WHARTON FINANCIAL ANALYTICS

Data Lab

RATERx

Michael R. Roberts

April 2021

Wharton University of Pennsylvania

About Wharton Financial Analytics

The **Wharton Financial Analytics** (**WFA**) initiative was created in 2020 for the purpose of empowering people to make better financial decisions through cutting edge research, innovative education programs, and industry engagement. In doing so, WFA serves the Wharton School, the University of Pennsylvania, and the broader global community.

For more information about Wharton Financial Analytics and its products and services, please visit our website at:
http://finance.wharton.upenn.edu/~mrrobert/
or contact
mrrobert@wharton.upenn.edu

Michael R. Roberts is the William H. Lawrence Professor of Finance at The Wharton School of the University of Pennsylvania. He developed this data lab solely for the purpose of class discussion. The lab is not intended to serve as an endorsement, source of data, or illustration of effective or ineffective decision making. Although based on real events and despite occasional references to actual companies, this lab is fictitious and any resemblance to actual persons or entities is coincidental.

Lewis Riccardo was excited but nervous. He had just completed a first-round Zoom interview for an analytics position at RATERx LLC. RATERx was a financial analytics consulting firm providing credit risk assessment and forecasting services to institutional investors. RATERx was also Lewis' first choice for full-time employment after he graduated three months from now in May of 2021. However, he knew his performance on the second round would largely determine his chances of getting an offer.

The second-round interview at RATERx consisted of a "project" for which candidates were given one week to complete. Each project was different so there was no way for past or current candidates to share information with one another. Further, candidates were given no time to prepare as progression to the second round began immediately after the first-round interview.

Specifically, candidates were given a task and data, similar in spirit to what they might face on the job though less challenging given the time constraints and lack of support. The deliverables at the end of the week were a 20-minute slide presentation given to three executives and all computer code used to generate any analysis and visuals. Lewis was asked to develop a machine learning model that could distinguish between investment-grade and below-investment-grade firms.

The clock was ticking.

RATERx

RATERx LLC. was founded on February 18, 2011 by Roni Brav and Isaac Garrison. Between the two of them, they had over 40 years of experience in the fixed income space at companies including Lehman Brothers, Pacific Investment Management Company (PIMCO), Morgan Stanley, and J.P. Morgan. Venturing out on their own was a natural progression for the former college roommates, both of whom appreciated the analytical elements of the fixed income space more than the trading or portfolio management aspects.

Headquartered in Mountain View, California, RATERx was a consulting firm providing cloud-based software solutions for assessing and forecasting credit risk. Their services included:

- **Portfolio management** to effectively structure and trade their fixed income portfolios.

- **Risk management** to quantify and monitor risk exposures.
- **Regulatory management** to ensure efficient regulatory compliance.

Their clients consisted of financial institutions, such as banks and insurance companies, and institutional investors, such as pension funds, endowments, and hedge funds; most of whom were obtained by leveraging the network of professionals Roni and Isaac had developed during their time on Wall Street and through word-of-mouth.

During the early years of the company, revenue was driven by custom software deployments and on-site consulting. Recognizing the limitations of this business model for scale, RATERx began transitioning towards a software as a service (SaaS) model in 2014. The company developed a core set of software tools that could be hosted in the cloud and easily customized for specific client needs. This transition dramatically lowered costs – financial and personal. On-site engagements were less frequent, and each engagement was leveraged to expand the core software set and reduce future customization work.

From 2016 to 2020, the company grew rapidly and smartly. Revenue growth hovered around 20% per annum while margins increased slightly as the company began taking advantage of scale economies. The onset of the Covid-19 crisis brought with it turmoil in the credit markets, as well as an explosion in demand for RATERx services. The company could not keep up with demand and wanted to avoid too rapid of an expansion and subsequent contraction when things returned to "normal." At the time of Lewis' interview, RATERx had over 250 employees in the Mountain View and, as of 2018, New York offices. It was an exciting time to work for the company given the spotlight and credit markets and demand for their services.

Credit Risk

Credit risk refers to the possibility that a borrower will not fulfill their financial obligation. This risk can be formalized in the context of the present value relation relating the value of a claim to its future cash flows and discount rates:

$$\text{Present Value}=\frac{E(CF_1)}{(1+r_1)}+\frac{E(CF_2)}{(1+r_2)}+\ldots+\frac{E(CF_T)}{(1+r_T)},$$

where $E(CF_t)$ is the **expected** cash flow in period t, and r_t is the expected return in period t. The expectation is important because it distinguishes what investors expect to receive from what they are promised, CF_t. For a fixed income instrument, such as a corporate bond, the promised cash flows include interest and principal payments. Each period there is some probability, p, that the borrower will default (i.e., not make the promised payment). Thus,

$$E(CF_t) = (1-p) \cdot CF_t + p \cdot L_t$$

where L_t is the liquidating payment to the creditor should the borrower default. This last payment is sometimes expressed as a fraction of the principal owed on the debt, α, referred to as the recovery rate. Thus, credit risk affects creditors' claims by affecting expected cash flows through (1) the probability of default or any loss in default.

