

Babu Banarasi Das

University



Predictive Analytics

(BCADS15301)

LAB FILE

SUBMITTED TO:
Mr. AYUSHMAN BHADURIA

SUBMITTED BY:-
Name: Vishal chaudhary
Roll No: 1230258472
Class: BCADS36

SPSS Modeler Practical Workbook: Data Integration and Analysis

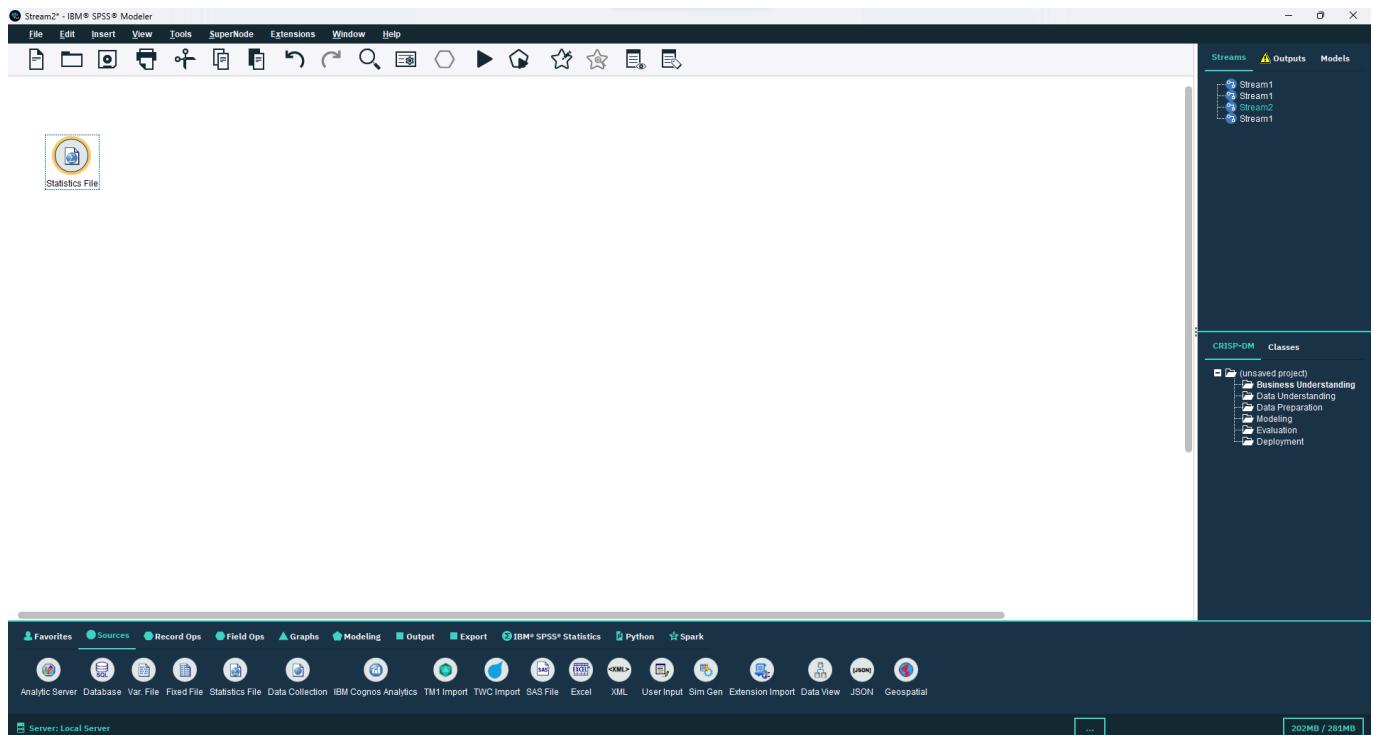
Definition: You work for a telecommunications firm and have to combine a number of datasets into a single dataset for analyses and modeling later.

Outcomes/Learning: Learning how to join different data sets
How to join records of different data sets
How to sample data from data sets

Required Tool: IBM SPSS Modeler Tool

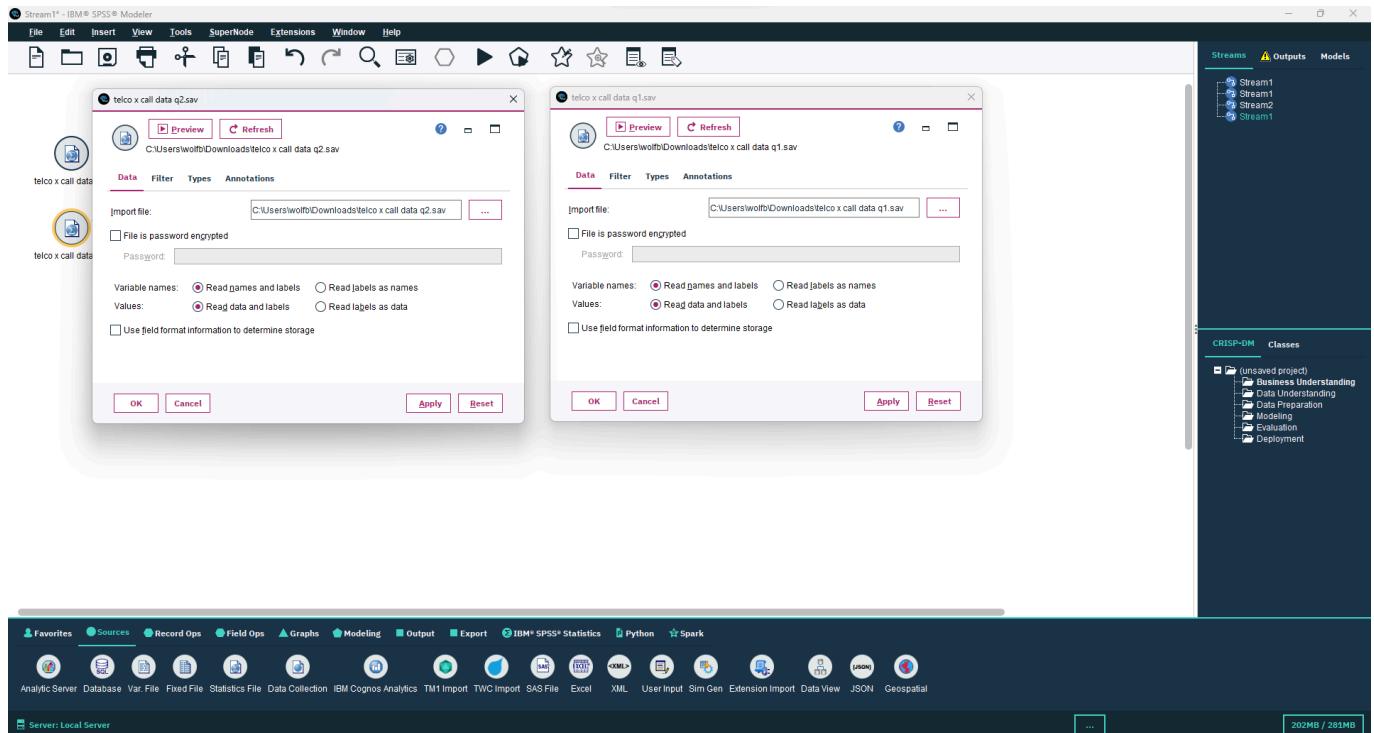
Working: Using Merge, Append, Sample nodes to merge two data sets, to append records from two or more data sets, To sample data from whole data sets.

Step 1: Open SPSS Modeler tool then on Source category select statistics File node (we are selecting Statistics node because the data set we are using is an sav file.)
Double clicking on statistics node will make it appear on canvas

