Inter-American Development Bank
Training: A Climate Risk Assessment Tool for Financial Institutions

Part 1
## Agenda

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Motivation</td>
</tr>
<tr>
<td>2</td>
<td>Credit Risk</td>
</tr>
<tr>
<td>3</td>
<td>Climate Variable</td>
</tr>
<tr>
<td>4</td>
<td>Tool Development Steps</td>
</tr>
<tr>
<td>5</td>
<td>Final Considerations</td>
</tr>
</tbody>
</table>
Section 1

Motivation
Supported by the Inter-American Development Bank (IDB), the Adapta Sertão / Proadapta Sertão project was developed in order to find feasible strategies and technologies for adapting to climate change in family agriculture, with an initial focus in the interior of Bahia State, Brazil.

The project aims to help small farmers cope with climate change in a resilient way by providing better farming practices in areas suffering from lack of rain, for example. Among these practices, we can mention the use of suitable seeds for each type of climate and mainly by the use of resilience tools such as artesian wells, which in case of drought can make the crop viable.
The Inter-American Development Bank (IDB) hired PwC Brasil to create an unprecedented credit analysis methodology for small farmers in the semi-arid region of Bahia. This new technique adds to the traditional credit scoring model the impact of climate change on agricultural production. The University of San Diego in California worked with PwC in the analysis and correlation of climatic data to create the climatic variable that composes the final model.
The main objective of the project is to build a financial tool that allows financial institutions to identify and quantify their exposure to climate risk in relation to agricultural credit.

The original study area is part of the semi-arid state of Bahia, where the PROADAPTA Project is located. PROADAPTA is a project to help small farmers combat severe weather conditions through the use of resilience tools such as artesian wells and dams.

The database was provided by Banco do Brasil, one of the main financial institutions in Brazil and the one with the largest participation in rural credit in the country.

Due to the small numbers of observations in the PROADAPTA region, the data area was increased to cover the entire semi-arid area of Bahia, since the climatic and social aspects are very similar to the PROADAPTA region.

Note: The dataset contains 187 municipalities in Bahia, which 15 belong to the PROADAPTA region.
Impacts in the project scope region

In 39 years, the backlands of Bahia lost 399 millimeters (mm) of rain.

The territory of the Jacuípe Water Basin lost 511 mm.

The joint and accumulated direct losses in milk production, cattle, sheep, goats and maize production in the drought of 2012 vary between R $ 443 and R $ 582 million (MM) for the semi-arid region of Bahia and R $ 18 to R $ 33 million to the Jacuípe Water Basin.
Section 2

Credit Risk
What is credit risk?

"Credit risk is defined as the potential that a borrower or counterparty in not fulfilling their obligations under the terms agreed in the credit agreement."

Losses include loss of principal and interest, a decrease in cash flow and an increase in collection costs, which occur in various circumstances, such as:

- A consumer does not make the payment of a mortgage loan, credit card bill or other line of credit.
- A company does not make the payment due on a mortgage, line of credit or other loan.
- An insolvent insurance company does not honor an obligation.
- A business or consumer does not pay a commercial invoice when it is due.
- An issuer of public or private securities does not make the payment of the coupon or payment of the principal when due.
- An insolvent bank does not return funds to a depositor.
- A court grants bankruptcy protection to a consumer or insolvent company.
Why is credit risk important?

The expansion of credit availability and GDP growth are positively correlated.
For each $1.00 in additional credit, historically, there was an increase of US $0.70 in GDP.

Expansion and contraction in the supply of credit is highly correlated to the economic cycle.

The "Great Recession" shows a larger than usual decline in credit supply and a weaker recovery than usual, leading to a relationship breakdown.

Fonte: St Louis Fed; NBER
What is counterparty risk?

Counterparty credit risk is understood as the possibility of losses arising from non-compliance with obligations relating to the settlement of transactions involving bilateral flows, including the trading of financial or derivative assets.

BACEN Resolution 4,557 of 2017

Risk that arises when the probability of default of the counterparties is positively correlated with general market risk factors. It is therefore the risk represented by the possibility that the counterparty shall not fulfill their part of the contract once it has weakened at the moment of crisis, for example.
Approaches to Internal Risk Models (IRB)

Minimum Requirements

• **Use of internal ratings**
  - Effective use of the "ratings" system and credit risk parameters in credit approval and credit risk management.

