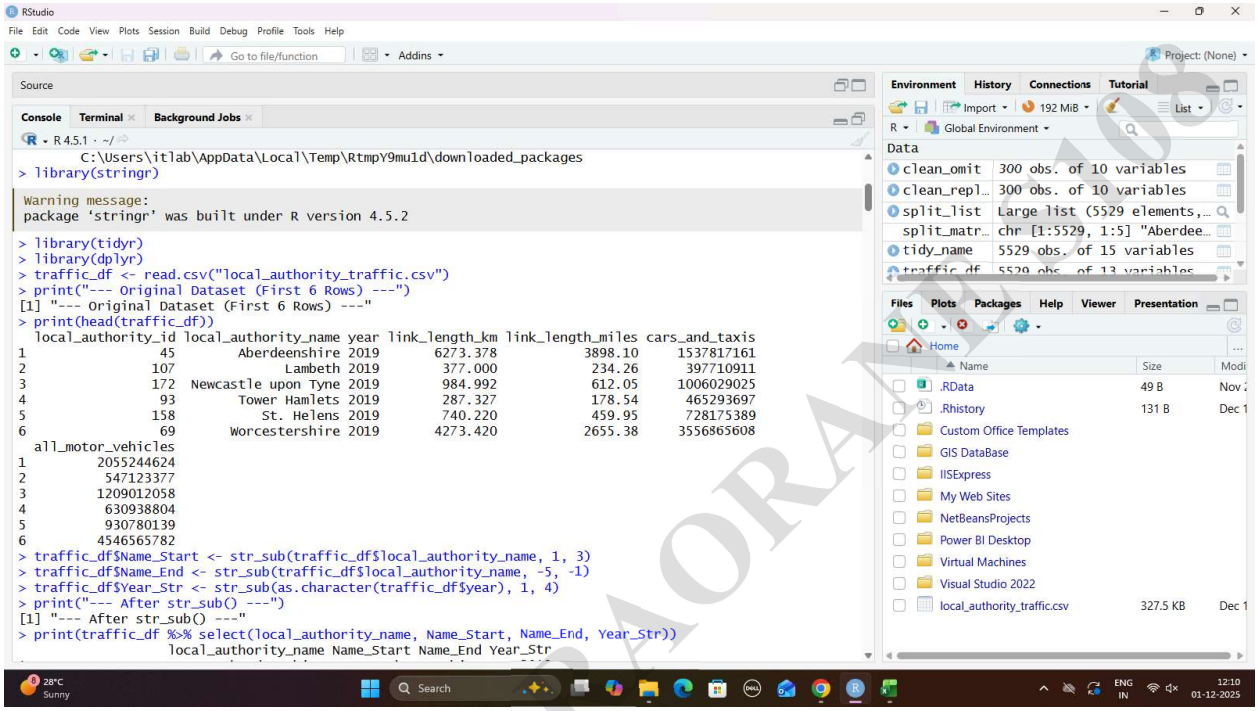
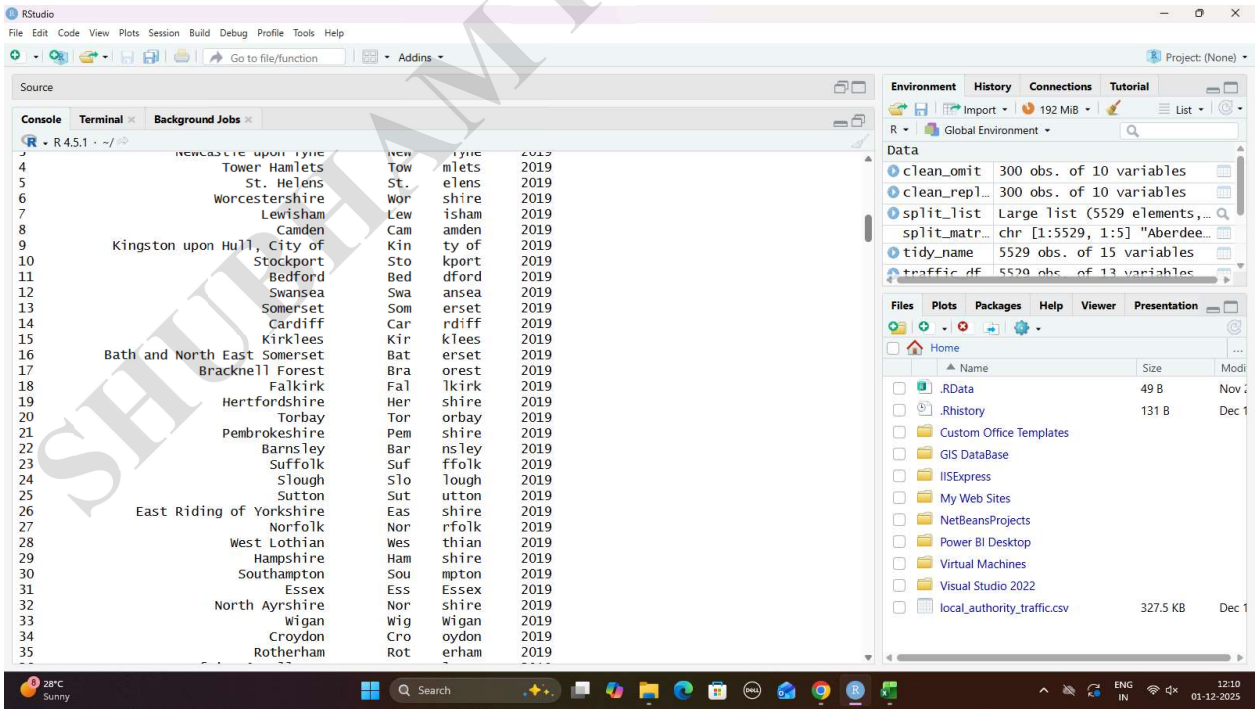
SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE
SUBJECT : Data Analysis with SAS / SPSS /R

