

BABU BANARASI DAS UNIVERSITY
School of Computer Application



CASE STUDY REPORT
Predictive Analytics: Predicting Customer Response using
SPSS Modeler

SUBMITTED TO:

Mr. Ayushman Bhadauria

SUBMITTED BY:

Name: Anshuman Singh

Roll No: 1230258087

Section: BCA DS 32

Definition

The main objective of this case study is to use predictive analytics techniques in IBM SPSS Modeler to analyze and predict customer responses in a marketing campaign. The project demonstrates how data-driven insights can help improve business decisions.

Outcomes / Learning

After completing this project, the student will be able to perform predictive analysis using SPSS Modeler and interpret model outputs to inform marketing decisions.

Required Tool

IBM SPSS Modeler – A visual data science and machine learning tool for data preparation, modeling, and deployment.

Working

SPSS Modeler is used to import data, clean and filter records, define variable roles, train classification models, evaluate their performance, and export predicted results for deployment.

Step 1: Data Overview (Understanding the Data)

Import the dataset using the Var. File or Database node. Use a Table node to inspect records and fields. Identify key fields such as Response (target), Gender, Recency, Frequency, and Monetary value.

The screenshot displays the IBM SPSS Stream2 interface. The main window shows a table with 12 fields and 30,000 records. The table is titled "Table (12 fields, 30,000 records) #7". The fields are: customer_id, gender, email_address, postal_code, monetary_value_01_01_2011, frequency_01_01_2011, recency_01_01_2011, has_received_test_mailing, and response_to_t. The table is displayed in a grid format with 20 rows visible. The interface includes a menu bar (File, Edit, Insert, File, Edit, Generate) and a toolbar with icons for file operations, editing, and generating reports. A sidebar on the left shows a diagram of the data flow, including a "Table" node and a "Data has 12 columns and 30000 rows" node. The bottom of the interface features a navigation bar with tabs for Favorites, Sources, Record Ops, Field Ops, Graphs, Modeling, Output, Export, IBM® SPSS® Statistics, Python, and Spark. The bottom right corner shows the status bar with "Server: Local Server" and a memory usage indicator "164MB / 200MB".

	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_t
1	723.000	male	name7502@tinet.fr	1018BO	2 medium	3 high	2 medium	yes	F
2	724.000	female	name25405@vnmmail.org	1132DG	1 low	3 high	1 low	yes	F
3	725.000	male	name15549@vnmmail.de	1003VT	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@tinet.jp	1711ON	1 low	3 high	1 low	yes	F
6	728.000	female	name20637@vnmmail.es	1055FG	2 medium	3 high	1 low	yes	T
7	729.000	female	name20637@vnmmail.es	1254MR	1 low	3 high	1 low	yes	F
8	730.000	female	name10414@tinet.inc	1721DG	2 medium	3 high	1 low	yes	F
9	731.000	male	name23372@vnmmail.inc	1713AQ	3 high	2 medium	1 low	yes	F
10	732.000	male	name20635@vnmmail.es	1264EC	3 high	2 medium	3 high	yes	T
11	733.000	female	name5356@tinet.jp	1648BT	3 high	2 medium	1 low	yes	F
12	734.000	female	name17502@vnmmail.de	1285XV	3 high	1 low	3 high	yes	F
13	735.000	female	name6308@tinet.fr	1292MB	1 low	2 medium	2 medium	yes	F
14	736.000	male	name10409@tinet.inc	1799IT	3 high	2 medium	1 low	yes	F
15	737.000	female	name13649@tinet.uk	1602DO	2 medium	3 high	1 low	yes	F
16	738.000	male	name25473@vnmmail.org	1071MK	1 low	3 high	1 low	yes	F
17	739.000	male	name13648@tinet.uk	1361RL	2 medium	3 high	1 low	yes	F
18	740.000	female	name23366@vnmmail.inc	1164TN	3 high	2 medium	1 low	yes	F
19	741.000	female	name3108@molbe.cat	1767TN	3 high	1 low	1 low	yes	F
20	742.000	male	name1606@lomejor.es	1681HP	1 low	3 high	1 low	yes	F

Step 1: data overview (understanding the data).

Step 2: Test Mailing Customers (Filtering)

Apply a Filter node to select only customers who received the test mailing (has_received_test_mail = 1) to ensure relevant training data.

