Manoel Fernando Alonso Gadi
Associate Professor
mfalonso@faculty.ie.edu

BDBA – Introd. to Business and Social Analytics

Business Banking Analytics - Ratios and Rating

¿What is Rating?

Rating scale: letter grading

Bravo Rating	Moodys	S&P	Fitch
AAA	Aaa	AAA	AAA
AA+	Aa1	AA+	AA+
AA	Aa2	AA	AA
AA-	Aa3	AA-	AA-
A+	A1	A+	A+
A	A2	A	A
A-	A3	A-	A-
BBB+	Baa1	BBB+	BBB+
BBB	Baa2	BBB	BBB
BBB-	Baa3	BBB-	BBB-
BB+	Ba1	BB+	BB+
BB	Ba2	BB	BB
BB-	Ba3	BB-	BB-
B+	B1	B+	B+
B	B2	B	B
B-	B3	B-	B-
CCC+	Caa1	CCC+	CCC+
CCC	Caa2	CCC	CCC
CCC-	Caa3	CCC-	CCC-
CC	Ca	CC	CC
C	C	C	C
DDD	C	DDD	DDD
DD	C	DD	DD
D	C	D	D

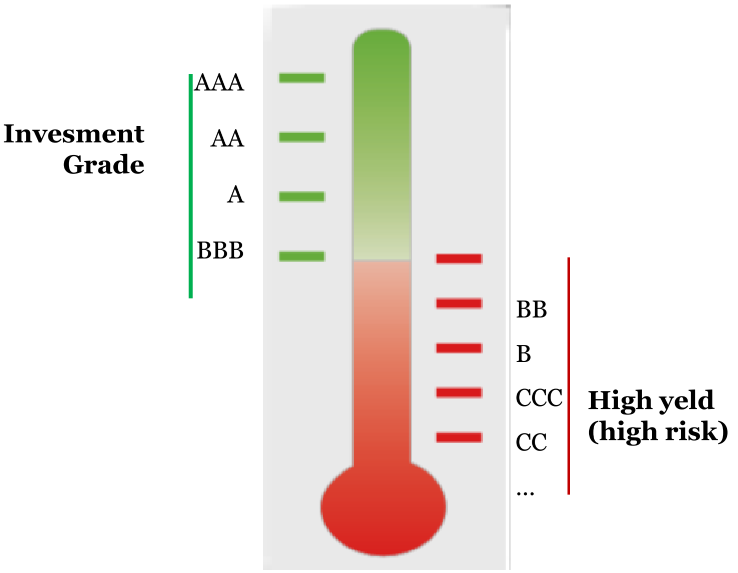
Investment Grade

High yeld

¿What is Rating?

