

## Using analytics to prevent fraud allows HDI to have a fast and real time approval for Claims

### SAS® Global Forum 2017

Rayani Melega, HDI Seguros

SAS® Real Time Decision Manager (RTDM) combines business strategy and analytics delivering real time recommendations and decisions to interactive customer channels, such as call centers or, in this case, claim services channel.

Even though it was designed for marketers, its applications can be used in businesses that demand a quick (or even online) and interactive response for their customers after an analysis of the customer needs and behavior. The main benefits from the use of RTDM are the reduction of dependency on IT resources and automatization of accurate decision processes, boosting profitability - with a friendly interface.

SAS® Fraud Framework (SFF) consists of many SAS® products. Between them, there are two fraud-specific products: SAS® Financial Crimes Monitor (FCM) and, the one applied in the project of this paper, SAS® Social Network Analysis Server (SNA).

The SAS® SNA is an investigator interface responsible for viewing and managing alerts. These alerts interact with business rules set in RTDM with the creation of, for example, criteria such as if the rule is solved for one specific customer data, then an alert is given. Otherwise, nothing is shown.

SNA has also the capability of displaying network diagrams, showing the relationships associated with the selected alert.

SAS® Enterprise Guide (SAS® Guide) is the software used to access historical stored data from RTDM. Besides this, it is also used to prepare data bases for statistical modeling prediction. It is designed for the use of business analysts, programmers and statisticians.

SAS® Enterprise Miner Client (SAS® Miner) enables users to understand data through creating predictive and descriptive analytical models using large datasets in a very friendly interface. The data prepared on SAS® Guide uploaded and accessed as metadata in Miner will be able to go through a statistical modeling process (with different techniques), cluster analysis, text mining and many other analysis in just a few clicks. Statistical models developed in Miner can be “packed” into a .spk extension file and published in SAS® Decision Manager to interact with rules in RTDM.

SAS® Decision Manager takes data, business rules and models to provide automated operations decisions in an organized and efficient way. It is responsible for integrating statistical models with business rules in the decision process.

### ABSTRACT

The objective of this paper is to introduce the project for **Insurance Fraud Detection**, developed in HDI Seguros | Brazil, that aims at improving fraud prevention in claim processes, implementing SAS® Anti-Fraud platform, composed by SAS® SFF and SAS® RTDM solutions. The main goals are:

- To support online decision making;
- To reduce false-positives and improve the accuracy of claim classification (the decision to set a claim as “to be investigated” or not);
- To design an interactive dashboard to manage claims to be investigated;

- To use relationship networks to visualize and investigate suspicious connections;
- To identify the best analysts to analyze each claim, depending on one's specialization;
- To include the business expertise into business rules;
- To select claims with high propensity to be a fraud with models developed by HDI Analytics team;
- To consult multiple datasets simultaneously;
- To manage a list of suspects, avoiding new activities, such as new policy issues or improper payments;
- To integrate with external databases, avoiding manual searches;
- Self-management of administrator functions, such as implement new business rules or new predictive models, depending on business demands;
- To manage all the alerts and investigations in one unique platform.

## INTRODUCTION

HDI Seguros company is part of the German Talanx Group and its main shareholder is Talanx International AG. Situated in Hannover and acting over 150 countries, Talanx Group is one of the major insurance groups in Europe, with an income from premiums of 31,8 billion euros in 2015.

Operating in Brazil over more than 35 years, HDI Seguros | Brazil has today a structure of 55 branches, 15 sales offices, 50 HDI "Bate-Prontos" (which are appraisal centers strategically located prepared to assist policy holders in case of motor claims) and 6 "Bate-Prontos" Mobile units, also for assistance attendance.

HDI Seguros | Brazil acts through national territory, mostly in massify insurance for Motor and Homeowners, holding 1.686.857 vehicles and 448.687 residences. It is the fifth motor insurance company in national market share ranking, holding a 8,8% of market share (data from 2016, December). The company ended 2016 with a gross written premium of R\$3 billion and a result before goodwill impairments of R\$154 million.