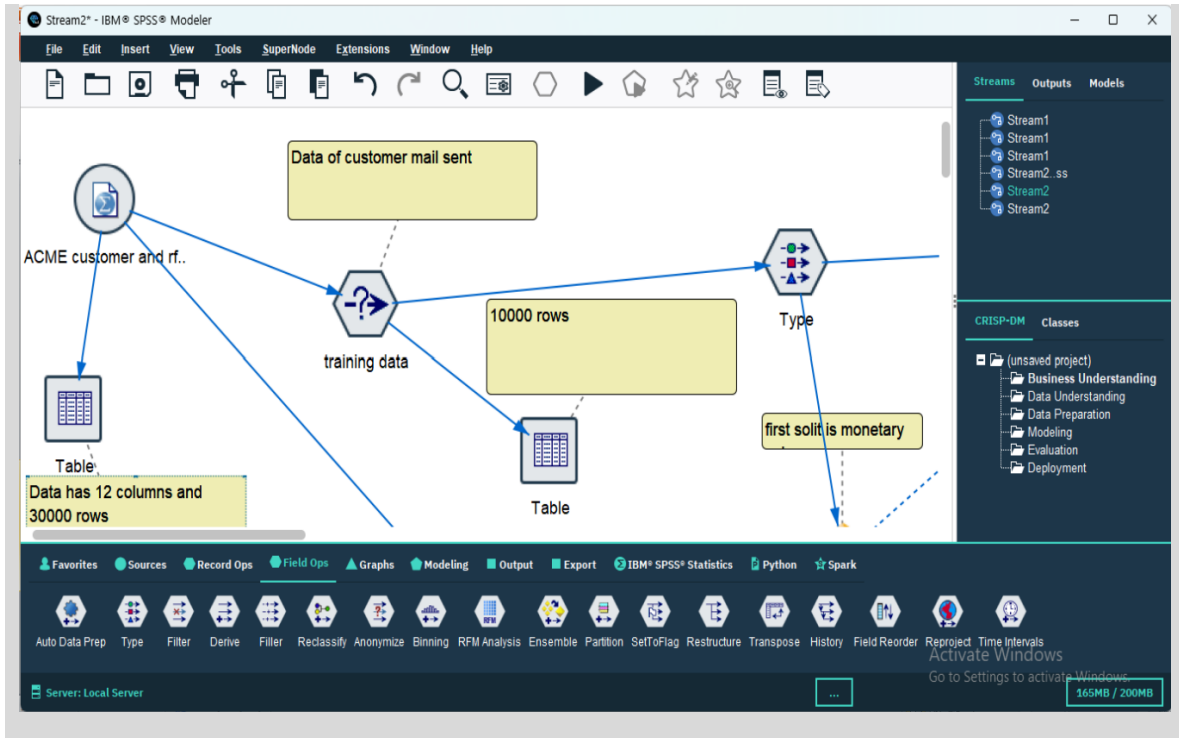
The screenshot displays the IBM SPSS Modeler interface. On the left, a workflow diagram shows a 'Table' node connected to a 'Filter' node, which is then connected to another 'Table' node. A yellow callout box indicates 'Data has 12 columns and 30000 rows'. On the right, a data table is shown with the following columns: customer_id, gender, email address, postal code, monetary_value_01_01_2011, frequency_01_01_2011, recency_01_01_2011, and has_rece. The table contains 20 rows of data. The bottom of the interface shows a toolbar with various icons and a status bar indicating 'Server: Local Server' and '159MB / 200MB'.

	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@ttnet.fr	1818BO	2 medium	3 high	2 medium	yes
2	724.000	female	name25485@vnmmail.org	1132DG	1 low	3 high	1 low	yes
3	725.000	male	name15543@vnmmail.de	1803IT	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@ttnet.jp	1711ON	1 low	3 high	1 low	yes
6	728.000	female	name20637@vnmmail.es	1055FG	2 medium	3 high	1 low	yes
7	729.000	female	name20637@vnmmail.es	1254MR	1 low	3 high	1 low	yes
8	730.000	female	name10414@ttnet.inc	1723DG	2 medium	3 high	1 low	yes
9	731.000	male	name23372@vnmmail.inc	1713AQ	3 high	2 medium	1 low	yes
10	732.000	male	name20635@vnmmail.es	1264EC	3 high	2 medium	3 high	yes
11	733.000	female	name5356@ttnet.jp	1648BT	3 high	2 medium	1 low	yes
12	734.000	female	name17582@vnmmail.de	1285XV	3 high	1 low	3 high	yes
13	735.000	female	name6388@ttnet.fr	1282NB	1 low	2 medium	2 medium	yes
14	736.000	male	name10409@ttnet.inc	1799IT	3 high	2 medium	1 low	yes
15	737.000	female	name13849@ttnet.uk	1802UO	2 medium	3 high	1 low	yes
16	738.000	male	name25473@vnmmail.org	1971NK	1 low	3 high	2 medium	yes
17	739.000	male	name13848@ttnet.uk	1361RL	2 medium	3 high	1 low	yes
18	740.000	female	name23366@vnmmail.inc	1164VN	3 high	2 medium	1 low	yes
19	741.000	female	name3188@molbe.cat	1767YH	3 high	1 low	1 low	yes
20	742.000	male	name1606@lomejor.es	1681HP	1 low	3 high	1 low	yes

step 2: test mailing customers (filtering).