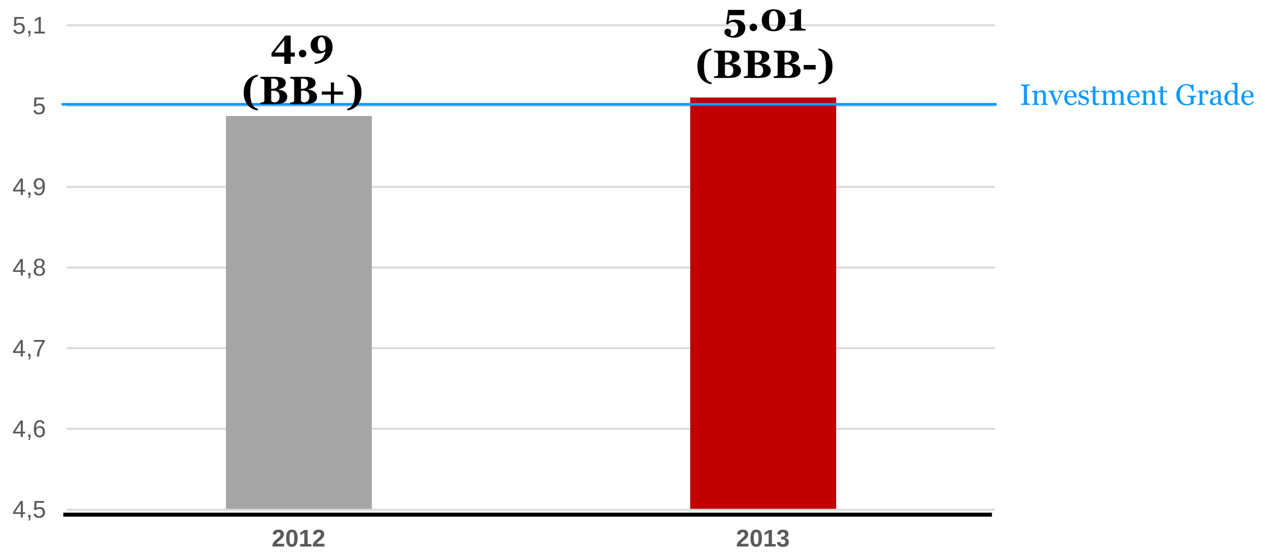
Numeric rating: scale from 0 to 10

Bravo Rating	Escala numérica
AAA	9.9 ≤ r < 10
AA+	9.6 ≤ r < 9.9
AA	9.2 ≤ r < 9.6
AA-	8.5 ≤ r < 9.2
A+	8 ≤ r < 8.5
A	7.6 ≤ r < 8
A-	6.9 ≤ r < 7.6
BBB+	6.2 ≤ r < 6.9
BBB	5.6 ≤ r < 6.2
BBB-	5.0 ≤ r < 5.6
BB+	4.4 ≤ r < 5.0
BB	3.9 ≤ r < 4.4
BB-	3.3 ≤ r < 3.9
B+	2.7 ≤ r < 3.3
B	2.2 ≤ r < 2.7
B-	1.6 ≤ r < 2.2
CCC+	1.4 ≤ r < 1.6
CCC	1.2 ≤ r < 1.4
CCC-	1.0 ≤ r < 1.2
CC	0.8 ≤ r < 1.0
C	0.6 ≤ r < 0.8
DDD	0.4 ≤ r < 0.6
DD	0.2 ≤ r < 0.4
D	0 ≤ r < 0.2



What is the Spanish Rating Distribution

Average Rating of the Spanish companies

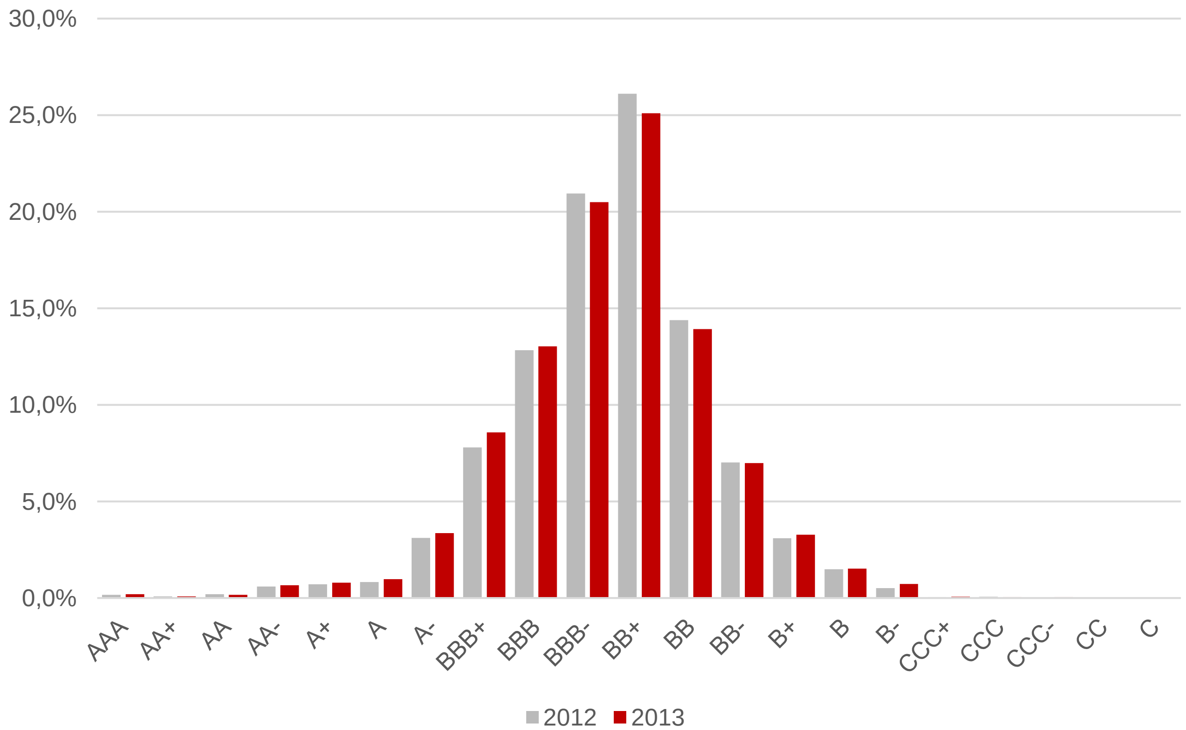


Spanish companies get an insvetment grade in 2013, after 4 years of high yeld

Source: Bravo Capital, with data from Registro Mercantil using 26.401 companies having turn over above 5 million €.

