

BABU BANARASI DAS UNIVERSITY

NAME : SHRASHTI KATIYAR

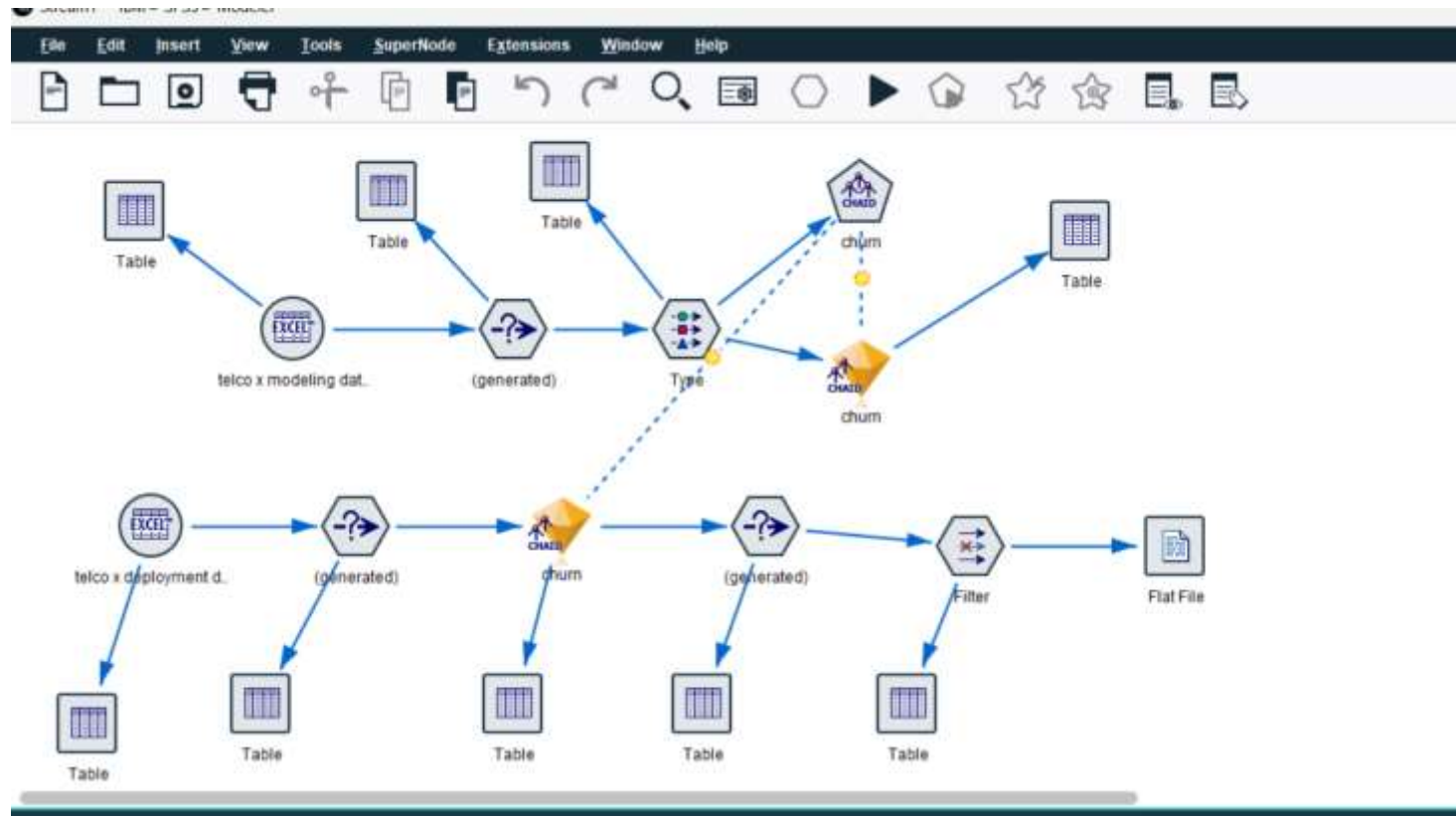
ROLL NO. : 1230258412

SUBMIT TO : MR. VIKASH SIR

ASSIGNMENT- PREDICTIVE ANALYTICS

- ▶ CREATING A DATA MINING PROJECT OF TELECOMMUNICATION DATASET FOR PREDICTING CHAID CHURN USING IBM SPSS MODELER.