• **Quantification of credit risk**
  - Definition of "default", estimation of parameters (PD, LGD and EAD).

• **Validation of internal estimates**
  - Constant evaluation of the accuracy and consistency of the ratings system as well as the parameters of credit risk quantification.
Credit losses can be measured by several variables

What is the probability that the counterparty will default?
- Quality of the debtor
- Loan Maturity
- Country risk
- Economic factors

What is the value of risk exposure?
- Current exposure
- Unused Credit Lines
- Product type
- Legal form of the contract
- Economic factors

What is value of the recovered amount?
- Quality of the guarantees
- Liquidity of guarantee
- Economic factors

Main challenges:
- Each of the models must be well calibrated for an integrated view of credit risk
- Taking business measures to mitigate losses from the credit decision-making process to recovery
**Credit losses can be measured by several variables**

**What is the probability that the counterparty will default?**
- Quality of the debtor
- Loan Maturity
- Country risk
- Economic factors

**Probability of Default models**
- Current exposure
- Unused Credit Lines
- Product type
- Legal form of the contract
- Economic factors

**PD: Probability of Default**

**What is the value of risk exposure?**

**Risk Exposure models**
- Quality of the guarantees
- Liquidity of guarantee
- Economic factors

**EAD: Exposure at Default**

**What is value of the recovered amount?**

**Collection Score models**
- Quality of the guarantees
- Liquidity of guarantee
- Economic factors

**LGD: Loss Given Default**

**Main challenges:**
- Each of the models must be well calibrated for an integrated view of credit risk
- Taking business measures to mitigate losses from the credit decision-making process to recovery
A Probabilidade de Default (PD) é a probabilidade de um mutuário (ou uma porcentagem de um grupo) ser inadimplente em um horizonte de tempo específico

- The probability of default is an element of various credit risk models.
- Data sources vary, and are subject to the type of customer and information provided.
- Different classification methodologies can be applied:
  - Specific systems or expert judgment
  - Scoring models based on statistics
  - Simulation techniques
  - Combination of the above (hybrid)

**Example. Bank Scorecard Rating System**

<table>
<thead>
<tr>
<th>Analysis guidelines</th>
<th>Weight</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product/ Demand/ Market</td>
<td>20%</td>
<td>5.0</td>
</tr>
<tr>
<td>Shareholders/ Management</td>
<td>15%</td>
<td>5.2</td>
</tr>
<tr>
<td>Access to credit</td>
<td>10%</td>
<td>5.0</td>
</tr>
<tr>
<td>Profitability</td>
<td>15%</td>
<td>2.0</td>
</tr>
<tr>
<td>Cash Flow Generation</td>
<td>25%</td>
<td>2.0</td>
</tr>
<tr>
<td>Solvency</td>
<td>15%</td>
<td>9.0</td>
</tr>
<tr>
<td>Risk assessment of internal PD</td>
<td>100%</td>
<td>4.4</td>
</tr>
<tr>
<td>Equity rating of Moody's</td>
<td></td>
<td>Ba3</td>
</tr>
</tbody>
</table>
Decision support model to complement the scoring analysis

- Attempts to predict the probability of default.
- It may include:
  - Statistical models based on financial / accounting data
  - Measures based on market prices
  - Specialized Systems
- It introduces important qualitative analyzes
- It combines the gap between the risk of the borrower and the institution
- Essential in certain markets
- The final evaluation reflects a better combination of judgment and modeling result with more granularity
Credit risk scoring and analysis models are essential tools, but in the end, they are only tools and not solutions. Organizations need strategic dashboards to relate/link business strategy to risk management strategy and risk appetite.

Business mission and sales objectives drive the activities of an organization and therefore are the most critical risks.

Risk strategy and risk appetite determine the nature and level of risks that are acceptable to the organization.

The goal of a CRM department is not to eliminate credit risk, but rather to manage risk in the portfolio, maximizing profits and fulfilling the company’s mission.

The credit risk management environment has evolved from simple credit risk classification to an active credit risk strategy structure based on profit.