1. Company distribution by Rating (grade)

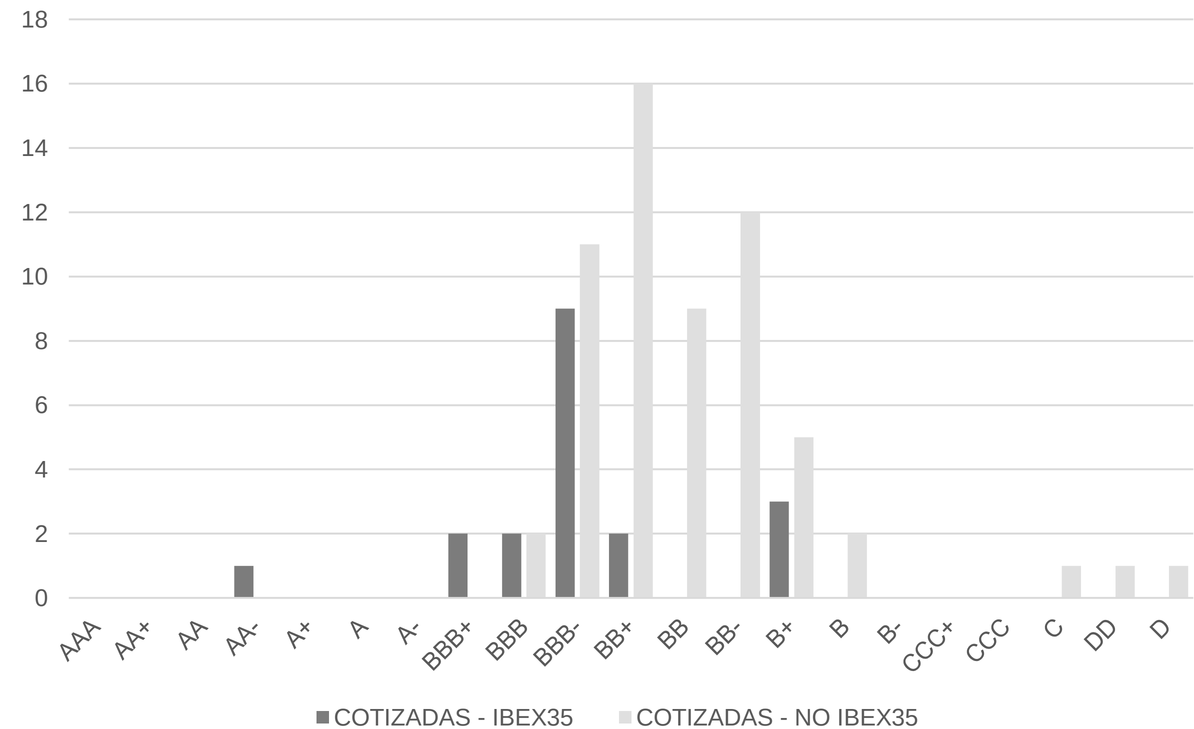
Spanish companies Rating Grade distribution



45% of Spanish companies are between BBB- (limit of investment grade) and BB+

Source: Bravo Capital, with data from Registro Mercantil using 26.401 companies having turn over above 5 million €.

Comparison between IBEX 35 and listed non-Ibex 35 comanies (empresas del mercado continuo)

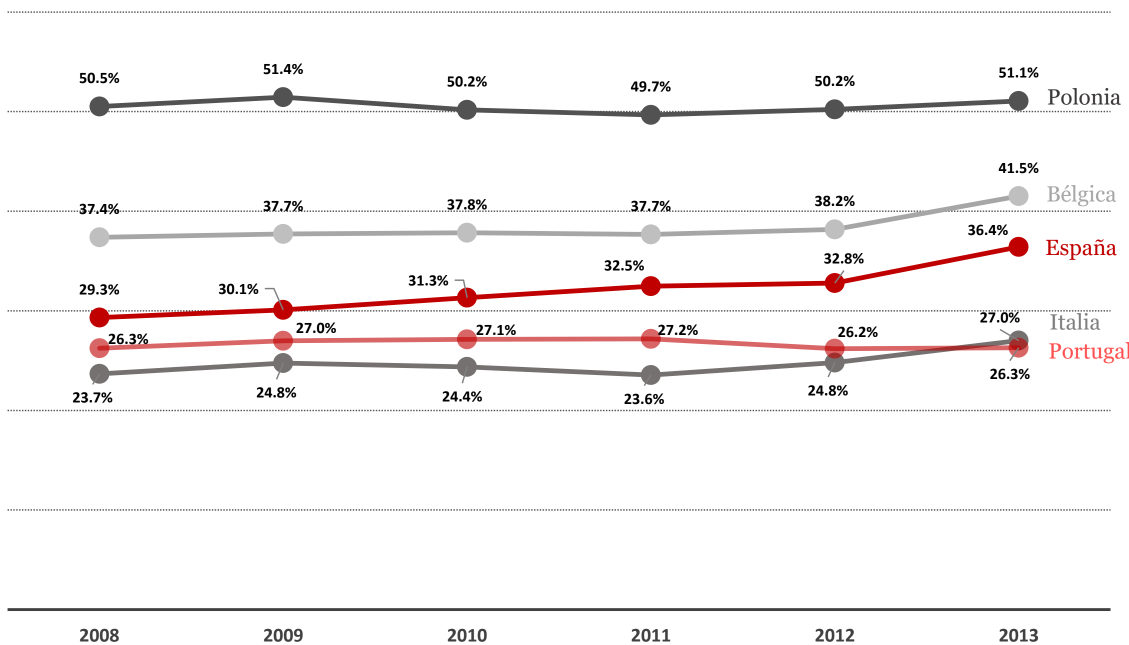


IBEX 35 have better Rating than the Spanish average, while other listed companies are worse tan the Spanish average.

Source: Registro Mercantil, Bravo Capital.

4. The economy

Comparison of the Ratio of Equity / Total Assets (%) by country



Source: Bravo Capital.

