BABU BANARASI DAS UNIVERSITY

School Of Computer Application

Academic Session 2023 – 2024

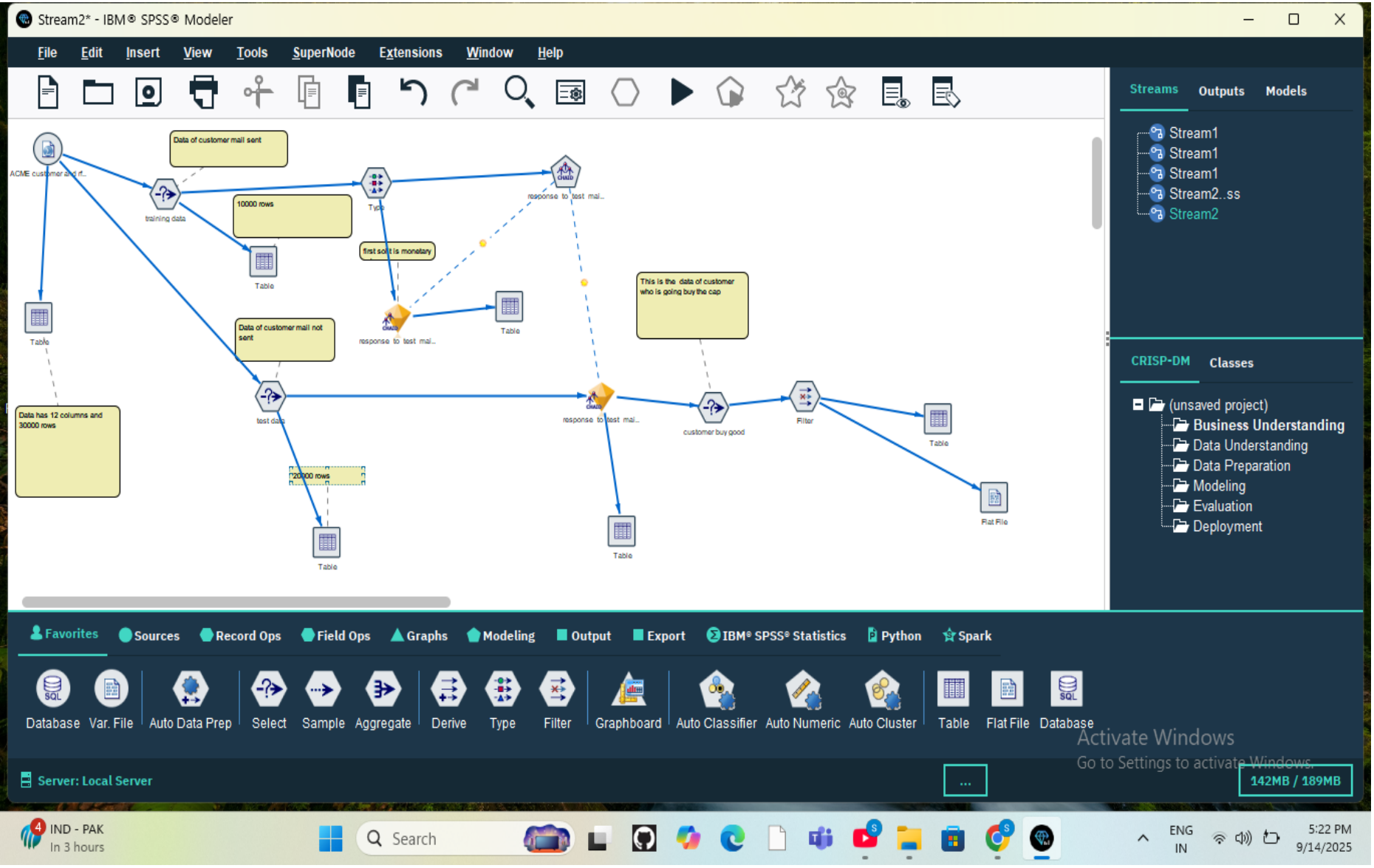
Name :Tejas pandey
Section :BCADS36
Roll No:1230258452
Assignment:Predictive Analytics
Semester:5th
Date:29 october 2025
Submit To :Mr. Ayushman bhadhuria sir

ACME CASE STUDY: Predicting Customer Response

Definition: Predictive Modelling is a statistical technique that uses Historical data to predict future outcomes.It help in identify patterns and relationship within data to make informed business decisions .