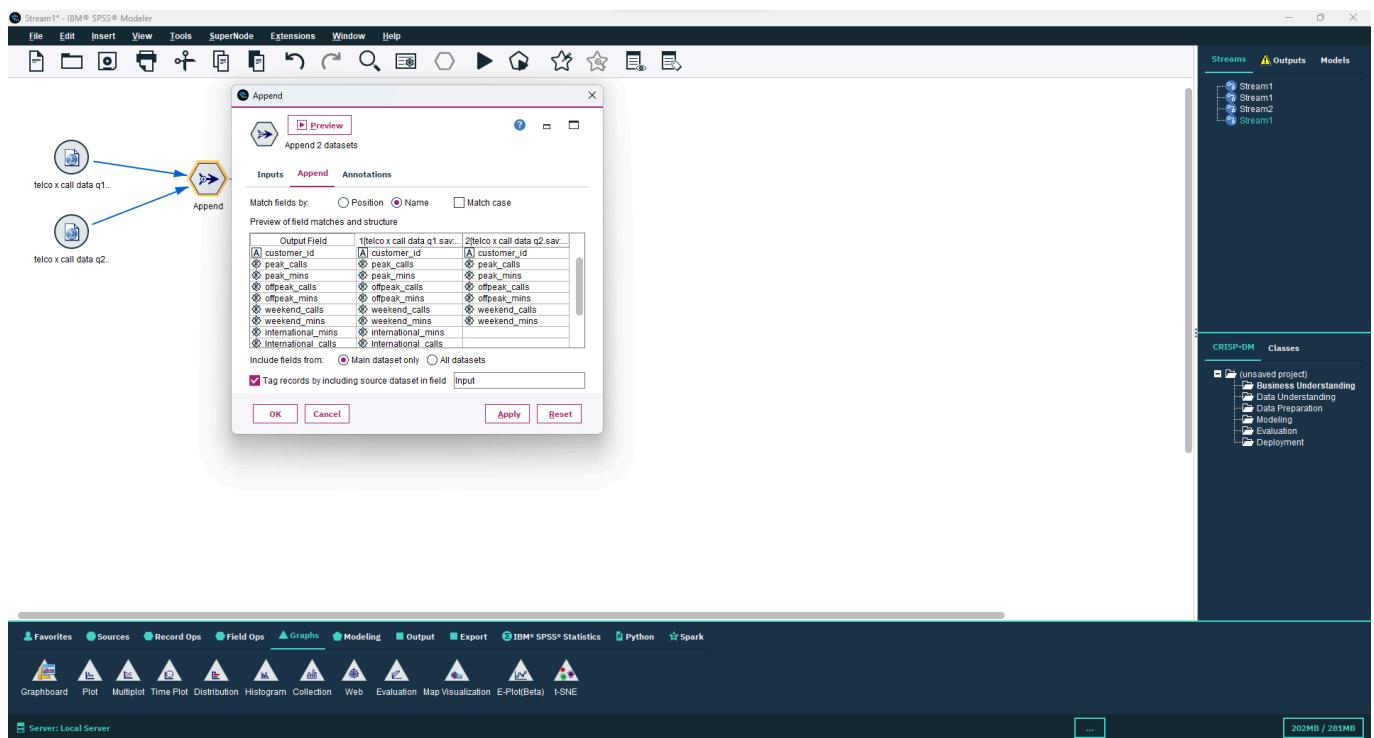


Step 2: Now we import a data set using the import option which can be accessed by double clicking on Statistics icon on the canvas.
We import a data set telco x call data q1.sav then import one more statistics file tele x call

data q2.sav.

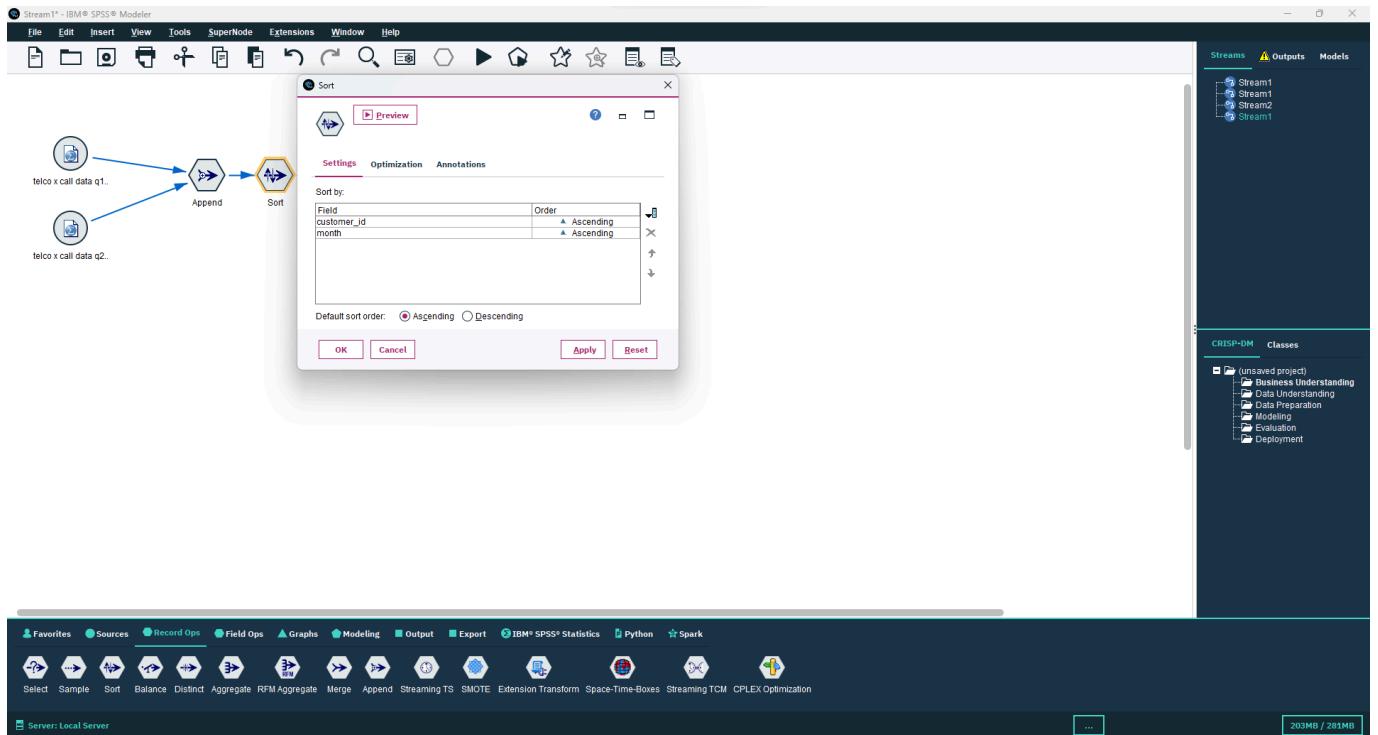


Step 3: We connect both of these files to Append node from Records Ops. click on apply and ok.

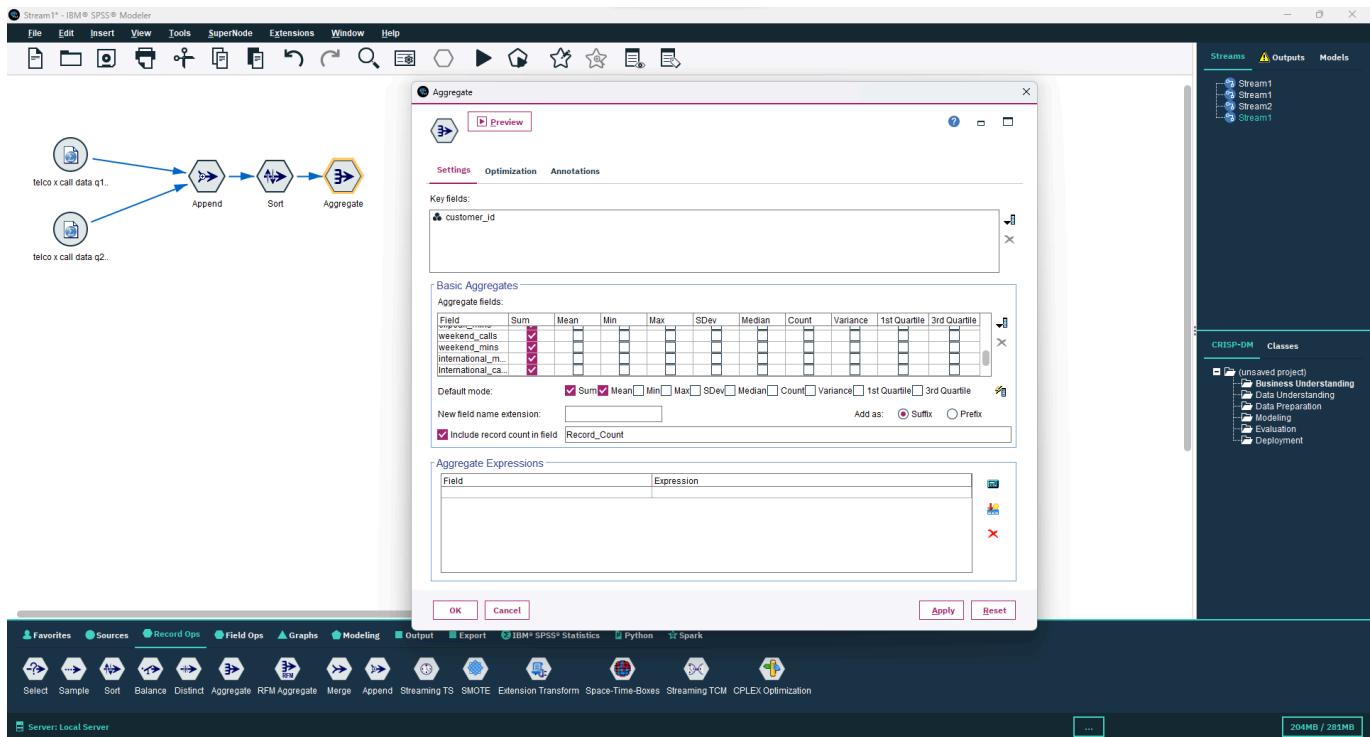


Step 4: Now we add sort node from Record Ops and connect it to Append node.