In 2016, 291.427 motor claims were assisted:

- 57,3% in HDI "Bate-Prontos" with service time under 30 minutes;
- 42,7% in call centers, with quality service indicators as expected.

In HDI Seguros | Brazil, that has approximately 25.000 motor claims per month all the claims are pre-analyzed and 3% are referred to investigation. From that, only 50% are proven as fraud or with any irregularity that decline the claim. It means that the fraud rate is only 1,5%. The analysis require a huge task force from analysts and depends on personal skills and experience.

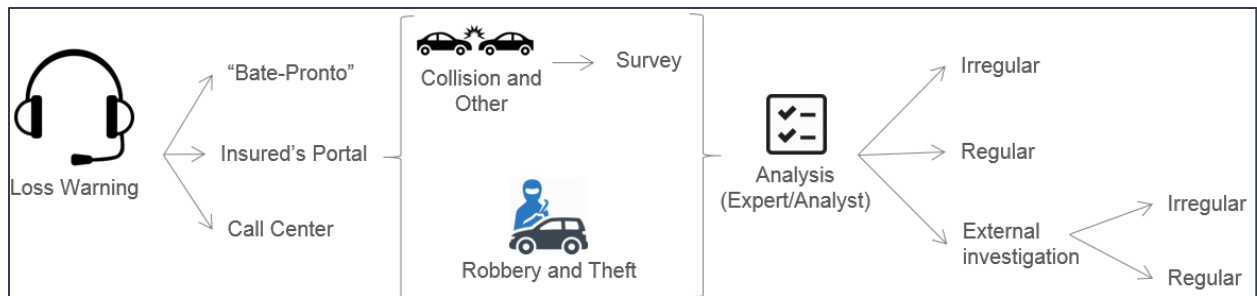
Therefore, the investment in fraud detection is important to reduce unnecessary costs, increase productivity and guarantee the profitability of our businesses especially in an economic national crisis scenario.

## THE SOLUTION

The objective of the compilation of such solutions in this project was to develop a robust tool for fraud and irregularities prevention, detection and investigation. Besides that, the purpose is also to improve productivity, avoiding the pre-analysis of 100% of the claims.

### Process "as it is":

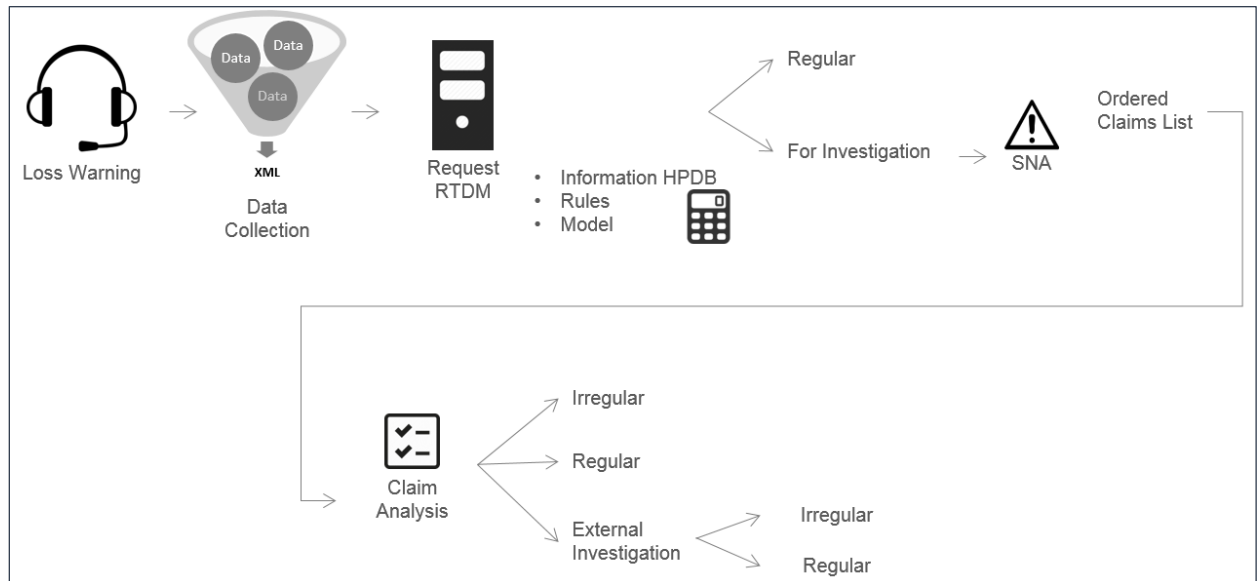
Before SAS® solutions implementation, the claim process was composed by the following steps, where analysis and submission to inquiry were only based on claim analyst criteria.



