Data Collection for NASA NEOs Corsework

Surya Ramesh

12/02/2021

Contents

0.1 Get data from API

- assemble the url with start and end dates as modifiable.
- check for any http error and set error message in case it does
- Save the JSON content of response as text for readability
- Format the content - flatten, stack the data frame inside and merge with content,remove the original data frame
- Write to CSV for sharing

```
# modify the start_daterange and end_daterange to the dates required
#smaller range shown here

start_daterange <- as.Date("2021-02-12", "%Y-%m-%d")
end_daterange <- as.Date("2021-02-18", "%Y-%m-%d")

date <- start_daterange

MycombinedData <- data.frame() # empty dataframe for the data

# Loop to get data from API and format it.

while (date <= end_daterange)

{
  url <- paste(
    "https://api.nasa.gov/neo/rest/v1/feed?start_date=",
    format(date, "%Y-%m-%d"),"&end_date=",format(date + 7, "%Y-%m-%d"),
    "&detailed=true&api_key=N0spzeNK1l22GJ1u8mAKqm0XHDE5dAbhhGvWRrFo",
    sep = "",collapse = NULL)


  #print(url)                                # for troubleshooting

  response <- GET(url)                      # the GET command for API

  #Check for http error response from API

  if (http_error(response))
```

```
  {

    stop("The API request failed")
  }

  # Return the response's content which contains the data needed

  #content(response)                        # for troubleshooting


  # Parse returned content as text

  my_json <- content(response, as = 'text')


  #write to individual JSON files for sharing

  write(my_json, paste(format(date, "%Y-%m-%d"), "_NASANeoWS.json"))

  # Format the data

  my_json_flat <- fromJSON(my_json, flatten = TRUE)

  NEOs <- my_json_flat$near_earth_objects

  NEOs_list <- list.stack(NEOs)

  close_approach_data_list <-list.stack(NEOs_list$close_approach_data)

  #merge NEOs_list list and flattened close_approach_data_list

  merged_df_original <-data.frame(NEOs_list, close_approach_data_list)

  #remove original close_approach_data column

  merged_df_original$close_approach_data <- NULL

  # print(nrow(merged_df_original)) # used for troubleshooting

  # bind the rows to MycombinedData

  MycombinedData <- rbind(MycombinedData, merged_df_original)


  date <- date + 7

  #Sys.sleep(120)     # sys sleeps for 120secs to avoid too many requests
}
```

0.2 Write to CSV file

```
write.csv(MycombinedData , "MyNASAData.csv")
```

```
#nrow(MycombinedData)  # used for troubleshooting
#View(MycombinedData)  # used for troubleshooting
```

End of Data Collection section