STEP 1:



STEP 2:

The screenshot displays the IBM SPSS Modeler interface. On the left, a 'Table' window shows a dataset with 25 fields and 31,769 records. The 'Table' tab is active, displaying a list of records with columns: customer_id, data_known, gender, age, tariff, dropped_calls, handset, and peak_mins. The 'Annotations' tab is also visible. The main workspace shows a workflow diagram with a 'Table' icon connected to a 'Filter' icon, which is then connected to a 'Flat File' icon. The right sidebar shows the 'Streams' tab with a 'churn' stream. The bottom toolbar includes various icons for data sources, operations, and outputs.

	customer_id	data_known	gender	age	tariff	dropped_calls	handset	peak_mins
1	KI00010	yes	Male	4...	CAT 50	1,000	SCF10	36,131
2	KI00020	yes	Male	2...	CAT 50	0,000	SCF10	39,437
3	KI00030	yes	Male	3...	CAT 50	2,000	SCF20	72,600
4	KI00040	yes	Male	2...	CAT 50	2,000	SCF10	72,600
5	KI00050	yes	Male	4...	CAT 50	0,000	SCF10	40,608
6	KI00060	yes	Male	2...	CAT 50	1,000	SCF10	46,260
7	KI00070	yes	Male	3...	CAT 50	1,000	SCF20	54,370
8	KI00080	yes	Male	2...	CAT 50	0,000	SCF20	51,043
9	KI00090	yes	Male	3...	CAT 50	1,000	SCF20	99,000
10	KI00100	yes	Male	4...	CAT 50	0,000	SCF20	85,400
11	KI00110	yes	Male	3...	CAT 50	2,000	SCF20	85,400
12	KI00120	yes	Male	3...	CAT 50	0,000	SCF20	68,123
13	KI00130	yes	Male	4...	CAT 50	2,000	SCF20	63,352
14	KI00140	yes	Male	2...	CAT 50	1,000	SCF20	41,447
15	KI00150	yes	Male	3...	CAT 50	2,000	SCF20	94,000
16	KI00160	yes	Male	2...	CAT 50	9,000	SCF20	41,342
17	KI00170	yes	Male	2...	CAT 50	1,000	SCF10	100,200
18	KI00180	yes	Male	3...	CAT 50	1,000	SCF10	41,918
19	KI00190	yes	Male	3...	CAT 50	2,000	SCF10	58,792
20	KI00200	yes	Male	4...	CAT 10	0,000	SCF10	74,600

STEP 3:

The screenshot displays the IBM SPSS Studio interface. On the left, a window titled 'Table (23 fields, 9,994 records) #1' shows a data table with the following columns: customer_id, data_known, gender, age, tariff, dropped_calls, handset, and peak_mins. The table contains 20 rows of data. On the right, a workflow diagram shows a 'Table' icon connected to a 'Stop Execution' button. Below the workflow, a 'State' dropdown menu is visible, showing a list of 'Complete' states. At the bottom, a status bar indicates 'Server: Local Server' and 'Records In: 25,544 Records Out: 43,456'.

