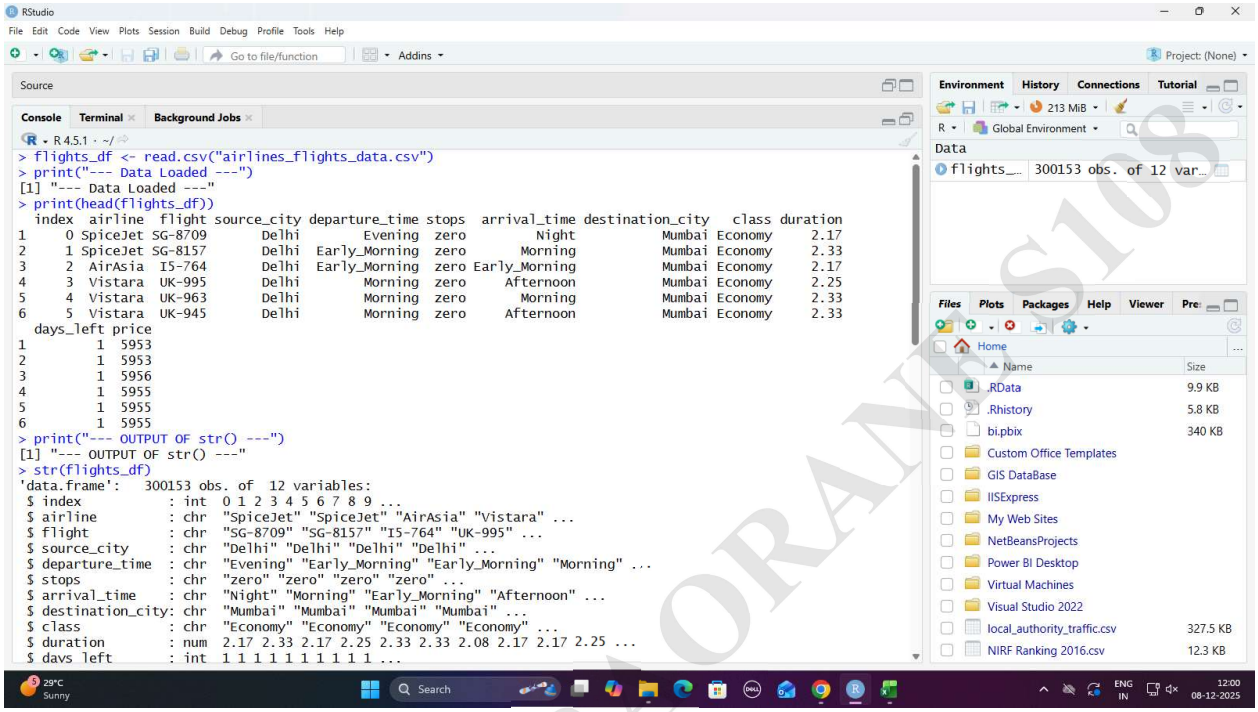
SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