AIM : Performing text manipulation using str_sub(), str_split() (R). import dataset.

OUTPUT

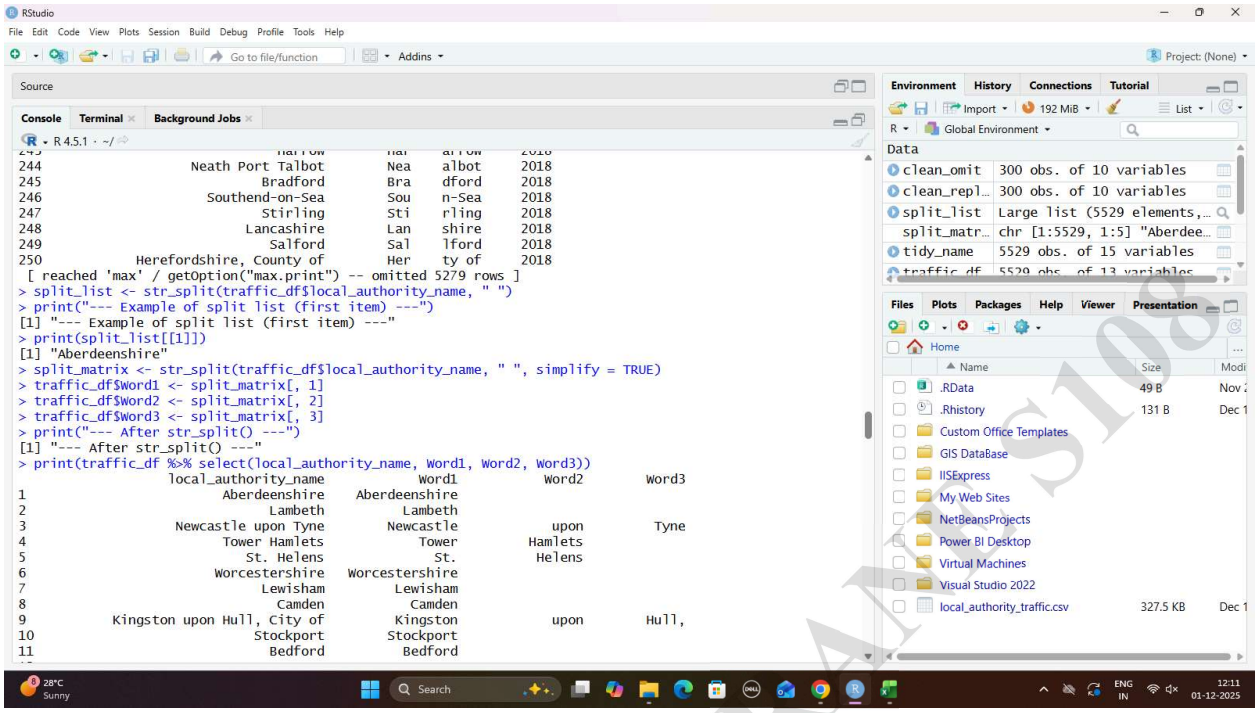




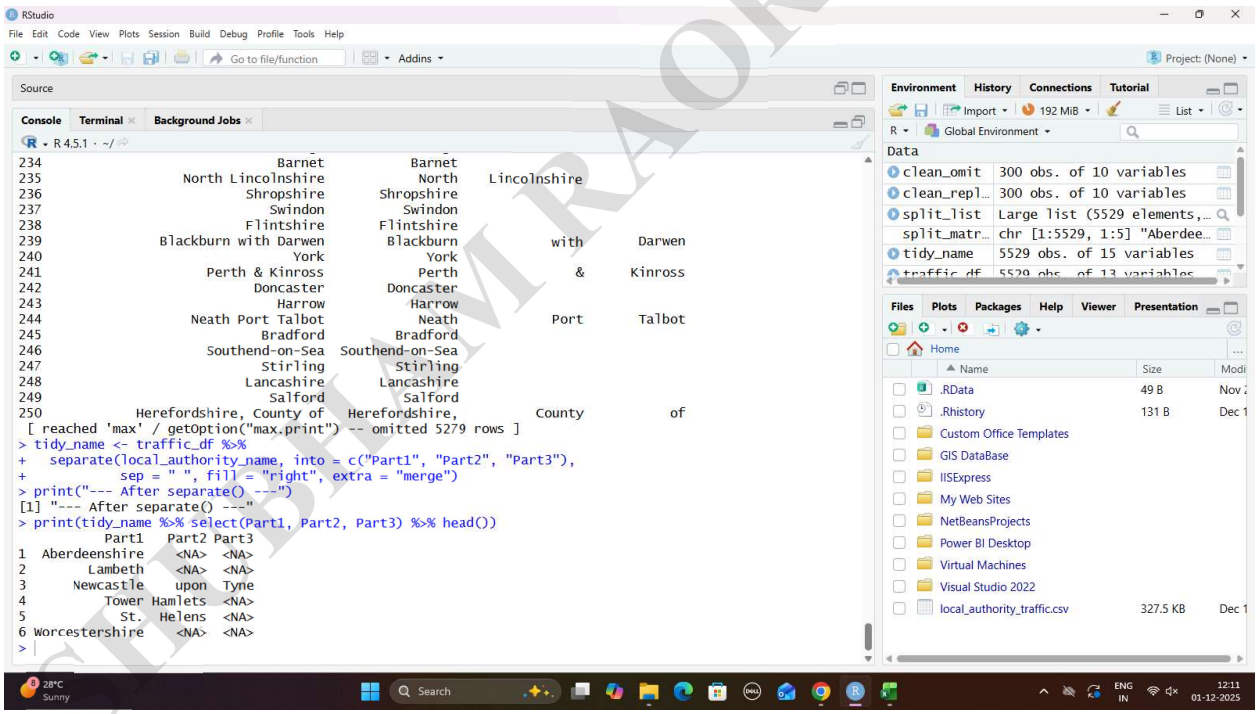
SHUBHAM RAORANE S108

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

SUBJECT : Data Analysis with SAS / SPSS /R



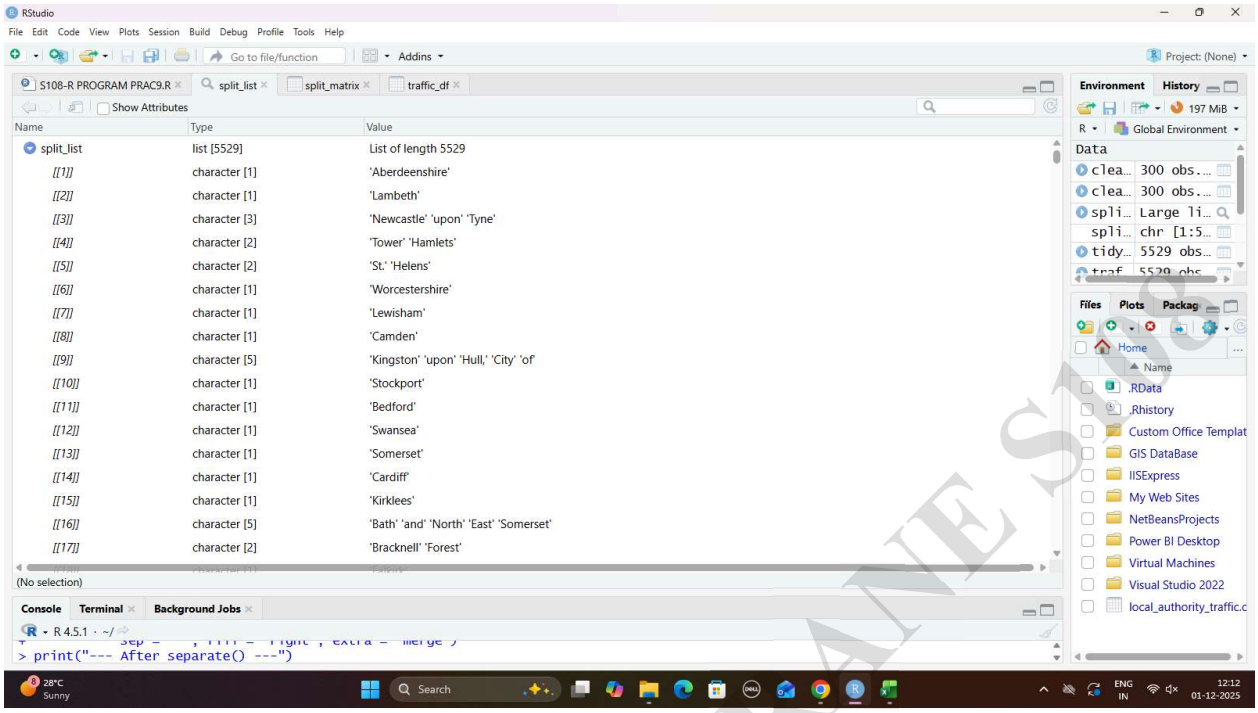
```
244                    Neath Port Talbot        Nea    albot     2018
245                             Bradford        Bra    dford     2018
246                      Southend-on-Sea        Sou    n-Sea     2018
247                             Stirling        Sti    rling     2018
248                           Lancashire        Lan    shire     2018
249                              Salford        Sal    lford     2018
250              Herefordshire, County of        Her    ty of     2018
 [ reached 'max' / getOption("max.print") -- omitted 5279 rows ]
> split_list <- str_split(traffic_df$local_authority_name, " ")
> print("--- Example of split list (first item) ---")
[1] "--- Example of split list (first item) ---"
> print(split_list[[1]])
[1] "Aberdeenshire"
> split_matrix <- str_split(traffic_df$local_authority_name, " ", simplify = TRUE)
> traffic_df$Word1 <- split_matrix[, 1]
> traffic_df$Word2 <- split_matrix[, 2]
> traffic_df$Word3 <- split_matrix[, 3]