Let's calculate some of the measures in the Case Study using Python

IE_MBD_FA_s07n8_YahooFinanceKeyStats4RatingCalc.ipynb

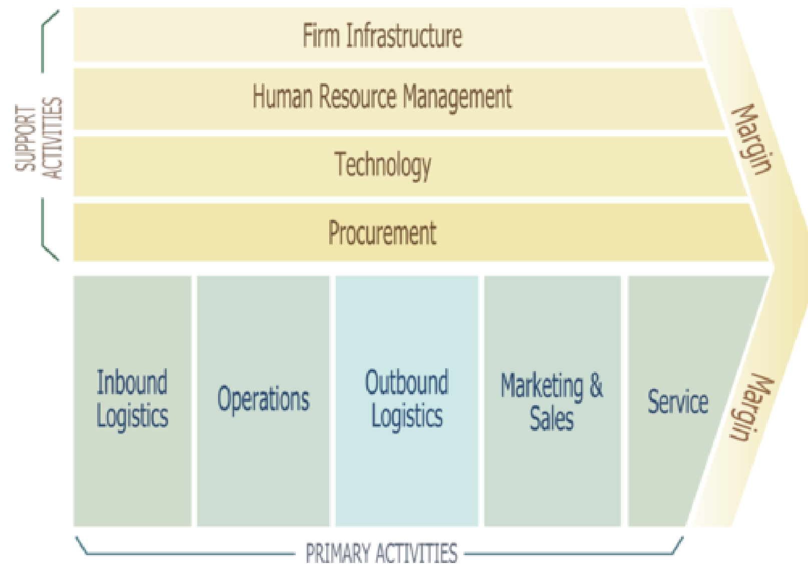
(available on campus)

Business Banking

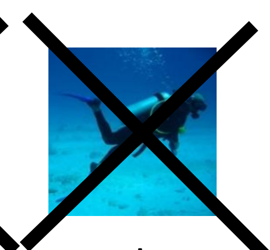
- It is possible to model using social demographics
- Probability of default
- Probability of accepting an offer
- Probability of fraud

- It is very hard to model directly.
- Model the way it operates
(balance sheet – expected ratios similar sector/size)