**Figure 1. Claim process before SAS® solutions implementation**

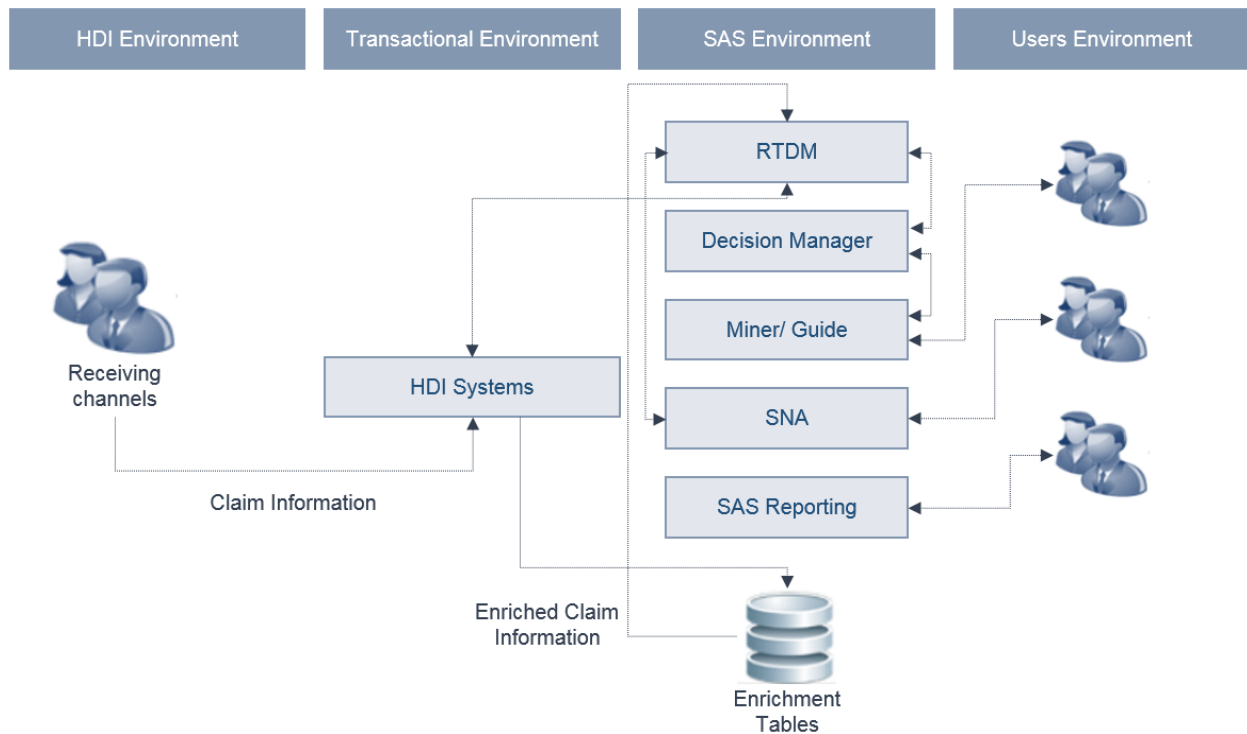
Process “as it will be”:

After the project, the claim process will be structured as follows, where analysis of claims will be selected by the business rules and statistical models.



**Figure 2. Claim process after SAS® solutions implementation**

The diagram below lists the architecture of the solution and the processes fluxes that were introduced to support the SAS® solution in HDI Seguros.



