Lithuanian Actuarial Society meet-up Early warnings and transaction data insights 2021-04-21

Agenda

- Intro of the team
- Agile ways of working and terminology
- Big data and tools for it
- Early warnings model
- Usage cases of transaction data
- Q&A



Intro to the team

Risk Data Scientists chapter

Early Warning and Transaction Data Insights squad



Today in the call

Danske Bank

Danske Bank

22,376 employees in 13 countries



Market shares as of end-2020

What are the key structures of the BWoW?



• Collection of Squads that work to address a similar business/ customer purpose ô

ô

ê

ô

ê

ê

ê

ô



- Empowered
- Persistent
- Cross-functional
- Aligned with business/ customer purpose



SEGMENT

- Formulate priorities at theme level
- Provide segment customer insight to tribes for prioritizations
- Articulate local segment needs

CHAPTER

- Deep expertise embedded fully into the Feature and Component Squads
- Carry the standards of expertise mastery

CENTER OF EXCELLENCE

• Scarce resources with deep expertise not embedded permanently in the Feature and Component Squads

The agile way of working



Why? (1/2)

Traditional data sources are backwards looking

[Company Name]	© Corporate Finance Institute®. All rights reserved						
Balance Sheet							
USD \$ millions]							
	2014	2015	2016	2017	201		
Assets							
Current assets:							
Cash	167,971	181,210	183,715	211,069	239,550		
Accounts Receivable	5,100	5,904	6,567	7,117	7,539		
Prepaid expenses	4,806	5,513	5,170	5,998	5,682		
Inventory	7,805	9,601	9,825	10,531	11,342		
Total current assets	185,682	202,228	205,277	234,715	264,112		
Property & Equipment	45,500	42,350	40,145	38,602	37,521		
Goodwill	3,580	3,460	3,910	3,870	3,850		
Total Assets	234,762	248,038	249,332	277,187	305,483		
Accounts Payable	3,902	4,800	4,912	5,265	5,671		
Current liabilities:							
Accrued expenses	1 220	1 5 4 1	1,662	1 965	1 900		
Unearned revenue	1 540	1,541	1,853	1,005	1 724		
Total current liabilities	6,762	7,901	8,427	9,082	9,294		
Long-term debt	50.000	50.000	30,000	30,000	30,000		
Other long-term liabilities	5,526	5,872	5,565	6,051	5,909		
Total Liabilities	62,288	63,773	43,992	45,133	45,203		
Shareholder's Equity							
Equity Capital	170,000	170,000	170,000	170,000	170,000		
Retained Earnings	2,474	14,265	35,340	62,053	90,280		
Shareholder's Equity	172,474	184,265	205,340	232,053	260,280		
Total Liabilities & Shareholder's Equity	234,762	248,038	249,332	277,187	305,483		
Check	0.000	0.000	0.000	0.000	0.000		

Age		Age		Age	
x	lx	x	l _x	x	l _x
0	100000	37	96921	74	66882
1	99305	38	96767	75	64561
2	99255	39	96600	76	62091
3	99222	40	96419	77	59476
4	99197	41	96223	78	56721
5	99176	42	96010	79	53833
6	99158	43	95782	80	50819
7	99140	44	95535	81	47694
8	99124	45	95268	82	44475
9	99110	46	94981	83	41181
10	99097	47	94670	84	37837
11	99085	48	94335	85	34471
12	99073	49	93975	86	31114
13	99057	50	93591	87	27799
14	99033	51	93180	88	24564
15	98998	52	92741	89	21443
16	98950	53	92270	90	18472
17	98891	54	91762	91	15685
18	98822	55	91211	92	13111
19	98745	56	90607	93	10773
20	98664	57	89947	94	8690
21	98577	58	89225	95	6871
22	98485	59	88441	96	5315
23	98390	60	87595	97	4016
24	98295	61	86681	98	2959
25	98202	62	85691	99	2122
26	98111	63	84620	100	1477
27	98022	64	83465	101	997
28	97934	65	82224	102	650
29	97844	66	80916	103	410
30	97750	67	79530	104	248
31	97652	68	78054	105	144
32	97549	69	76478	106	81
33	97441	70	74794	107	43
34	97324	71	73001	108	22
35	97199	72	71092	109	11
26	97065	72	22002	110	