Step 3: Defining Variables / Building Model

Use a Type node to set variable roles (Target/Input). Train a classification model such as Decision Tree or Logistic Regression to predict customer response.



3: defining variables / building model.

Step 4: Model Output (Checking Results)

Review the model nugget and output table showing predicted responses and confidence scores. Use evaluation nodes to check accuracy and confusion matrix.

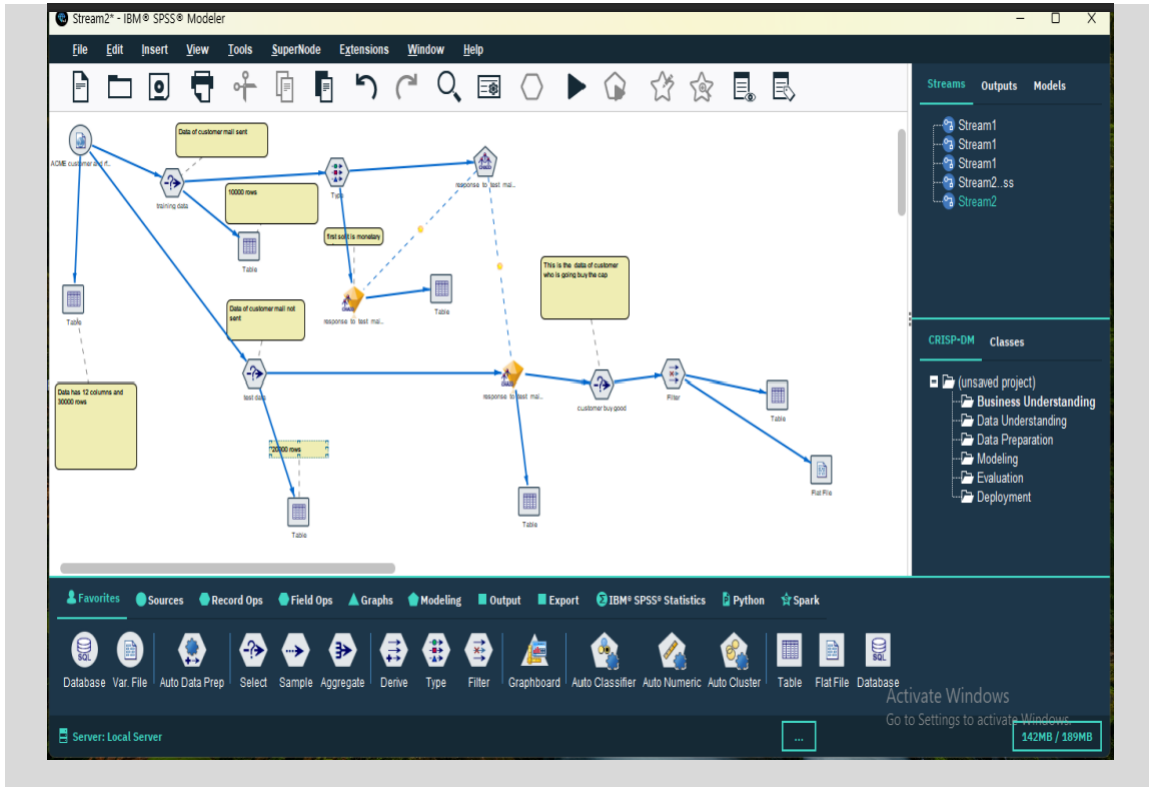
The screenshot displays the Alteryx Designer interface. On the left, a workflow is visible with nodes for data input, processing, and output. The main window shows a table of data with columns: 'id', 'gender', 'email address', 'postal code', 'monetary_value_01_01_2011', and 'frequency_01_01_2011'. The table contains 20 rows of data. The bottom status bar indicates 'Server: Local Server' and '470MB / 200MB'.

