Business Intelligence and Databases 2023

‘Insuring the Unpredictable: A Data-Driven Approach to Natural Disaster Risk Management with FEMA Data’

Version 2.0
Date 26-04-2023
Status Final

Team 25

Names [redacted] Gracjan Chmielnicki
Arsen Ordokov
Tom Hansult
Felix van Delden

Table of contents

1 Introduction

1.1 Overview

Insurance companies need to assess risks associated with natural disasters in order to calculate insurance premiums and determine whether to offer insurance in certain regions. However, this requires proper insight into natural disasters in terms of frequency, severity, impact, and geographic location. Moreover, Insurance companies need to consider the financial risk for offering insurance in areas highly affected by natural disasters. This decision-making process needs to be highly data driven. Over the course of this project, the team will use the FEMA data on natural disasters to derive relevant KPI’s and develop a dashboard to provide insurance companies an overview of relevant natural disaster data for their decision-making process.

1.2 Case description

The underlying case of the project is the insurance case. This case provides data on disasters and catastrophes in the United States and is published and maintained by the Federal Emergency Management Agency (FEMA) of the United States. Therefore, it also provides information about money that has been released by FEMA and the government of the US in response to emergencies. Insurance companies need to assess risks for natural disasters to calculate premiums and decide on whether insurance can be offered in a certain region. The general aim of this project is to transform and use the data provided by FEMA in a way that can help insurance companies in this decision-making process and derive further insights.

1.3 Main goal

The key objective of this project is to provide helpful insights to insurance companies for their premium calculation, risk assessment and product innovation in the field of natural disasters. These data insights will be structured in KPI’s and visualized in a dashboard on various aspects of natural disasters in the United States that are relevant to insurance companies.

1.4 Approach

To develop insightful dashboard, the group first considered the various tables provided in the FEMA dataset and over the course of thorough group discussions normalized this data and designed the database that will be used for deriving and visualizing KPI’s. Next, the team designed the balance scorecard and defined relevant KPI’s for the insurance companies. Using these KPI’s the team designed the dashboard, finding the best visualization for each KPI and implementing drill-down and filtering functionality. This was done while keeping in mind the general value proposition to the insurance companies. Lastly, the. group used RapidMiner to try to identify patterns in the data at hand that might be of interest to insurance companies.

2 Database

2.1 Overview

The Insurance case provided multiple datasets, five to be exact. To normalize the data at hand the group regarded each dataset or table in an isolated manner. Likewise, they are going to be discussed in different paragraphs of the following outline of the normalization process.

2.2 Normalization process

The general procedure that was followed for the normalization process, was firstly asking the question, which part of the data is truly insightful and helpful to the general idea that this project is trying to convey in this project, every attribute or table, that the group found to be useless to the general aim was thus eliminated. However, as it was not yet clear exactly what data is used, this procedure was subject to cautiousness in terms of eliminating datapoints. Following which, the tables were adapted to the 1^{st}, 2^{nd}, 3^{rd} NF.

2.3 Design of ER-Diagrams

The Chen notation was followed. As best practice suggests, Entities were modeled in yellow, attributes in blue, relationships in green, primary keys were underlined. Furthermore, foreign keys were colored in purple, to enhance the expressiveness of the diagram. In most cases the relationship is associated with an entity that maps the primary key of the respective table against one another as foreign keys, this helped relating tables and thus the normalization process.

2.4 Database design

The Federal Emergency Management Agency presents a unique case as it offers several databases, which differ from other cases. In this regard, the insurance dataset does not provide internal data from an insurance company, but rather presents information from the FEMA, a US government agency. Therefore, a comprehensive discussion was necessary to determine the project's objectives and what aspects to depict on the dashboard. The primary focus was on the frequency and cost of various disasters and programs. To achieve this goal, the team identified two tables of interest and eliminated the remaining ones. The normalization process began.