	Develop	ment ye	ear									
U/W year	1	2	3	4	5	6	7	8	9	10	11	12
2001	29	88	155	228	280	309	333	355	373	371	376	372
2002	30	67	112	164	203	239	267	274	298	304	310	
2003	29	74	128	198	274	311	374	421	424	428		
2004	41	111	231	387	477	580	613	591	642			
2005	30	92	137	205	291	340	374	393				
2006	11	37	67	102	173	233	250					
2007	10	40	90	129	195	250						
2008	24	49	85	121	181							
2009	28	71	110	176								
2010	18	34	54									
2011	25	212										
2012	42											
link ratios:	3.18	1.77	1.53	1.35	1.20	1.10	1.04	1.06	1.01	1.02	0.99	

Balance sheet

Mortality tables

Claims triangles

Why? (2/2)

Forward looking data used for portfolio approach and not single names



Macrof	orece	et Sca	ndinavia										
IVIACI U I	Year	GDP ¹	Private cons. ¹	Public cons.1	Fixed inv. ¹	Ex- ports ¹	lm- ports ¹	Infla- tion ¹	Wage growth ¹	Unem- ploym ³	Public budget ⁴	Public debt ⁴	Curre acc.
Denmark	2020 2021 2022	-3.7 2.3 3.4	-3.1 3.0 3.6	-1.2 1.6 0.4	0.8 3.1 3.8	-8.8 3.6 7.2	-6.6 4.7 6.4	0.4 0.8 1.2	2.5 2.0 1.8	4.7 4.7 3.8	-3.5 -1.9 -1.2	43.4 41.9 41.5	7.4 6.9 7.7
Sweden	2020 2021 2022	-3.2 3.3 2.6	-4.7 3.8 3.0	-0.2 2.4 1.5	-1.4 4.1 2.7	-5.5 6.9 3.9	-6.8 7.6 3.9	0.6 1.2 0.8	2.0 1.9 1.9	8.3 8.3 7.3	-3.0 -2.8 -1.0	39.0 40.0 39.0	5.3 5.1 5.1
Norway	2020 2021 2022	-3.3 4.0 3.3	-7.5 8.0 5.5	1.7 2.0 2.0	-5.6 1.0 2.0	-3.0 5.8 4.5	-12.0 5.0 6.0	1.3 2.5 2.1	2.0 2.3 2.9	5.0 3.3 2.6	-	-	-
Macro f	oreca	st. Euro	oland										
	Уеаг	GDP ¹	Private cons. ¹	Public cons.1	Fixed inv. ¹	Ex- ports ¹	lm- ports ¹	Infla- tion ¹	Wage growth ¹	Unem- ploym ³	Public budget ⁴	Public debt ⁴	Currer acc.4
Euro area	2020 2021 2022	-7.0 4.9 3.4	-8.8 4.5 5.3	-0.8 3.2 1.8	-13.7 -1.2 2.1	-9.1 13.1 3.5	-10.5 10.6 4.3	0.3 1.1 1.1	-1.0 1.0 1.0	8.0 8.9 8.6	-8.8 -6.2 -4.4	101.7 102.1 102.3	2.6 2.6 2.8
Germany	2020 2021 2022	-5.3 4.0 2.2	-5.9 3.2 3.3	4.5 4.1 1.4	-3.7 2.8 2.1	-10.2 13.8 3.3	-8.8 10.9 4.3	0.5 1.2 1.3	-0.1 1.5 0.0	4.2 4.2 3.7	-6.0 -4.0 -2.5	71.2 70.1 69.0	6.0 6.3 6.1
Finland	2020 2021 2022	-3.3 2.2 2.8	-4.0 3.5 2.7	0.0 1.5 1.0	-2.5 2.0 3.0	-10.0 5.0 6.0	-8.0 5.0 5.0	0.3 1.0 1.5	1.9 2.5 2.5	7.8 8.0 7.2	-8.0 -4.1 -2.4	68.6 71.0 70.8	-0.6 -0.6 -0.6
Macro f	oreca	st. Glol	bal										
	Уеаг	GDP ¹	Private cons. ¹	Public cons. ¹	Fixed inv. ¹	Ex- ports ¹	lm- ports ¹	Infla- tion ¹	Wage growth ¹	Unem- ploym ³	Public budget ⁴	Public debt ⁴	Curre acc.4
USA	2020 2021 2022	-3.4 3.3 3.8	-3.8 4.6 4.3	0.9 1.4 2.1	-2.4 4.6 4.3	-13.7 1.7 2.0	-10.7 6.6 3.8	1.2 1.6 1.6	4.6 2.0 2.3	8.1 6.5 5.3	-16.0 -12.2 -7.2	126.4 132.3 131.9	-2.1 -2.1 -2.1
China	2020 2021 2022	1.7 9.2 5.5	1.5 9.0 7.0	-	3.0 10.0 4.0	-	-	3.0 2.0 2.5	7.5 7.0 7.0	-	-11.9 -11.8 10.9	-	0.6 0.4 0.4
ик	2020 2021 2022	-11.1 5.2 6.9	-14.4 4.2 8.3	-9.9 4.0 2.1	-12.8 6.2 11.9	-12.5 6.4 5.9	-21.1 6.1 8.2	0.9 1.6 1.9	2.0 1.2 1.6	4.4 4.9 4.6	-19.0 -7.7 -4.5	109.1 109.3 108.3	-2.0 -3.8 -3.6
Japan	2020 2021 2022	-5.6 2.7 2.5	-7.0 2.4 2.7	1.9 1.8	-5.7 -1.0	-12.8 9.1	-7.5 1.5	-0.2 0.3	-	2.8 2.7 2.5	-	-	-

