Understanding the Portugal D&B Failure Score

THIS DOCUMENT IS INTENDED TO ADDRESS THE FOLLOWING QUESTIONS:

- What is the D&B Failure Score?
- What does the D&B Failure Score predict?
- What is the availability of the Informa Failure Score?
- How is the D&B Failure Score calculated?
- How does the Informa Failure Score perform?
- What is the Relationship between the Informa Failure Score and Failure Rates?

INTRODUCTION

The Failure Score provides an overall statistical assessment on the future prospects of an active Public or Private business facing a situation of failure over the next 12 months, and is applicable to all firms with minimum information available in the Dun & Bradstreet Data Cloud.

Based on this assessment, it is possible to rank businesses based on each one's probability of failure, i.e. becoming "Bad". Such ranking may be retrieved from the outputs provided by each of the following risk classifications: a Score ranging from 1,001 to 1,999; a 1 - 100 Percentile; a 1 - 4 Risk Class. The difference between classes in each of these classifications is based on the likelihood of a business experiencing the above definition of "Bad" performance over the next 12 months. The risk of failure of over 600,000 active Portuguese businesses is accordingly assessed, based on the available demographic, financial, trade and negative information about each one.

The performance of the Portuguese Failure Score confirms that this is a highly effective solution to predict the potential insolvency of your existing and prospective customers. The solution allows you to:

- Automate decisions for increased efficiency
- Allow faster processing of large volumes of transactions
- Free up resources to look at the time-intensive borderline decisions
- Enable more consistent decisions across the entire organization
- Reduce the costs associated with full-scale application and annual risk reviews
- Apply scores across an entire portfolio to quickly identify risk and opportunity
- Satisfy regulatory needs for timely, consistent and objective review of decisions at the account level

This document explains in greater detail how the Portuguese Failure Score was developed.

PORTUGAL FAILURE SCORE

WHAT THE FAILURE SCORE PREDICTS

The Failure Score predicts the likelihood of a firm facing a financial crisis over the next 12 months, materialized in the assignment of bankruptcy, insolvency proceedings, or dissolution of activity together with unsolved legal incidents.

The Portuguese Failure Score was developed by the Portuguese Analytics Team, and has met the strict quality standards set by Dun & Bradstreet. This model contains several improvements and updates relative to the previous version, providing a better understanding of the key variables that influence the future financial health of a business.

The Score was statistically derived combining variables from complementary type of variables that at each moment characterize a business, from where the most predictive explanatory variables were identified. These include not only information from the financial statements, but also information on trade experiences, as well as demographic and negative information.

The legal events which constitute failure in Portugal include:

- Insolvency proceedings
- Bankruptcy
- Dissolution with legal incidents (i.e., pending courts judgments and insolvency courts, unpaid Social Security or Business Income Tax) in the last 36 months.

The expression "Bads" used in this document denotes businesses that have registered one or more of the above legal events. In turn, "Goods" means businesses that do not have any of the above legal events.

Any business that is classified as "Bad" is always identified with the respective situation of "Bad". However, as long as businesses are still active and observe one of the previous situations of "Bad" (e.g., Insolvency Proceedings), they have a risk class, which is the highest risk class. All cases that are already out of business, regardless of being or not being "Bad", are classified as NQ (Not Qualified).

AVAILABILITY OF THE FAILURE SCORE

The Failure Score is available for over 600,000 Portuguese registered businesses. This number can be split into circa 350,000 firms and approximately 250,000 sole traders. In the case of firms, the number corresponds to the vast majority of Portuguese businesses registered in the Chamber of Commerce.

The exceptions where the Score and risk class are not available include the following categories, which fall out of the scope of the assessed Universe:

- Activities not considered: Financial Activities and Financial Intermediation (Central Bank and banks in general; Leasing activities; Factoring activities; Trusts, funds and similar financial entities; Activities of credit institutions; Activities of credit purchase financing firms; Life insurance; Non-life insurance; Pension funds and supplementary professional schemes; …); Holdings.
- Legal forms not considered: Foreign Entity; Private Institution of Public Interest; Local Authority; Undivided Inheritance Fund; Trade Union; Association; Foundation; Body of the Public Administration; Embassy; Consulate; Entities assimilated to legal persons; Religious Legal Person.
- Businesses that are currently classified as "Bad".
- Businesses not active.
- Entities that are in the stage of liquidation.
- New businesses (age below 12 months).
- Businesses without minimum information.