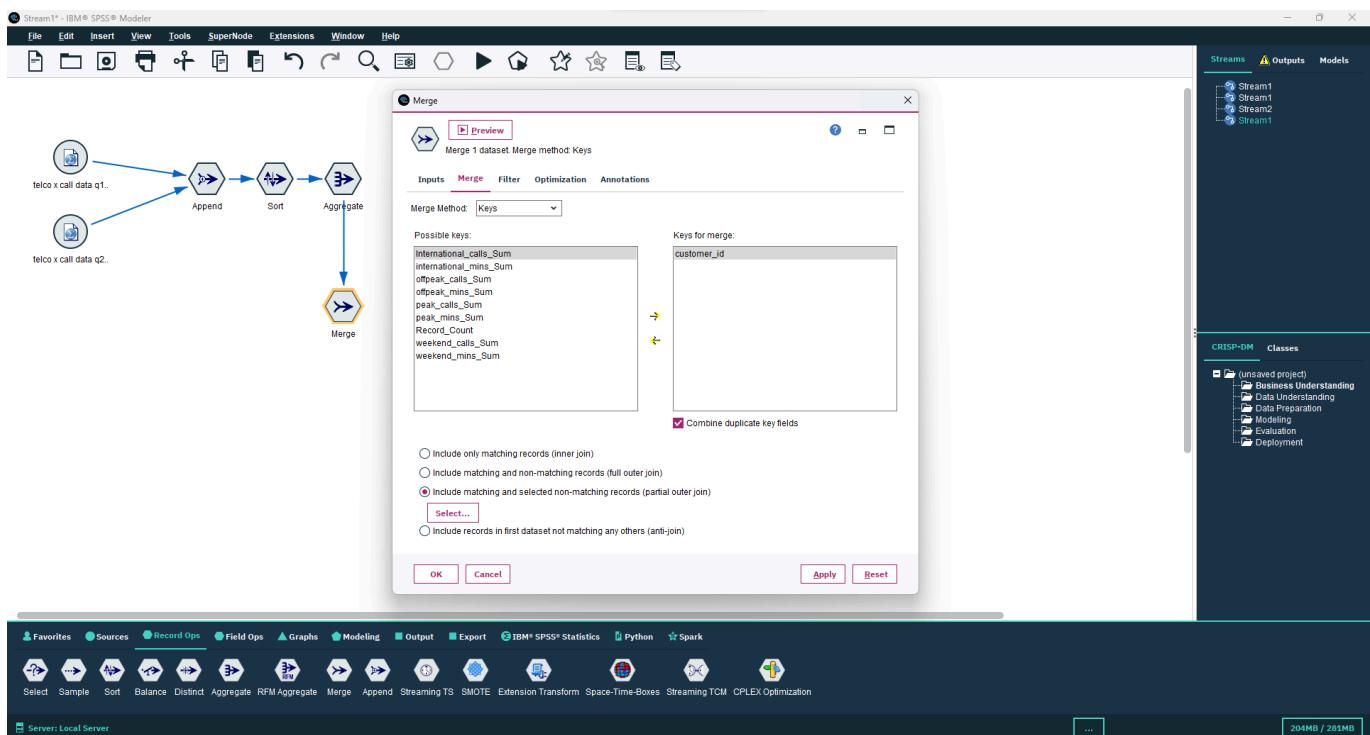
By double clicking on sort node a window opens in which under Sort by we select customer_id and month fields and sort them in ascending order.



Step 5: We add Aggregate node from Record ops and connect it to sort node. We customer_id as Key Field and all other fields should be selected in Aggregate fields and all should have only sum box as checked.



Step 6: Now add Merge node from Record ops and connect it to Aggregate node. Select customer_id as the key for merge and partial outer join by double clicking on Merge node.



Step 7: We connect table node to see any change in data set.

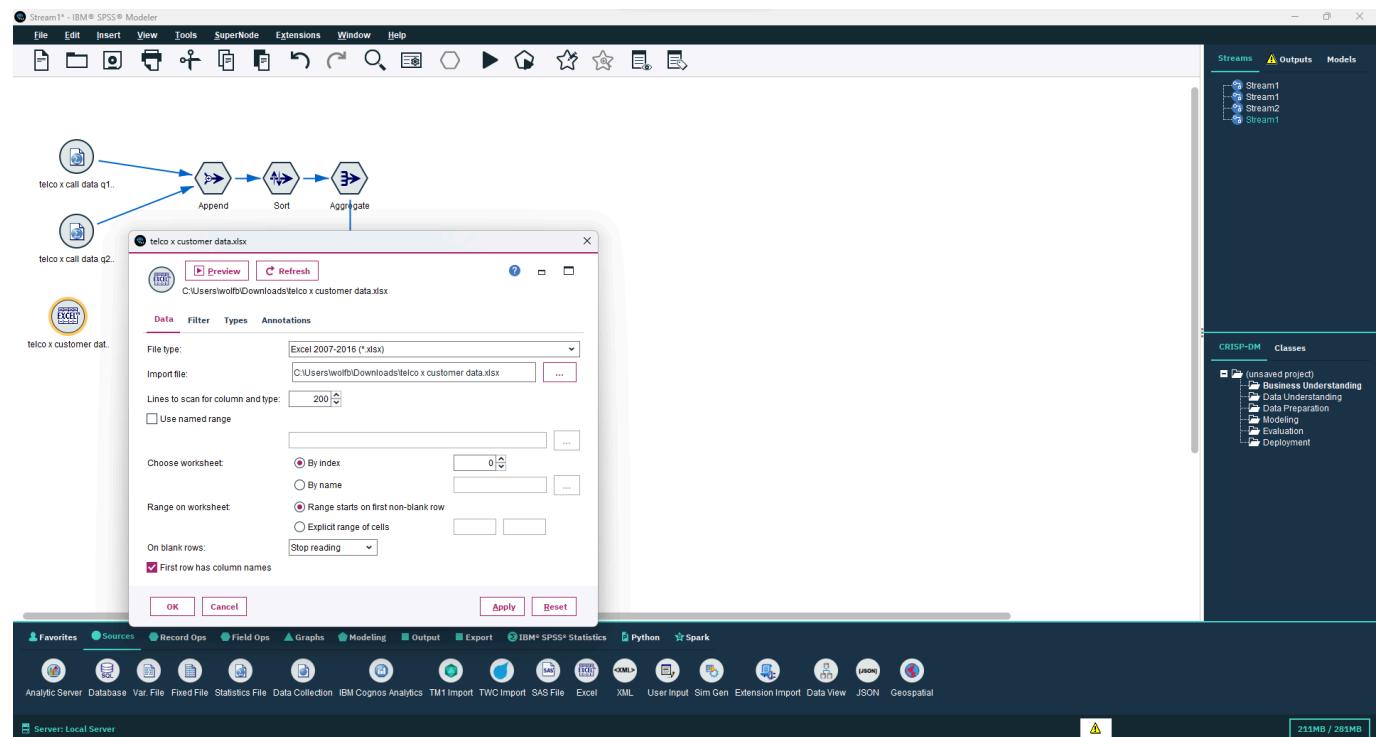
Table (10 fields, 31,769 records)