	customer_id	data_known	gender	age	tariff	dropped_calls	handset	peak_mins
1	B247888	yes	Female	3...	CAT 50	0.000	BS110	593.600
2	B376450	yes	Female	4...	CAT 50	6.000	CAS30	134.200
3	B105000	yes	Female	6...	CAT 100	5.000	ASAD170	497.600
4	B188070	yes	Male	2...	CAT 200	0.000	550	1335.600
5	B201040	yes	Female	2...	CAT 200	2.000	ASAD170	844.600
6	B144070	yes	Female	3...	Play 300	10.000	ASAD170	334.600
7	B201620	yes	Male	1...	CAT 200	1.000	550	1243.600
8	B408820	yes	Female	1...	CAT 200	12.000	550	2346.000
9	B170490	yes	Male	2...	CAT 200	5.000	550	976.600
10	B348730	yes	Male	3...	CAT 100	1.000	CAS30	491.200
11	B173200	yes	Female	1...	CAT 200	2.000	550	927.600
12	B262270	yes	Male	4...	Play 100	0.000	BS210	64.200
13	B207600	yes	Female	1...	Play 300	2.000	550	486.600
14	B204790	yes	Male	2...	Play 100	0.000	ASAD90	122.400
15	B204390	yes	Male	2...	CAT 100	0.000	BS110	537.000
16	B302950	yes	Male	5...	Play 300	5.000	CAS40	204.600
17	B273340	yes	Male	5...	Play 300	1.000	550	304.600
18	B310150	yes	Female	1...	CAT 200	0.000	WC95	718.600
19	B313900	yes	Male	4...	CAT 200	2.000	550	637.200
20	B271610	yes	Male	5...	Play 100	1.000	550	36.600

STEP 4:

The screenshot displays the IBM SPSS Modeler software interface. On the left, a 'Table' window shows a dataset with 21 fields and 10,001 records. The table includes columns for customer_id, data_known, gender, age, tariff, dropped_calls, handset, and peak_mins. The main workspace shows a workflow diagram with a 'Table' node connected to a 'Filter' node, which is then connected to a 'Flat File' node. The right sidebar contains a 'Streams' tab with a list of tables and their record counts, and a 'Classes' tab with a project tree. The bottom toolbar includes various tool icons for data manipulation and analysis.

	customer_id	data_known	gender	age	tariff	dropped_calls	handset	peak_mins
1	N247580	yes	Female	3...	CAT 50	0.000	BS110	393.600
2	N376450	yes	Female	4...	CAT 50	6.000	CAS30	136.200
3	N109980	yes	Female	6...	CAT 100	8.000	ASAD170	657.600
4	N109980	no		6...		Smalls	Smalls	
5	N188070	yes	Male	2...	CAT 200	0.000	BS0	1335.600
6	N201040	yes	Female	2...	CAT 200	2.000	ASAD170	844.600
7	N164670	yes	Female	3...	Play 300	10.000	ASAD170	336.600
8	N201620	yes	Male	1...	CAT 200	1.000	BS0	1243.600
9	N400830	yes	Female	1...	CAT 200	12.000	BS0	2346.000
10	N170490	yes	Male	2...	CAT 200	9.000	BS0	974.600
11	N348730	yes	Male	3...	CAT 100	1.000	CAS30	481.600
12	N172280	yes	Female	1...	CAT 200	2.000	BS0	927.600
13	N262370	yes	Male	4...	Play 100	0.000	BS210	64.200
14	N207500	yes	Female	1...	Play 300	2.000	BS0	486.600
15	N204780	yes	Male	2...	Play 100	0.000	ASAD090	122.400
16	N286290	yes	Male	2...	CAT 100	0.000	BS110	537.000
17	N382950	yes	Male	5...	Play 300	9.000	CAS60	204.600
18	N273340	yes	Male	5...	Play 300	1.000	BS0	304.600
19	N310150	yes	Female	1...	CAT 200	0.000	WC95	718.600
20	N311300	yes	Male	4...	CAT 200	2.000	BS0	827.200

STEP 5:

table (24 fields, 4,518 records)

