Name :   Eshaan R. Joshi
Roll Number :   18007
Subject :   Statistical Computing using R & Python

Reading a CSV File

Reading the file using the `read.csv()` function

```
Reading a CSV File

[1]: data <- read.csv('monthly_crude_oil_processed.csv',header=TRUE,sep=',')
```

Output

```
[2]: print(head(data))
        Month Year         OIL.COMPANIES Quantity..000.Metric.Tonnes. last_updated
1    February 2023              CPCL-TOTAL                     927.88   2023-08-19
2    February 2023 NRL-NUMALIGARH, ASSAM                      244.06   2023-08-19
3       March 2023 IOCL-PANIPAT, HARYANA                     1297.96   2023-08-19
4       March 2023    IOCL-GUWAHATI,ASSAM                      96.81   2023-08-19
5       March 2023    IOCL-DIGBOI,ASSAM                        59.40   2023-08-19
6       March 2023              BORL-BINA                      735.71   2023-08-19
```

```
[3]: head(data)
```

A data.frame: 6 × 5

	Month	Year	OIL.COMPANIES	Quantity..000.Metric.Tonnes.	last_updated
	<chr>	<int>	<chr>	<dbl>	<chr>
1	February	2023	CPCL-TOTAL	927.88	2023-08-19
2	February	2023	NRL-NUMALIGARH, ASSAM	244.06	2023-08-19
3	March	2023	IOCL-PANIPAT, HARYANA	1297.96	2023-08-19
4	March	2023	IOCL-GUWAHATI,ASSAM	96.81	2023-08-19
5	March	2023	IOCL-DIGBOI,ASSAM	59.40	2023-08-19
6	March	2023	BORL-BINA	735.71	2023-08-19

Reading an Excel file

Reading the file using the `read.xlsx()` function from the `xlsx` library

```
Reading an Excel file (xlsx)

[4]: library(xlsx)

[5]: data_xl <- read.xlsx("importexport202223.xlsx",sheetIndex = 1)
```

Output

```
[6]: print(head(data_xl))
     IMPORT.EXPORT..Quantity.in..000.Metric.Tonnes.  PRODUCT APRIL    MAY   JUNE
   1                                                  IMPORT CRUDE OIL 21626 19644 19441
   2                                                  IMPORT       LPG  1605  1363  1264
   3                                                  IMPORT        MS     0    30   127
   4                                                  IMPORT   Naphtha    35    30     2
   5                                                  IMPORT       ATF     0     0     0
   6                                                  IMPORT       SKO     0     0     0
     JULY AUGUST SEPTEMBER OCTOBER NOVEMBER DECEMBER JANUARY FEBRUARY MARCH
   1 20624  17637     16772   18123    19003    19618   20058    19285 20729
   2  1417   1574      1448    1374     1778     1718    1709     1647  1410
   3    63      0       190     327      211      120       0        0     0
   4   145     90        76     135       28       88      67      121   117
   5     0      0         0       0        0        0       0        0     0
   6     0      0         0       0        0        0       0        0     0
     TOTAL
   1 232561
   2  18309
   3   1069
   4    936
   5      0
   6      0
```

```
[7]: head(data_xl)
```

A data.frame: 6 × 15

IMPORT.EXPORT..Quantity.in..000.Metric.Tonnes.	PRODUCT	APRIL	MAY	JUNE	JULY	AUGUST	SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER	JANUARY	FEBRUARY	MARCH	TOTAL
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	IMPORT	CRUDE OIL	21626	19644	19441	20624	17637	16772	18123	19003	19618	20058	19285	20729	232561
2	IMPORT	LPG	1605	1363	1264	1417	1574	1448	1374	1778	1718	1709	1647	1410	18309
3	IMPORT	MS	0	30	127	63	0	190	327	211	120	0	0	0	1069
4	IMPORT	Naphtha	35	30	2	145	90	76	135	28	88	67	121	117	936
5	IMPORT	ATF	0	0	0	0	0	0	0	0	0	0	0	0	0
6	IMPORT	SKO	0	0	0	0	0	0	0	0	0	0	0	0	0

Reading a SPSS file (.sav)

Reading the file using the `read.spss()` function from the `foreign` library

```
Reading a SPSS data file .sav

[8]: library(foreign)

[9]: data_sps <- read.spss('Orders.sav',to.data.frame = TRUE)
```

Output

```
[10]: print(head(data_sps))

      Order_ID Employee_ID Customer_ID  Order_Date Shipped_Date Shipper_ID
    1       30           9          27 13356662400  13357267200          2
    2       31           3           4 13357094400  13357267200          1
    3       32           4          12 13357267200  13357267200          2
    4       33           6           8 13357958400  13358044800          3
    5       34           9           4 13358563200  13358649600          3
    6       35           3          29 13358908800  13359081600          2
                    Ship_Name     Ship_Address    Ship_City Ship_State_Province
    1 Karen Toh             789 27th Street  Las Vegas                       NV
    2 Christina Lee         123 4th Street   New York                        NY
    3 John Edwards          123 12th Street  Las Vegas                       NV
    4 Elizabeth Andersen    123 8th Street   Portland                        OR
    5 Christina Lee         123 4th Street   New York                        NY
    6 Soo Jung Lee          789 29th Street  Denver                          CO
      Ship_ZIP_Postal_Code Ship_Country_Region Shipping_Fee Taxes Payment_Type
    1                99999                 USA          200     0  Check
    2                99999                 USA            5     0  Credit Card
    3                99999                 USA            5     0  Credit Card
    4                99999                 USA           50     0  Credit Card
    5                99999                 USA            4     0  Check
    6                99999                 USA            7     0  Check
        Paid_Date Notes Tax_Rate Tax_Status Status_ID
    1 13356662400            0         NA         3
    2 13357094400            0         NA         3
    3 13357267200            0         NA         3
    4 13357958400            0         NA         3
    5 13358563200            0         NA         3
    6 13358908800            0         NA         3
```

```
[11]: head(data_sps)
```

A data.frame: 6 × 20

	Order_ID	Employee_ID	Customer_ID	Order_Date	Shipped_Date	Shipper_ID	Ship_Name	Ship_Address	Ship_City	Ship_State_Province	Ship_ZIP_Postal_Code	Ship_Country_Region	Shipping_Fee	Taxes	Payment_Type	Paid_Date	Notes	Tax_Rate	Tax_Status	Status_ID
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<dbl>	<dbl>	<chr>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
1	30	9	27	13356662400	13357267200	2	Karen Toh	789 27th Street	Las Vegas	NV	99999	USA	200	0	Check	13356662400		0	NA	3
2	31	3	4	13357094400	13357267200	1	Christina Lee	123 4th Street	New York	NY	99999	USA	5	0	Credit Card	13357094400		0	NA	3
3	32	4	12	13357267200	13357267200	2	John Edwards	123 12th Street	Las Vegas	NV	99999	USA	5	0	Credit Card	13357267200		0	NA	3
4	33	6	8	13357958400	13358044800	3	Elizabeth Andersen	123 8th Street	Portland	OR	99999	USA	50	0	Credit Card	13357958400		0	NA	3
5	34	9	4	13358563200	13358649600	3	Christina Lee	123 4th Street	New York	NY	99999	USA	4	0	Check	13358563200		0	NA	3
6	35	3	29	13358908800	13359081600	2	Soo Jung Lee	789 29th Street	Denver	CO	99999	USA	7	0	Check	13358908800		0	NA	3