The minimum level of data requirement is composed by:

- Company name
- Company address
- A valid Standard Industry Code (CAE)
- Legal form

Cases which do not meet these criteria are not considered as part of the scoreable universe, and have a score of blank or null and the Risk Indicator will be null or 'dash'.

MODEL DEVELOPMENT PROCESS

The models built for the Portuguese Failure Score enhance the potential benefits from our extensive Data Cloud. In addition to its comprehensiveness, the Dun & Bradstreet Data Cloud is monitored according to strict quality processes, providing a timely and consistent Data Cloud. Such quality allowed the achievement of high performance for the Failure Score, which ultimately facilitates and leverages the use of data provided by Informa.

The predictive factors behind the Failure Score and the way they combine were chosen based on the use of robust statistical techniques, together providing mathematical equations composed by the selected variables and weights that are able to translate into a single number the risk of Failure of the businesses assessed.

The model development required the selection of the multiplicity of factors characterizing each business at a certain period in time. In order to derive a powerful statistical model, Informa retrieved predictive information (photos) of active firms in two different points in time (31.12.2012 and 31.12.2013) relative to their respective:

- Demographic data and activity signs: includes business legal form, age since constitution, age of balance sheet, number of employees, physical address, industry type and other related information of a particular firm.
- Financial information: includes financial data about the balance sheet and income statement, whenever available.
- Trade data: comprises information from trade experiences.
- Negative data: contains information pertaining to legal incidents related to potential payment difficulties.

The dataset with the two previous photos was merged with the information about the failure status 12 months after each photo. In order to maintain the logic close to a hazard model, any case in failure at the end of the first photo (31.12.2012), was excluded from the second photo, even if at the time it remained active.

With the data of the two photos, a total of over 1.3 million cases were analyzed, out of which near 16,000 cases were "Bad".

From the various potential predictive factors initially considered, the selection was based on the individual predictive power of each one and the way they combine, so as to avoid duplicating similar information, but also not to leave behind relevant information.

Appendix A contains a more comprehensive list of data elements which are used in calculating the score.

SCORING SYSTEM AND MODEL SELECTION

In order to identity the best combination of predictors, a forward and backward stepwise selection of variables was used. The final model was the one that best results revealed in terms of:

- Discriminatory power
- Lack of multicollinearity problems
- Comprehensiveness of information about each assessed entity
- Economic meaning of the relation between variables
- Robustness to out of sample and out of time tests

SCORING OUTPUTS – SCORE VALUES

The Failure Score assigns the following measurements of risk:

- **A "Raw Score" of 1001 - 1999,** where 1001 is applicable to businesses that have the highest probability of failure, and 1999 represents businesses with the lowest probability of failure. This Score provides a direct relationship between the score and the level of risk. The marginal odds of being good doubles for each 40 point increase. For example, a score of 1250 <=> odds of being good = 8, 1290 <=> odds of being good = 16, 1330 <=> odds of being good = 32, etc. This score enables a customer to utilize more granular cutoffs to drive the automated decision-making process.

- **A "Percentile Score" of 1 - 100,** where 1 represents businesses that have the highest probability of failure, and 100 represents businesses with the lowest probability of failure. This Percentile shows how the risk of a business compares to other businesses in the Dun & Bradstreet Data Cloud, and is most effectively used by customers to rank order their portfolios, from highest to lowest risk of business failure.

- **A "Risk Class" of 1 - 4,** which is a segmentation of the universe into four distinct risk groups, where 1 denotes businesses that have the lowest probability of failure, and 4 denotes businesses with the highest probability of failure. 3 stands for businesses near the national average probability of failure. This Risk Class enables a customer to quickly segment its new and existing accounts into various risk groups for high-level analysis and reporting.