File Edit Generate

Annotations

customer_id	data_known	gender	age	tariff	dropped_calls	handset	peak_mins
N376450	yes	Female	4...	CAT 50	6.000	CAS30	136.200
N166070	yes	Male	2...	CAT 200	0.000	S50	1335.600
N201620	yes	Male	1...	CAT 200	1.000	S50	1243.800
N408820	yes	Female	1...	CAT 200	12.000	S50	2346.000
N348730	yes	Male	3...	CAT 100	1.000	CAS30	481.200
N173280	yes	Female	1...	CAT 200	2.000	S50	92.2346.0
N207500	yes	Female	1...	Play 300	2.000	S50	486.600
N204780	yes	Male	2...	Play 100	0.000	ASAD90	122.400
N214430	yes	Female	4...	Play 100	2.000	S50	54.000
N299310	yes	Female	2...	CAT 100	8.000	CAS30	405.000
N239620	yes	Male	5...	Play 100	2.000	BS110	234.600
N294430	yes	Female	2...	CAT 100	0.000	CAS30	356.400
N407710	yes	Male	5...	CAT 200	13.000	BS110	1433.400
N104380	yes	Female	3...	CAT 100	8.000	S50	622.800
N376800	yes	Male	1...	Play 100	1.000	ASAD90	213.600
N146120	yes	Female	4...	CAT 200	0.000	ASAD90	993.000
N369160	yes	Female	2...	CAT 200	6.000	CAS30	818.400
N213340	yes	Female	2...	Play 100	0.000	CAS30	43.085
N291760	yes	Female	2...	CAT 100	0.000	CAS30	392.400
N156940	yes	Female	2...	Play 100	1.000	S50	168.600

OK

Table

Table

Table

Table

Table

Table

Filter

Flat File

Database Var. File Auto Data Prep Select Sample Aggregate Derive Type Filter Graphboard Auto Classifier Auto Numeric Auto Cluster Table Flat File Database

Server: Local Server

STEP 6:

The screenshot displays the IBM SPSS Modeler interface. On the left, a 'Table' window shows a dataset with 23 fields and 31,789 records. The 'Annotations' tab is active, showing a table with columns: customer_id, data_known, gender, age, tariff, dropped_calls, handset, and peak_mins. The data is organized into rows, with some rows highlighted. Below the table, there are four 'Table' icons in a row.

On the right, a workflow diagram is shown. It starts with a 'churn' node (a yellow diamond) connected to a 'Table' node (a blue square). The 'Table' node is connected to a 'Filter' node (a blue hexagon). The 'Filter' node is connected to a 'Flat File' node (a blue square). The 'Filter' node also has a 'churn' node (a yellow diamond) connected to it. The 'Flat File' node is connected to a 'Table' node (a blue square).

The bottom of the interface shows a toolbar with various icons for data preparation and analysis, including 'Database', 'Var. File', 'Auto Data Prep', 'Select', 'Sample', 'Aggregate', 'Derive', 'Type', 'Filter', 'Graphboard', 'Auto Classifier', 'Auto Numeric', 'Auto Cluster', 'Table', 'Flat File', and 'Database'.

	customer_id	data_known	gender	age	tariff	dropped_calls	handset	peak_mins
1	K100010	yes	Male	4...	CAT 50	1.000	SOP10	36.131
2	K100020	yes	Male	2...	CAT 50	0.000	SOP10	39.437
3	K100030	yes	Male	3...	CAT 50	2.000	SOP20	72.600
4	K100040	yes	Male	2...	CAT 50	2.000	SOP10	72.600
5	K100050	yes	Male	4...	CAT 50	0.000	SOP10	40.608
6	K100060	yes	Male	2...	CAT 50	1.000	SOP10	46.260
7	K100070	yes	Male	3...	CAT 50	1.000	SOP20	56.370
8	K100080	yes	Male	2...	CAT 50	0.000	SOP20	51.043
9	K100090	yes	Male	3...	CAT 50	1.000	SOP20	99.000
10	K100100	yes	Male	4...	CAT 50	0.000	SOP30	45.400
11	K100110	yes	Male	3...	CAT 50	2.000	SOP20	45.400
12	K100120	yes	Male	3...	CAT 50	0.000	SOP20	48.123
13	K100130	yes	Male	4...	CAT 50	2.000	SOP20	43.352
14	K100140	yes	Male	2...	CAT 50	1.000	SOP20	41.447
15	K100150	yes	Male	3...	CAT 50	2.000	SOP20	56.000
16	K100160	yes	Male	2...	CAT 50	9.000	SOP20	41.342
17	K100170	yes	Male	2...	CAT 50	1.000	SOP10	100.200
18	K100180	yes	Male	3...	CAT 50	1.000	SOP10	41.916
19	K100190	yes	Male	3...	CAT 50	2.000	SOP10	58.782
20	K100200	yes	Male	4...	CAT 50	0.000	SOP10	84.600