	id	gender	email address	postal code	monetary_value_01_01_2011	frequency_01_01_2011
1	23.000	male	name7502@tinet.fr	1818BO	2 medium	3 high
2	24.000	female	name25485@vnmall.org	1132DG	1 low	3 high
3	25.000	male	name15543@vnmall.de	1803YT	3 high	1 low
4	26.000	male	name28335@rigzag.be	1205WR	3 high	1 low
5	27.000	female	name5354@tinet.jp	1711ON	1 low	3 high
6	28.000	female	name20637@vnmall.es	1055FG	2 medium	3 high
7	29.000	female	name20636@vnmall.es	1254HR	1 low	3 high
8	30.000	female	name10414@tinet.inc	1723DG	2 medium	3 high
9	31.000	male	name23372@vnmall.inc	1713AQ	3 high	2 medium
10	32.000	male	name20635@vnmall.es	1264EC	3 high	2 medium
11	33.000	female	name5356@tinet.jp	1648BT	3 high	2 medium
12	34.000	female	name17582@vnmall.de	1285KV	3 high	1 low
13	35.000	female	name6388@tinet.fr	1282NB	1 low	2 medium
14	36.000	male	name10409@tinet.inc	1799IT	3 high	2 medium
15	37.000	female	name13849@tinet.uk	1802VO	2 medium	3 high
16	38.000	male	name25473@vnmall.org	1971NR	1 low	3 high
17	39.000	male	name13848@tinet.uk	1361RL	2 medium	3 high
18	40.000	female	name23366@vnmall.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

step 4: model output (checking results).

Step 5: Applying the Model

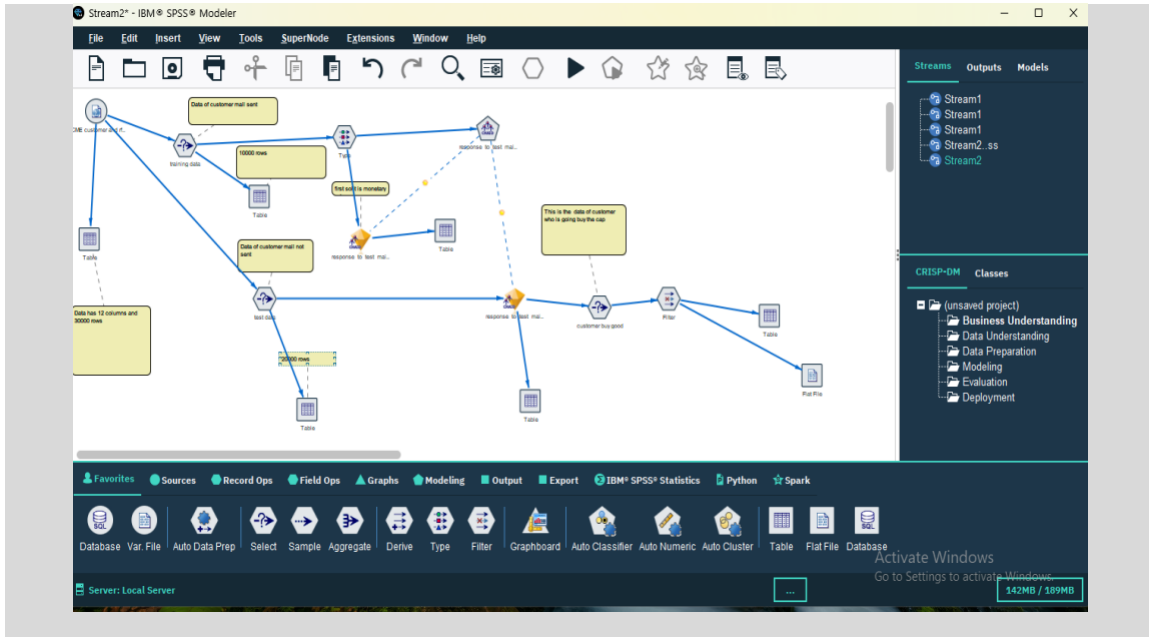
Apply the trained model to the holdout/test dataset to generate predictions for customers who did not receive the mail.



step 5: applying the model.

Step 6: Exporting Results (Deployment)

Export the predicted positive customers to a flat file (customer_to_contact.txt) including fields: customer_id, predicted_category, confidence_score.



step 6: exporting results (deployment).

Result / Discussion

The model identified key predictors of customer response such as recency and frequency. Predictions with high confidence should be prioritized for outreach. Model performance metrics (accuracy, precision, recall) should be reviewed and documented once the model is trained.

Conclusion

This case study demonstrated end-to-end predictive modeling using IBM SPSS Modeler. The workflow from data import to deployment provides a template for future marketing analytics tasks.