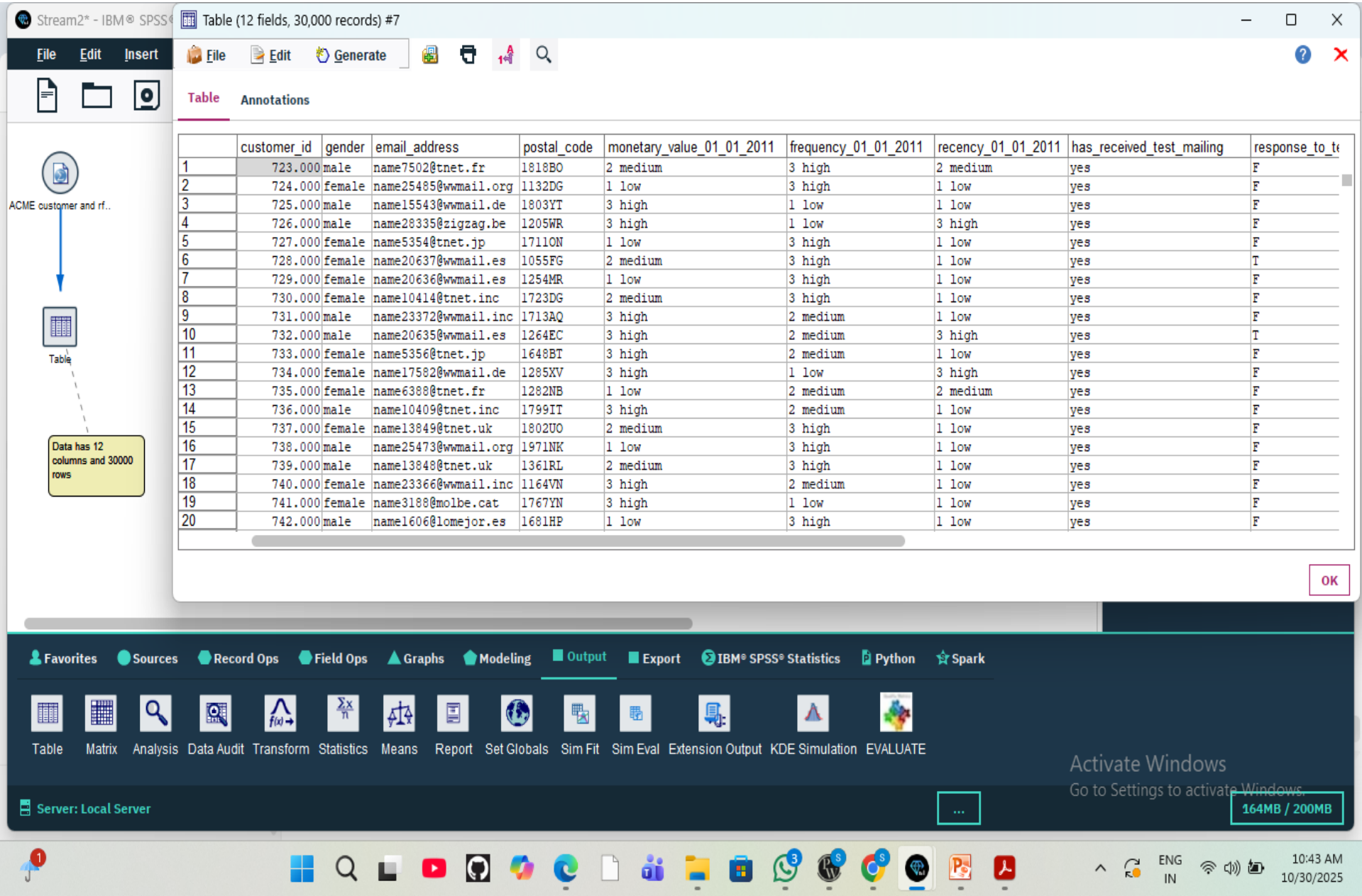
Outcome: Learned how to import and explore data in **SPSS MODELER** Understand how to identify important variables.

