Reading JSON

What is JSON?

- Javascript Object Notation

- Format of Data used for many APIs

- Data can be stored as
  - Numbers
  - Strings
  - Arrays
  - Objects

JSON Data

- JSON data is written as name/value pairs.

- A name/value pair consists of a field name (in double quotes), followed by a colon, followed by a value:

{ "name":"C-DAC" }

- Valid data types:
  - a string
  - a number
  - an object (JSON object)
  - an array
  - a boolean
  - null

Resemblance of XML & JSON by syntax

XML

```
<employees>
  <employee>
    <firstName>John</firstName>
          <lastName>Doe</lastName>
  </employee>
  <employee>
    <firstName>Anna</firstName>
          <lastName>Smith</lastName>
  </employee>
  <employee>
    <firstName>Peter</firstName>
          <lastName>Jones</lastName>
  </employee>
</employees>
```

JSON

```
{"employees":[
   { "firstName":"John",
    "lastName":"Doe" },
   { "firstName":"Anna",
    "lastName":"Smith" },
   { "firstName":"Peter",
    "lastName":"Jones" }
]}
```

How is JSON better than XML?

- JSON is shorter

- JSON can use arrays

- XML has to be parsed by XML parser whereas JSON can be parsed using a standard Javascript function

Reading JSON

- There can be two functions namely read_json and fromJSON to read the JSON data into R

Syntax : fromJSON(txt, simplifyVector = TRUE,...)

Where

txt : a JSON string, URL or file

simplifyVector : simplifies nested lists into vectors

Example

```json
{
    "firstName": "John",
    "lastName": "Smith",
    "age": 25,
    "address":
    {
        "streetAddress": "21 2nd Street",
        "city": "New York",
        "state": "NY",
        "postalCode": "10021"
    },
    "phoneNumber":
    [
        {
            "type": "home",
            "number": "212 555-1234"
        },
        {
            "type": "fax",
            "number": "646 555-4567"|
        }
    ]
}
```

```
> library(jsonlite)
> jsonData <- fromJSON("D:/Data Science Training – Level III/Datasets/contacts.txt")
> class(jsonData)
[1] "list"
> names(jsonData)
[1] "firstName"   "lastName"    "age"         "address"     "phoneNumber"
>
> #nested objects
> jsonData$phoneNumber
  type       number
1 home 212 555-1234
2  fax 646 555-4567
> jsonData$phoneNumber$number
[1] "212 555-1234" "646 555-4567"
```
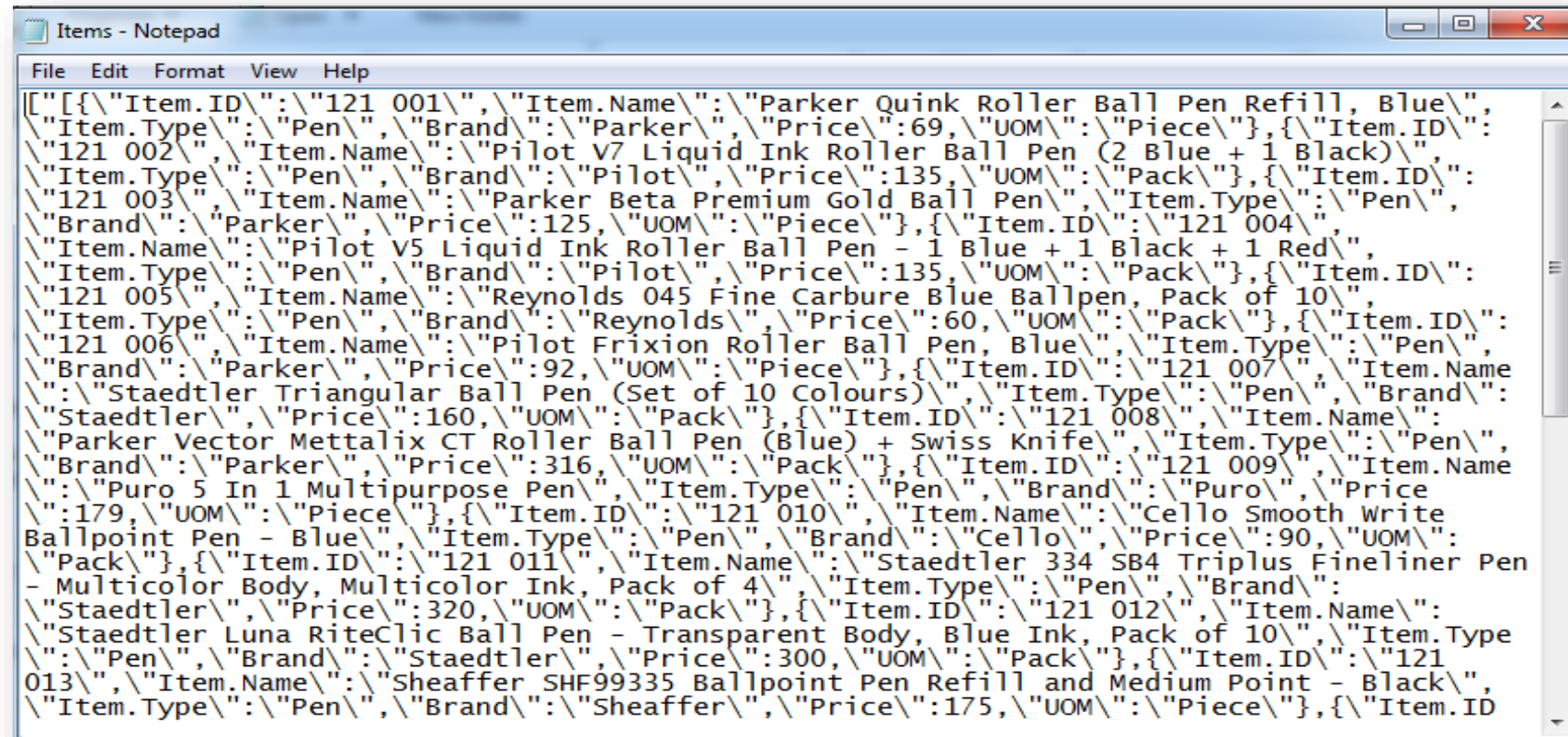
Reading JSON

