TECHINAL ASSESSMENT - DATA SCIENCTIST POSITION 006191

Hannes Moritz

Content

Tables

Figures

Recommendations

The following analysis was performed given the provided data.

Thanks to resource constraints, only a limited number of companies can be analysed in depth. Therefore, in the following report, certain closer analysis suggestions were developed based on each metric, while a general conclusion with one suggestion is done under the subsection of *Recommendation Summary*.

Some companies seemingly reported inaccurate numbers, visually identified by significant outliers, for which reason the data is cleaned. Those obvious errors were corrected without closer contact with the subject matter experts. Details with regards to the cleaning can be seen at the end of this section under *Performed data cleaning*. Additionally, potential linear relationships are briefly shown and describes in the appendix.

Recommendation based on GWP

As the GWP is a key indicator for the size and therefore the potential importance of a company, it should also be a key indicator for potential further analysis. Therefore, the overall percentage of share based on GWP is calculated. Based on this calculation, the cumulative portion is calculated.

Based on the GWP, the following three companies comprise around 30% of the overall GWPs collected and are therefore recommended for a closer check.

Company	GWP %	Cumulative %
Firm 210	12.88	12.88
Firm 4	11.03	23.92
Firm 26	5.55	29.47

Table 1 GWP and cumulative sum in %

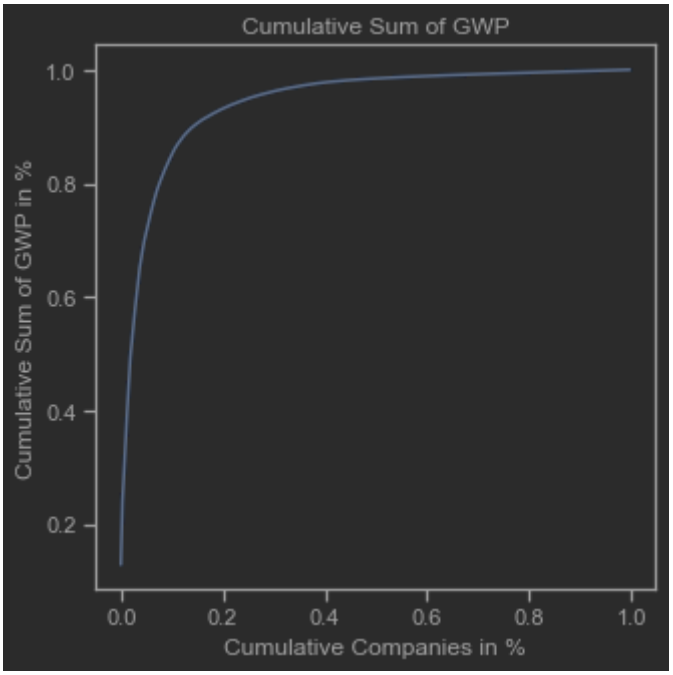

Figure 1 Cumulative sum of GWP vs. cumulative companies (both in %)

Recommendation based on NWP

As the NWP is a key indicator for the risk profile of the company, it might be used as a proxy for the riskiness of their insurance portfolio. Therefore, the overall percentage of share based on NWP is calculated. Based on this calculation, the cumulative portion is calculated.

Some companies show significantly large positive and negative values, which might be again related to data inconsistencies, which would be needed to be clarified with a subject matter expert. Assuming that in general values between 0 and 1 are valid, and the higher a value is, the less business was reinsured, a focus should lie here upon those companies, which hold most of the risk in their own books instead of passing it to reinsurance companies. Based on the stated assumptions, the following companies are suggested to be closer analysed in the context of the NWP.

Company	NWP in %
Firm 276	99.99
Firm 228	
Firm 35	
Firm 189	
Firm 250	

Table 2 Companies with high NWP

Recommendation based on SCR coverage ratio

The SCR shows whether the company can meet its legal obligations. Assuming that a number of company's reported unplausible data (e.g., SCR equals 0 or is negative), the following two companies could not meet the requirements in the last year and therefore probably need an urgent analysis about the causes and especially management's strategy regarding getting back into the legal requirements and frameworks.

Company	SCR in %
Firm 213	69.21
Firm 148	96.35

Table 3 Companies below SCR requirements

The following three companies narrowly met the SCR requirements and should therefore be further analysed or should be monitored more closely in order to ensure the successful management of risks to avoid potential undercapitalization.

Company	SCR in %
Firm 296	101.41
Firm 199	106.65
Firm 100	107.71

Table 4 Companies narrowly above SCR requirements

Recommendation based on Net Combined Ratio

The net combined ratio indicates the profitability of a company. As the profitability might be an indicator for longer term development of the SCR, monitoring significant negative deviations might be of interest for a closer analysis. The ratio for companies 228, 166 and 284 might be a due to data invalidity.