Required tools: SPSS MODELER



This image show the complete predictive modelling workflow in SPSS MODELER , illustrating the process from data import and preapartion to model building,testing,and deployment for customer response prediction

Step 1:



	customer_id	gender	email_address	postal_code	monetary_value_01_01_2011	frequency_01_01_2011	recency_01_01_2011	has_received_test_mailing	response_to_te
1	723.000	male	name7502@tnet.fr	1818BO	2 medium	3 high	2 medium	yes	F
2	724.000	female	name25485@wwmail.org	1132DG	1 low	3 high	1 low	yes	F
3	725.000	male	name15543@wwmail.de	1803YT	3 high	1 low	1 low	yes	F
4	726.000	male	name28335@zigzag.be	1205WR	3 high	1 low	3 high	yes	F
5	727.000	female	name5354@tnet.jp	1711ON	1 low	3 high	1 low	yes	F
6	728.000	female	name20637@wwmail.es	1055FG	2 medium	3 high	1 low	yes	T
7	729.000	female	name20636@wwmail.es	1254MR	1 low	3 high	1 low	yes	F
8	730.000	female	name10414@tnet.inc	1723DG	2 medium	3 high	1 low	yes	F
9	731.000	male	name23372@wwmail.inc	1713AQ	3 high	2 medium	1 low	yes	F
10	732.000	male	name20635@wwmail.es	1264EC	3 high	2 medium	3 high	yes	T
11	733.000	female	name5356@tnet.jp	1648BT	3 high	2 medium	1 low	yes	F
12	734.000	female	name17582@wwmail.de	1285XV	3 high	1 low	3 high	yes	F
13	735.000	female	name6388@tnet.fr	1282NB	1 low	2 medium	2 medium	yes	F
14	736.000	male	name10409@tnet.inc	1799IT	3 high	2 medium	1 low	yes	F
15	737.000	female	name13849@tnet.uk	1802UO	2 medium	3 high	1 low	yes	F
16	738.000	male	name25473@wwmail.org	1971NK	1 low	3 high	1 low	yes	F
17	739.000	male	name13848@tnet.uk	1361RL	2 medium	3 high	1 low	yes	F
18	740.000	female	name23366@wwmail.inc	1164VN	3 high	2 medium	1 low	yes	F
19	741.000	female	name3188@molbe.cat	1767YN	3 high	1 low	1 low	yes	F
20	742.000	male	name1606@lomejor.es	1681HP	1 low	3 high	1 low	yes	F

Data Overview (Understanding the data).

a. Open SPSS Modeler, import the training dataset (Excel/CSV statistical).

b. Connect a Table node to summarize.
Check how many rows .Check how many columns

c. Identify important fields:
Response →Target (Yes/No: Customer bought or not)
Gender, recency – 01.01.2011, frequency – 01.01.2011, monetary-value – 01.01.2011 → input

Step 2: Test Mailing Customer (Filtering)

a. Use a filter node.

b. b. Keep only customers who were in the test mailing (has_received_test_mail = 1).