Business and risk strategy should also include both the internal environment and the external regulatory environment and the competitive environment.
Risk rating is a powerful **credit risk management tool**, but is often misinterpreted or misused

### Risk Rating is ...
- A management information tool
- A tool to differentiate the degree of credit risk
- Inputs for many credit risk management decisions, including:
  - Credit approval
  - Pricing
  - Administration requirements
  - Portfolio Management
- A platform to develop a common language for identifying and reporting counterparty risks
- An essential building block in advancing risk management capabilities
- A means of comparing portfolio risks with external benchmarks

### Risk Rating is not ...
- A substitute for the experience and judgment of experienced bankers/experts
- A tool for decision-making
- Always an automated process
- A barrier to new loans
- A "final solution" in improving credit risk management
Classification of operations (Credit Risk)

Categorizing customers accurately results in adequate provisioning for credit operations, as well as enabling better portfolio management.

“Understanding the composition of the credit portfolio and the behavior of the clients is of great value for the financial institutions because it allows to know the best actions to be taken.”
A rigorous credit analysis requires an analysis of different aspects, including commercial, financial and structural risks.

<table>
<thead>
<tr>
<th>Macroeconomic analysis</th>
<th>Management analysis</th>
<th>Financial analysis</th>
<th>Type of borrower</th>
<th>Type of loan</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Proactive or reactive</td>
<td>Liquidity</td>
<td>Holding</td>
<td>With collateral</td>
</tr>
<tr>
<td>Inflation</td>
<td>Strategies</td>
<td>Profitability</td>
<td>Main operating subsidiaries</td>
<td>No collateral</td>
</tr>
<tr>
<td>Demographic Trends</td>
<td>Motivation</td>
<td>Leverage</td>
<td>Secondary operating subsidiaries</td>
<td>Long term</td>
</tr>
<tr>
<td>Business Cycle</td>
<td>Experience</td>
<td>Interest and debt coverage</td>
<td>Local or foreign – domicile</td>
<td>Short term</td>
</tr>
<tr>
<td>Political stability</td>
<td>Integrity</td>
<td>Volume of business</td>
<td></td>
<td>Subordinate</td>
</tr>
<tr>
<td>Regulatory environment</td>
<td>Corporate governance</td>
<td>Other relevant indices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legal environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry Trends</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Consumer
- Employment situation
- FICO Score, credit rating
- Guarantee (real estate, equipment, stocks and bonds, etc.)
- Non-payment of other obligations

### Corporate
- Background (industry, size, number of employees)
- Guarantee (inventory, investment, inventory, equipment, patent, etc.)
- Profitability (income stream, revenue trend, etc.)
Section 3

Climatic variable - climate as an influential variable in agricultural production
Climate resilience: a new component for credit risk assessment

The need to revise credit risk models

In the Brazilian financial system, credit risk has historically been focused on the profile of the borrower, not holistically considering external variables.

Industries are affected by climate change, but none of them is as strongly impacted as agriculture. Therefore, effective credit risk models will be a powerful tool to help this market.

The increase in data technology, "Big Data", made it possible to assess the risk of other perspectives, such as external factors such as climate change.

Technology has advanced in terms of resilience tools and therefore, if its effectiveness in the financial world is proven, both society and banks will benefit from its use.

Business models are adapting due to severe weather changes.

Is climate a relevant factor?

The climate has a relevant statistical correlation with default, where the economic activity is dependent on the climate and the drought predominates.

This correlation is significant enough to justify updating credit scoring methods by commercial banks, including the climate variable, especially for agricultural producers in climate-sensitive regions, such as the region covered by this project, the semi-arid state of Bahia.
Impacts of climate change

75% of CEOs believe their companies have developed more sustainable products and services in response to the risks of climate change. 19th Annual Global CEO Survey

"Potential effects of climate change in Northeast Brazil suggest a 4º C increase in the surface temperature." - INPE

"Reduction of 22% in precipitation in Northeast Brazil"

"Cocoa production in the state of Bahia could be affected due to the probable increase in temperature and rainfall."

Severe weather events in Brazil (from 2002 to 2012):

- 355 billion BRL of losses for the country
- 13 thousand occurrences
- 46 millions affected people
- 3745 people were killed

19th Annual Global CEO Survey
Section 4

Tool Development Steps
Approach to the summary model

The main objective of this project is to build a tool that incorporates climate change into the traditional model of credit scoring, enabling the understanding of credit clients more accurately.