File Edit Generate

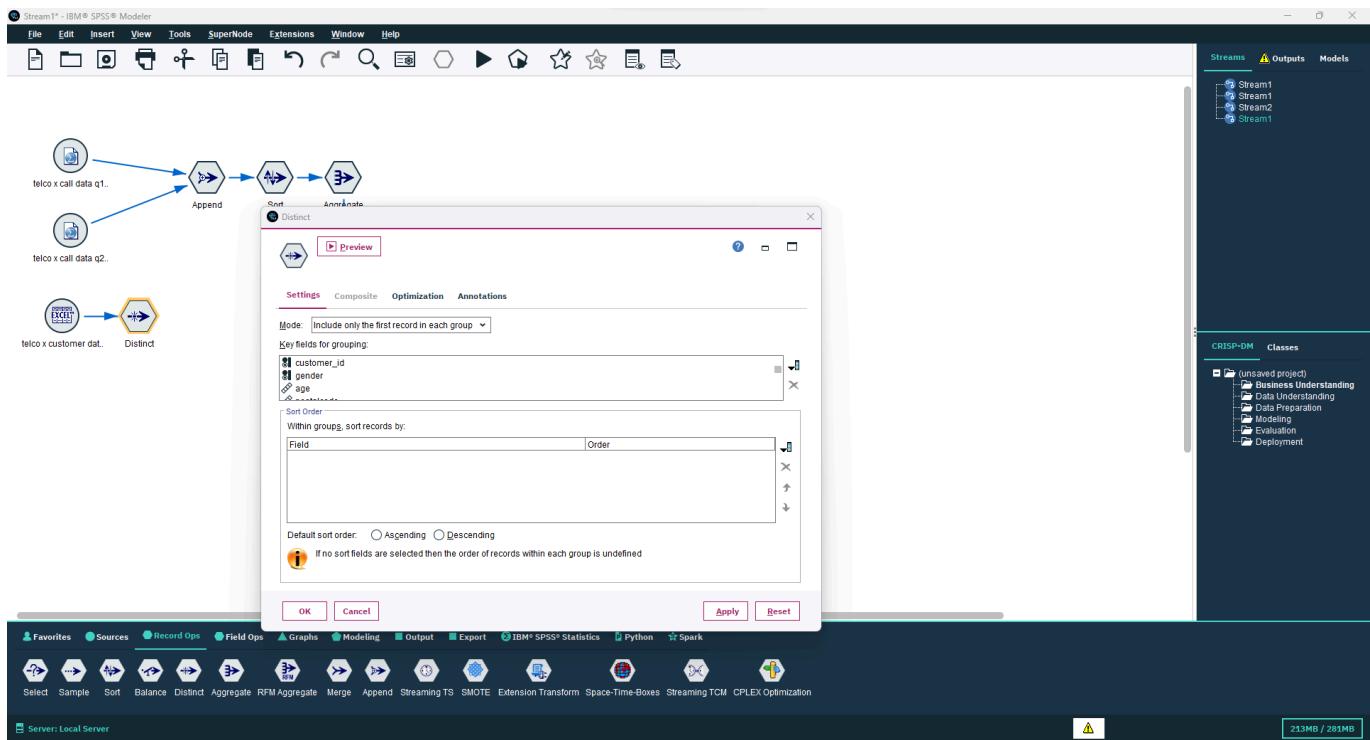
Table Annotations

	customer_id	peak_calls_Sum	peak_mins_Sum	offpeak_calls_Sum	offpeak_mins_Sum	weekend_calls_Sum	weekend_mins_Sum	international_mins_Sum	International_calls_Sum	Record_Count
1	KL00010	14.000	36.131	10.000	7.973	24.000	14.533	0.705	7.000	6
2	KL00020	54.000	39.437	34.000	21.153	0.000	0.000	4.609	0.000	6
3	KL00030	44.000	72.600	1.000	27.600	22.000	37.200	7.700	0.000	6
4	KL00040	44.000	72.600	1.000	27.600	22.000	37.200	8.621	1.000	6
5	KL00050	32.000	40.608	14.000	18.824	1.000	1.234	2.494	0.000	6
6	KL00060	56.000	46.260	6.000	11.085	8.000	10.233	0.858	0.000	6
7	KL00070	68.000	56.370	6.000	4.012	1.000	4.012	3.189	0.000	6
8	KL00080	40.000	51.043	13.000	34.802	5.000	14.230	6.734	0.000	6
9	KL00090	54.000	99.000	34.000	53.100	0.000	0.000	12.404	0.000	6
10	KL00100	37.000	65.400	17.000	29.400	0.000	0.000	2.074	0.000	6
11	KL00110	37.000	65.400	17.000	29.400	0.000	0.000	2.062	0.000	6
12	KL00120	61.000	68.123	34.000	33.880	8.000	8.334	6.031	0.000	6
13	KL00130	73.000	63.352	31.000	26.803	1.000	3.046	12.062	0.000	6
14	KL00140	40.000	41.447	23.000	16.133	16.000	14.209	8.364	4.000	6
15	KL00150	54.000	96.000	9.000	34.800	14.000	22.800	25.436	2.000	6
16	KL00160	15.000	41.342	17.000	23.841	6.000	13.189	3.675	0.000	6
17	KL00170	62.000	100.200	17.000	31.500	5.000	9.000	2.547	1.000	6
18	KL00180	57.000	41.918	32.000	21.827	6.000	11.988	10.687	0.000	6
19	KL00190	64.000	58.782	12.000	28.749	2.000	3.533	9.348	0.000	6
20	KL00200	50.000	84.600	15.000	27.900	21.000	42.000	6.155	1.000	6
21	KL00210	50.000	84.600	15.000	27.900	21.000	42.000	6.355	0.000	6

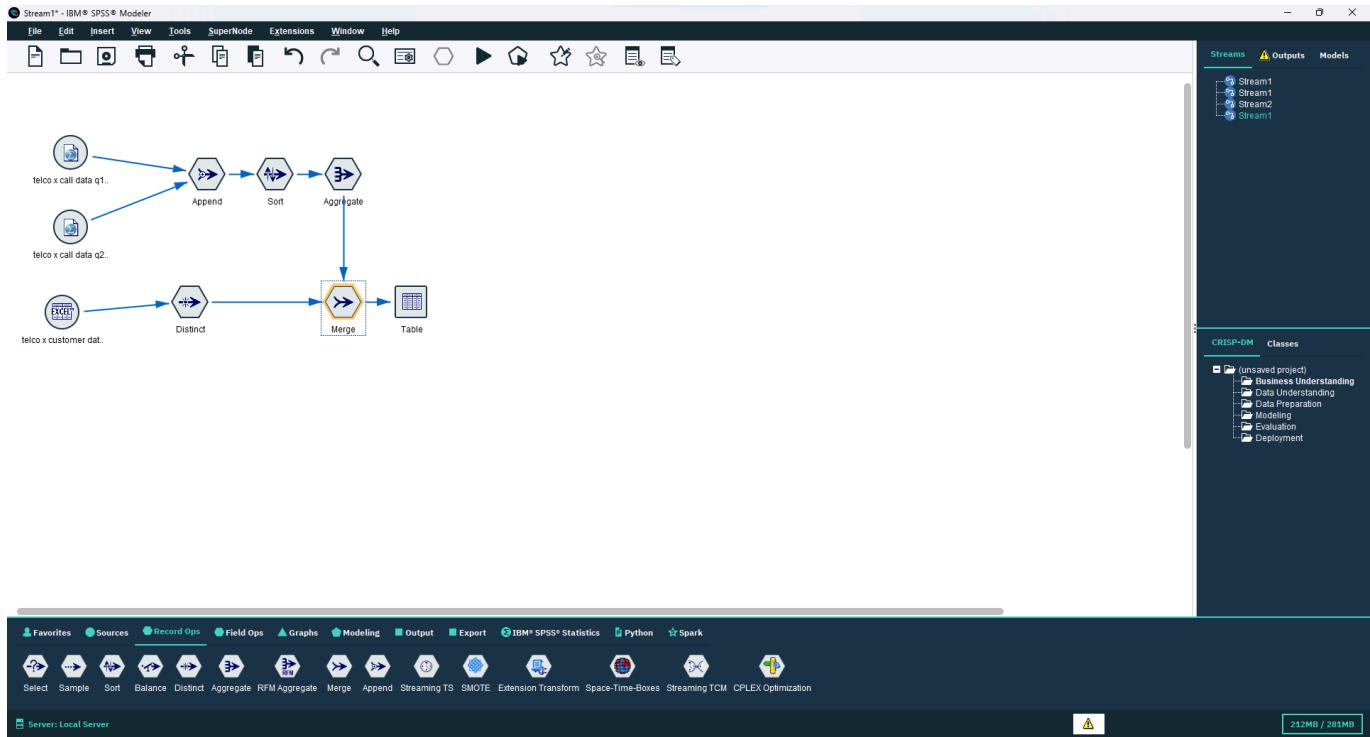
Step 8: Now we import a new data set telco x customer data.xlsx which is an excel file.



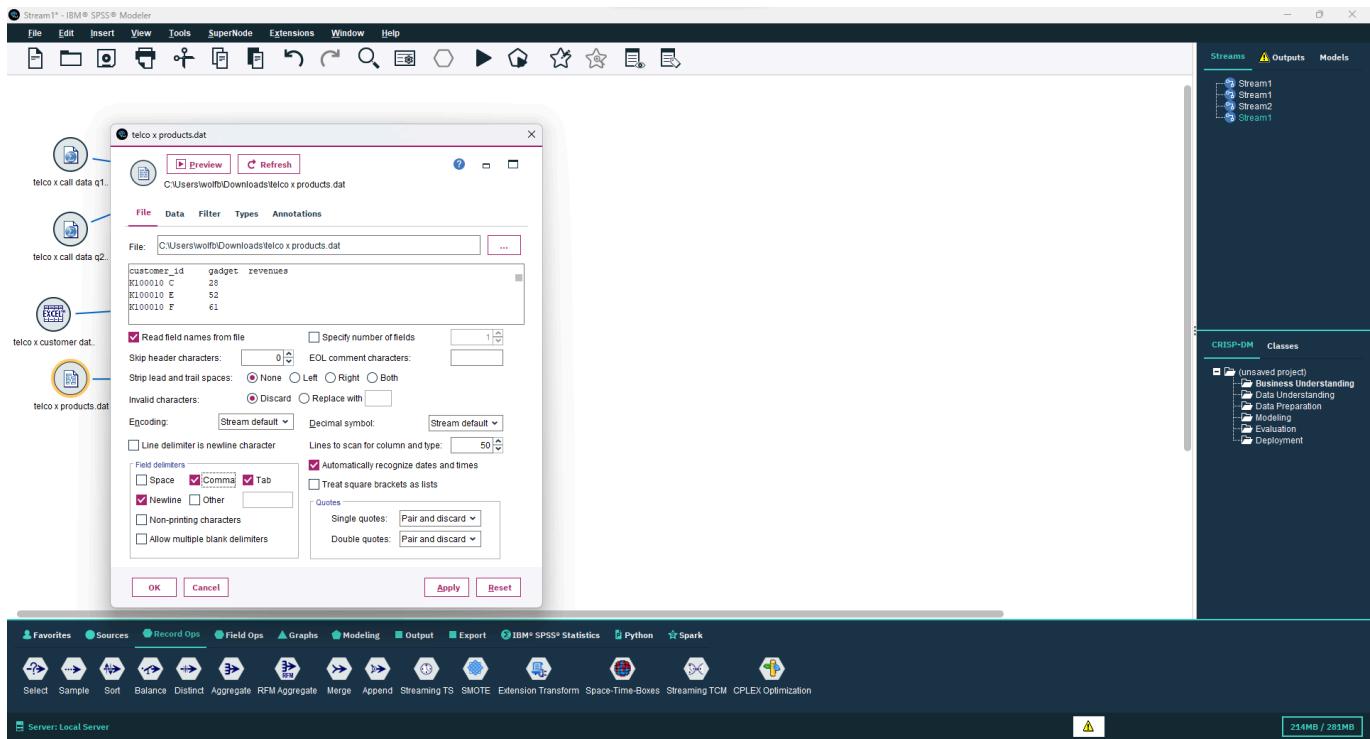
Step 9: Connect a Distinct node from Record ops and take all the fields of the data set as key fields in Distinct node by double clicking the Distinct node. And select Include only first record from each group.