Not as Big Data

Data Flow - Overall



韓 hadoop SQL héloop AA AA Classifier models TE table Raw Data DK/NO/SE/FI (labelling (payment systems) SQL Aggregations (transaction enrichment) models) These are smaller so they can be viewed in tableau **Transaction** level (billions of transactions) Same as TE but **country and inflow/outflow** specific These have **labels** (over 15 per country) Since **2016** Daily productionized table Labels can be social benefits, salary, pension, Only DB transactions gambling etc Inflow classifiers for DK/FI/NO/SE retail customers C.a 40 payment types AA also developed an outflow classifier, but we can only use the data after model owners have been Caveats: appointed No labels Lacklustre counterparty coverage (43% DK, 19% FI, • 31% NO, 9% SE)

Caveats:

- Only coverage for **Danske Bank** clients
- Changes to enrich data **applied back in time**
- Correction transactions applied

- The labels come from a machine learning model, so there might be **inaccuracies** (eg what is included)
- Same caveats from TE table apply

In essence, what can we do?

We want to combine transaction (and external data) with all the other slices of the pie and reveal potential warning signs





Building models which fit into designed framework...

Danske Bank



Model Infrastructure

From development to deployment



Early Warning

Early Warning Model



Target and features

Downgrade of PD score was chosen as target variable for capturing deteriorating customers

Training data	
12 months	6 months

To avoid seasonal bias training dataset is constructed from different time periods



Selected model - XGBOOST (gradient boosting decision tree) algorithm



Abbreviation	Model name	Mean ROC-AUC score
LR	Logistic regression	0.628083
KNN	K Neighbors Classifier	0.551441
DT	Decision Tree Classifier	0.578554
RF	Random Forest Classifier	0.668644
XGB	Gradient Boosting Classifier	0.800645

- Boosting is an ensemble technique where new models are added to correct the errors made by existing models

- Difference between Gradient boosting and random forest: GBT build trees one at a time, where each new tree helps to correct errors made by previously trained tree. RFs train each tree independently, using a random sample of the data.