Limited space for behaviour variation

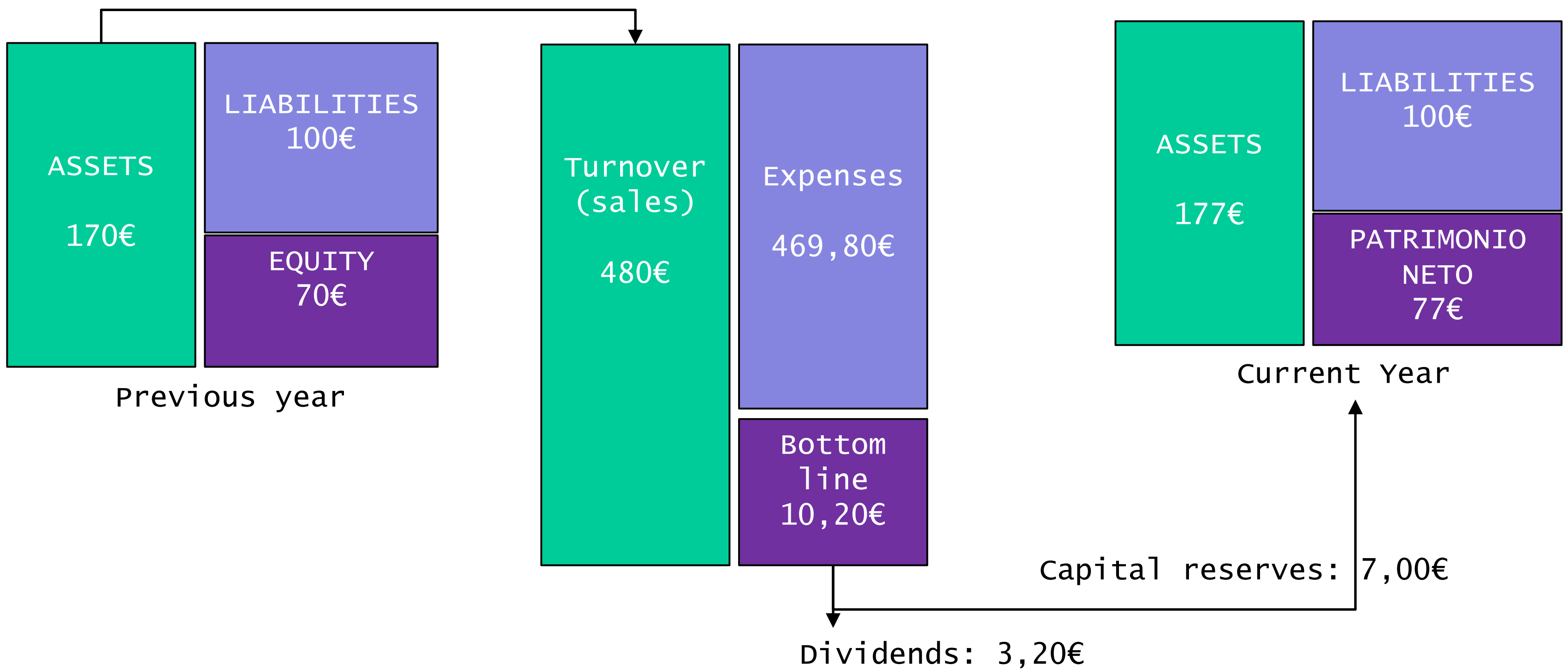


Plastic Surgery



Scuba Diving

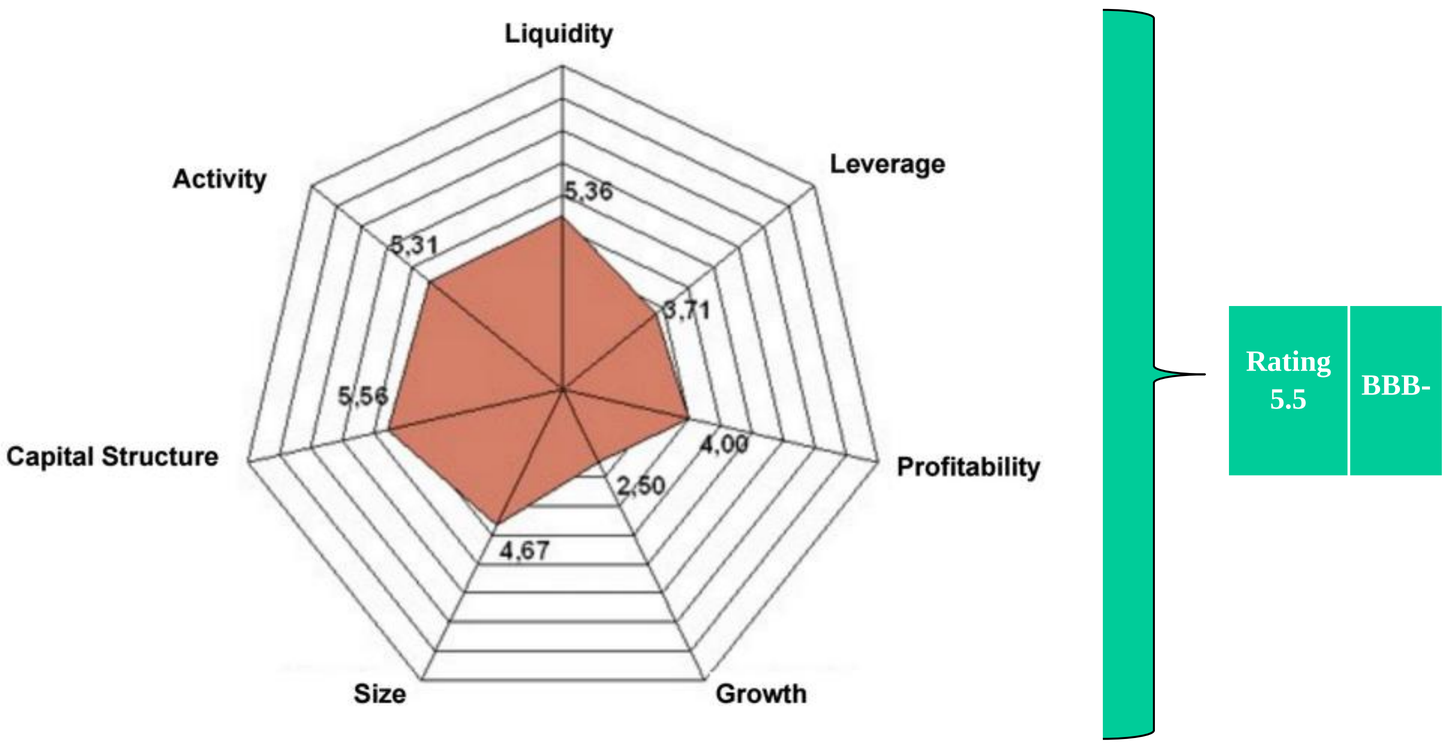
Balance sheet
and
Profit and Loss account



Rating areas

- Rating
  - Quantitative Ratios
    - Activity
    - Liquidity
    - Leverage
    - Profitability
    - Capital Structure
    - Size Appropriateness
    - Grow Ratios
  - Qualitative Overlay
    - Area 1. Product Demand / Market
    - Area 2. Shareholders / Management
    - Area 3. Access to Credit
    - Area 4. Profits
    - Area 5. Generation of Resources
    - Area 6. Solvency

After creating each ratio our objective is to transform it into a single grading, a Rating.



Rating (Letters)

RATING
AAA
AA+
AA
AA-
A+
A
A-
BBB+
BBB
BBB-
BB+
BB
BB-
B+
B
B-
CCC+
CCC
CCC-
CC
C
DDD
DD
D

Let's review some Ratios examples in Python

ie_mbd_s07n8_calc_ratios_n_rating.ipynb

(available on campus)

Using ratios in modelling:

Ratio:

Inventory turnover in days = 365 * Stock / Sales [in days]
(12200) / (40100)

Inventory turnover in days = 365 * 1.877.282 € / 8.405.131 € = 81,52 days

Cons of using the ratio as it is:

- Drives dominance when combining ratios with very different scales.
- It overestimates differences between very good and average.

Transforming ratios – part 1

Binning

- **Data :** 0, 4, 12, 16, 16, 18, 24, 26, 28
- **Equal width**
  - Bin 1: 0, 4 [-,10)
  - Bin 2: 12, 16, 16, 18 [10,20)
  - Bin 3: 24, 26, 28 [20,+)
- **Equal frequency**
  - Bin 1: 0, 4, 12 [-, 14)
  - Bin 2: 16, 16, 18 [14, 21)
  - Bin 3: 24, 26, 28 [21,+)