> print("--- After str_split() ---")
[1] "--- After str_split() ---"
> print(traffic_df %>% select(local_authority_name, Word1, Word2, Word3))
                   local_authority_name           Word1          Word2        Word3
1                          Aberdeenshire    Aberdeenshire
2                                Lambeth         Lambeth
3                  Newcastle upon Tyne         Newcastle           upon         Tyne
4                        Tower Hamlets           Tower         Hamlets
5                           St. Helens             St.          Helens
6                       Worcestershire   Worcestershire
7                             Lewisham        Lewisham
8                              Camden          Camden
9           Kingston upon Hull, City of        Kingston           upon         Hull,
10                           Stockport       Stockport
11                             Bedford         Bedford
```



```
234                            Barnet          Barnet
235                 North Lincolnshire           North    Lincolnshire
236                         Shropshire      Shropshire
237                            Swindon         Swindon
238                         Flintshire      Flintshire
239                Blackburn with Darwen       Blackburn           with       Darwen
240                                York            York
241                      Perth & Kinross           Perth              &      Kinross
242                             Doncaster       Doncaster
243                              Harrow          Harrow
244                    Neath Port Talbot           Neath           Port       Talbot
245                             Bradford        Bradford
246                      Southend-on-Sea  Southend-on-Sea
247                             Stirling        Stirling
248                           Lancashire      Lancashire
249                              Salford         Salford
250              Herefordshire, County of  Herefordshire,         County         of
 [ reached 'max' / getOption("max.print") -- omitted 5279 rows ]
> tidy_name <- traffic_df %>%
+   separate(local_authority_name, into = c("Part1", "Part2", "Part3"),
+            sep = " ", fill = "right", extra = "merge")
> print("--- After separate() ---")
[1] "--- After separate() ---"
> print(tidy_name %>% select(Part1, Part2, Part3) %>% head())
        Part1   Part2 Part3
1 Aberdeenshire    <NA>  <NA>
2       Lambeth    <NA>  <NA>
3     Newcastle    upon  Tyne
4   Tower Hamlets  <NA>  <NA>
5        St.  Helens  <NA>  <NA>
6 Worcestershire    <NA>  <NA>
>
```