Table 1 shows the distribution of the failure Risk Cass Failure Score in the Portuguese database, as well as the percentile ranking and Failure Score.

Table 1: Distribution of Failure Score in the Portuguese database

FAILURE RISK CLASS	FAILURE PERCENTILE	FAILURE SCORE	% OF BUSINESSES WITHIN EACH CLASS
1	75 - 100	1495 - 1999	24.67%
2	38 - 74	1447 - 1494	44.05%
3	12 - 37	1326 - 1446	21.84%
4	1 - 11	1001 - 1325	9.43%

Note: The '% of Businesses within Each Class' column is prior to applying any overrides.

MODEL PERFORMANCE

Informa took into consideration different indicators to evaluate the model performance in terms of its discriminatory power

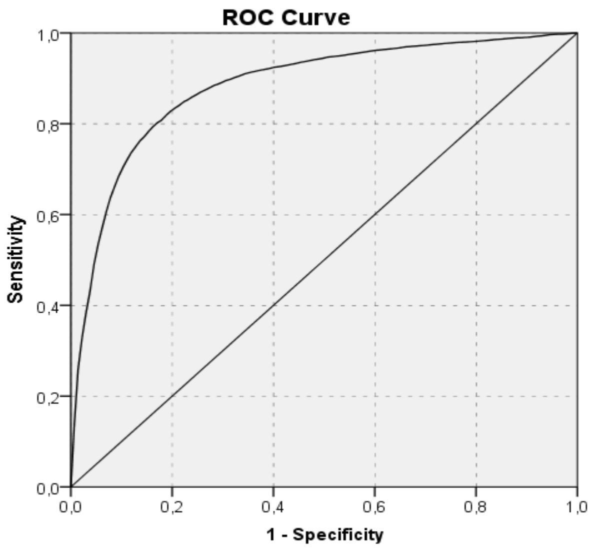
- Ranking accuracy by model, decile or quintile
- Close match between predicted and actual bad rates
- The Kolmogorov-Smirnoff statistic distance between cumulated distribution of good and bad cases as rank ordered by the model
- Predictive Index or Cumulative Accuracy Profile assessment of model gains compared to a random classifier
- The lift Gain chart with emphasis on showing the improvement in capturing BADS at the 10th and 20th scores

One way to measure model performance is by examining a trade-off curve. A trade-off curve or a Receiver Operating Characteristic curve is a plot of ascending accumulation of "Good" businesses vs. "Bad" businesses. It is useful for illustrating model performance both at a particular score and across the spectrum of score distribution.

The trade-off curve in Graph 1 illustrates the effectiveness of the Failure Score by identifying the failure captured within population groups. At approximately 20% of the population, the Failure Score scores identified approximately 80% of the "Bads". This means that if a business focused on the worst scoring 20% of their portfolio using the Failure Score, they would capture 80% of the "Bads" in that group.

The trade-off curve in Graph 1 illustrates the effectiveness of the Failure Score.

Graph 1: Failure Score performance



Scorecards are developed assuming that the relationships observed between past business characteristics and subsequent performance will hold true on future businesses. Because of this assumption, development statistics should be viewed as estimates, and not precise forecasts, of future performance at a given score.

SCORE PERFORMANCE MONITORING

Informa is committed to provide the highest quality risk assessment of businesses in Portugal. In that sense, regular performance monitoring of the scorecards and backtesting exercises are used to ensure the maintenance of high performance standards of the scores in discriminating the failure risk of businesses in Portugal. Whenever required, adjustments and recalibration are applied to keep up the performance of scores.

RELATIONSHIP BETWEEN THE FAILURE SCORE AND THE PROJECTED FAILURE RATES

A Risk Class is designed as a high-level segmentation tool defined into 4 classes. A Risk Class is statistically determined by the similarity of risk within the classes in contrast to other classes. Cases with the lowest Failure Risk fall in Class 1, whereas cases with the highest risk are in Class 4. Risk Class 3 is close to the national average.