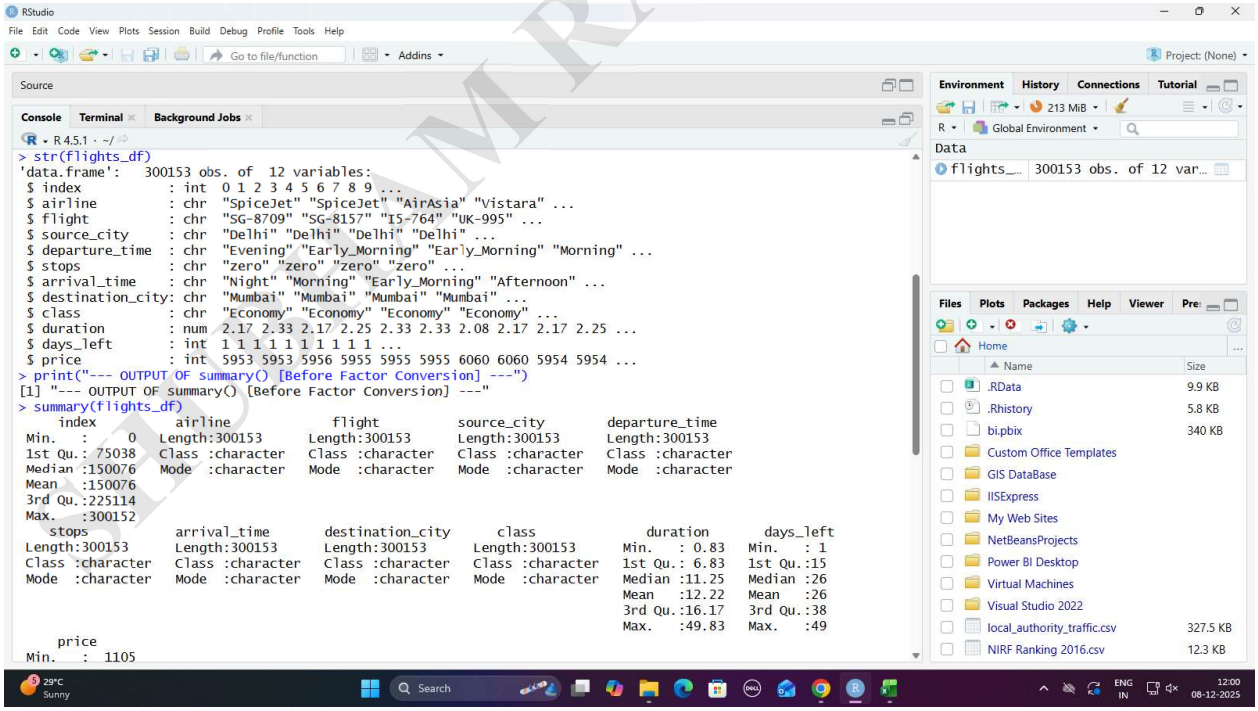
SUBJECT : Data Analysis with SAS / SPSS /R

AIM : Generating basic summaries using str() or summary() (R).