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

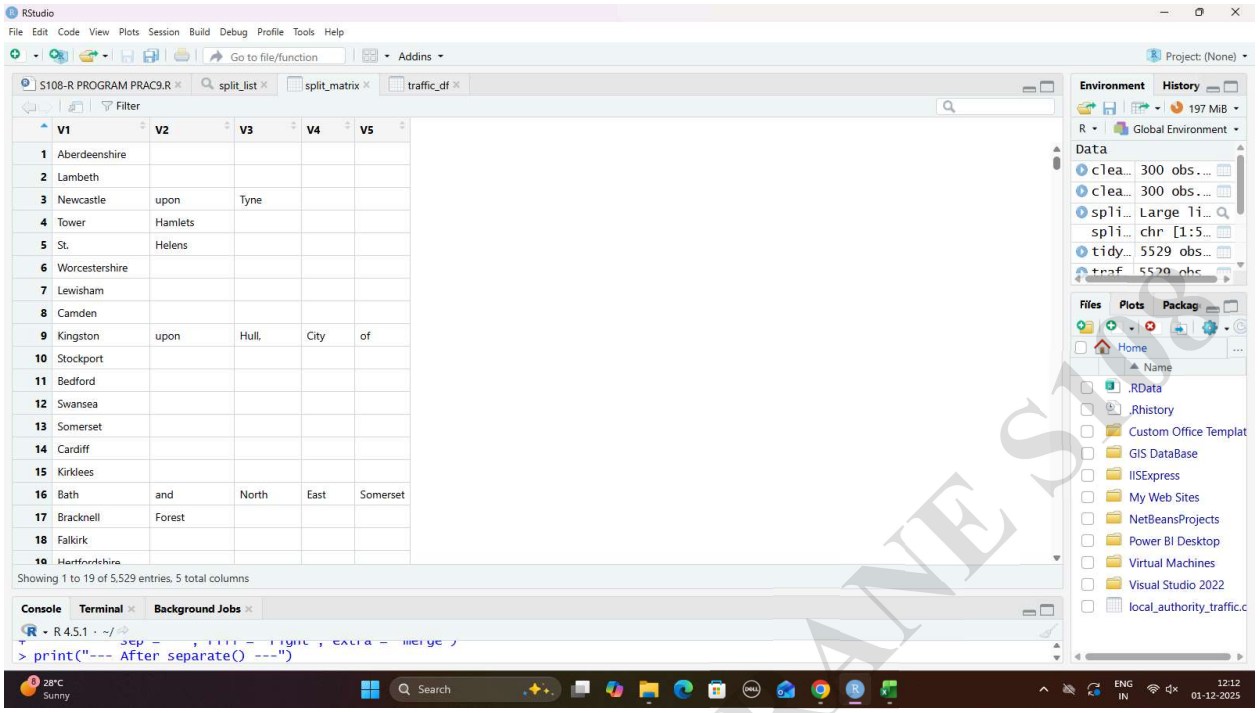
SUBJECT : Data Analysis with SAS / SPSS /R