Stream2* - IBM® SPSS® Modeler

File Edit Insert View Tools SuperNode

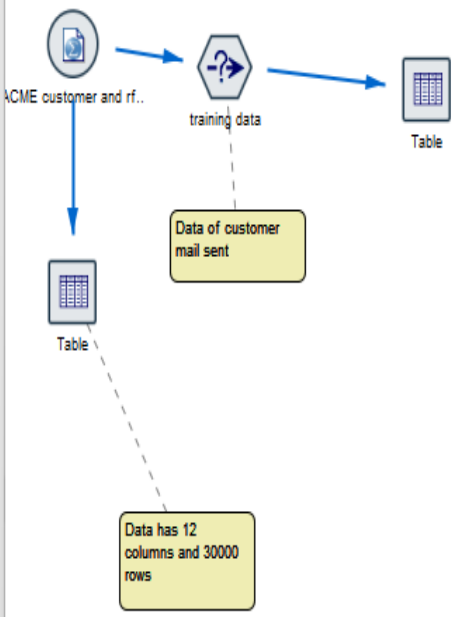


Table (12 fields, 10,000 records) #2

File Edit Generate

Table Annotations

	customer_id	gender	email_address	postal_code	monetary_value_01_01_2011	frequency_01_01_2011	recency_01_01_2011	has_rece
1	723.000	male	name7502@tnet.fr	1818BO	2 medium	3 high	2 medium	yes
2	724.000	female	name25485@wwmail.org	1132DG	1 low	3 high	1 low	yes
3	725.000	male	name15543@wwmail.de	1803YT	3 high	1 low	1 low	yes
4	726.000	male	name28335@zigzag.be	1205WR	3 high	1 low	3 high	yes
5	727.000	female	name5354@tnet.jp	1711ON	1 low	3 high	1 low	yes
6	728.000	female	name20637@wwmail.es	1055FG	2 medium	3 high	1 low	yes
7	729.000	female	name20636@wwmail.es	1254MR	1 low	3 high	1 low	yes
8	730.000	female	name10414@tnet.inc	1723DG	2 medium	3 high	1 low	yes
9	731.000	male	name23372@wwmail.inc	1713AQ	3 high	2 medium	1 low	yes
10	732.000	male	name20635@wwmail.es	1264EC	3 high	2 medium	3 high	yes
11	733.000	female	name5356@tnet.jp	1648BT	3 high	2 medium	1 low	yes
12	734.000	female	name17582@wwmail.de	1285XV	3 high	1 low	3 high	yes
13	735.000	female	name6388@tnet.fr	1282NB	1 low	2 medium	2 medium	yes
14	736.000	male	name10409@tnet.inc	1799IT	3 high	2 medium	1 low	yes
15	737.000	female	name13849@tnet.uk	1802UO	2 medium	3 high	1 low	yes
16	738.000	male	name25473@wwmail.org	1971NK	1 low	3 high	2 medium	yes
17	739.000	male	name13848@tnet.uk	1361RL	2 medium	3 high	1 low	yes
18	740.000	female	name23366@wwmail.inc	1164VN	3 high	2 medium	1 low	yes
19	741.000	female	name3188@molbe.cat	1767YN	3 high	1 low	1 low	yes
20	742.000	male	name1606@lomejor.es	1681HP	1 low	3 high	1 low	yes

OK

Favorites Sources Record Ops Field Ops Graphs Modeling Output Export IBM® SPSS® Statistics Python Spark

Table Matrix Analysis Data Audit Transform Statistics Means Report Set Globals Sim Fit Sim Eval Extension Output KDE Simulation EVALUATE

Server: Local Server

...