OUTPUT



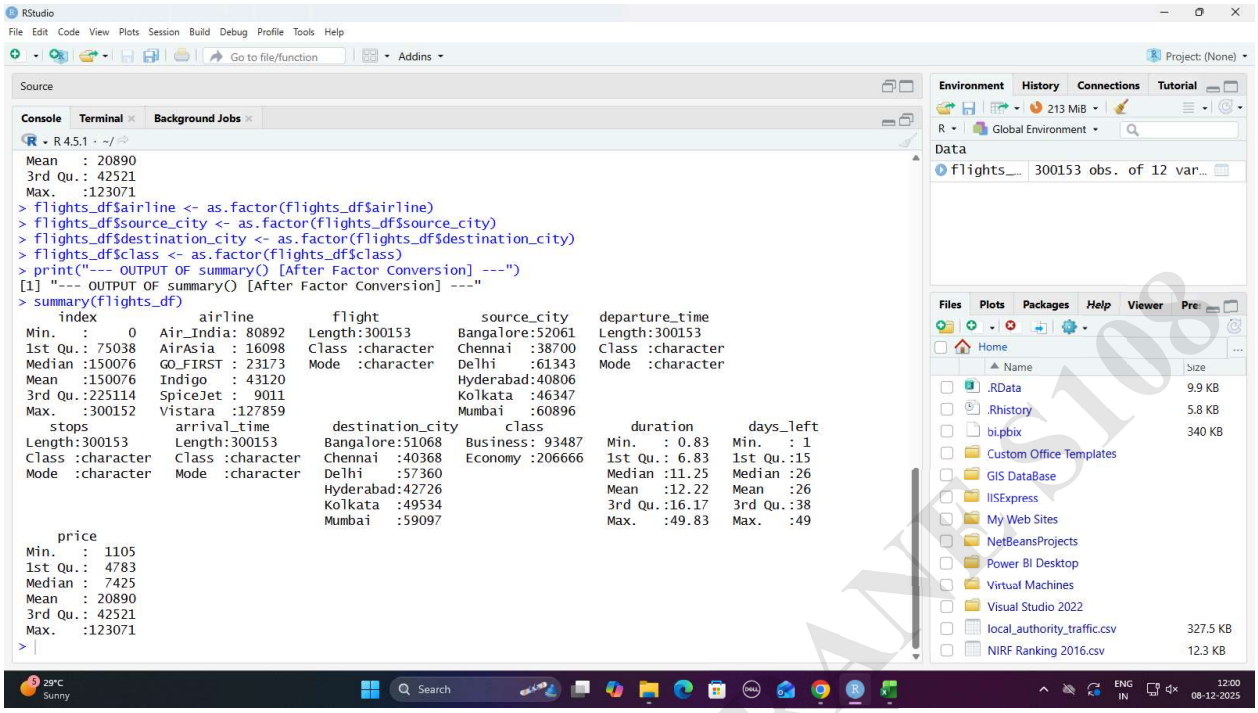
```
> flights_df <- read.csv("airlines_flights_data.csv")
> print("--- Data Loaded ---")
[1] "--- Data Loaded ---"
> print(head(flights_df))
  index  airline  flight source_city departure_time stops  arrival_time destination_city   class duration
1     0 SpiceJet SG-8709       Delhi        Evening  zero         Night           Mumbai Economy     2.17
2     1 SpiceJet SG-8157       Delhi  Early_Morning  zero       Morning           Mumbai Economy     2.33
3     2  AirAsia  I5-764       Delhi  Early_Morning  zero Early_Morning           Mumbai Economy     2.17
4     3  Vistara  UK-995       Delhi        Morning  zero     Afternoon           Mumbai Economy     2.25
5     4  Vistara  UK-963       Delhi        Morning  zero       Morning           Mumbai Economy     2.33
6     5  Vistara  UK-945       Delhi        Morning  zero     Afternoon           Mumbai Economy     2.33
  days_left price
1         1  5953
2         1  5953
3         1  5956
4         1  5955
5         1  5955
6         1  5955
> print("--- OUTPUT OF str() ---")
[1] "--- OUTPUT OF str() ---"
> str(flights_df)
'data.frame':   300153 obs. of  12 variables:
 $ index           : int  0 1 2 3 4 5 6 7 8 9 ...
 $ airline         : chr  "SpiceJet" "SpiceJet" "AirAsia" "Vistara" ...
 $ flight          : chr  "SG-8709" "SG-8157" "I5-764" "UK-995" ...
 $ source_city     : chr  "Delhi" "Delhi" "Delhi" "Delhi" ...
 $ departure_time  : chr  "Evening" "Early_Morning" "Early_Morning" "Morning" ...
 $ stops           : chr  "zero" "zero" "zero" "zero" ...
 $ arrival_time    : chr  "Night" "Morning" "Early_Morning" "Afternoon" ...
 $ destination_city: chr  "Mumbai" "Mumbai" "Mumbai" "Mumbai" ...
 $ class           : chr  "Economy" "Economy" "Economy" "Economy" ...
 $ duration        : num  2.17 2.33 2.17 2.25 2.33 2.33 2.08 2.17 2.17 2.25 ...
 $ days_left       : int  1 1 1 1 1 1 1 1 1 1 ...
```