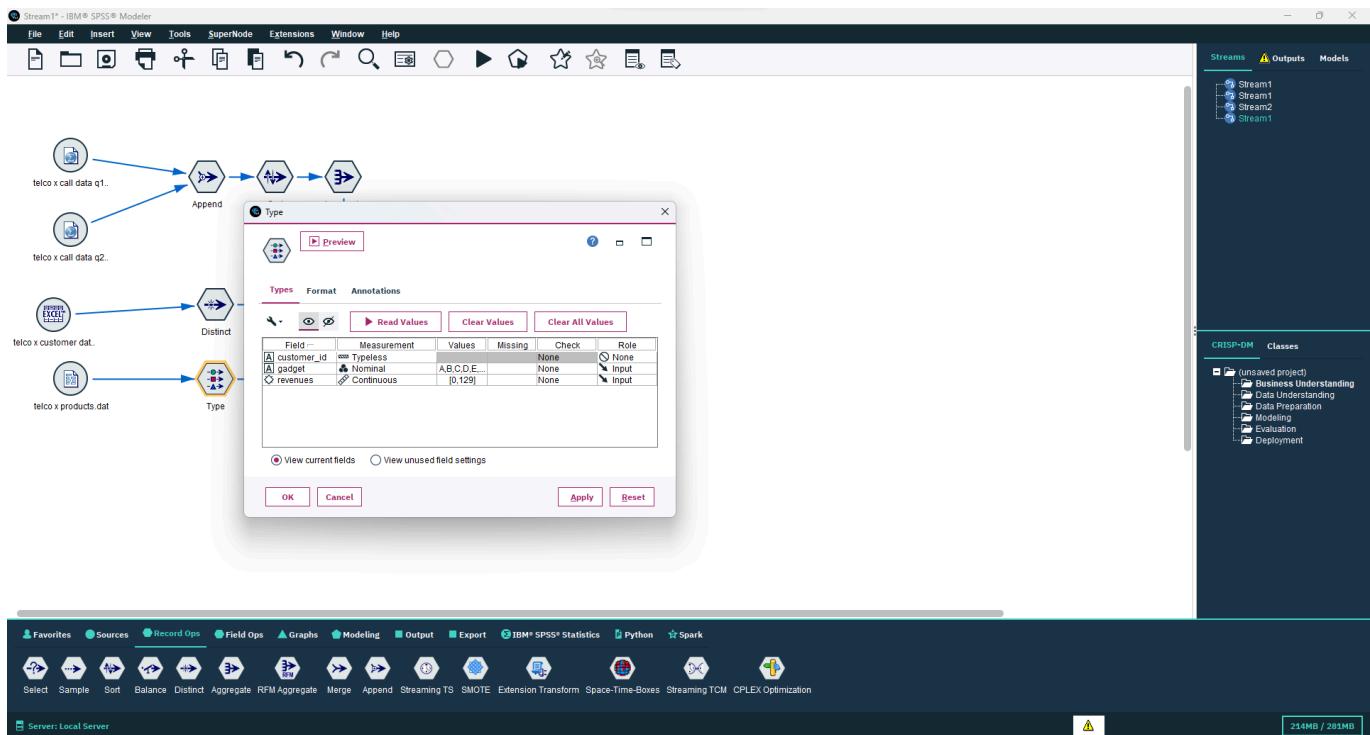
Step 10: Connect the Distinct node to Merge node.



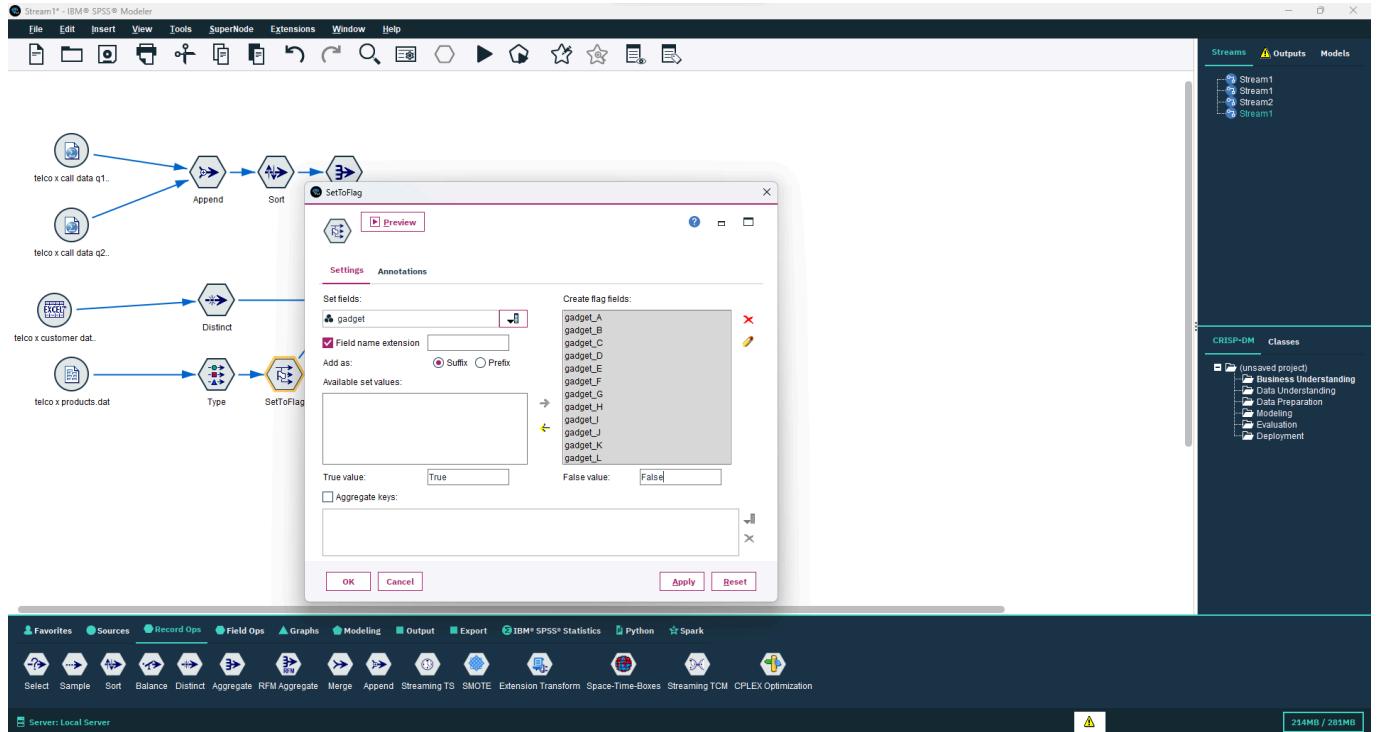
Step 11: Take a var file node from Sources Category and import telco x products.dat.