Algorithm Comparison

Most important model features

		ID	Predictor	Estimation period (average)		
Customer information data		1	Number of transactions with bad customers	12 months		
		2	Days in the bank	Last month		
		3	Average exposure	Last 12 months		
Credit, risk data		4	Volume of transactions with bad customers	Last 12 months		
		5	Excess amount	Last 12 months		
Behavioural data		6	Exposure amount	Last month		
		7	Volatility of monthly number of outgoing transactions	Last 12 months		
		8	Utilization share	Last 12 months		
	, ,	9	Average leasing exposure	Last 12 months		
Profit data				10	Volatility of monthly amount of outgoing transactions	Last 12 months
		11	Average return on allocated capital	Last 12 months		
		12	Delinquency (0 – 5 days)	Last 12 months		
		13	Excess amount	Last month		
Transaction data		14	Volatility of excess amount	Last 12 months		
		15	Number of accounts	Last 12 months		

Features contribution to the prediction

- To explain the prediction produced by the model we have used LIME to show features contributions
- Features can have either positive (increase the probability) or negative (decrease the probability) to the prediction

Case 1: prediction = 0,2563

Feature name	Actual value of this features	Impact to the prediction
Delinquent amount 31-60 days	22194,27	0,10977
Average monthly amount of incoming transactions (6 months)	5143,6	0,08064
Average monthly amount of incoming transactions (1 months)	11,5	0,04427
Delinquent amount 91+ days	0	-0,02133
Average exposure utilization (6 months)	0,199	-0,01756

0,2563 (prediction) = -0,0587 (negative contribution) + 0,30624 (positive contribution) + 0,008801 (bias)



Some Analytical User Cases

Some examples of the use of transaction data in reports and analytics



Using the labels from inflow classifier models to identify salaried employees



in incoming transactions







Identifying subportfolios and companies that are UK dependent



Using machine learning to develop an early warning model based on transaction data



Looking into clients that are receiving/applying for social benefits

Monitoring spending patterns of customers on interest only Exposure by ratio of overspending



Identifying whether weak segments have a propensity for overconsumption or saving

Exports

Cross Border Incoming Transactions

Quarterly Moving Average: 2018-present

Reporting and monitoring

Monitoring of cross boarder transaction as a proxy for exports has continuously been included in many different internal setups.

For now monitoring is mostly stand-alone one-pagers and not fully incorporated in the reports as exports are deemed to have a significant interest present due to the pandemic.

Data validation and usability



Data has been validated against national statistics and shows similar trends in most industries. The cross country transaction amounts are different from the reported export levels why this can only be used as a proxy for export development and not reflecting exact export amounts.

This also means, that the further we drill down into specific segments, sub industries or companies uncertainty will increase risk of false conclusions.



Most industry cross boarder transactions are normalizing after dip in summer 2020 Quarterly Moving Average: 2018-present

Sectors which have not recovered

All Industries



Sectors with minimal impact from Covid lockdown



Sector representative of all industries: good recovery from 2020 lockdown





Key messages

- The total volume of transactions using a quarterly moving average. The overall trend was increasing from 2018 into the first few months of 2020. There was then a sharp decline that had its nadir in August 2020. Since then there has been a very robust recovery
- March sees a recovery in exports, with increases in all Nordic countries. Monthly developments are higher in all industries as a result. Both cross border transactions and non-cross border transactions have seen a pick-up, a lot of it due to Easter holidays increasing consumption
- Consumer Goods and Retailing had a smaller dip in Q3 2020, but have since experienced strong growth that has propelled them far beyond pre lockdown levels - perhaps a result of more online trading
- Transportation is very representative of the portfolio as a whole. A sharp decline in Q3 2020 , but a good recovery since
- 2 industries that still seem to be operating at lower levels are Hotels, Restaurants & Leisure, and Shipping, Oil & Gas. These industries saw a typical sharp decline in Q3 2020, but have since failed to recover to pre lockdown levels. Hotels , Restaurants & Leisure have seasonal variations with peak business being in the summer months, and this trend still prevailed in 2020, but at a lower base level.

Tax

*Note: Figures shown below are insights from Danske Bank transaction data on portfolio level. We use transactions with skat or moms in the text field as a proxy. The year on year volume and change is shown in figures below.

Tax Analysis

Value of Identified Tax Outflows, Volume, DKK (percentages represent YoY change in volume)



CVRs identified as having had tax outflows, Nr, DKK (percentages represent YoY change in volume)



Overall tax outflow development <u>Key Messages</u>

Country and Portfolio Level:

Figures shown below are insights from Danske Bank transaction data on portfolio level. We use transactions with skat or moms in the text field as a proxy.