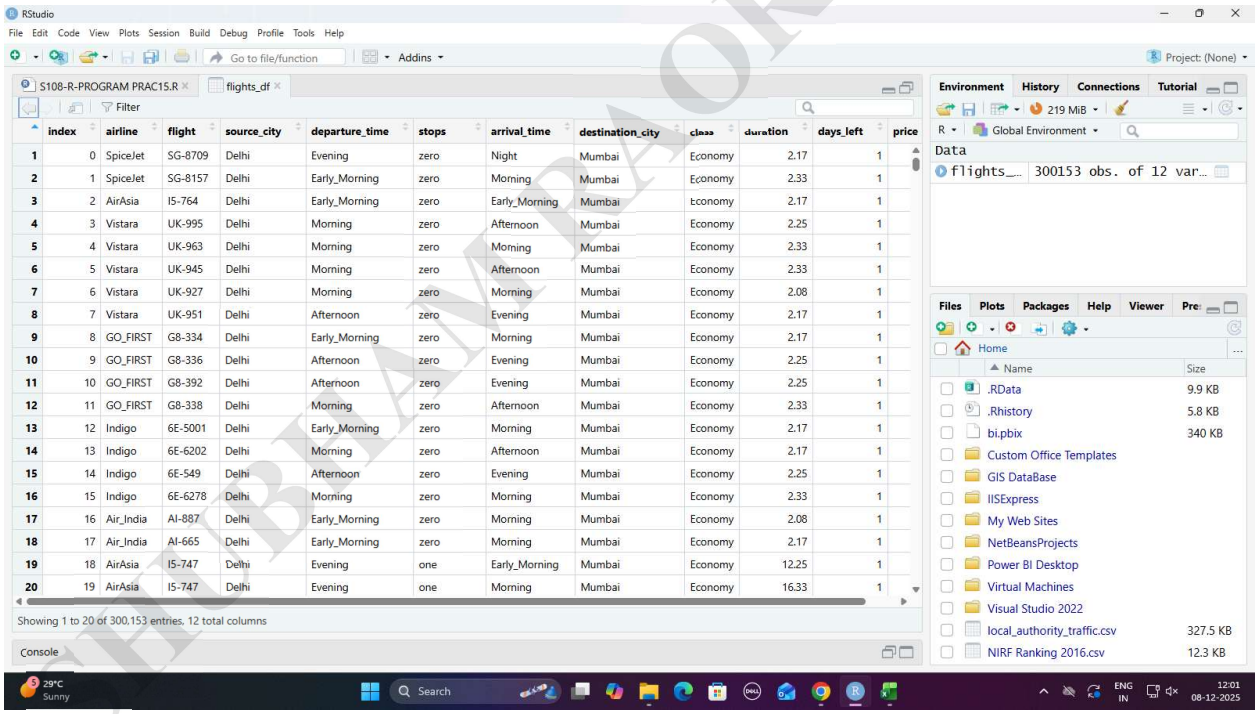
```
> str(flights_df)
'data.frame':   300153 obs. of  12 variables:
 $ index           : int  0 1 2 3 4 5 6 7 8 9 ...
 $ airline         : chr  "SpiceJet" "SpiceJet" "AirAsia" "Vistara" ...
 $ flight          : chr  "SG-8709" "SG-8157" "I5-764" "UK-995" ...
 $ source_city     : chr  "Delhi" "Delhi" "Delhi" "Delhi" ...
 $ departure_time  : chr  "Evening" "Early_Morning" "Early_Morning" "Morning" ...
 $ stops           : chr  "zero" "zero" "zero" "zero" ...
 $ arrival_time    : chr  "Night" "Morning" "Early_Morning" "Afternoon" ...
 $ destination_city: chr  "Mumbai" "Mumbai" "Mumbai" "Mumbai" ...
 $ class           : chr  "Economy" "Economy" "Economy" "Economy" ...
 $ duration        : num  2.17 2.33 2.17 2.25 2.33 2.33 2.08 2.17 2.17 2.25 ...
 $ days_left       : int  1 1 1 1 1 1 1 1 1 1 ...
 $ price           : int  5953 5953 5956 5955 5955 5955 6060 6060 5954 5954 ...
> print("--- OUTPUT OF summary() [Before Factor Conversion] ---")
[1] "--- OUTPUT OF summary() [Before Factor Conversion] ---"
> summary(flights_df)
     index          airline             flight          source_city        departure_time
 Min.   :     0   Length:300153      Length:300153      Length:300153      Length:300153
 1st Qu.: 75038   Class :character   Class :character   Class :character   Class :character
 Median :150076   Mode  :character   Mode  :character   Mode  :character   Mode  :character
 Mean   :150076
 3rd Qu.:225114
 Max.   :300152
   stops          arrival_time       destination_city     class              duration       days_left
 Length:300153      Length:300153      Length:300153      Length:300153      Min.   : 0.83   Min.   : 1
 Class :character   Class :character   Class :character   Class :character   1st Qu.: 6.83   1st Qu.:15
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Median :11.25   Median :26
                                                                             Mean   :12.22   Mean   :26
                                                                             3rd Qu.:16.17   3rd Qu.:38
                                                                             Max.   :49.83   Max.   :49
     price
 Min.   :  1105
```