- **Climate Variable**
  - Historical data
  - Future climate data
  - Area and cultivation season

- **Financial data**
  - Data preparation
  - Definition of default
  - Cleaning the database

- **Development of the model**
  - Analysis of Variables
  - Binary process - assignment of values "0 or 1" (dummy variables)
  - Logistic regression

- **Results evaluation**
  - Validation tests:
    - Operational characteristic of the receiver (ROC)
  - Test of Kolmogorov Smirnov (KS)
  - Area under the curve (AUC)
Assumptions and limitations

Premissas

• Five years of financial data;
• The Credit Scoring methodology is widely used by financial institutions to check credit risk for contractual loans;
• A sample containing information on small farmers in the interior of the State of Bahia

Data availability

The analyzed database contains information on loans (BdB) for a period of five years (2012-2017) and climate data (IBGE) only for 1990-2014, which reduced the final database for the analysis.

Overcoming actions

In order to obtain a more precise analysis, different scenarios were observed and compared with the final result of the tool.

Reduced crop information

The analyzed database did not include crop information related to each policyholder.

Overcoming actions

In order to deal with this issue, proxies were used, with correlated analysis of the individual tools used for each crop, municipality and year of production.

Accuracy of climate data

The PROADAPTA region contains only a few municipalities, making it difficult to understand climate change.

Overcoming actions

The expansion of the project region allowed minimizing the impact on data accuracy.

Limitations
**Project steps**

1. **Climate Variable**
   - Data extraction
   - Preparation of data
   - Creation of the climatic variable

2. **Financial Data**
   - Obtaining financial data
   - Preparation of data
   - Definition of default
   - Variables analysis

3. **Model Development**
   - Sampling
   - Model development

4. **Calibration**

**Outcome valuation**

**End**
Project Approach

**Climate Variable**

1. **Extraction of data**
   - Extracting temperature data (gridded surface temperature climatologies and anomalies at 1 ° x 1 ° spatial resolution globally)
   - Extracting precipitation data (satellite-derived data available as total precipitation (in mm) at 5' x 5' spatial resolution and are updated roughly monthly)
   - Extracting crop data over time

2. **Extraction of data**
   - To reconstruct the daily temperature, the anomaly was added to the climatology.
   - To add precipitation to monthly periods, the values were summed to find the total of each month.
   - These data are made available through global scales, which makes it possible to understand the temperature / precipitation value at a specific point and period.

3. **Climate Variable**
   - Correlating the three sets of data cited above, a variable is obtained that correlates temperature, precipitation and crop level, which makes it possible to understand, for example, how the crop was better or worse in a given region according to climate change in this place.
   - This new variable will be used in the credit scoring model as a predictive variable.

<table>
<thead>
<tr>
<th>ID</th>
<th>Year</th>
<th>Month</th>
<th>Precipitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1990</td>
<td>1</td>
<td>238,880,542,683,152</td>
</tr>
<tr>
<td>2</td>
<td>1990</td>
<td>1</td>
<td>236,439,206,364,179</td>
</tr>
</tbody>
</table>

¹ Berkeley Earth Surface Temperature (BEST);
² Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS);
**I – Climate Variable: Data Extraction**

In order to test the correlations between climate and any result, it is necessary to have the ability to link time-series measurements of environmental conditions at specific locations, for example: location, municipality or type of product to be grown.

In this way, it is possible to create a data set at the level of individual loans or a group of creditors from the same municipality and to observe correlations between climatic variables in relation to the results.

Thus, validated historical meteorological data and future climate projections were used, as well as the correlation between the two.

*All data sources described here are free for public use; therefore, there are no anticipated costs of climatic data for the acquisition of climatic data for this type of analysis*
I - Climate Variable: Extraction of data
I - Climate Variable: Data preparation

For a rich characterization of the time variation, it was necessary to decompose the climate variation using monthly data before joining with the annual data of the agricultural harvest. The climatic factor is therefore composed of the following factors:

1\(^{st}\) - Shock conditions - unpredictability factor (unexpected variation) of the climate;
2\(^{nd}\) - Seasonal variation - expected variation of the climate by region and time;
3\(^{rd}\) - Climate differences - adaptation of farmers to climate change and
4\(^{th}\) - Transverse measures - that is the variation of the climatic factors between the municipalities.

\[ C_{my} = (\hat{C}_{my} - \hat{C}_{im}) + (\hat{C}_{im} - \hat{C}_{my}) + (\hat{C}_{my} - \hat{C}_{i1}) + \hat{C}_{i1} \]

The main result for the analysis of the climatic variable is therefore the yield of an agricultural crop in a municipality and year given the climatic conditions presented in the region under analysis.