Transforming ratios – part 2

2.1 Percentile (Rank or Quantile)

2.2 Normalization

2.3 Bucketing Normalization

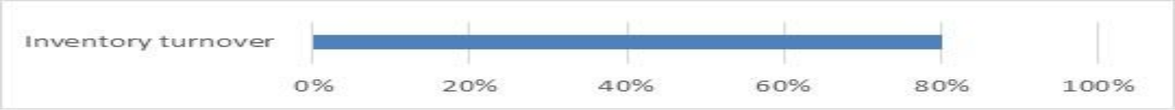
Transforming ratios – part 2.1

2.1 Percentile (Rank or Quantile)

Inventory turnover in days = 365 * Stock / Sales [in days]
(12200) / (40100)

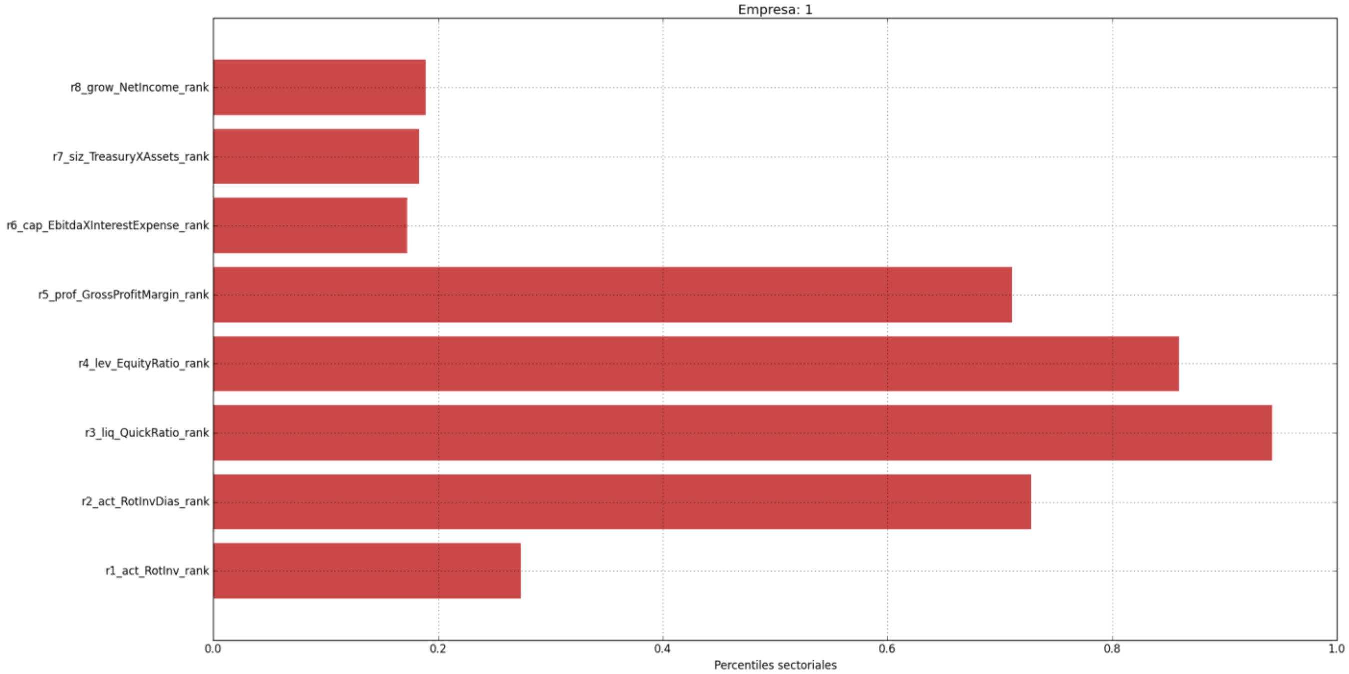
Inventory turnover in days = 365 * 1.877.282 € / 8.405.131 € = 81,52 days

Variable transformation (equivalent to categorization for individuals):
option 1): Percentile



Easy to understand, but requires constant update of the ranking (MEAN: as it works with the order it eliminates problems with outliers driving any mean calculation), it underestimates differences between very good and average.

Ranking in comparison with other companies



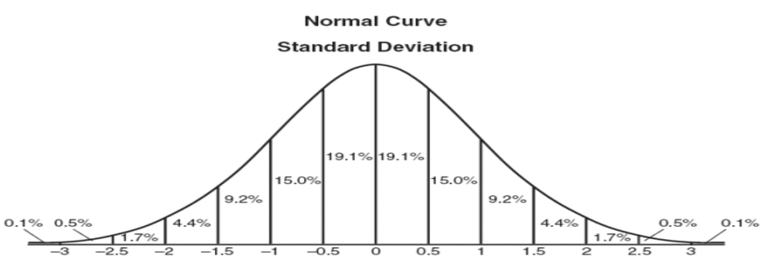
Transforming ratios – part 2.2

2.2 Normalization

Inventory turnover in days = 365 * Stock / Sales [in days]
(12200) / (40100)

Inventory turnover in days = 365 * 1.877.282 € / 8.405.131 € = 81,52 days

Variable transformation (equivalent to categorization for individuals):
option 1): Percentile

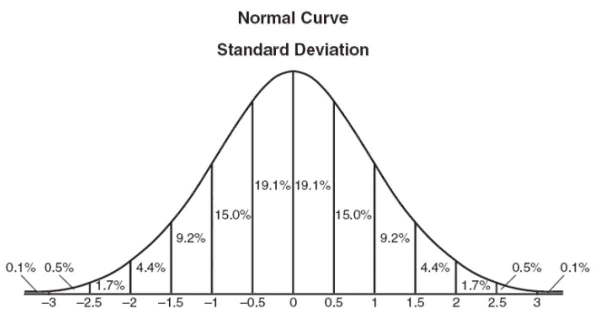


A little bit harder to visualise,
but it does not require constant update. (RISK: outliers must be eliminated)