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS /R



```
 Mean   : 20890
 3rd Qu.: 42521
 Max.   :123071
> flights_df$airline <- as.factor(flights_df$airline)
> flights_df$source_city <- as.factor(flights_df$source_city)
> flights_df$destination_city <- as.factor(flights_df$destination_city)
> flights_df$class <- as.factor(flights_df$class)
> print("--- OUTPUT OF summary() [After Factor Conversion] ---")
[1] "--- OUTPUT OF summary() [After Factor Conversion] ---"
> summary(flights_df)
     index           airline         flight           source_city    departure_time
 Min.   :     0   Air_India: 80892   Length:300153      Bangalore:52061   Length:300153
 1st Qu.: 75038   AirAsia  : 16098   Class :character   Chennai  :38700   Class :character
 Median :150076   GO_FIRST : 23173   Mode  :character   Delhi    :61343   Mode  :character
 Mean   :150076   Indigo   : 43120                      Hyderabad:40806
 3rd Qu.:225114   SpiceJet :  9011                      Kolkata  :46347
 Max.   :300152   Vistara  :127859                      Mumbai   :60896
    stops           arrival_time       destination_city      class          duration       days_left
 Length:300153      Length:300153      Bangalore:51068   Business: 93487   Min.   : 0.83   Min.   : 1
 Class :character   Class :character   Chennai  :40368   Economy :206666   1st Qu.: 6.83   1st Qu.:15
 Mode  :character   Mode  :character   Delhi    :57360                     Median :11.25   Median :26
                                       Hyderabad:42726                     Mean   :12.22   Mean   :26
                                       Kolkata  :49534                     3rd Qu.:16.17   3rd Qu.:38
                                       Mumbai   :59097                     Max.   :49.83   Max.   :49
     price
 Min.   :  1105
 1st Qu.:  4783
 Median :  7425
 Mean   : 20890
 3rd Qu.: 42521
 Max.   :123071
>
```



	index	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	duration	days_left	price
1	0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2.17	1	
2	1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.33	1	
3	2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2.17	1	
4	3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2.25	1	
5	4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai	Economy	2.33	1	
6	5	Vistara	UK-945	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2.33	1	
7	6	Vistara	UK-927	Delhi	Morning	zero	Morning	Mumbai	Economy	2.08	1	
8	7	Vistara	UK-951	Delhi	Afternoon	zero	Evening	Mumbai	Economy	2.17	1	
9	8	GO_FIRST	G8-334	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.17	1	
10	9	GO_FIRST	G8-336	Delhi	Afternoon	zero	Evening	Mumbai	Economy	2.25	1	
11	10	GO_FIRST	G8-392	Delhi	Afternoon	zero	Evening	Mumbai	Economy	2.25	1	
12	11	GO_FIRST	G8-338	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2.33	1	
13	12	Indigo	6E-5001	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.17	1	
14	13	Indigo	6E-6202	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2.17	1	
15	14	Indigo	6E-549	Delhi	Afternoon	zero	Evening	Mumbai	Economy	2.25	1	
16	15	Indigo	6E-6278	Delhi	Morning	zero	Morning	Mumbai	Economy	2.33	1	
17	16	Air_India	AI-887	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.08	1	
18	17	Air_India	AI-665	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.17	1	
19	18	AirAsia	I5-747	Delhi	Evening	one	Early_Morning	Mumbai	Economy	12.25	1	
20	19	AirAsia	I5-747	Delhi	Evening	one	Morning	Mumbai	Economy	16.33	1	

Showing 1 to 20 of 300,153 entries, 12 total columns

SHUBHAM RAORANE S108