL.ZS?date=2019&format=json>

scheme version 2 of indicator *(unemployment - % of* format *(JSON)*

the Indicators API *total labor force(modeled ILO estimate)*

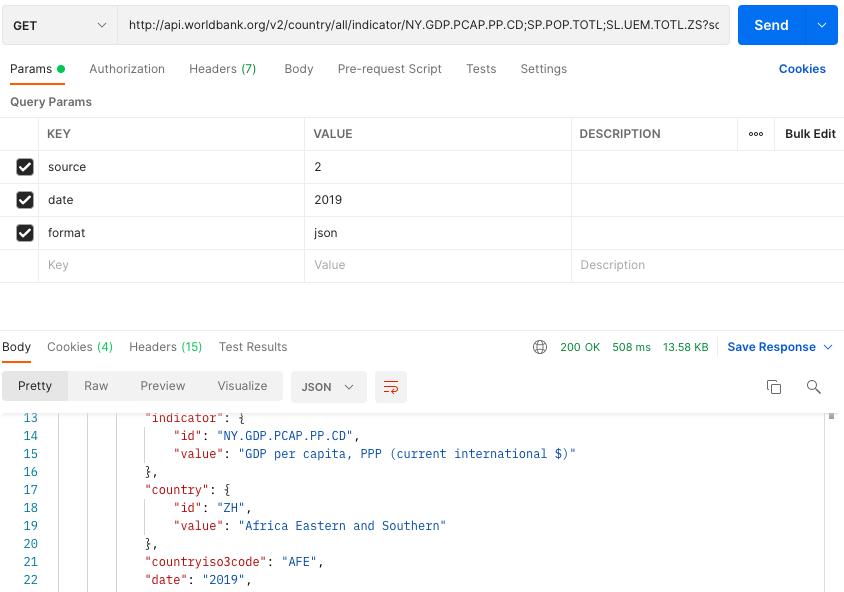
for multiple indicators:

indicators *(GDP per capita, total population)*

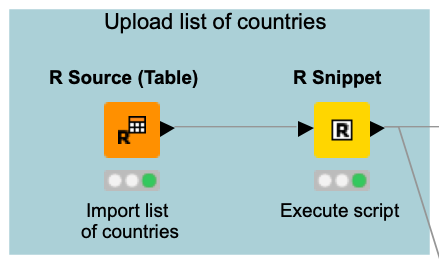
<http://api.worldbank.org/v2/country/all/indicator/NY.GDP.PCAP.PP.CD;SP.POP.TOTL?source=2&date=2019&format=json>

dataset source

We used Postman to check the validity of request. Please refer below for Postman result:



In order to build our data pipeline in KNIME we firstly loaded list of countries using R script (WDI package). Grouped rows, such as European Union, Arab World, East Asia and others were excluded from the list of countries using R script. After importing countries to KNIME with R Source node, R Snippet node was used in order to execute that R script.



Further we wrote the URL for World Bank API using String Manipulation, which is join("http://api.worldbank.org/v2/country/",$knime.in.code$,"/indicator/SL.UEM.TOTL.ZS?format=json&date=2019") for unemployment

and

join("http://api.worldbank.org/v2/country/",$knime.in.code$,"/indicator/NY.GDP.PCAP.CD;SP.POP.TOTL?source=2&format=json&date=2019") for GDP and Population.

The URLs for each country were added in new column, we send a get request, extracted values of indicators using JSON Path and converted JSON into table. We received data table, the next step was to clean that data. We filtered out all rows with missing values and unnecessary columns, remaining with country name, country code and indicators.