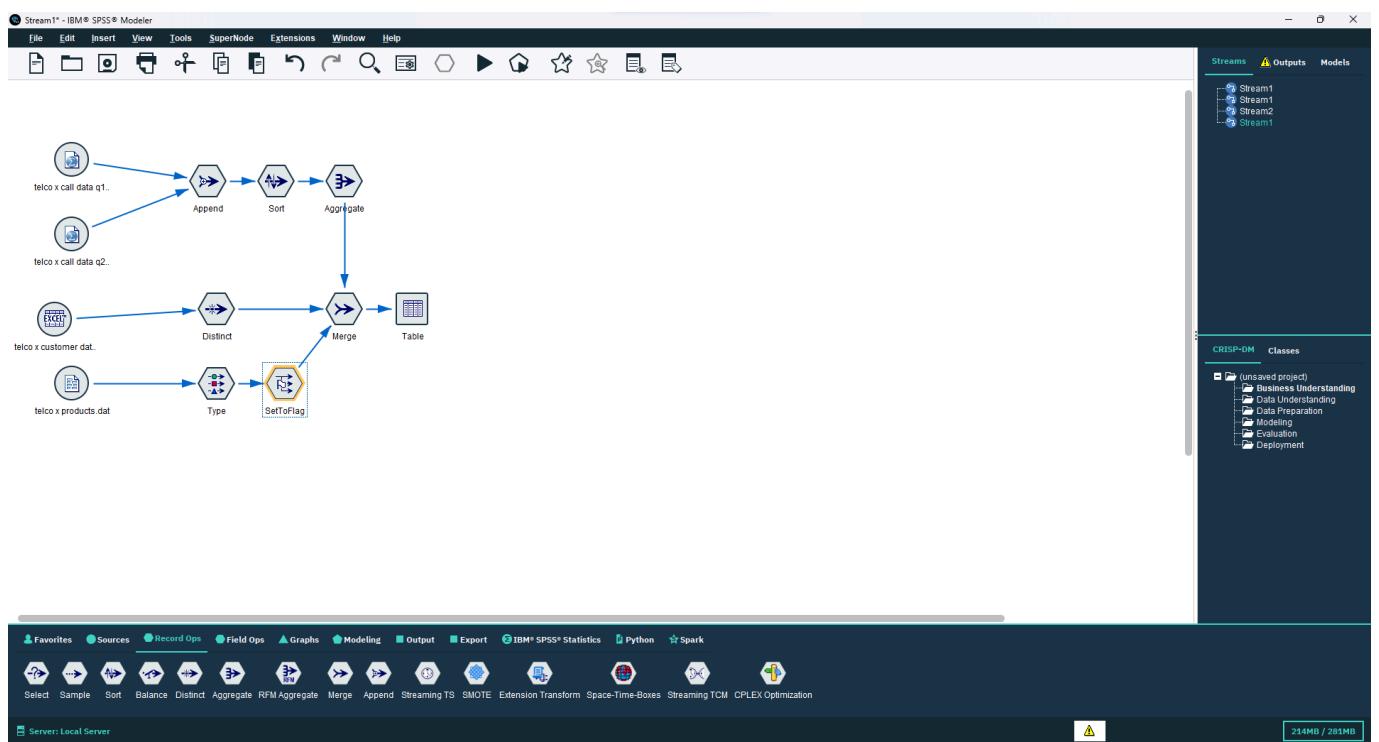
Step 12: Connect Type node to the new var file and get the specific category of all values by clicking on Read Values button in Type node window.



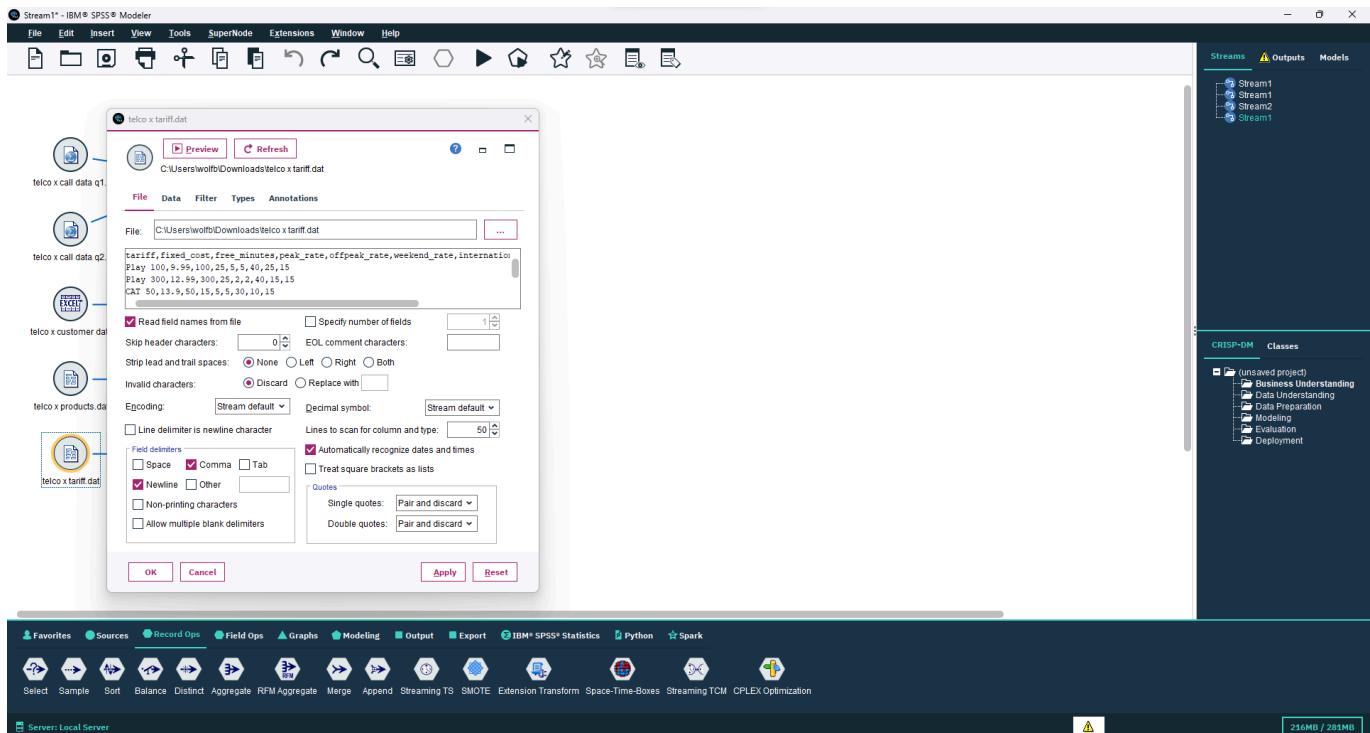
Step 13: Now we connect type node to Set to flag node and select gadget in Set fields section in Set to flag window and all the values should be selected and sent to Create Flag Fields.



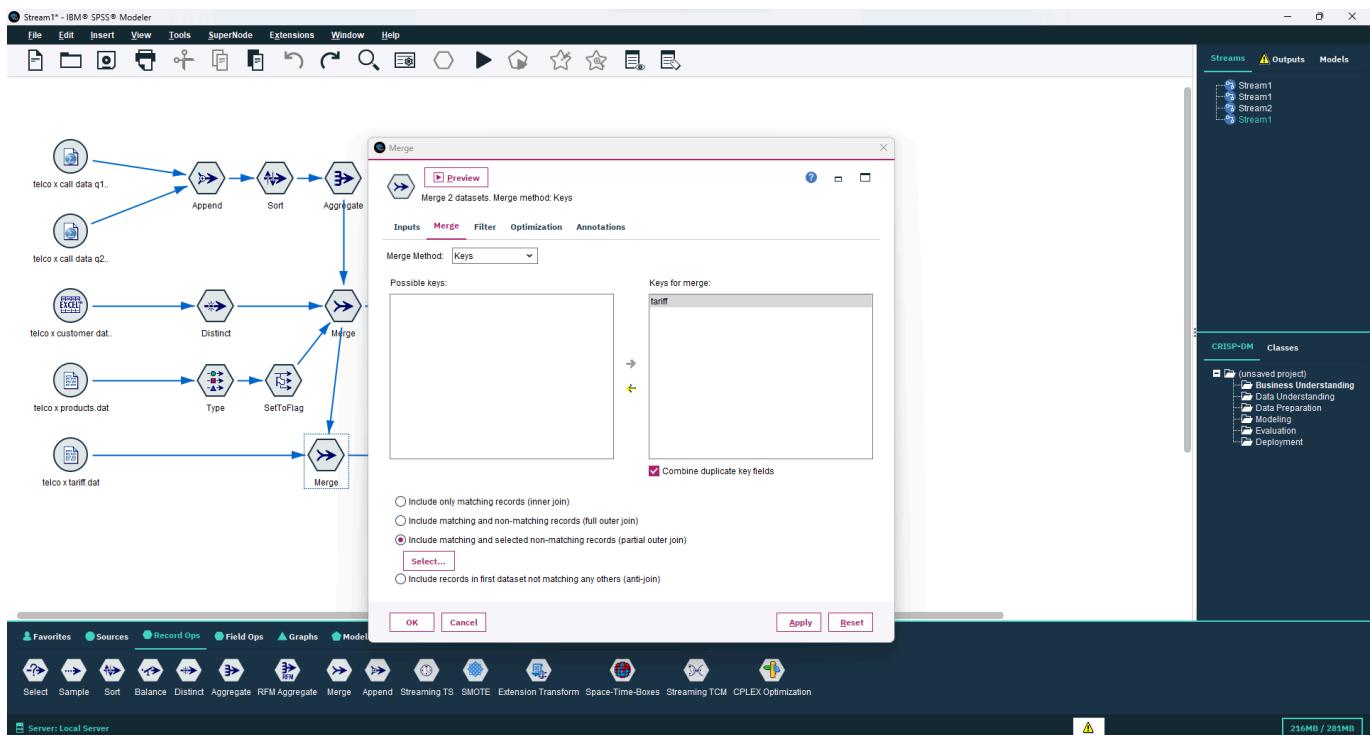
Step 14: Connect the Set to Flag node to Merge node.