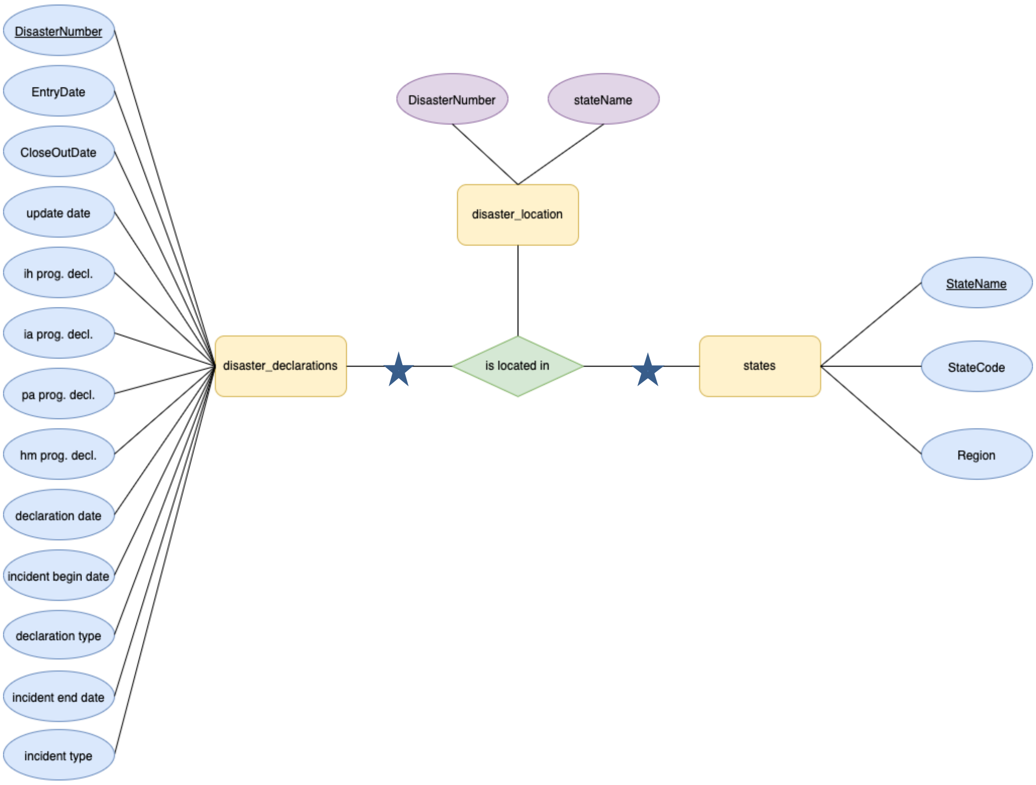

Figure 1: disasterdeclarations ER-Diagram

Firstly, the team focused on 'FEMAWebDisasterDeclaration' table. Redundant attributes such as 'hash', 'id', or 'lastRefresh' were eliminated. The geographical dimension was defined as states, rendering the 'fipsCountyCode' and 'placeCode' unnecessary. 'stateName' and 'stateCode', which were dependent variables, were extracted, and a new table 'State' was created. An intermediate table was necessary to connect the 'states' and 'disaster_declarations' tables, and primary and foreign keys were established. 'DisasterNumber' became the primary key in the 'disaster_declarations' table, while 'stateName' was assigned the same role in the 'states' table. The 'disaster_location' table had two foreign keys to connect to the other tables.

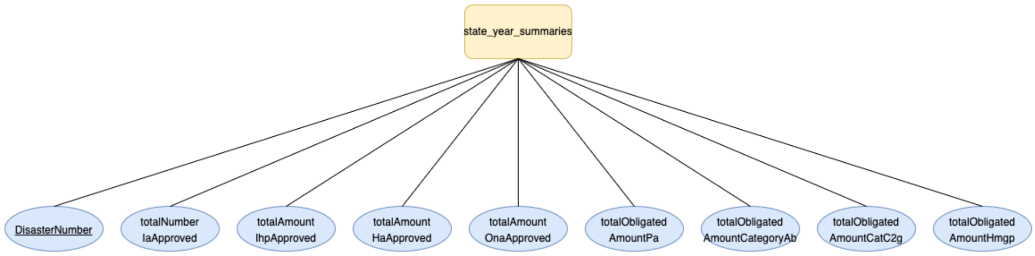

Figure 2: state_year_summaries ER-diagram

Next, the team examined the 'FEMAWebDisasterSummaries' table. Normalizing this table did not present any significant challenges since it lacked dependent variables. It mainly involved removing redundant attributes such as 'hash', 'id', 'lastrefresh', 'paLoadDate', and 'iaLoadDate'. 'DisasterNumber' was chosen to be a primary key as it met all required conditions.