Credit risk can also affect investors claims by affecting the expected return, r, if default risk is systematic.[1] Table 1 shows estimated corporate debt betas by rating and by maturity (for investment-grade rated credits).

Table 1. Average Debt Betas by Rating and Maturity

By Rating	A and Above	BBB	BB	B	CCC
Average Beta	< 0.05	0.10	0.17	0.26	0.31

By Maturity		1-5 Year	5-10 Year	10-15 Year	> 15 Year
Average Beta (Investment grade)		0.01	0.06	0.07	0.14

Source: Peter DeMarzo and Jonathan Berk, 2020, *Corporate Finance* 5th Ed., Pearson.

The table shows that market betas are increasing in corporate credit risk, as measured by credit ratings, and maturity.

Credit Ratings

Credit ratings measure the creditworthiness of individual financial instruments or entities, such as corporations, municipalities, and sovereign nations. Ratings are provided by ratings agencies, of which

[11] See Edwin, Elton, Martin Gruber, Deepak Agrawal, and Christopher Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance* 56, 247 – 277.

there are many companies. However, approximately 95% of the rating business is controlled by three agencies: Moody's, Standard & Poor's (S&P), and Fitch. In fact, the Moody's and S&P control approximately 80% of the global market, with Fitch responsible for 15% of the remainder.

Table 2 shows the ratings scales used by the Moody's, S&P, and Fitch to rate corporate debt.[2]

Table 2. Credit Ratings Scales

Moody's	S&P	Fitch	
Aaa	AAA	AAA	Investment grade
Aa1	AA+	AA+	
Aa2	AA	AA	
Aa3	AA-	AA-	
A1	A+	A+	
A2	A	A	
A3	A-	A-	
Baa1	BBB+	BBB+	
Baa2	BBB	BBB	
Baa3	BBB-	BBB-	
Ba1	BB+	BB+	Speculative grade (Below investment grade, High yield, Junk)
Ba2	BB	BB	
Ba3	BB-	BB-	
B1	B+	B+	
B2	B	B	
B3	B-	B-	
Caa1	CCC+	CCC+	
Caa2	CCC		
Caa3	CCC-		
Ca	CC	CC	
C	C	C	
	D	D	

Source: Moody's, Standard & Poor's, and Fitch.

Credit risk is increasing as one moves down the columns of Table 2. Referring to the previous discussion, creditors are less likely to receive their promised payments because the probability of default is increasing and the recovery in default is decreasing as one moves from the top of the scale to the bottom. Table 3 presents descriptions of what each rating means.

[2] Each company employs many ratings scales for other, more refined purposes, such as assessing short-term payment risk or the risk of specialized investment vehicles.

Table 3. Credit Ratings Descriptions

Moody's/S&P Rating	Description
Aaa/AAA	Highest quality and lowest level of credit risk.
Aa/AA	High quality. Together with Aaa these obligations constitute "high-grade" obligations.
A/A	Upper-medium grade and subject to low credit risk.
Baa/BBB	Medium-grade and subject to moderate credit risk. May contain speculative characteristics.
Ba/BB	Speculative and subject to substantial credit risk.
B/B	Speculative and subject to high credit risk.
Caa/CCC	Speculative of poor standing and subject to very high credit risk.
Ca/CC	Highly speculative and near or in default with some prospect of principal and interest recovery.
C/C,D	Lowest rated and typically in default with little prospect of recovery of principal or interest.

Source: Moody's and Standard & Poor's.

While credit risk is a central concern for lenders to corporations, default among firms with rated debt is a relatively rare event as seen in Table 4. The table also shows a strong cyclical component to defaults and the sharp distinction in credit risk between the investment-grade and speculative-grade groups.

How exactly credit ratings are determined is, unsurprisingly, proprietary. However, given the importance of credit ratings to financial markets, there is a great deal of academic research, as well as literature from the ratings agencies themselves, on their determination and implications for asset pricing.[3]

One Week

One week is not a lot of time to execute a machine learning project, especially since Lewis could not put his academics completely on hold. Lewis knew from conversations with alumni at RATERx that the project was meant to be all-encompassing and actionable. Candidates were expected to execute a complete data science workflow including:

- Extract, transform, load (ETL);
- Exploratory data analysis (EDA); and
- Modeling.

[3] See Robert Kaplan and Gabriel Urwitz, 1979, "Statistical Models of Bond Ratings: A Methodological Inquiry," *The Journal of Business* 52, 231 - 261; Darrell Duffie and Kenneth Singleton, 2012, "Credit Risk: Pricing, Measurement, and Management," Princeton University Press; Standard & Poor', "Guide to Credit Rating Essentials"; and Moody's, "Moody's Credit Rating Prediction Model."