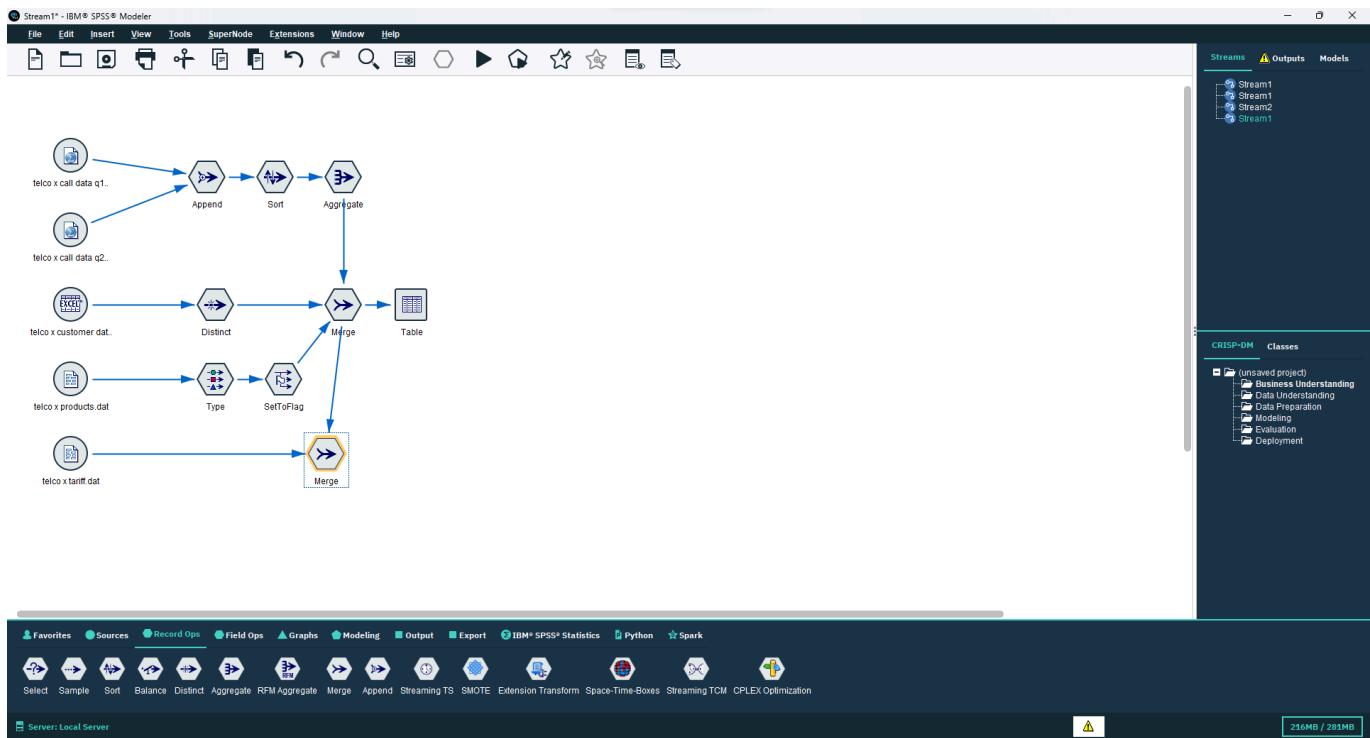


Step 15: Now we add another var file telco x tariff.dat



Step 16: Connect Merge node to the new var file and take tariff as key field and partial outer join in the Merge node window then connect the previous Merge Node to the new one Which is connected to telco x tariff.dat.



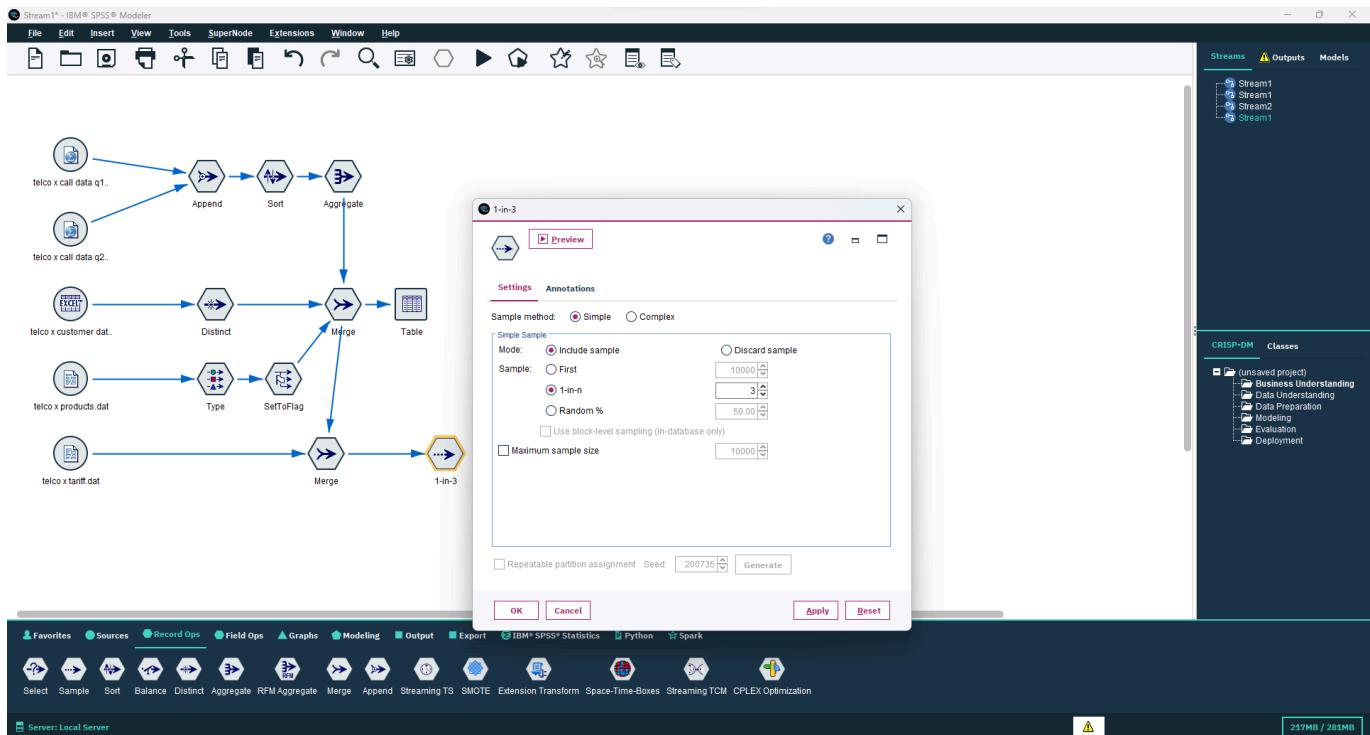


Step 17: Then connect the sample node from the Record ops to the Merge node connected to telco x tariff.dat.

Sample node is used to take a part of data from a data set to test the data set.

Open the Sample node window by double clicking on it and now select 1 in n rule now we choose a integer in place of n.

If we choose 3 in place of n the result will show the first record and then the third record skipping the record on the second place.



Step 18: Connect Table node to Sample node to view the final result which will show sample data of the four joint data sets.

Table (43 fields, 50,463 records)