**Figure 3. Solution design**

Macro functional description:

✓ Online score calculation

- The channels request the web service RTDM, sending all the claim information that is necessary for rules screening. Inside the RTDM Online Calculator, business rules and prediction models are applied generating a final fraud score.
- Depending on this final score calculated through RTDM Online Calculator, the claim analysis will be set as "Regular claim" or "In Analysis claim"
- The calculated scores and the results of rules and prediction models are forwarded to the requester channels via RTDM and stored in a table inside solution data base.
- The result of the analysis is displayed for the requester channels and also stored in solution data base.

✓ Claim investigation

- The claims can be investigated with the usage of network diagrams in the SNA interface
- The analysis generated by RTDM are displayed in the same interface
- The interface also receives additional information that will enrich the investigation process
- All the information is compiled and displayed as alerts
- The creation of alerts is based on the fraud score and identification of suspicious behavior
- The alerts are exhibited in the SNA interface, sorted by fraud score and date. Older claims with higher score will be at the top of the list of alerts

- The alerts are analyzed by claim analysts, that will register their conclusions into the solution data base. These conclusions are then forwarded to integrate other HDI systems.

In HDI the set of rules was splitted into 2 sets:

- One of 48 business rules, designed by business specialists based on their expertise and experience, focused on identifying anomalies and irregularities. These rules combine claim information in mathematical conditions that, when assisted, cause the rule to be violated. For each rule this violation assign the claim a score that is summed to all the other rules scores for that specific claim. These rules can assign the claims a score of 0, 20 or 100, depending on the rule calculated.
- One of 5 rules, composed by 9 statistical prediction models. For each claim, only one statistical model is processed assigning the claim a score that is also summed to the scores of the other set of rules. These rules assign the claims a score proportional to the claim fraud probability, which goes from 0% to 100%.

At the end of this process, if the sum of all the score rules reaches the value of 100, then the claim becomes an alert at SNA list of suspicious claims.

This process is repeated at least 4 times through the RTDM requests bellow:

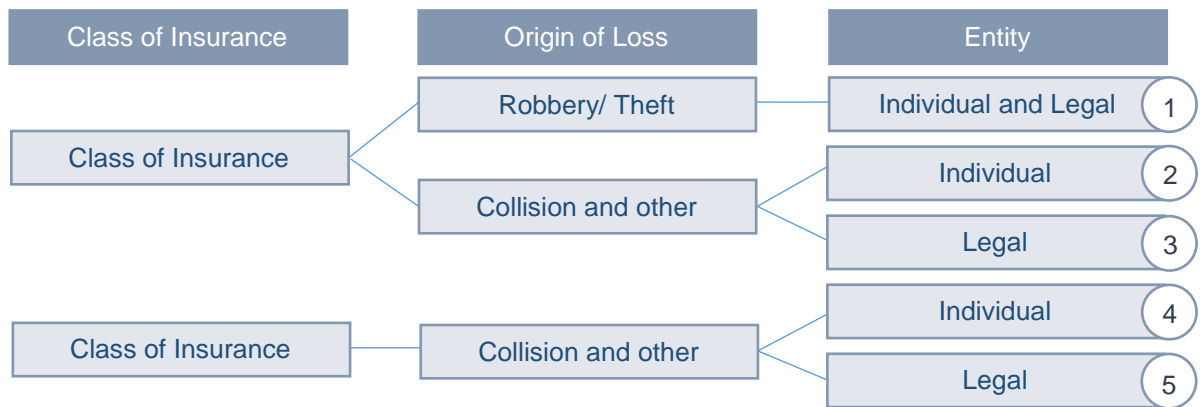
1. Claim Opening (T1)
2. After Claim Inspection (T2)
3. Regulation (T3)
4. Closing (T4)

Additional requests are made automatically in case of updates in claim information, producing new sums of scores and, possibly, new alerts.