2.5 Remarks

2.5.1 Remark 1

The team encountered a decision-making dilemma concerning the connection between the 'disaster_declarations' and 'state_year_summaries' tables. Despite both tables sharing the same primary key, the team chose not to connect them. While this option was available, the team opted for separate normalization to enhance the tables' simplicity and flexibility. This decision permits future joins of these tables as needed. Although the team recognizes that connecting the tables would have been a viable alternative, careful discussion and analysis led to the determination that separate normalization would yield the best outcome.

2.5.2 Remark 2

The table 'year_state_summary' was created to aggregate the total amounts for each program declared in each year in each state to support the map visualization and drill-down functionality of the dashboard. It is not part of the formal database design, but can rather be seen as an intermediary step between the formal database design and the final data visualization. The fact that this table is not part of the formal database design means that it is not optimized for efficiency or normalized according to database design principles, to be precise it does not have a primary key. Instead, the focus is on aggregating and summarizing the data in a way that supports the specific visualization requirements.

3 Key Performance Indicators

3.1 Balanced Score Card

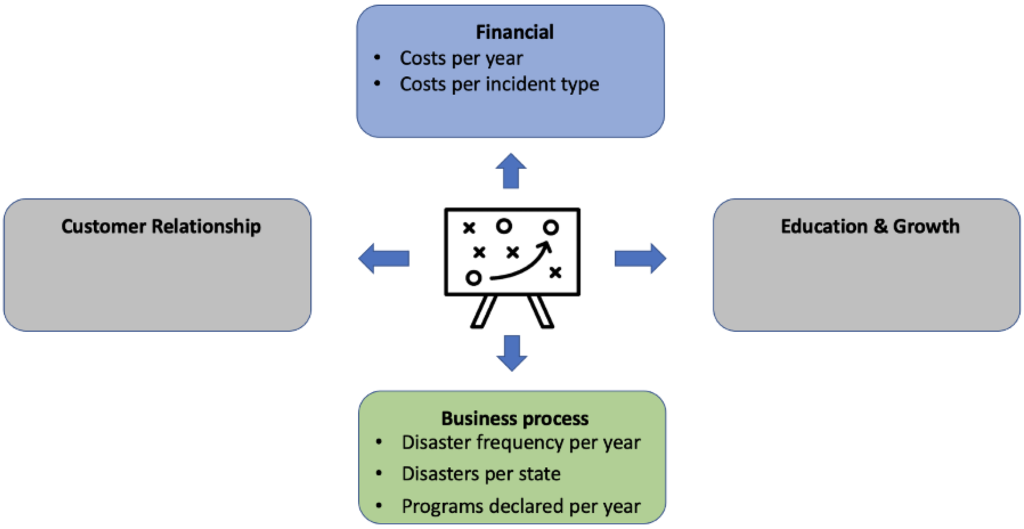

Figure 3: Balanced Score Card

Normally the balanced scorecard aligns the strategic goals of the company to Key Performance Indicators using internally available data on the general state of the company (sales, profitability etc.). The insurance case however does not provide internal data of an insurance company, but rather presents data from the Federal Emergency Management Agency, a US government agency. Because of this, the KPI's and their alignment to the goals of insurance companies require a slightly different approach and definition. The Balanced Scorecard will thus be divided into two, instead of the traditional four dimensions (Financial, Customer Relationship, Internal Process and Education & Growth) because it is not feasible to derive KPI's in the dimensions of Customer Relationship and Education & Growth with the available data. Furthermore, the financial perspective considers the financial impact of disasters rather than express the financial wellbeing of the company, as there is no data available on that.