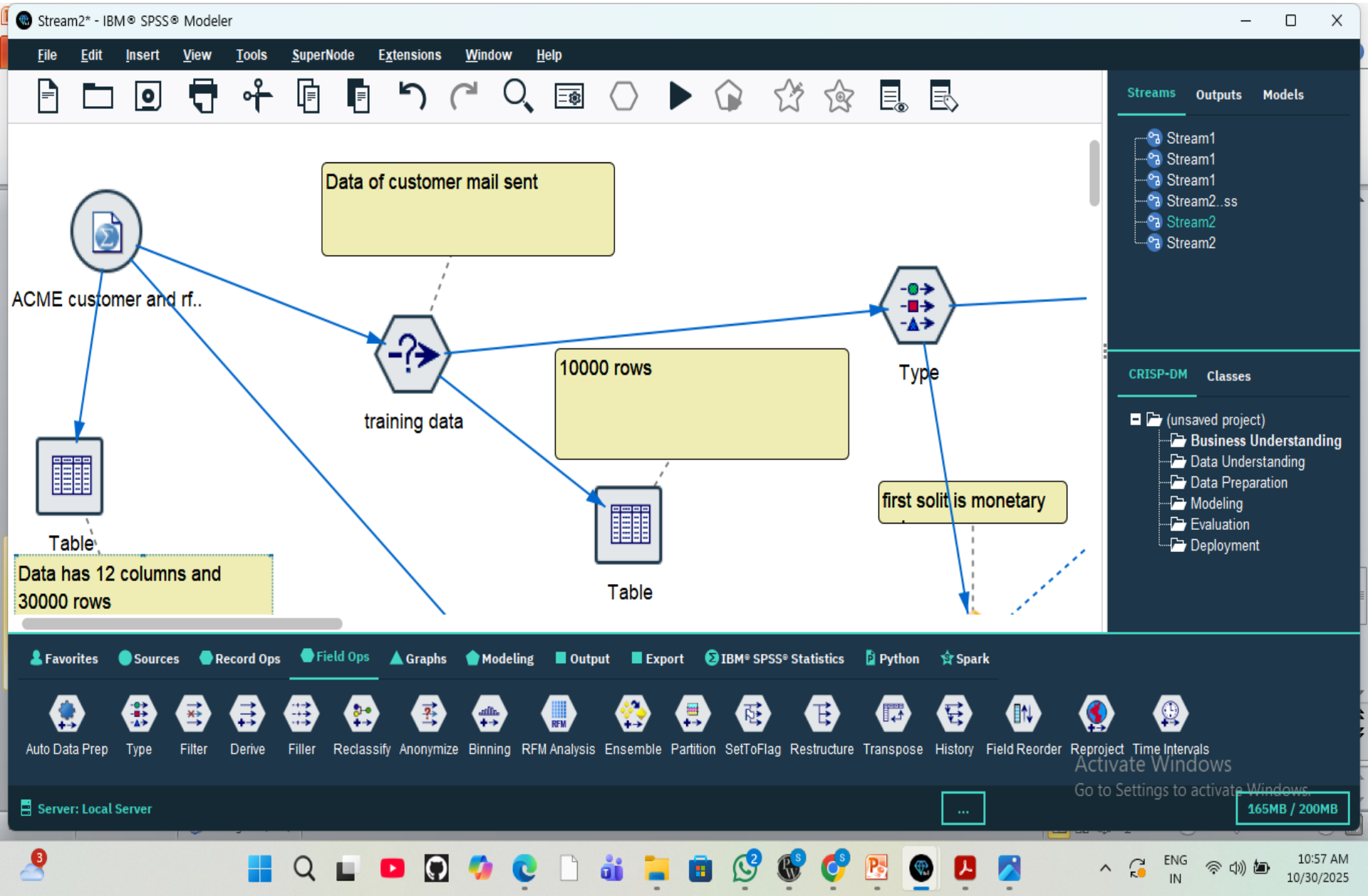
159MB / 200MB

Search

ENG IN

10:38 AM 10/30/2025

Step 3:


Stream2* - IBM® SPSS® Modeler
File Edit Insert View Tools SuperNode Extensions Window Help
Streams Outputs Models
Stream1
Stream1
Stream1
Stream2..ss
Stream2
Stream2
Data of customer mail sent
ACME customer and rf..
training data
10000 rows
Type
Table
Data has 12 columns and 30000 rows
Table
first solit is monetary
CRISP-DM Classes
(unsaved project)
Business Understanding
Data Understanding
Data Preparation
Modeling
Evaluation
Deployment
Favorites Sources Record Ops Field Ops Graphs Modeling Output Export IBM® SPSS® Statistics Python Spark
Auto Data Prep Type Filter Derive Filler Reclassify Anonymize Binning RFM Analysis Ensemble Partition SetToFlag Restructure Transpose History Field Reorder Reproject Time Intervals
Server: Local Server
165MB / 200MB
Activate Windows
Go to Settings to activate Windows.
ENG IN
10:57 AM
10/30/2025

Predictive Modelling.

a. Add a Type node – Define roles:
   Response → Target
   Other customer details → Input
b. Train the model to predict response.—

Step 4: Model Output (Checking results)

a. Connect a Table node downstream of the trained model nugget.

b. Look at the new fields created by the model:

Predicted field → model's prediction

Confidence → probability score for each prediction

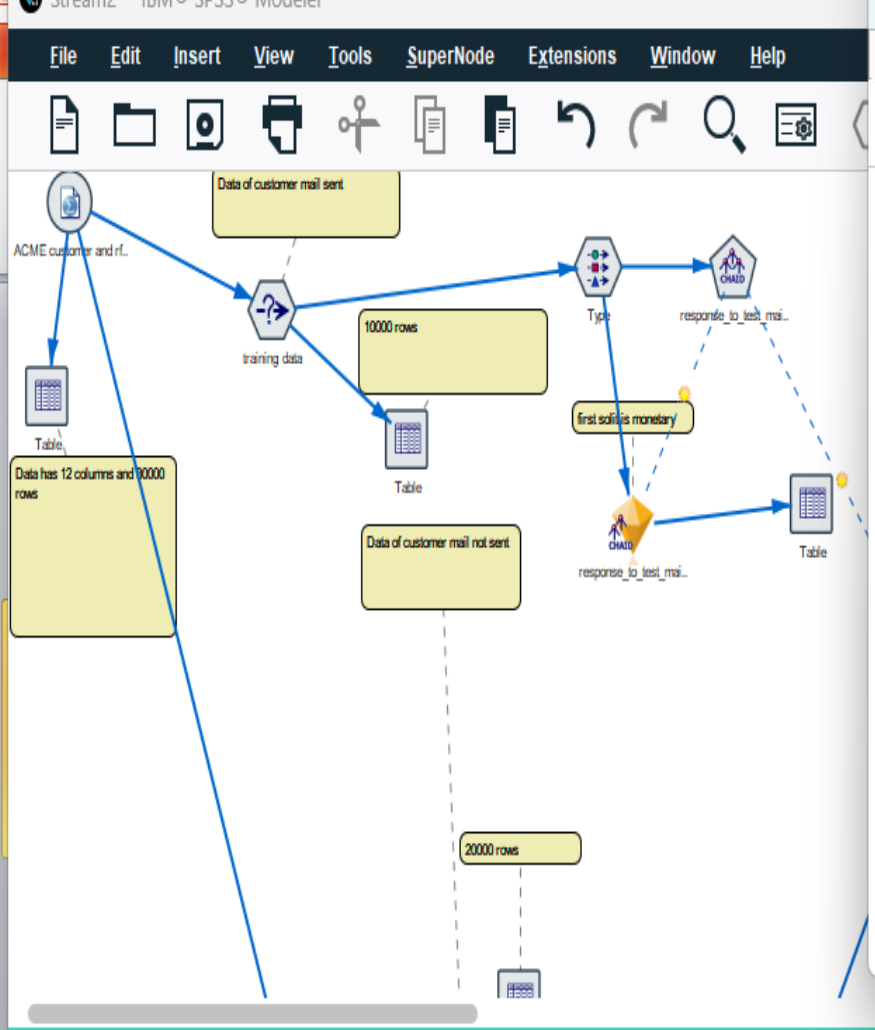


Table (14 fields, 10,000 records) #2

File Edit Generate

Table Annotations

	her_id	gender	email_address	postal_code	monetary_value_01_01_2011	frequency_01_01_2011
1	23.000	male	name7502@tnet.fr	1818BO	2 medium	3 high
2	24.000	female	name25485@wwmail.org	1132DG	1 low	3 high
3	25.000	male	name15543@wwmail.de	1803YT	3 high	1 low
4	26.000	male	name28335@zigzag.be	1205WR	3 high	1 low
5	27.000	female	name5354@tnet.jp	1711ON	1 low	3 high
6	28.000	female	name20637@wwmail.es	1055FG	2 medium	3 high
7	29.000	female	name20636@wwmail.es	1254MR	1 low	3 high
8	30.000	female	name10414@tnet.inc	1723DG	2 medium	3 high
9	31.000	male	name23372@wwmail.inc	1713AQ	3 high	2 medium
10	32.000	male	name20635@wwmail.es	1264EC	3 high	2 medium
11	33.000	female	name5356@tnet.jp	1648BT	3 high	2 medium
12	34.000	female	name17582@wwmail.de	1285XV	3 high	1 low
13	35.000	female	name6388@tnet.fr	1282NB	1 low	2 medium
14	36.000	male	name10409@tnet.inc	1799IT	3 high	2 medium
15	37.000	female	name13849@tnet.uk	1802UO	2 medium	3 high
16	38.000	male	name25473@wwmail.org	1971NK	1 low	3 high
17	39.000	male	name13848@tnet.uk	1361RL	2 medium	3 high
18	40.000	female	name23366@wwmail.inc	1164VN	3 high	2 medium
19	41.000	female	name3188@molbe.cat	1767YN	3 high	1 low
20	42.000	male	name1606@lomejor.es	1681HP	1 low	3 high

OK

Favorites Sources Record Ops Field Ops Graphs Modeling Output Export IBM® SPSS® Statistics Python Spark

Auto Data Prep Type Filter Derive Filler Reclassify Anonymize Binning RFM Analysis Ensemble Partition SetToFlag Restructure Transpose History Field Reorder Reproject Time Intervals

Server: Local Server

170MB / 200MB

ENG 11:06 AM

Step 5: Applying the Model

a. Import the testing dataset (customers who did not receive the mailing).

b. Apply the trained model to this test dataset.

c. The model will generate predictions for test customers.

d. Check how many customers are predicted to respond positive (predicted = T).