Overall, in the portfolio, we see that last year had less transaction volume pertaining to transactions we identified as tax, and less companies have them. The decline is moderate and it has a lot to do with activity slowdown (less revenue resulting in lower tax paid).

The postponements have a lot to do with permissive regulation and we have observed that even traditionally well performing companies have taken advantage of the tax postponement to garner extra liquidity.

Volume-wise, the development in terms of lowered tax outflow is mainly visible in capital goods, commercial property, hotels, restaurants and leisure, pharma and medical devices and shipping, oil and gas.

In terms of number of companies paying tax less, apart from utilities and infrastructure, commercial property and pharma, the development is visible in every industry, primarily in automotive, retailing, transportation, consumer goods, hotels, restaurants and leisure.

Overall CVR with tax outflow development

-1,6%

2020

SME

SME Transaction Data Insights



Transaction volumes per industry

		2019			5050			Diff in Total
	CBS Inflow Bn	Total Inflow Bn	CBS%	CBS Inflow Bn	Total Inflow Bn	CBS%	IIIIOV 909	1111000 909
Education	11.0	429.3	2.6%	13.2	435.7	3.0%	17.9%	1.5%
Professional services	37.0	164.9	22.4%	54.6	188.0	29.0%	29.4%	14.1%
Consumer staples	8.8	126.2	6.9%	8.8	132.5	6.7%	-4.1%	5.0%
Machinery	30.3	111.6	27.2%	28.3	105.8	26.7%	-1.7%	-5.1%
IT-services	12.0	68.0	17.7%	13.1	81.0	16.1%	-8.6%	19.1%
Commercial services	0.9	61.5	1.5%	0.9	64.3	1.5%	-2.1%	4.6%
Non-profit associations	3.9	63.1	6.2%	4.3	51.0	8.5%	36.4%	-19.1%
Consumer discretionary	12.2	53.5	22.7%	10.9	45.1	24.3%	6.7%	-15.7%
Trade Consumer discretionary	5.9	42.0	14.0%	6.9	51.0	13.6%	-2.6%	21.5%
Passenger transport	1.4	46.9	3.0%	0.5	44.9	1.0%	-65.6%	-4.4%
Media	2.4	25.4	9.3%	3.3	28.4	11.5%	23.6%	11.5%
Leisure	3.1	22.9	13.4%	1.8	16.3	11.2%	-16.4%	-28.9%
Health and social care	0.1	17.6	0.8%	0.1	17.7	0.7%	-12.5%	0.7%
Restaurants	0.1	14.1	0.7%	0.2	11.8	2.0%	197.2%	-16.4%
Hotels	1.5	8.5	18.1%	0.2	5.0	3.9%	-78.3%	-41.6%

• Note: The analysis covers SME clients in Denmark. Outliers are not excluded.

1] We take DK SMEs with exposure below DKK100m as of Jan 2021 that fit the SME definition and are not part of Other Commercials, Financials, Personal Customers or Public Institutions.

• 2] For the tax analysis, we take into account clients that have had a payment between April 1st 2019 and end of March 2020, but they have not exhibited that behaviour in the following year.

- 3) Transactions with bad customers are transactions where the counterparty is a Danske Bank client rated B. Netflow represents inflow minus outflow.
- 4) Export dependent companies are companies that on a monthly basis have had at least 30% of their inflow coming from cross border activities.

Insights from transaction data

- Industries receiving the biggest payouts (top 15) from Erhvervstyrelsen include Hotels, Restaurants and Leisure, Services, Retailing, Social Services, Consumer Goods, Transportation. If we are to break it down further: the top 5 sub-industries are Restaurants, Leisure, Trade Consumer Discretionary, Professional Services, and Hotels.
- Looking at the aforementioned industries, we observe the most visible drop in inflow is in Hotels, Leisure, Non Profit Associations, Restaurants and Consumer Discretionary. Cross border inflow has declined the most within Hotels, Passenger Transport, Leisure, Media, Health and Social Care

Danish employment insights

Customers receiving "dagpenge" experiencing an increase in December

Key messages

- Number of customers receiving monthly salary has stabilized following the lowest point in May and increasing since in trend. Figures do not include seasonal adjustment
- Dagpenge has lowered since the highest point in July, but we have seen a sudden increase in December.
- In terms of government packages, there has been a slowdown since summer, but a slight upswing in December. Industries most favoured by these packages have been Hotels, Restaurants and Leisure, Services, and Retailing.