This regression is estimated for each crop separately and may include climatic variations relative to temperature, precipitation or both at the same time.
I - Climate Variable: Data preparation

To assess recent and future impacts on climate performance (or, more fundamentally, to understand the impact of climate on crop yields and agricultural profits), climate indicators have been defined that accurately reflect the spatial and temporal extent of exposure. Data was available free of charge on global agricultural and national land survey services and satellite data. Then, specific climatic variables were constructed based on the growing area and the growing seasons in the region of interest.

Example of indicator creation:

```stata
/*decomposition da precipitação em 4 termos: p_hat, p_deviation, p_seasonal and p_shock*/
egen p_muni_month_mean = mean(precipitation), by(id month);
gegen p_seasonal = p_muni_month_mean - p_hat;
gegen p_shock = precipitation - p_muni_month_mean;
gegen p_abs_shock = abs(p_shock);
gegen p_deviation = p_hat - p_hat;
gegen p_abs_shock_sq = p_abs_shock ^ 2;
gegen p_abs_seasonal = abs(p_seasonal);
gegen p_abs_seasonal_sq = p_abs_seasonal ^ 2;
gegen p_deviation = abs(p_deviation);
gegen p_deviation_sq = p_deviation ^ 2;
save "$dir\Brazil Monthly Municipality Precipitation 1990-2017 Long redux_v3", replace;
```
I - Climate Variable: Calculus

With any of these three terms, it is possible to estimate the direct linear slope effect of the climate result. Merging the data of agriculture and climate it is possible to obtain a relation between municipality and climatic variation for each crop in Brazil, estimating then the following regression:

\[
Y_{ict} = \beta_0 + \beta_1 C_i + \beta_2 C^2_i + \beta_3 C_{ict} + \beta_4 C^2_{ict} + \beta_5 \epsilon_{ict} + \epsilon_{ict}
\]

The variable, therefore, correlates climate change data, such as temperature and precipitation, in the semi-arid region of Bahia and the evolution of agricultural production in the same region. In this way, this variable shows a direct relationship between climate and production, demonstrating crops that have grown better or worse in periods of drought or rain, for example.
Project Approach

**Obtaining the data**
- Information on loans from 107,185 borrowers from 186 municipalities in the semi-arid region of Bahia;

**Preparing the data**
- Completeness;
- Special treatment for lost values and outliers;
- Exclusions from inconsistent contracts;
- Analysis of variables;
- Segregation of products into 4 groups.

**Default definition**
- In default: customers with a delay of 90 days or more;
- Non-default: Customers with a maximum delay of less than 90 days.

**Variables analysis**
- Univariate analysis: analysis of model variables individually;
- Bivariate analysis: analysis of the correlation between the variable and the client’s status.
II - Financial data: Obtaining the data

The tool was developed based on the data provided by Banco do Brasil (BdB). The data contains loan information from 107,185 borrowers from 186 municipalities in the semi-arid region of Bahia, and contains 7,838,904 observations, referring to 297,115 contracts for a five-year observation period (2012-2017).

Main analyzes carried out in the database

Completeness: the integrity of the database was checked with various checks and granular analyzes;

Special treatment for missing values and outliers: once missing values and outliers were identified, they were analyzed to ensure the meaning of each of them in the data (e.g.: removal of contracts with very high contracted values, default day assignment - 0 or 90 - based on days past due)

Exclusions:
• Contracts related to operations other than credit for agricultural and livestock activities (agricultural funds): only the contracts related to this type of operations (indicated as "PR_FAMILIAR" in the database) were maintained;
• Contracts prior to July 31, 2012: the contracts for which the contracting date is prior to July 2012, as this is the first transaction date sent by the bank;
• Contracts of less than 50 BRL: hiring values too low, which may be due to a database error;
• Contracts not related to the cultivation of maize, beans, manioc and banana: municipalities that do not produce these crops were not analyzed;
• Transactions over 24 months from the date of contracting: all operations were observed over a 24-month horizon (12 months referring to the "crop year" and 12 months referring to the schedule of payment of contracts with a single parcel).
II - Financial data: Preparation of the data

The model development was limited due to the lack of data related to contracts at the granular level.