Table 2 National average failure rate by Failure Risk Class

Table 2: National average failure rate by Risk Class

FAILURE RISK CLASS	% OF INFORMA FILE REPRESENTED	PROJECTED BAD RATE WITHIN RISK CLASS	PROJECTED CUMULATIVE % OF FAILURES IDENTIFIED
1	24.67%	0.11%	2.22%
2	44.05%	0.25%	11.35%
3	21.84%	1.03%	29.22%
4	9.43%	8.82%	100.00%

APPENDIX A

SELECTED ELEMENTS USED IN THE MODEL

Following is a list of some of the predictors used in the Failure Score Model:

Demographic Information and Activity Signs

VARIABLE	IMPACT ON MODEL
Sector of activity	Some sectors of activity have a higher exposure to economic crises than others, and so their business risk and risk of failure is also greater.
Legal form	Some legal forms are related to higher risk than others, especially when this information is combined with the type of shareholders.
Region	Different regions have different levels of risk which are reflected in the failure model
Age of company	Recently launched businesses are related to a very low risk of failure, as the underlying problems that normally lead to the failure of the business may not have emerged yet. Also, in general, more established businesses have greater stability, and hence their risk is reduced.
Age of balance sheet (whenever available)	The more outdated is the financial information of a firm, the fewer are the signs that it is actively trading; thus, the greater are the signs of the risk of an incoming situation of failure.
Number of employees	Businesses with heavy structures of human resources, i.e., higher weight of fixed costs, tend to have a higher risk of failure when compared to lighter structures.

Financial Information

VARIABLE	IMPACT ON MODEL
Existence of financial information	The disclosure of financial information is associated to lower opacity of a business and therefore to a lower risk.
Solvability	A higher solvability ratio (Equity/Debt) means a lower exposure to external debt and therefore o lower risk.
Net Return on Assets	A higher Net Return On Assets shows a more economically sound and stable business and, therefore, a lower risk.
Retained Earnings over Assets	The higher is the reinvestment of the firm, the better will be its future prospects of not having a situation of failure.
Revenues	A higher value of revenues is normally associated to synergy effects that help a business continue working, and hence reduces the risk of failure. Even so, rather small business, normally related to family businesses, also have a lower probability of failure.
Relative Weight of Debts to Public Entities	The higher the weight the higher is the risk.
Equity	Negative values of equity are related to a higher probability of failure.

Payment Information

VARIABLE	IMPACT ON MODEL
Recent Paydex®	Good payments on the most recent month indicates a lower level of risk.
Negative Data	The lack of historical negative data against a business is a strong indicator of low likelihood of closure within 12 months with unpaid debts.
Age, type and value of negative data	The higher the value of legal demands related to payment disputes and the more recent they are, the higher is the risk of failure. The risk also varies depending on the type of legal demands (e.g., Tax debts).

APPENDIX B

PROJECTED PERFORMANCE TABLES

The following Summary and Detailed Projected Performance Tables are based on a representative sample and actual performance may vary based on individual customer portfolios.

SUMMARY PROJECTED PERFORMANCE TABLES

CUMULATIVE FAILURE SCORE PERFORMANCE						
RISK CLASS	SCORE RANGE	PERCENTILE RANGE (APPROX.)	% OF BUSINESSES (APPROX.)	FAILURE RATE	% OF FAILURES IDENTIFIED	GOOD-BAD RATIO
1	1495 - 1999	75 - 100	24.67%	0.11%	2.22%	894.65
2	1447 - 1999	38 - 100	68.73%	0.20%	11.35%	495.39
3	1326 - 1999	12 - 100	90.57%	0.40%	29.22%	251.10
4	1001 - 1999	1 - 100	100.00%	1.22%	100.00%	81.00

FAILURE SCORE PERFORMANCE WITHIN RANGE				
SCORE RANGE	PERCENTILE RANGE (APPROX.)	% WITHIN RANGE (APPROX.)	FAILURE RATE	% OF FAILURES IDENTIFIED
1495 - 1999	75 - 100	24.67%	0.11%	2.22%
1447 - 1494	38 - 74	44.05%	0.25%	9.13%
1326 - 1446	12 - 37	21.84%	1.03%	17.86%
1001 - 1325	1 - 11	9.43%	8.82%	70.78%

EXPLANATIONS

CUMULATIVE FAILURE SCORE PERFORMANCE

- **% of Businesses:** To set an approval rate, select the appropriate percentile range that yields the desired approval rate. For example, to develop a credit policy that approves a projected 90.57% of all customers requires accepting businesses scoring at or above 1326 (or 12 - 100 percentiles). Businesses scoring below the cutoff score 1326 are reviewed, declined, etc.