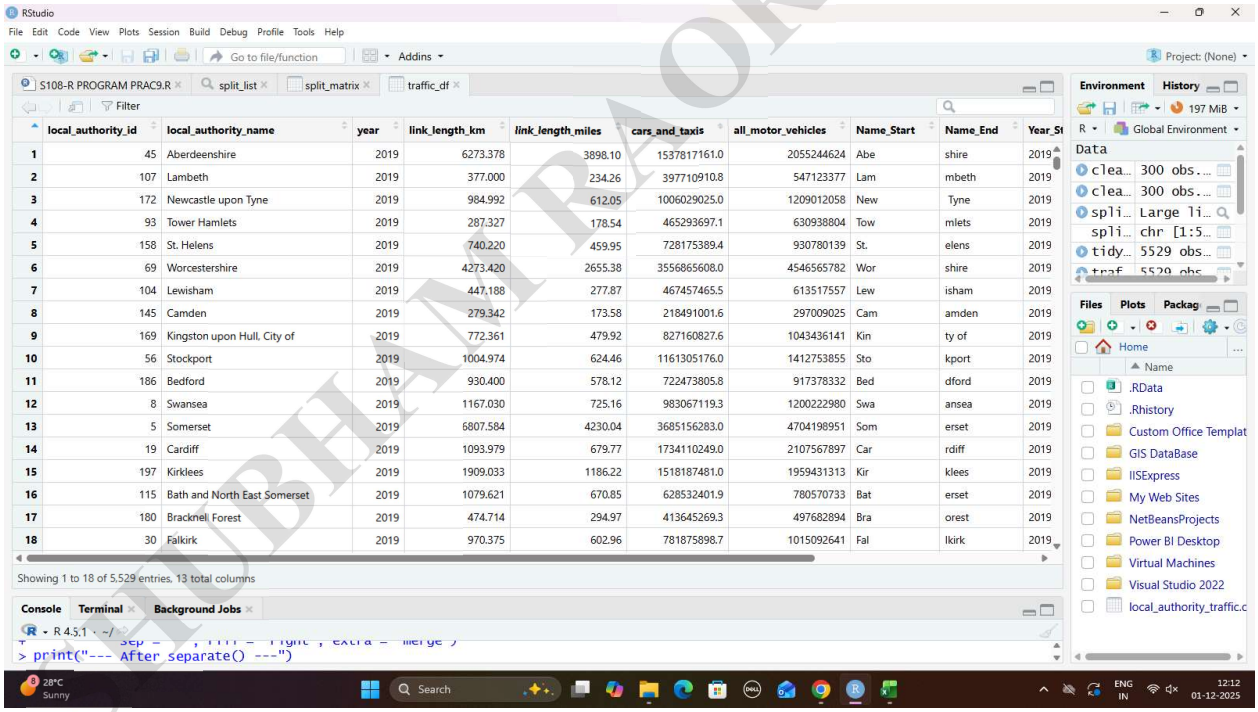
SHUBHAM RAORANE S108

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

SUBJECT : Data Analysis with SAS / SPSS /R



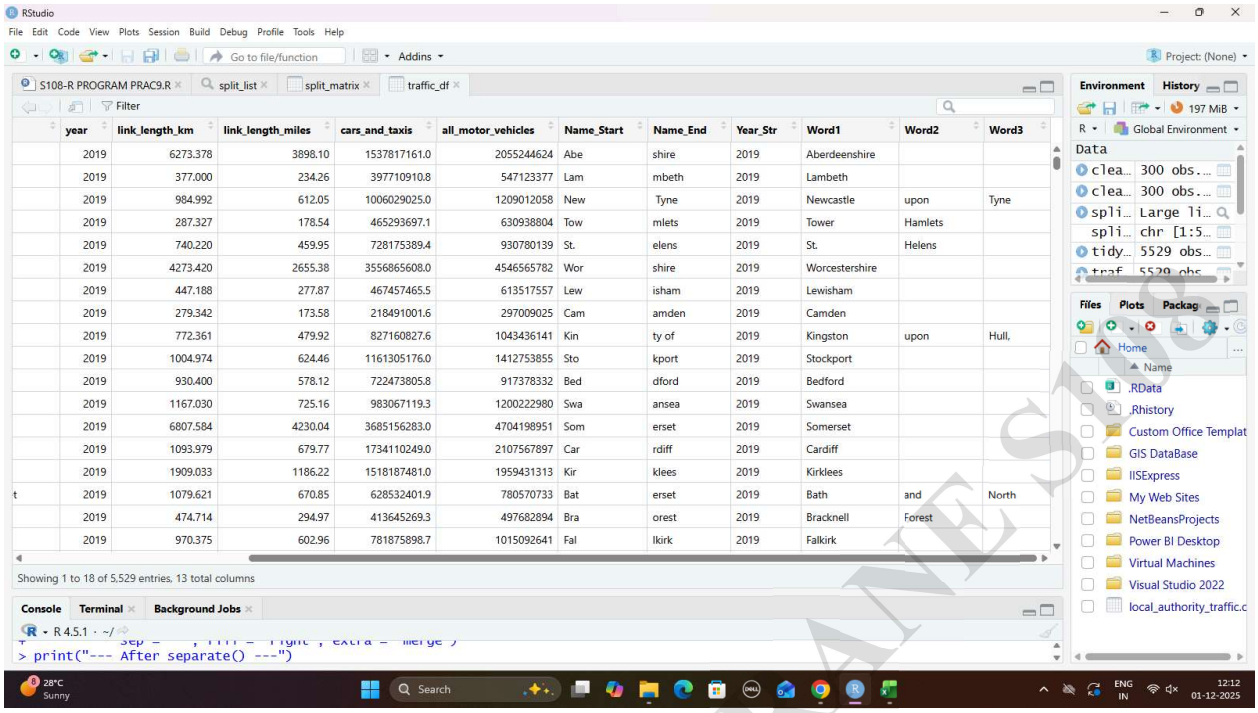
	V1	V2	V3	V4	V5
1	Aberdeenshire				
2	Lambeth				
3	Newcastle	upon	Tyne		
4	Tower	Hamlets			
5	St.	Helens			
6	Worcestershire				
7	Lewisham				
8	Camden				
9	Kingston	upon	Hull,	City	of
10	Stockport				
11	Bedford				
12	Swansea				
13	Somerset				
14	Cardiff				
15	Kirklees				
16	Bath	and	North	East	Somerset
17	Bracknell	Forest			
18	Falkirk				



	local_authority_id	local_authority_name	year	link_length_km	link_length_miles	cars_and_taxis	all_motor_vehicles	Name_Start	Name_End	Year_St
1	45	Aberdeenshire	2019	6273.378	3898.10	1537817161.0	2055244624	Abe	shire	2019
2	107	Lambeth	2019	377.000	234.26	397710910.8	547123377	Lam	mbeth	2019
3	172	Newcastle upon Tyne	2019	984.992	612.05	1006029025.0	1209012058	New	Tyne	2019
4	93	Tower Hamlets	2019	287.327	178.54	465293697.1	630938804	Tow	mlets	2019
5	158	St. Helens	2019	740.220	459.95	728175389.4	930780139	St.	elens	2019