## THE STATISTICAL PREDICTION MODELS

HDI Analytics team developed the statistical models, using a database containing the target variable and 1.057 predictor variables, structured with historical data between April 2014 and April 2016. Another database was also prepared in the same structure, but with historical data from May 2016 to August 2016 to be used as an Out of Time database, or in other words, to test the statistical models developed. Since that fraud is a very rare event, with a rate of 1,5%, it becomes a case of inflated zeros. The solution to model this data was to balance the number of Frauds and Non Frauds in a proportion of 20%. All the results were then produced over this balanced data.

The reason that there are 5 rules on the models rules set is the identification of 5 main profiles of claims, or segments, combining class of insurance (insured or third part), origin of loss (robbery/ theft or collision and other causes) and entity (individual or legal).



**Figure 4. Segmentation to allow fraud prediction for different profiles and types of claims**

It is common that in the first request (T1) there is missing information about the driver and repair shop involved at the collision (or other causes) claims. For this reason, 2 models were developed for each of the 2, 3, 4 and 5 segments, that will be processed depending on the available information at the time of RTDM request, totalizing 9 statistical models in the solution. This strategy was applied because statistical models that count on this information produces better prediction results than those that does not.

	Request 1	Request 2	Request 3	Request 4
Segment 1	Unique model	Unique model	Unique model	Unique model
Segment 2	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)
Segment 3	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)
Segment 4	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)
Segment 5	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)	Model (I) or (II)

Model (I): Considering driver and car repair shop information  
 Model (II): Not considering driver and car repair shop information

**Figure 5. Scheme for the statistical models processed considering segment and available information at the time of RTDM request.**

The statistical models were developed with the usage of SAS® Miner, having one specific diagram for each segment and model. The database had its variables categorized by Interactive Grouping or Interactive

Binning nodes and then splitted into 'training' and 'validation' datasets using the Data Partition node to avoid overfitting. Next, the Regression node was applied. The best technique used differs from model to model, but for most of them, the best regression type was the logistic regression, with probit or logit link function and stepwise or backward methods for the choice of predictive variables, with entry and stay significance levels of 5%.