- **Failure Rate:** The failure rate represents those businesses that score between the lowest value in the score range (or percentile) and 1999 (or 100 percentile). For example, the failure rate for a credit policy which approves all businesses with a score at or above 1326 (or 12 - 100 percentile) is expected to be 0.4%.

- **% of Failures Identified:** The percentage of total failed businesses that score between 1001 and the cutoff point for the approval rate. For example, approving businesses with a score at or above 1326 (or 12 - 100 percentile) is expected to eliminate 70.78% of the "Bad" businesses.

- **Good-Bad Ratio (Odds):** The ratio of "Good" businesses to "Bad" businesses among those businesses that score between the lowest value in the score range and 1999 (or 100 percentile). For example, a credit policy that approves all businesses scoring at or above 1326 (or 12 - 100 percentiles) should result in a portfolio with 251.10 "Good" businesses for every "Bad" business in the portfolio.

FAILURE PERFORMANCE WITHIN RANGE

- **Failure Rate within Range:** The failure rate for those businesses that score within the score range. For example, the failure rate for businesses scoring between 1001 - 1325 (or 1 - 11 percentile) is expected to be 8.82%.

- **% Of Failures Identified:** The percentage of total failed businesses within the score range. For example, 70.78% of failed businesses are expected to score between 1001 - 1325 (or 1 - 11 percentile).

DETAILED PROJECTED PERFORMANCE TABLE

CUMULATIVE FAILURE SCORE PERFORMANCE						FAILURE SCORE PERFORMANCE WITHIN RANGE			
SCORE RANGE	PERCENTILE RANGE (APPROX.)	% OF BUSINESSES (APPROX.)	FAILURE RATE	% OF FAILURES ELIMINATED	GOOD-BAD RATIO	SCORE RANGE	PERCENTILE RANGE (APPROX.)	FAILURE RATE	% OF FAILURES IDENTIFIED
1554 - 1999	96 - 100	5%	0.05%	0.18%	2103.4	1554 - 1999	96 - 100	0.05%	0.18%
1532 - 1999	91 - 100	10%	0.06%	0.46%	1258.8	1532 - 1553	91 - 95	0.08%	0.28%
1515 - 1999	86 - 100	15%	0.08%	0.85%	894.8	1515 - 1531	86 - 90	0.11%	0.39%
1506 - 1999	81 - 100	20%	0.10%	1.55%	717.4	1506 - 1514	81 - 85	0.14%	0.70%
1498 - 1999	76 - 100	25%	0.11%	2.08%	623.0	1498 - 1505	76 - 80	0.16%	0.53%
1490 - 1999	71 - 100	30%	0.12%	2.98%	536.8	1490 - 1497	71 - 75	0.19%	0.90%
1485 - 1999	66 - 100	35%	0.14%	4.09%	488.2	1485 - 1489	66 - 70	0.20%	1.11%
1480 - 1999	61 - 100	40%	0.15%	5.59%	450.9	1480 - 1484	61 - 65	0.22%	1.50%
1475 - 1999	56 - 100	45%	0.17%	6.99%	409.9	1475 - 1479	56 - 60	0.24%	1.40%
1470 - 1999	51 - 100	50%	0.18%	8.37%	375.9	1470 - 1474	51 - 55	0.27%	1.38%
1465 - 1999	46 - 100	55%	0.19%	9.38%	344.1	1465 - 1469	46 - 50	0.29%	1.01%
1456 - 1999	41 - 100	60%	0.19%	10.60%	311.0	1456 - 1464	41 - 45	0.32%	1.22%
1441 - 1999	36 - 100	65%	0.21%	11.93%	247.0	1441 - 1455	36 - 40	0.40%	1.33%
1425 - 1999	31 - 100	70%	0.22%	13.63%	189.9	1425 - 1440	31 - 35	0.52%	1.70%
1413 - 1999	26 - 100	75%	0.25%	16.06%	146.7	1413 - 1424	26 - 30	0.68%	2.43%
1390 - 1999	21 - 100	80%	0.29%	19.64%	117.5	1390 - 1412	21 - 25	0.84%	3.58%
1348 - 1999	16 - 100	85%	0.35%	25.28%	58.7	1348 - 1389	16 - 20	1.67%	5.64%
1319 - 1999	11 - 100	90%	0.44%	32.77%	33.9	1319 - 1347	11 - 15	2.87%	7.49%
1270 - 1999	6 - 100	95%	0.67%	52.91%	17.4	1270 - 1318	6 - 10	5.43%	20.14%
1001 - 1999	1 - 100	100%	1.22%	100.00%	6.6	1001 - 1269	1 - 5	13.24%	47.09%