The database presents data from 2012 to 2017, however climatic information is available from 1990 to 2014. Therefore, the intersection between both allows the development of a model with a data frame between 2012 and 2014.

Due to the lack of information on the relationship between crop type and farmers in the BdB database, a correlation was made with the data provided by the BdB (contract information) and the climatic variable that relates climate change to information on agricultural crops in the municipalities of Bahia.

Only a few borrowers were identified with information related to the resilience tools.

<table>
<thead>
<tr>
<th>Product</th>
<th>With resilience tool</th>
<th>No resilience tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beans</td>
<td>2,873</td>
<td>47,897</td>
</tr>
<tr>
<td>Cassava</td>
<td>2,294</td>
<td>39,087</td>
</tr>
<tr>
<td>Maize</td>
<td>2,410</td>
<td>38,049</td>
</tr>
<tr>
<td>Banana</td>
<td>1,692</td>
<td>36,448</td>
</tr>
</tbody>
</table>
The definition of default was based on the payment information of the borrower in the last 24 months from the first date of transaction in the database, checking the days of maximum delay of each contract.

**In default**: Customers with a delay of 90 days or more

**Non-default**: Customers with a maximum delay of less than 90 days

Total customers in default: 1,404

Total non-default customers: 54,241

Default rate: 2.52%
II - Financial data: Analysis of variables

As the data is received, the continuous variables are analyzed and discretized so that it is possible to determine the discriminatory power of each of them. For this, the information value and weight of evidence of each one of them is calculated, which will evidence the relevant variables for the model.

Once the most important variables are identified, they go through the binning process and are transformed into dummy variables, which are classified as "1" and "0", 1 for "yes" and 0 for "no".

Univariate analysis

Fundamental analysis for creating a good credit scoring model.

It consists of an evaluation of all variables available for the development of the model, which must be understood as inconsistent values, frequency of distribution of variables, outliers, missing values, etc., as well as the best way to treat them.

Bivariate analysis

The bivariate analysis verifies the relationship between each predictive variable and the client's status. A predictive variable is one that affects the response and can be measured.

Customer status is good or bad characterization.

The main purpose of this analysis is to analyze the power of the predictive variable, the correlation between them and strange behaviors.
**Project Approach**

- **Definition of methodology:** Credit Scoring;
- **Statistical tool:** Logistic regression;
- **Software used:** R;
- **Calculation of Probability of Default**;
- **Results analysis**

---

**Sampling**

- Divide the database into two parts, 30% and 70%;
- The 30% part is used for validation (Sampling);

**Model Development**

- Divide the database into two parts, 30% and 70%;
- The 70% part is used for the development of the model;

---

"This process (segregation of the database by 30% and 70%) allows us to evaluate the power of the model to discriminate between good (non-default clients) and bad (in-default clients)"
### III - Development of the model: Sampling

Due to the different responses of each agricultural production to climate change, the data were divided into four different groups, considering the main crops produced by small farmers and each level of productivity in the semi-arid region of Bahia.

**Distribution of the product (culture) in the database:**

- **91.2%** of farmers are located in municipalities that produce **Beans**;
- **74.4%** of farmers are located in municipalities that produce **Cassava**;
- **72.7%** of farmers are located in municipalities that produce **Maize**;
- **68.5%** of farmers are located in municipalities that produce **Banana**.

*Dados obtidos do IBGE (2016)

Note: Some other crops, such as sugarcane, have also been identified. However, most of the time they are produced by large producers, which is not part of the project scope.
**III - Development of the model: Sampling**

The sample was randomly **divided into two parts**, one containing **70%** of the contracts, which were used for the development of the model and the other containing the remaining **30%**, which were used for validation.