*Note: Government packages are included in the analysis when: they are sent by Erhvervstyrelsen through Erhvervstyrelsen's Danske Bank accounts. The analysis does not include other institutions or Erhvervstyrelsen's transfers through non-Danske Bank accounts.

*Note: Government packages is including all types of government packages; Salary compensation, fixed cost etc.

Danske Bank

Spending Monitor

08 April 2021

Spending Monitor Ash Wednesday breaks spending records

- · Danish card and MobilePay spending, up to the week ending on 4 April shows spending at around 5% above normal levels. However, there are major distortions to the data with this release. First, last week was Easter, which led to shops having closed for the holidays (Thursday, Friday, Sunday and this weeks Monday). Compared to Easter 2019 spending last week was up by 11%. However, last week also had pay day - unlike the Easter week of 2019.
- · Overall spending over Easter looks strong, considering that many restaurants, hotels, hairdressers, department stores and the likes where still forced closed. This is largely due to very strong sales in retailing from Monday to Wednesday, with Wednesday last week showing the strongest retail sales yet, when discounting Black Friday in both 2019 and 2020.
- · Spending got a further lift from the pay out of holiday back-pay, which started to roll out last week. So far, DKK30bn has been applied for (before taxes).
- · The next few weeks the reopening will move along, having been kicked off on Tuesday with the reopening of hairdressers, massage parlours etc. on 21 April we get a further reopening of restaurants for out-door service, and shopping malls. Despite restrictions, we expect this to give a further lift to spending.

Notes on the spending data

The spending data is based on transactions online and offline, both domestically and abroad, with cards and MobilePay in stores for around 1m Danske Bank Danish personal customers with active accounts. All data is anonymised and non-referable.

The spending data is used as a proxy for private consumption. However, it does not include cash spending and account transfers. Hence, notably, fixed costs and spending on housing are not included. Changes in cash spending patterns will also affect the data.

The charts show spending in 2020 and 2021 compared with the same week in 2019.

Data is not adjusted for price developments. This edition of the Spending Monitor contains minor revisions of historical data.



Spending in physical stores subsides somewhat, and online continues down Source: Donake Bank

Figure 2: Evolution of aggregate spending in Denmark and Sweden. The figure shows the evolution of aggregate spending in Denmark and Sweden from 2 January to 5 April 2020. Red lines show the evolution of spending in 2020 as a percentage of daily average spending in 2019. Grey lines show the same series for the same weekday in 2019, i.e., 364 days earlier. The dash vertical line denotes 11 March, when the Danish government announced the lockdown. Shaded red regions highlight the drop in spending in both countries at this point in time.

%Daily average spending (2019), by country:



22

Pandemic, Shutdown and Consumer Spending: Lessons from Scandinavian Policy Responses to COVID-19*

Asger Lau ANDERSEN (University of Copenhagen and CEBI) Emil Toft HANSEN (University of Copenhagen and CEBI) Niels JOHANNESEN (University of Copenhagen and CEBI) Adam SHERIDAN (University of Copenhagen and CEBI)

May 12, 2020

Abstract

This paper uses transaction data from a large bank in Scandinavia to estimate the effect of social distancing laws on consumer spending in the COVID-19 pandemic. The analysis exploits a natural experiment to disentangle the effects of the virus and the laws aiming to contain it: Denmark and Sweden were similarly exposed to the pandemic but only Denmark imposed significant restrictions on social and economic activities. We estimate that aggregate spending dropped by around 25 percent in Sweden and, as a result of the shutdown, by 4 additional percentage points in Denmark. This implies that most of the economic contraction is caused by the virus itself and occurs regardless of whether governments mandate social distancing or not. The age gradient in the estimates suggest that social distancing reinforces the virus-induced drop in spending for individuals with low health risk but attenuates it for individuals with high health risk by lowering the overall prevalence of the virus in the society.

+45 45 12 85 31 louhan@danekebank.dk