EXPLANATIONS

CUMULATIVE FAILURE SCORE PERFORMANCE

- **Approval Rate:** To use, select the appropriate projected score or percentile cutoff that yields the desired approval rate. Approved businesses are companies scoring between the lowest value in the score range (or percentile) and 1999 (or 100 percentile). For example, a credit policy that approves 70% of all businesses requires accepting businesses between 1425 - 1999 (or 31 - 100 percentiles). Businesses scoring below the cutoff (1001 - 1424) are reviewed, declined, etc.

- **Failure Rate:** Represents those businesses that score between the lowest value in the score range and 1999. For example, the failure rate for a credit policy which approves all businesses with a score at or above 1425 (or 31 - 100 percentiles) is expected to be 0.22%.

- **% of Failures Identified:** The percentage of total failed businesses that score between 1001 and the cutoff point for the approval rate. For example, approving businesses with a score at or above 1425 (31 - 100 percentile) is expected to eliminate 100%–13.63%=86.37% of the "bad" businesses.

- **Good-Bad Ratio (Odds):** The ratio of "Good" businesses to "Bad" businesses among those businesses that score between the lowest value in the score range and 1999 (or 100 percentile). For example, a credit policy which approves all businesses scoring at or above 1425 (or 31 - 100 percentiles) should result in a portfolio with 189.9 "Good" businesses for every "Bad" business in the portfolio.

FAILURE SCORE PERFORMANCE WITHIN RANGE:

- **Failure Rate:** The incidence of failure for those businesses that score within the score range. For example, the failure rate for companies scoring between 1413 - 1424 (or 26 - 30 percentiles) is expected to be 0.68%.

- **% of Failures Identified:** The percentage of total failed businesses within the score range. For example, 2.43% of all failed companies are expected to score between 1413 and 1424 (or 26 - 30 percentiles).

APPENDIX C

GLOSSARY OF SCORING TERMS

TERM	EXPLANATION
Informa Failure Score	Risk Score predicting likelihood of Failure
Raw Score	Score with a direct relationship to Probability of Failure. The Failure form of the raw score is a 4 digit score
Percentile Score	Lesser granularity of the Failure Score: Values range between 1 and 100, where 1 is the highest probability of failure
Risk Class	Lowest granularity of Failure Score; Segmentation of the Failure Score is done into 4 risk segments, where 1 denotes the lowest probability of failure
Scoreable Universe	Includes all records in the Data Cloud which meet criteria for score assignment. Examples of records excluded from the Scoreable Universe include Out of Business records, Foreign Companies etc.
Scored Universe	Includes all cases which have a score assigned
Observation Point	Date at which the data sample of active businesses is extracted and where data elements observed at that point in time are evaluated as potential predictors
Performance Window	Period where the data sample is monitored to classify businesses as GOOD and BAD
Failure BAD definition	List of Legal Events that define the targeted risk behavior
BAD	A business which meets the Bad definition, i.e., a business which has been subject to one or more of the legal events defined as failure.
GOOD	A business which does not have any information listed within the BAD definition, i.e. a business which has not been subject to any of the legal events defined as failure.
Out of Business	Businesses that are no longer trading