Syntax : read_json(path, simplifyVector, …)

Where

      path : file on disk

      simplifyVector : simplifies nested lists into vectors and   data frames

Example

```
["[{\"Item.ID\":\"121 001\",\"Item.Name\":\"Parker Quink Roller Ball Pen Refill, Blue\",
\"Item.Type\":\"Pen\",\"Brand\":\"Parker\",\"Price\":69,\"UOM\":\"Piece\"},{\"Item.ID\":
\"121 002\",\"Item.Name\":\"Pilot V7 Liquid Ink Roller Ball Pen (2 Blue + 1 Black)\",
\"Item.Type\":\"Pen\",\"Brand\":\"Pilot\",\"Price\":135,\"UOM\":\"Pack\"},{\"Item.ID\":
\"121 003\",\"Item.Name\":\"Parker Beta Premium Gold Ball Pen\",\"Item.Type\":\"Pen\",
\"Brand\":\"Parker\",\"Price\":125,\"UOM\":\"Piece\"},{\"Item.ID\":\"121 004\",
\"Item.Name\":\"Pilot V5 Liquid Ink Roller Ball Pen - 1 Blue + 1 Black + 1 Red\",
\"Item.Type\":\"Pen\",\"Brand\":\"Pilot\",\"Price\":135,\"UOM\":\"Pack\"},{\"Item.ID\":
\"121 005\",\"Item.Name\":\"Reynolds 045 Fine Carbure Blue Ballpen, Pack of 10\",
\"Item.Type\":\"Pen\",\"Brand\":\"Reynolds\",\"Price\":60,\"UOM\":\"Pack\"},{\"Item.ID\":
\"121 006\",\"Item.Name\":\"Pilot Frixion Roller Ball Pen, Blue\",\"Item.Type\":\"Pen\",
\"Brand\":\"Parker\",\"Price\":92,\"UOM\":\"Piece\"},{\"Item.ID\":\"121 007\",\"Item.Name
\":\"Staedtler Triangular Ball Pen (Set of 10 Colours)\",\"Item.Type\":\"Pen\",\"Brand\":
\"Staedtler\",\"Price\":160,\"UOM\":\"Pack\"},{\"Item.ID\":\"121 008\",\"Item.Name\":
\"Parker Vector Mettalix CT Roller Ball Pen (Blue) + Swiss Knife\",\"Item.Type\":\"Pen\",
\"Brand\":\"Parker\",\"Price\":316,\"UOM\":\"Pack\"},{\"Item.ID\":\"121 009\",\"Item.Name
\":\"Puro 5 In 1 Multipurpose Pen\",\"Item.Type\":\"Pen\",\"Brand\":\"Puro\",\"Price
\":179,\"UOM\":\"Piece\"},{\"Item.ID\":\"121 010\",\"Item.Name\":\"Cello Smooth Write
Ballpoint Pen - Blue\",\"Item.Type\":\"Pen\",\"Brand\":\"Cello\",\"Price\":90,\"UOM\":
\"Pack\"},{\"Item.ID\":\"121 011\",\"Item.Name\":\"Staedtler 334 SB4 Triplus Fineliner Pen
- Multicolor Body, Multicolor Ink, Pack of 4\",\"Item.Type\":\"Pen\",\"Brand\":
\"Staedtler\",\"Price\":320,\"UOM\":\"Pack\"},{\"Item.ID\":\"121 012\",\"Item.Name\":
\"Staedtler Luna RiteClic Ball Pen - Transparent Body, Blue Ink, Pack of 10\",\"Item.Type
\":\"Pen\",\"Brand\":\"Staedtler\",\"Price\":300,\"UOM\":\"Pack\"},{\"Item.ID\":\"121
013\",\"Item.Name\":\"Sheaffer SHF99335 Ballpoint Pen Refill and Medium Point - Black\",
\"Item.Type\":\"Pen\",\"Brand\":\"Sheaffer\",\"Price\":175,\"UOM\":\"Piece\"},{\"Item.ID
```

```
> items_json <- read_json("D:/Data Science Training - Level III/Datasets/Items.json",
+                         simplifyVector = T)
> class(items_json)
[1] "character"
> items <- fromJSON(items_json)
> class(items)
[1] "data.frame"
```