Variable = (121 days – median)/ standard deviation

Normalization is not perfect, but very useful

Inventory turnover in days = 365 * Stock / Sales [in days]
(12200) / (40100)



Assumes it has a normal distribution

Rotación de inventario en días Transformad	% del total ejemplo real
<-0.5	80.74%
-0.5-0	15.60%
0-0.5	1.36%
0.5-1.0	0.59%
1.0-1.5	0.41%
1.5-2	1.30%
Total general	100.00%

In practice we normally do not have variables that have normal distributions, but this transformation still ranks very well and places all ratios in quite equivalent size and dimension.

You guys should know now how to develop a model in Python

In [2]:

```
import pandas
from pandas import pivot_table
import math
import numpy as np
import re
import statsmodels.api as sm

#################################################################################
#       READING DATA
#################################################################################
print "READIND DATA..."
df = pandas.read_csv("https://dl.dropboxusercontent.com/u/28535341/ciff_mbd_s03_empresas_calc_ratios_v2.csv", sep=';')
print df
```

```
READIND DATA...
      empresa  ejercicio         10000         12000         12200  \
0           1       2012    828.515249    709.049699    233.375710
1           1       2013    748.989022    648.524431    361.432523
2           2       2012  11987.277250   6787.263455   3279.256266
3           2       2013  12902.113200   9654.128274   3272.684230
4           3       2012  56205.467260   9251.156238     76.095916
5           3       2013  56070.996800   8740.391038     83.138436
6           4       2012   2665.685637   2270.297290   1444.226016
7           4       2013   2459.033333   2095.130623   1241.165583
8           5       2012   1399.372247   1347.519100    432.412547
9           5       2013   1242.397076   1211.066694    350.840458
10          6       2013   1502.658905    694.364201    108.799022
```

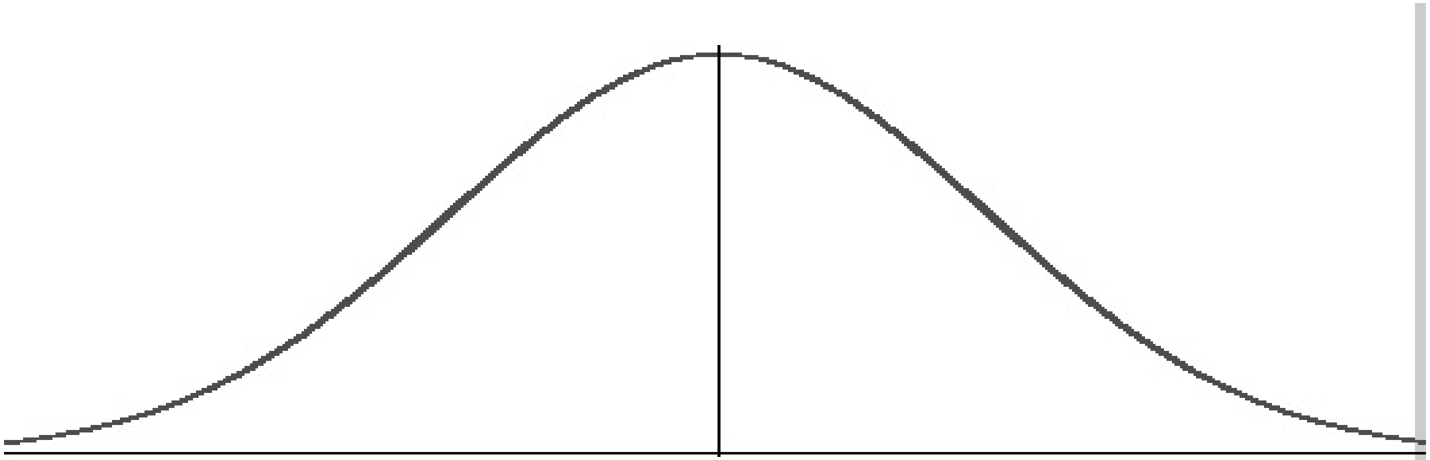
Example of a quick and "dirty" model

```
Model Family:                Binomial   Df Model:                           15
Link Function:                  logit   Scale:                             1.0
Method:                          IRLS   Log-Likelihood:                -328.32
Date:                Thu, 07 May 2015   Deviance:                       656.64
Time:                        12:54:16   Pearson chi2:                 1.57e+03
No. Iterations:                    14
================================================================================================================
                                               coef    std err          t      P>|t|      [95.0% Conf. Int.]
----------------------------------------------------------------------------------------------------------------
r1_act_RotInv_normalised                     0.1413      0.090      1.573      0.116        -0.035     0.317
r2_act_RotInvDias_normalised               185.8368     31.401      5.918      0.000       124.292   247.381
r3_liq_QuickRatio_normalised                40.8477     28.449      1.436      0.151       -14.910    96.606
r4_lev_EquityRatio_normalised               -0.0800      0.702     -0.114      0.909        -1.457     1.297
r5_prof_GrossProfitMargin_normalised        -0.0543      0.127     -0.427      0.669        -0.303     0.195
r6_cap_EbitdaXInterestExpense_normalised     0.0093      0.086      0.109      0.913        -0.159     0.177
r7_siz_TreasuryXAssets_normalised            1.0017      0.245      4.090      0.000         0.522     1.482
r8_grow_NetIncome_normalised                -0.2880      0.200     -1.443      0.149        -0.679     0.103
r1_act_RotInv_rank_normalised                2.1922     39.724      0.055      0.956       -75.666    80.050
r2_act_RotInvDias_rank_normalised            2.6069     39.723      0.066      0.948       -75.248    80.462
r3_liq_QuickRatio_rank_normalised            0.1416      0.157      0.901      0.368        -0.166     0.450
r4_lev_EquityRatio_rank_normalised          -1.3355      0.196     -6.798      0.000        -1.721    -0.950
r5_prof_GrossProfitMargin_rank_normalised   -0.1482      0.123     -1.207      0.227        -0.389     0.092
r6_cap_EbitdaXInterestExpense_rank_normalised 0.6150     0.148      4.165      0.000         0.326     0.904
r7_siz_TreasuryXAssets_rank_normalised      -0.0102      0.156     -0.066      0.948        -0.316     0.295
r8_grow_NetIncome_rank_normalised           -0.1693      0.110     -1.538      0.124        -0.385     0.046
================================================================================================================
```