Step 6: Exporting Results (Deployment)

a. Use a filter node to keep only customers predicted as positive.

b. Export results using a Flat File Node → save as customer-to-contact.txt.

c. Include only required fields:

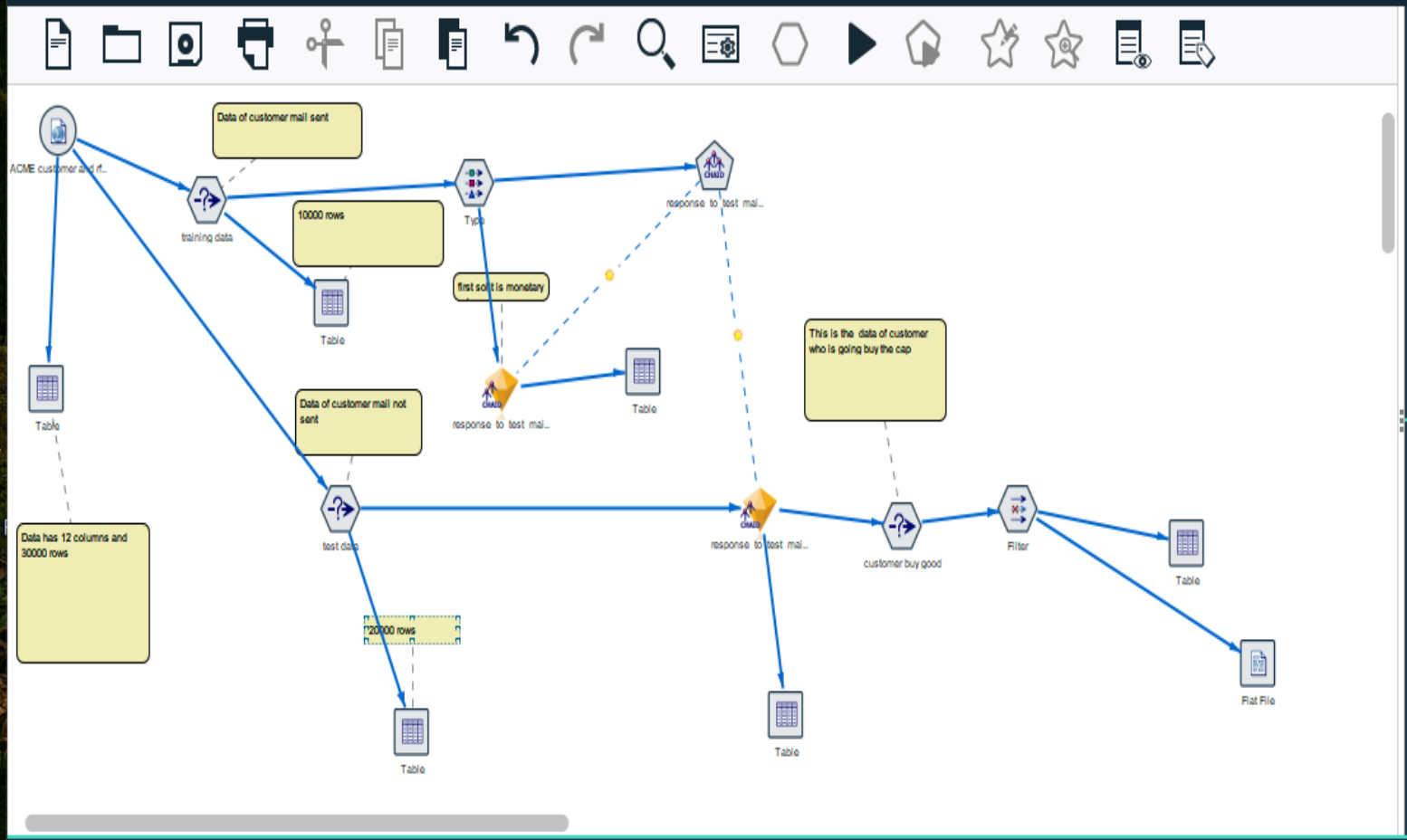
Predicted category (rename to predicted_category)

Customer_id

Confidence score (rename to confidence_score)

Stream2* - IBM® SPSS® Modeler

File Edit Insert View Tools SuperNode Extensions Window Help



Streams Outputs Models

- Stream1
- Stream1
- Stream1
- Stream2..ss
- Stream2

CRISP-DM Classes

- (unsaved project)
  - Business Understanding
  - Data Understanding
  - Data Preparation
  - Modeling
  - Evaluation
  - Deployment

Favorites Sources Record Ops Field Ops Graphs Modeling Output Export IBM® SPSS® Statistics Python Spark

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

142MB / 189MB