In doing so, it is **possible to evaluate the power of the model** in discriminating good (non-default) and bad (in-default) clients in a more precise way, such as in terms of **stability and calibration**.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Development</th>
<th>Validation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize</td>
<td>28,312</td>
<td>12,134</td>
<td>40,446</td>
</tr>
<tr>
<td>Beans</td>
<td>35,529</td>
<td>15,227</td>
<td>50,756</td>
</tr>
<tr>
<td>Cassava</td>
<td>28,961</td>
<td>12,412</td>
<td>41,373</td>
</tr>
<tr>
<td>Banana</td>
<td>26,690</td>
<td>11,439</td>
<td>38,129</td>
</tr>
</tbody>
</table>
III - Development of the model: Definition of Methodology

Methodologies for the quantification of credit risk

Credit risk modeling

- Classification models
  - Decision trees
  - Bayesian Networks
  - Logistic regression
  - Neural networks
- Actuarial models
  - CreditRisk+
  - Credit Metrics
  - Merton
  - KMV
- Models based on asset values
  - Credit Portfolio View
- Macroeconomic models
- Models of intensity

Selected Methodology

- Expert-based decision making
- Operates with causal relationships quantified by conditional probability values
- Variable to be explained (dependent) as a function of independent variables
- Computational technique inspired by the neural structure of intelligent organisms
- Focuses on default event, not considering rating migration
- Estimate the future distribution of changes in the value of the loan portfolio
- Consider the value of the assets and liabilities of the corporate balance sheet
- Improves Merton’s model through corporate balance sheet volatility
- Probability of borrowers default according to macroeconomic scenarios
- Monte Carlo process using Bernoulli distribution for default
III - Development of the model: Execution of the analyzes
Methodologies for the quantification of credit risk

Methodology Details

The Logistic Regression assigns the probability of default of a borrower according to the history of occurrences in the data sample.

- Logistic regression is a particular case of regression;
- It is easy to find the probability of occurrence of categorized or binary data;
- S-shaped curve;
- Lower limit 0 (zero) and upper limit 1 (one).

Summary of the analysis carried out

**Sample**

55.645

Loan registration

**Analysed crops**

Beans | Cassava | Maize | Banana

**Variables**

- Loan size
- Economic activity
- Rural property
- Formal education
- Civil Status
- Resilience tool
- Climate Variable

**Modeling - Scenarios**

1. Credit Scoring model, considering only loan information.
   **Available Data: 2012-2017**

2. Credit Scoring model, considering loan information and climate variable.
   **Data available: 2012-2014**
III - Development of the model: Execution of the analyzes

Governance - Traditional structure of credit analysis

[Diagram showing the flow of analysis, with roles and decision points]
III - Development of the model: Execution of the analyzes

Governance - Structure with climate risk analysis

Traditional credit assessment structure

Sustainability Analyst → Sustainability Manager → Superintendent of sustainability
Analyze historical data and projections of climatic variations

Risk Analyst → Risk Manager → Superintendent of risks
Incorporate climate data into internal risk models
For the development of the statistical model to calculate the Credit Scoring and aggregation of the climatic variable, the **R Studio software** was chosen. The program is free and features a large amount of libraries and modeling functions widely used both in academia and in the market.

The software can work satisfactorily with a large amount of data and generate graphs and correlations.
III - Development of the model: Execution of the analyzes

Technology

Console

Plots & help

Envirenment
"Analyzes (tests) of performance and calibration were performed to verify the reliability and adherence of the model"
### IV - Calibration and results: Evaluation of models

| 1. | Current Portfolio Analysis |
| 2. | Stability analysis of the portfolio: t-student |
| 3. | Stability analysis of the model: Mann-Whitney (Wilcoxon) |
| 4. | Stability analysis of the model: Kruskal-Wallis |
| 5. | Discrimination power curve: CAP and ROC |
| 6. | Confidence interval |
| 7. | Divergence Test (TD) |
| 8. | Information Value (IV) |
| 9. | Measures of entropy: CIER |
| 10. | Kolmogorov Test - Smirnov |
| 11. | Analysis of calibration curve fit |
| 12. | Binomial Testing |
| 13. | Binomial test using granularity adjustment (pooling) |
| 14. | Chi-Square Test (Hosmer-Lemeshow) |
| 15. | Matrix of CONFUSION |
| 16. | Stability analysis of variables - ICA (Sampling Characteristic Index) |
The Kolmogorov-Smirnov test is a non-parametric method that evaluates whether two samples have similar distributions or, rather, if they were extracted from the same population. If the differences observed between them are large they are probably not due to chance. It is a very sensitive method, detecting differences in relation to the central tendency, dispersion and symmetry. It can be used for data measured on the ordinal, interval or reason scales, with no requirement that they have normal distribution.
IV - Calibration and results: Evaluation of models

ROC Curve - Receiver Operating Characteristic

A visual method that can be constructed from two sample scores, one for abnormal cases, as defaulting debtors, and another for normal cases. Through its curve based on sensitivity and specificity definitions, it is possible to calculate the area under the curve (AUROC) that evaluates the discriminatory power of the Credit Scoring model.