File Edit Generate

Table Annotations

1	CAT	100	17.500	100	15	5	5	30	10	15	K1L10780	gender	Male	age	4544.00	2.000	2004-09-25	gnull16	1.000	Fre	Pay	\$80	Active	13.000	664.200	72.00
2	CAT	100	17.500	100	15	5	5	30	10	15	K1L10770	Male	3...	1112.00	1.000	2003-09-25	gnull16	6.000	Fre	Pay	\$50	Active	13.000	669.400	124.00	
3	CAT	100	17.500	100	15	5	5	30	10	15	K1L10770	Male	3...	1112.00	1.000	2003-09-25	gnull16	6.000	Fre	Pay	\$50	Active	13.000	669.400	124.00	
4	CAT	100	17.500	100	15	5	5	30	10	15	K1L10760	Female	4...	8287.00	4.000	2003-01-22	2004-03-19	2.000	Fre	Pay	ASAD90	Churned	8.000	438.600	157.00	
5	CAT	100	17.500	100	15	5	5	30	10	15	K1L10760	Female	4...	8287.00	4.000	2003-01-22	2004-03-19	2.000	Fre	Pay	ASAD90	Churned	8.000	438.600	157.00	
6	CAT	100	17.500	100	15	5	5	30	10	15	K1L10750	Male	3...	8980.00	4.000	2003-08-03	gnull16	2.000	Fre	Pay	\$80	Active	2.000	826.200	4.00	
7	CAT	100	17.500	100	15	5	5	30	10	15	K1L10740	Male	3...	3914.00	2.000	2004-05-27	gnull16	1.000	Fre	Pay	\$80	Active	1.000	461.400	2.00	
8	CAT	100	17.500	100	15	5	5	30	10	15	K1L10740	Male	3...	3914.00	2.000	2004-05-27	gnull16	1.000	Fre	Pay	\$80	Active	1.000	461.400	2.00	
9	CAT	100	17.500	100	15	5	5	30	10	15	K1L10740	Female	4...	6245.00	3.000	2005-05-25	2004-10-04	2.000	Fre	Pay	ASAD90	Churned	9.000	464.200	85.00	
10	CAT	100	17.500	100	15	5	5	30	10	15	K1L10720	Female	2...	4151.00	2.000	2005-06-27	gnull16	2.000	Fre	Pay	\$80	Active	2.000	473.400	34.00	
11	CAT	100	17.500	100	15	5	5	30	10	15	K1L10720	Female	2...	4151.00	2.000	2005-06-27	gnull16	2.000	Fre	Pay	\$80	Active	2.000	473.400	34.00	
12	CAT	100	17.500	100	15	5	5	30	10	15	K1L10720	Female	2...	4151.00	2.000	2005-06-27	gnull16	2.000	Fre	Pay	\$80	Active	2.000	473.400	34.00	
13	CAT	100	17.500	100	15	5	5	30	10	15	K1L10710	Male	3...	2147.00	1.000	2003-04-20	gnull16	0.000	Fre	Pay	ASAD170	Active	2.000	417.600	147.00	
14	CAT	100	17.500	100	15	5	5	30	10	15	K1L10700	Male	4...	9763.00	4.000	2005-05-19	gnull16	2.000	Fre	Pay	\$80	Active	3.000	719.400	90.00	
15	CAT	100	17.500	100	15	5	5	30	10	15	K1L10690	Male	4...	9915.00	4.000	2004-06-01	gnull16	2.000	Fre	Pay	\$80	Active	3.000	644.400	59.00	
16	CAT	100	17.500	100	15	5	5	30	10	15	K1L10690	Male	4...	9915.00	4.000	2004-06-01	gnull16	2.000	Fre	Pay	\$80	Active	3.000	644.400	59.00	
17	CAT	100	17.500	100	15	5	5	30	10	15	K1L10680	Female	4...	5066.00	3.000	2005-12-14	2008-05-21	1.000	Fre	Pay	ASAD90	Churned	8.000	522.000	63.00	
18	CAT	100	17.500	100	15	5	5	30	10	15	K1L10680	Female	4...	5066.00	3.000	2005-12-14	2008-05-21	1.000	Fre	Pay	ASAD90	Churned	8.000	522.000	63.00	
19	CAT	100	17.500	100	15	5	5	30	10	15	K1L10680	Female	4...	5066.00	3.000	2005-12-14	2008-05-21	1.000	Fre	Pay	ASAD90	Churned	8.000	522.000	63.00	
20	CAT	100	17.500	100	15	5	5	30	10	15	K1L10670	Female	5...	1000.00	1.000	2005-05-25	2005-05-25	1.000	Fre	Pay	ASAD90	Churned	10.000	459.600	75.00	
21	CAT	100	17.500	100	15	5	5	30	10	15	K1L10670	Female	5...	1000.00	1.000	2005-05-25	2005-05-25	1.000	Fre	Pay	ASAD90	Churned	10.000	459.600	75.00	
22	CAT	100	17.500	100	15	5	5	30	10	15	K1L10660	Female	5...	6800.00	4.000	2004-01-03	2006-01-21	1.000	Fre	Pay	ASAD90	Churned	3.000	637.200	39.00	
23	CAT	100	17.500	100	15	5	5	30	10	15	K1L10650	Female	5...	5225.00	2.000	2003-01-01	gnull16	1.000	Fre	Pay	ASAD170	Active	2.000	436.200	163.00	
24	CAT	100	17.500	100	15	5	5	30	10	15	K1L10650	Female	5...	5225.00	2.000	2003-01-01	gnull16	1.000	Fre	Pay	ASAD170	Active	2.000	436.200	163.00	
25	CAT	100	17.500	100	15	5	5	30	10	15	K1L10640	Male	3...	9303.00	4.000	2005-10-02	gnull16	8.000	Fre	Pay	ASAD170	Active	3.000	533.400	22.00	
26	CAT	100	17.500	100	15	5	5	30	10	15	K1L10640	Male	3...	9303.00	4.000	2005-10-02	gnull16	8.000	Fre	Pay	ASAD170	Active	3.000	533.400	22.00	
27	CAT	100	17.500	100	15	5	5	30	10	15	K1L10630	Female	3...	9087.00	4.000	2003-08-09	gnull16	2.000	Fre	Pay	\$80	Active	12.000	579.600	20.00	
28	CAT	100	17.500	100	15	5	5	30	10	15	K1L10620	Female	2...	4464.00	2.000	2006-12-26	gnull16	1.000	Fre	Pay	ASAD170	Active	4.000	603.000	34.00	
29	CAT	100	17.500	100	15	5	5	30	10	15	K1L10610	Female	2...	2218.00	1.000	2005-09-06	2009-09-13	0.000	Fre	Pay	ASAD90	Churned	9.000	549.000	19.00	
30	CAT	100	17.500	100	15	5	5	30	10	15	K1L10610	Female	2...	2218.00	1.000	2005-09-06	2009-09-13	0.000	Fre	Pay	ASAD90	Churned	9.000	549.000	19.00	
31	CAT	100	17.500	100	15	5	5	30	10	15	K1L10610	Female	2...	2218.00	1.000	2005-09-06	2009-09-13	0.000	Fre	Pay	ASAD90	Churned	9.000	549.000	19.00	
32	CAT	100	17.500	100	15	5	5	30	10	15	K1L10600	Male	3...	4057.00	6.000	2005-06-06	2008-01-12	2.000	Fre	Pay	ASAD90	Churned	12.000	532.200	8.00	
33	CAT	100	17.500	100	15	5	5	30	10	15	K1L10600	Male	3...	4057.00	6.000	2005-06-06	2008-01-12	2.000	Fre	Pay	ASAD90	Churned	12.000	532.200	8.00	
34	CAT	100	17.500	100	15	5	5	30	10	15	K1L10590	Male	3...	5627.00	3.000	2003-01-15	2004-11-22	2.000	Fre	Pay	ASAD90	Churned	14.000	682.800	19.00	
35	CAT	100	17.500	100	15	5	5	30	10	15	K1L10590	Male	3...	2247.00	1.000	2005-10-15	2008-07-14	4.000	Fre	Pay	\$80	Churned	12.000	678.600	28.00	
36	CAT	100	17.500	100	15	5	5	30	10	15	K1L10590	Male	3...	2247.00	1.000	2005-10-15	2008-07-14	4.000	Fre	Pay	\$80	Churned	12.000	678.600	28.00	
37	CAT	100	17.500	100	15	5	5	30	10	15	K1L10570	Female	3...	7611.00	4.000	2005-06-11	2009-09-16	2.000	Fre	Pay	ASAD90	Churned	11.000	540.000	99.00	
38	CAT	100	17.500	100	15	5	5	30	10	15	K1L10560	Female	4...	1580.00	1.000	2005-08-22	gnull16	1.000	Fre	Pay	\$80	Active	5.000	506.400	38.00	
39	CAT	100	17.500	100	15	5	5	30	10	15	K1L10550	Male	4...	5675.00	3.000	2005-12-18	2007-01-22	11.000	Fre	Pay	ASAD90	Churned	7.000	557.400	39.00	
40	CAT	100	17.500	100	15	5	5	30	10	15	K1L10540	Female	2...	6053.00	3.000	2005-04-22	2009-07-01	3.000	Fre	Pay	\$80	Churned	12.000	560.400	44.00	
41	CAT	100	17.500	100	15	5	5	30	10	15	K1L10540	Female	2...	6053.00	3.000	2005-04-22	2009-07-01	3.000	Fre	Pay	\$80	Churned	12.000	560.400	44.00	
42	CAT	100	17.500	100	15	5	5	30	10	15	K1L10540	Male	4...	2015.00	1.000	2003-04-14	gnull16	2.000	Fre	Pay	\$80	Active	11.000	672.000	0.00	
43	CAT	100	17.500	100	15	5	5	30	10	15	K1L10540	Male	4...	2015.00	1.000	2003-04-14	gnull16	2.000	Fre	Pay	\$80	Active	11.000	672.000	0.00	
44	CAT	100	17.500	100	15	5	5	30	10	15	K1L10520	Female	4...	4795.00	2.000	2003-05-19	gnull16	0.000	Fre	Pay	ASAD170	Active	8.000	676.800	178.00	
45	CAT	100	17.500	100	15	5	5	30	10	15	K1L10510	Male	4...	6705.00	4.000	2003-07-18	gnull16	10.000	Fre	Pay	\$80	Active	8.000	765.600	78.00	
46	CAT	100	17.500	100	15	5	5	30	10	15	K1L10500	Male	2...	5664.00	3.000	2003-01-19	2007-02-14	11.000	Fre	Pay	ASAD90	Churned	185.000	729.000	5.00	
47	CAT	100	17.500	100	15	5	5	30	10	15	K1L10490	Male	4...	4255.00	2.000	2006-06-25	gnull16	1.000	Fre	Pay	\$80	Active	206.000	630.000	106.00	
48	CAT	100	17.500	100	15	5	5	30	10	15	K1L10490	Male	5...	4255.00	2.000	2006-06-25	gnull16	1.000	Fre	Pay	\$80	Active	206.000	630.000	106.00	
49	CAT	100	17.500	10																						