Transforming ratios – part 2.3

2.3 Bucketing Normalization

ROA = Net Income / Total Assets



Percentage of companies	1.6%	3.2%	6.5%	12.9%	25.8%	25.8%	12.9%	6.5%	3.2%	1.6%
ROA to be in the group	< -29%	>= -29%	>= -17%	>= -7%	>= -2.3%	>= 0.25%	>= 1.6%	>= 4%	>= 7%	>= 15%
Bucket Associated	9	8	7	6	5	4	3	2	1	0
Translation of Bucket	D	C	CC	CCC	B	BB	BBB	A	AA	AAA

Example of a Bucket Model:

PLAYA REAL, SA

cif **A07025919** Province **Palma**

Rating Calculation (% ; times)

	2013 Consolidados			2014 Consolidados		
Debt Coverage *Gross Debt / EBITDA*	-87.3	7	CC	-30.9	7	CC
Interest Coverage *EBITDA / Financial Expenditure*	-0.5	6	CCC	-0.8	6	CCC
Current ratio *Current Assets / Current Liabilities*	0.6	8	C	0.6	8	C
Quick ratio *(Current assets - Inventories) / Current liabilities*	0.6	7	CC	0.6	7	CC
Cash Flow Liquidity *(Cash + Short Term Investments + Operating Cashflow) / Current*	13.1%	7	CC	33.0%	5	B
Gross Financial Debt/ Total Asse *Gross Financial Debt/ Total Assets*	80.7%	8	C	73.9%	8	C
Fondos Propios / Total Activos *Fondos Propios / Total Activos*	13.5%	7	CC	12.0%	7	CC
Assets Turnover *Sales / Total Assets*	10.4%	9	D	0.1%	9	D
Return on Assets (ROA) *Net Income / Total Assets*	-3.0%	6	CCC	-4.4%	6	CCC
EBITDA Margin *Ebitda / Turnover*	-8.9%	8	C	-2120.3%	9	D
Operating Profit Ma *Operating Profit / Turnover*	-18.5%	8	C	-2988.0%	9	D
Variation of EBITDA *EBITDA (Year N) / EBITDA (Year N-1*		9	D	-63.6%	9	D
Numeric Rating		**1.97**			**1.82**	
Rating Grade		**CC**			**CC-**	

Main steps in Machine Learning

(A) Sampling and Performance definition

- Bad definition (which is a default (1) and a non-default (0)? What is a fraud (1) and non-fraud (0)?
- Observation period and performance (days, months, 12 months, calendar years)

(B) Feature Engineering

- Combining or transforming existing explanatory features into ratios. For example, age instead of date of birth.
- Here, it plays a lot the creativity and experience of the analyst.

(C) Preprocessing

- Standardization of values (mean removal and variance scaling), Normalization, Binarization, Encoding categorical features, Imputation of missing values.

(D) Model Fit & Feature Selection

- Depending on the problem under analysis a technique may fit better.
- Methods: Logistic Regression / Decision Tree / Neural Networks
- Feature Selection: Forward, Backward, Stepwise Selection

(E) Model Evaluation

- Accuracy Score
- ROC curve and Gini Score
- F1 Score
- Mean Square Error
- R Squared

(F) Cross Validation

- Cross Validation is a technique used to increase the model generalization power, meaning it is useful not only to predict training data but also to be applied into the real world data. It does it by avoiding overfitting.

(D.2) Hyperparameter (normally wrongly skipped)

- Hyperparameter refer to parameters of a given method (example: Logistic Regression, Decision Tree, etc). In general a method have many hyperparameter and they come with a default set of values that have no reason to work well for your current application, so it is a good idea to do some Hyperparameter optimization specifically to your problem.

Manoel Fernando Alonso Gadi
Associate Professor
mfalonso@faculty.ie.edu

BDBA – Introd. to Business and Social Analytics

Business Banking Analytics - Ratios and Rating