The companies with the lowest profitability (namely negative) by far (at least the factor 4) are the following 5 firms:

Company	Net Combined Ratio (negative)
Firm 228	1076.15
Firm 166	989.15
Firm 284	906.30
Firm 72	49.51
Firm 178	21.00

Table 5 Companies with a negative net combined ratio

Recommendation based on Deviations

With regards to potentially significant deviations, the change rates on a yearly basis were calculated. This is followed by the variation calculation from one year to the other. Finally, for the last year, the average over all accounts is calculated. The following three companies are

proposed for closer examination, as they comprise for around 85% of the overall variations and are therefore recommended for a closer analysis.

Company	Variation	Cumulative %
Firm 185	85727	52.05
Firm 19	49045	81.83
Firm 197	6327	85.67

Table 6 Variation and cumulative variation in %

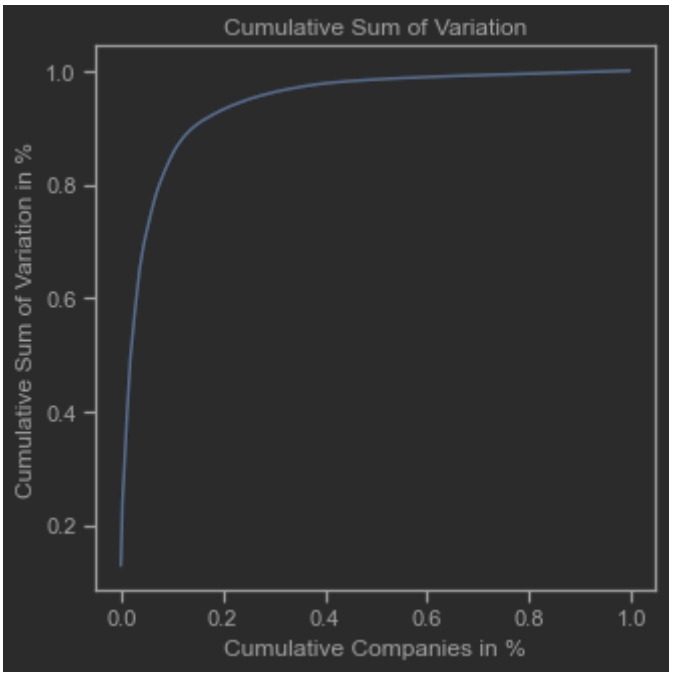

Figure 2 Cumulative sum of variation vs. companies (both in %)

Recommendation Summary

As demonstrated, depending on the categories, multiple companies might be closely checked based on the individual circumstances. Depending on the individual severity of the scores, the closer analysis for the following five companies is recommended:

Company	Reason
Firm 213	SCR < 1
Firm 148	SCR < 1
Firm 213	SCR < 1
Firm 210	GWP size
Firm 4	GWP size

Table 7 Recommendation Summary for closer analysis

If further resources are available, more companies could be added to the recommendation list.

Performed data cleaning

Data below/above the given threshold is set to the remaining mean for this analysis.

Underwritings Data Positive:

Account	Threshold
Gross expense ratio	> 0.25*1e6
Gross combined ratio	> 12000
Net expense ratio	> 0.25*1e6
Net combined ratio	> 0.25*1e6
Pure net claims ratio	> 0.25*1e6
Pure gross claims ratio	> 0.25*1e6

Table 8 Underwriting's data positive corrections

Underwritings Data Negative:

Account	Threshold
Gross expense ratio	< -2000
Gross combined ratio	< -2000
Net expense ratio	< -2000
Net combined ratio	< -2000
Pure net claims ratio	< -2000
Pure gross claims ratio	< -2000

Table 9 Underwriting's data negative corrections

Appendix

Linear Dependency Plot

As seen in the following plot along the diagonal, we can especially observe a substantial number of heavy tailed distributions, which look normal-alike, but clearly show skewness. In general, we can also observe multiple occurrences of clear linear dependencies as expected between some variables, e.g., the total assets vs. total liabilities. Sometimes, some time-enduring clusters can be observed, which therefore probably belong to the same company with changing values during time. As already seen earlier, significant clusters are generated often for comparably small values, probably due to the highly skewed distribution of the importance of different companies based on the GWP.