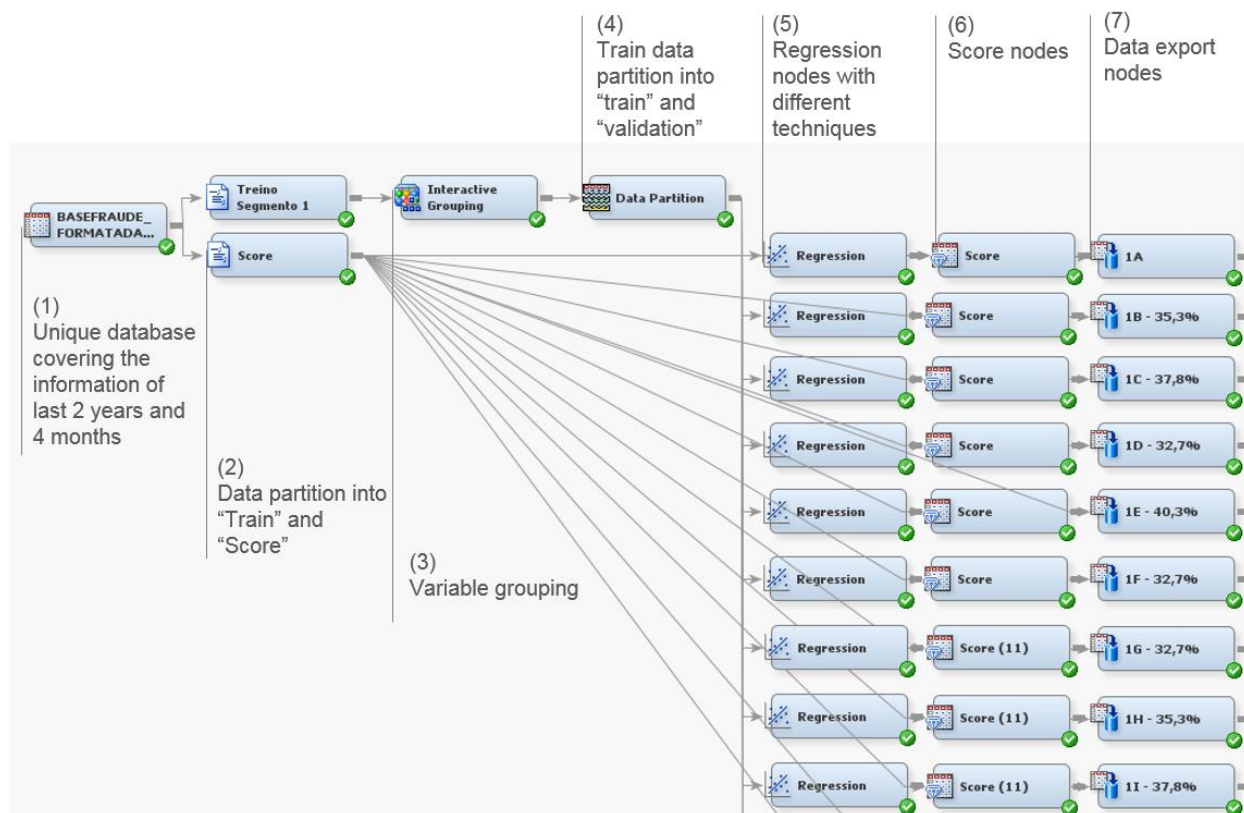


Figure 6. Example of diagram to find the best regression technique for fitting data for segment 1

The criteria to choose the best model fitted was the KS, considering the ranges and classifications bellow:

KS Values	Evaluation
Under 20%	Low
From 20% to 30%	Acceptable
From 30% to 40%	Good
> 40%	Excelent

Table 1. Ranges of KS and statistical model evaluation.

To evaluate the cut-offs for each statistical model, the Out of Time database was sorted from the highest to the lowest event probability (score) and separated into 10 deciles. The chosen cut-offs was those with accumulated fraud between 70% to 80%.

Decile	% Accumulated Target
1	36%
2	56%
3	67%
4	76%
5	82%
6	87%
7	89%
8	92%
9	94%
10	100%
<b>Total</b>	<b>100%</b>
<b>KS: 37,5%</b>	

**Table 2. Example of results for statistical model 1 for segment 2.**

For segments 1 to 3, the cut-off point chosen was the lower limit for decile 5 and, for segments 4 and 5, for decile 6. Every claim whose score is higher than the cut-off point of its statistical model becomes an alert in SNA and has to be analyzed. This way, in order to find 70% to 80% irregular claims, 54% of claims are going to be analyzed.

## CONCLUSION

The idea of this project is to reduce manual efforts on claim analysis, improve productivity and profitability of claim processes by being able to easily take fast and accurate decisions. In addition, the ability to adapt this solution for new needs, new customer and frauds behavior is a big deal.

The project now is in its pilot level, operating successfully over 10 “Bate-Prontos”. The expectation is to expand it for all the channels in Brazil in the course of 2017.

## REFERENCES

Talanx Group website: <http://www.talanx.com>

HDI Seguros | Brazil website: <http://www.hdi.com.br>



## ACKNOWLEDGMENTS

I would like to express the deepest appreciation to all people that have worked in this project.

In special, a thank you:

To HDI Seguros CEO, Murilo Riedel, for believing in the project, even when it was just ideas.

To HDI Seguros Executive Board, Fabio Leme, Frank Ohi, Marcelo Moura and Wellington Lopes, for supporting all phases of the project.

To Claim Business team, Luciana Santiago and Yoshico Shimono.

To IT team, Luiz Novaes, Marcio Heder and Paulo Sergio Sgarbi.

To SAS® Brazil team, Vitor Vicente and Vivian Vieira.

Finally, I would like to thank my Analytics team, Claudia Fidelis and Stephanie Ribeiro de Campos, for working all time with focus and determination.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Rayani Melega

HDI Seguros | Brazil

+55 11 5508-2164

[Rayani.melega@hdiseguros.com.br](mailto:Rayani.melega@hdiseguros.com.br)

<https://linkedin.com/in/rayani-melega-4b1a2483>