6	69	Worcestershire	2019	4273.420	2655.38	3556865608.0	4546565782	Wor	shire	2019
7	104	Lewisham	2019	447.188	277.87	467457465.5	613517557	Lew	isham	2019
8	145	Camden	2019	279.342	173.58	218491001.6	297009025	Cam	amden	2019
9	169	Kingston upon Hull, City of	2019	772.361	479.92	827160827.6	1043436141	Kin	ty of	2019
10	56	Stockport	2019	1004.974	624.46	1161305176.0	1412753855	Sto	kport	2019
11	186	Bedford	2019	930.400	578.12	722473805.8	917378332	Bed	dford	2019
12	8	Swansea	2019	1167.030	725.16	983067119.3	1200222980	Swa	ansea	2019
13	5	Somerset	2019	6807.584	4230.04	3685156283.0	4704198951	Som	erset	2019
14	19	Cardiff	2019	1093.979	679.77	1734110249.0	2107567897	Car	rdiff	2019
15	197	Kirklees	2019	1909.033	1186.22	1518187481.0	1959431313	Kir	klees	2019
16	115	Bath and North East Somerset	2019	1079.621	670.85	628532401.9	780570733	Bat	erset	2019
17	180	Bracknell Forest	2019	474.714	294.97	413645269.3	497682894	Bra	orest	2019
18	30	Falkirk	2019	970.375	602.96	781875898.7	1015092641	Fal	lkirk	2019

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

SUBJECT : Data Analysis with SAS / SPSS /R



SHUBHAM RAORANE S108