**AUROC Result:**

*Area below 0.7: low discrimination;*

*Area between 0.7 and 0.8: acceptable discrimination;*

*Area between 0.8 and 0.9: excellent discrimination;*

*Area above 0.9: exceptional discrimination.*
IV - Calibration and results: Evaluation of models

Comparison - Matrix of confusion

Based on the analysis of the confusion matrix, it is possible to observe a better **discriminatory power** in the credit scoring models that take into account the climatic variable, meaning that this variable is an important factor that contributes to the payment behavior of farmers in that region.

The financial tool, taking into account climate change, allows the FI to predict a **more accurate probability of default**, based on each crop and how they respond to climate change.

In the presence of climate shocks, less risky contracts are more prone to default than riskier ones. This means that if financial institutions begin to consider this variable in their models, it will be possible to invest in resiliency tools and, consequently, to obtain a better result from the borrowers in that region.
### IV - Calibration and results: Evaluation of models

#### Comparison - Matrix of confusion

<table>
<thead>
<tr>
<th></th>
<th>No climate variable</th>
<th>With climatic variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Real</td>
</tr>
<tr>
<td>Banana</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Maize</td>
<td>Model</td>
<td>Real</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beans</td>
<td>Model</td>
<td>Real</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cassava</td>
<td>Model</td>
<td>Real</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For statistical regression of the model, the R Software was used.
**IV - Calibration and results: Evaluation of results**

Another objective of this tool is to analyze the impacts of the maintenance of resilience tools (water storage technology) on small farmers’ payment capacities.

When shocks occur in temperature or precipitation, borrowers in the state of Bahia are more likely to be in default. However, this effect is significantly attenuated by the presence of water storage infrastructure of various types.

The conclusion of the analysis shows that water storage technologies are little used by small farmers and the use of such technologies could increase the productivity of the farmer.

Financial institutions could consider credit loans for individuals who offer some water supply infrastructure, as these clients would face the least difficulties in their agricultural activities.

**Impact of resilience tools on climate shocks**

**Comparison Between the Probability of Default**

<table>
<thead>
<tr>
<th>PD Banana (with climate variable)</th>
<th>PD Beans (with climate variable)</th>
<th>PD Cassava (with climate variable)</th>
<th>PD Maize (with climate variable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average PD</td>
<td>2.36%</td>
<td>2.44%</td>
<td>2.63%</td>
</tr>
<tr>
<td>Resilience tool</td>
<td>2.00%</td>
<td>2.27%</td>
<td>2.54%</td>
</tr>
<tr>
<td>No resilience tool</td>
<td>2.37%</td>
<td>2.37%</td>
<td>2.45%</td>
</tr>
</tbody>
</table>

Note: The table above shows the average probability of default (PD) for different resilience scenarios, with and without climate variables. The values represent a decrease or increase in PD compared to the baseline scenario.
Section 5

Final considerations
Final considerations

1) Adoption of the climate scoring model market

Having proven that climate can be a relevant component of credit scoring models for small farmers in climate-sensitive regions, Development Banks, such as the IDB, could expand this exercise to other areas and with more commercial banking partners. With more commercial banks involved, adopting this tool as a standard industry procedure will be more successful;

2) Improvements in financial data to be useful in climate models of credit scoring

Data availability is a problem for the development of a suitable climate credit scoring tool. It is therefore recommended that the IDB provide incentives for the financial market and governments to maintain longer track records in areas that need more government assistance, such as climate-sensitive regions. With longer records, more accurate analysis could be done, providing more reliable tools to reduce impacts on banks (default) and on society (less financial hardship for small farmers). In addition, the IDB could provide additional studies and guidance to banks on how to structure the specific financial data needed to create climate credit scoring models.

3) Involvement of insurers to provide financial options for climate risk in future projects

In order to increase the chances of success of this tool, other exercises can be carried out, such as the involvement of insurers to analyze possible products that could reduce the bank’s exposure, but also create alternatives for creditors that may default.
Questions and answers
Thank you