Table 4. Global Corporate Default Summary

	Total			Investment-grade		Speculative-grade	
Year	Defaults	Default Rate (%)	Debt Outstanding ($bil.)	Defaults	Default Rate (%)	Defaults	Default Rate (%)
1981	2	0.14	0.06	0	0.00	2	0.62
1982	18	1.19	0.90	2	0.18	15	4.41
1983	12	0.76	0.37	1	0.09	10	2.94
1984	14	0.91	0.36	2	0.17	12	3.27
1985	19	1.11	0.31	0	0.00	18	4.33
1986	34	1.72	0.46	2	0.15	30	5.70
1987	19	0.94	1.60	0	0.00	19	2.81
1988	32	1.38	3.30	0	0.00	29	3.86
1989	44	1.78	7.28	3	0.22	35	4.68
1990	70	2.73	21.15	2	0.14	56	8.12
1991	93	3.25	23.65	2	0.14	65	11.05
1992	39	1.49	5.40	0	0.00	32	6.10
1993	26	0.60	2.38	0	0.00	14	2.50
1994	21	0.63	2.30	1	0.05	15	2.11
1995	35	1.05	8.97	1	0.05	29	3.53
1996	20	0.51	2.65	0	0.00	16	1.81
1997	23	0.63	4.93	2	0.08	20	2.01
1998	56	1.28	11.27	4	0.14	48	3.67
1999	109	2.14	39.38	5	0.17	92	5.56
2000	136	2.48	43.28	7	0.24	109	6.23
2001	229	3.79	118.79	7	0.23	173	9.87
2002	226	3.60	190.92	13	0.42	159	9.49
2003	119	1.93	62.89	3	0.10	89	5.07
2004	56	0.78	20.66	1	0.03	38	2.02
2005	40	0.60	42.00	1	0.03	31	1.51
2006	30	0.48	7.13	0	0.00	26	1.19
2007	24	0.37	8.15	0	0.00	21	0.91
2008	127	1.80	429.63	14	0.42	89	3.70
2009	268	4.19	627.70	11	0.33	224	9.94
2010	83	1.21	97.48	0	0.00	64	3.02
2011	53	0.80	84.30	1	0.03	44	1.84
2012	83	1.14	86.70	0	0.00	66	2.59
2013	81	1.06	97.29	0	0.00	64	2.31
2014	60	0.69	91.55	0	0.00	45	1.44
2015	113	1.36	110.31	0	0.00	94	2.77
2016	163	2.08	239.79	1	0.03	143	4.23
2017	95	1.20	104.57	0	0.00	83	2.46
2018	82	1.03	131.65	0	0.00	72	2.09
2019	118	1.30	183.21	2	0.06	92	2.54

Source: Standard & Poor's, Default, Transition, and Recovery: 2019 Annual Global Corporate Default and Rating Transition Study.

To emphasize the importance of the first two elements, interviewers often spent a significant amount of time asking candidates about the data and how they handled certain aspects in their data wrangling and feature engineering. One friend who had recently interviewed with RATERx said that his data had contained errors and inconsistencies with the data dictionary that interviewers had purposely inserted to see if candidates were performing the appropriate due diligence before analyzing the data.

Beyond the analytics, the presentation had to emphasize the value of the project to RATERx. One of the biggest failures of data science investment in practice is that it often fails to produce value-accretive, actionable information. Rather, many engagements are often just exercises. It was critical that applicants recognize and be able to communicate precisely how what they are doing would help the company by serving its clients.

Appendix

The data are contained in an Apache Parquet file: raterx-data.pq. All financial data is measured in nominal millions of dollars.

Table 5: Data Dictionary

Position	Variable	Data Type	Description
			Company information
0	gvkey	int64	Unique company identifier
1	datadate	datetime64[ns]	Fiscal year end and date of financial reporting
2	conm	object	Company name
3	sich	float64	4-digit SIC code
4	credit_rating	category	Long term issuer credit rating.
			Balance Sheet Data
5	che	float64	Cash and short-term marketable securities
6	act	float64	Total current assets
7	ppent	float64	Plant, property, and equipment net of accumulated depreciation.
8	at	float64	Total assets
9	dlc	float64	Short-term debt and debt due within one year
10	lct	float64	Total current liablities
11	dltt	float64	Long-term debt
12	lt	float64	Total liabilities
13	txdb	float64	Deferred taxes
14	re	float64	Retained earnings
15	seq	float64	Total stockholders equity
16	mib	float64	Minority interest
			Income Statement Data
17	sale	float64	Net sales
18	cogs	float64	Cost of goods sold
19	xsga	float64	Selling, general, and administrative expenses
20	xrent	float64	Rental expense
21	oibdp	float64	Operating income before depreciation
22	oiadp	float64	Operating income after depreciation
23	xint	float64	Interest expense
24	intc	float64	Interest capitalized
25	ib	float64	Net income before extraordinary items
26	dvp	float64	Preferred dividends

Wharton
UNIVERSITY *of* PENNSYLVANIA
FINANCIAL ANALYTICS

Wharton Financial Analytics
The Wharton School, University of Pennsylvania
3620 Locust Walk, Suite 2461
Philadelphia, PA 19104

To connect with Wharton Financial Analytics, contact
mrrobert@wharton.upenn.edu or visit **http://finance.wharton.upenn.edu/~mrrobert/**