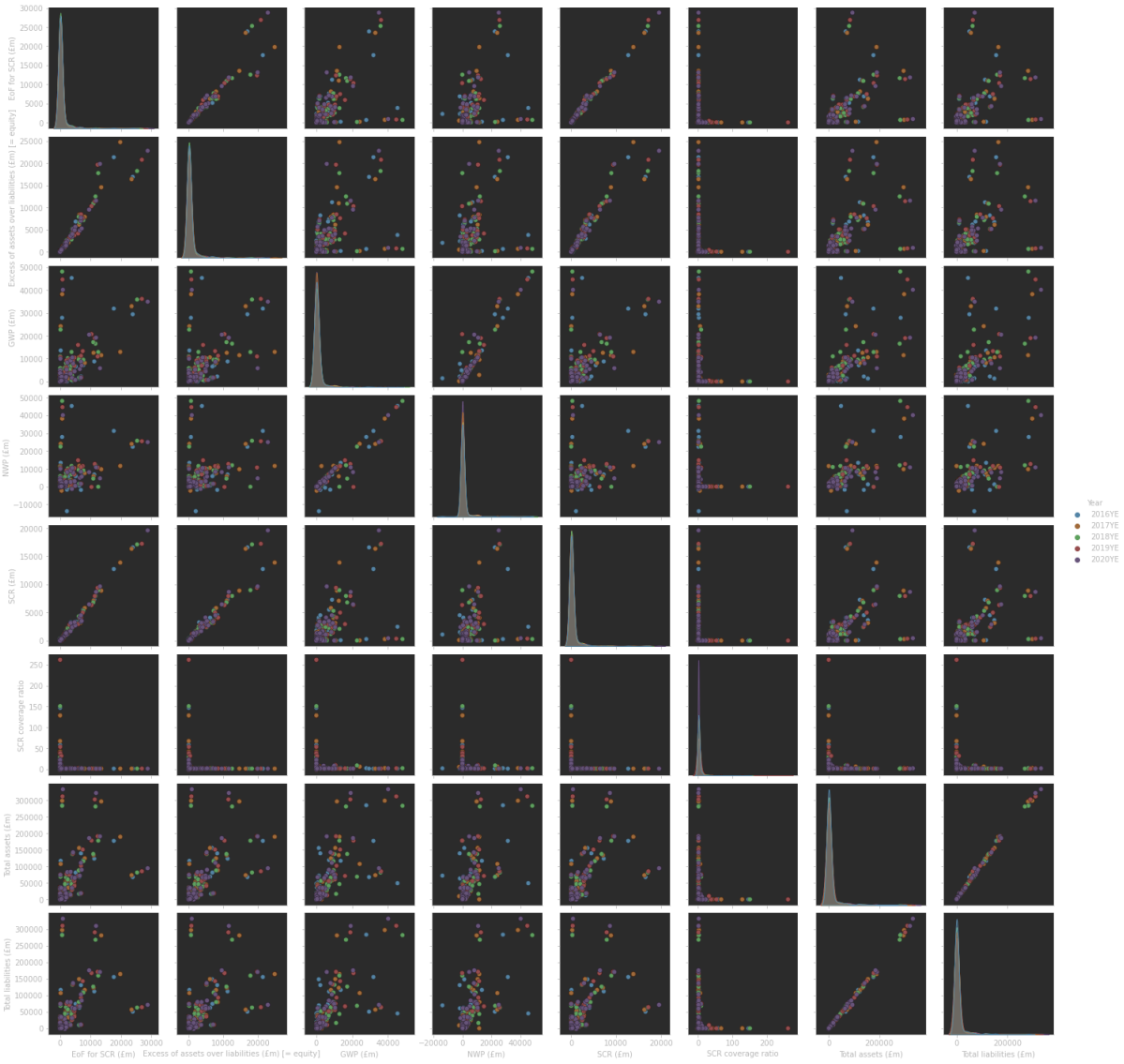

Figure 3 Dependencies of general accounts

Further Classification Proposal

Further classification beyond the shown traditional means can also be applied. In the following, the different companies are clustered into different groups, not necessarily based on the traditional booking accounts, but also based on e.g., their behaviour from year to year. This might give further insights, as these clusters might identify further insights for domain experts, especially when these cluster classifications change over time, which might be an indicator for a change in the business model of the corresponding company.

Following this analysis, the class distribution for 2020 is the following:

Automatically generated Class #	Number of Companies in this class
-1	65
0	254
2	1
3	1
5	1
6	1
7	1

Table 10 Classification classes distribution

As observed, the number of companies is not uniformly distributed at all. The following companies are in those classes, which only consist of one company for the year 2020.

Company	Class
Firm 134	2
Firm 297	6
Firm 300	5
Firm 76	7
Firm 80	3

Table 11 Single company classes 2020

Following this analysis, a further drilldown into the already independently available information beyond this analysis might be of interest before starting a full investigation, especially in combination together with subject matter experts.

Interestingly, by analysing the change rate from one year to another, one could also identify potential changes in the business model, leading to another classification. Nevertheless, the change rate itself is not useful in the way of interpreting, so in the following, only those changes are analysed, which are unequal to 0, meaning, the class has changed. When analysing these, most movements happen between two automatically generated classes, while the rest is again distributed among multiple classes.

The minority class change happened for the following companies:

Company	2019 Class	2020 Class
Firm 107	1	-1
Firm 278	5	-1
Firm 319	4	0
Firm 323	4	0

Table 12 Minority class changes 2019 to 2020

Again, this might be a reason to get into details of the potential business change if no further information is not already available.

Therefore, as shown in this section, for future use, it might be also of interest to classify companies based on automatically generated class labels. Nevertheless, detailed analysis of these classes will be required to deeply understand the class segmentations together with subject matter experts.