3.2 Financial perspective

The KPI's defined in the financial dimension are the costs per year and the costs per incident type. Costs per year helps identifying a general trend in the damages of disasters. For insurance companies increased damages done by disasters mean an increased number of claims and total costs, which is why it is crucial for them to monitor trends like these. Please note that damage or "expensiveness" is expressed in terms of the amount of money released by FEMA for reparations. Likewise, the costs per incident type helps understand the most devastating types of disasters and can therefore support in insurance product composition and premium calculation.

3.3 Business process perspective

The KPI's defined in the business process perspective are the Disaster frequency per year and disasters per state. The disaster frequency per year helps understanding what the most common types of disasters are (also on state-level) and if there is types of disasters that are getting more common. This aids in innovation of insurance products in the field of natural disasters. Preparation for example can have significant impact in lowering the impact of a disaster (Onuma et al., 2017). Therefore, from understanding what types of disasters occur in which areas and which types of disasters are getting more common new insurance products could be derived, such as bounding the insurance offer to a certain level of preparation for a common disaster in the area or offering reduced premiums for well-prepared households. Likewise, disaster per state aids in identifying areas highly affected by certain natural disasters, this can support the product innovation process in a similar manner.

Moreover, the business process perspective also contains the KPI Frequency of Programs declared per year, as this KPI gives insight into a general trend of what different types of programs become more frequent and what the overall development of them is. There are different program types including but not limited to the individuals and household program or the Public Assistance program. This helps better understand the needs of the customers in various domains and enables properly addressing these needs. Furthermore, it gives insight into possible domains of new customer acquisition, as it can be derived which domains are most affected by disasters. It also shows if there is a general growth trend emerging in these domains, hinting at an increased demand for relating insurance products.

3.4 Other KPI's

As mentioned in previous sections of this report, the data at hand is provided by the Federal Emergency Management Agency (FEMA), a government agency of the United States and not an insurance company itself. This requires a non-traditional approach to KPI definition, which has already been discussed. There are a variety of other KPI's (relating to natural disasters) that could be very helpful to insurance companies, for which unfortunately the data is not available. These KPI's mainly include those that describe the overall state of the company, for example in terms of financial health. Additional examples would be the average time a claim is processed and what percentage of the claims is rejected/ granted. This is a perspective that cannot be covered with the data at hand that is provided by the FEMA but would also be very insightful to insurance companies in terms of natural disasters.

3.5 Data accuracy for KPI calculation

The dataset at hand provides information on natural disasters recorded by the Federal Emergency Management Agency of the United States (FEMA). The data provides a comprehensive overview on the geographical location and occurrences of natural disasters and the data suffices for the calculation of the according KPI's. The KPI's in the financial perspective are based on the money released by the us government or FEMA in response to different disasters. While this is an indication of the severity of a disaster and the damages it caused, it is not a precise measurement on how much money in damages is caused by a disaster in reality, as this would have to take into account further data on insurance claims etc., which unfortunately is not publicly available. To summarize, the data is sufficient for a calculation of all the KPI's defined in Business Process perspective, whereas KPI's defined in the Financial perspective are not as accurate and would require including not publicly available data to enhance its accuracy.

4 Dashboard

4.1 Overview

The dashboard was designed according to the KPI's defined in the previous section of. This report. The center of the dashboard is a map of the United States, that is color-coded according to the distinct count of disaster numbers, or in other words how many different disasters have occurred in that state in the past (Red tone increases with the number of different disasters). When hovering over a state, further details like the total amount of money granted by FEMA for all the disasters in that state are shown. Below the map, there is a filter option, allowing for filtering after the type of the disasters. The color-code and details of the map will change accordingly. Furthermore, the map acts as a central filter for the dashboard, allowing for drill-down operations. When clicking a specific state, or choosing a selection of states, all the other KPI will change accordingly and display the KPI's in terms of the state-level that is indicated in the selection of states in the map.

The other KPI's are expressed in stacked bar charts as this was found to be the best visualization, because there are various components to display and express within one diagram. Each of the diagrams has a legend indicating what color represents what value. At the same time the legend also acts as a filter and highlighter, if a more detailed view is desired.

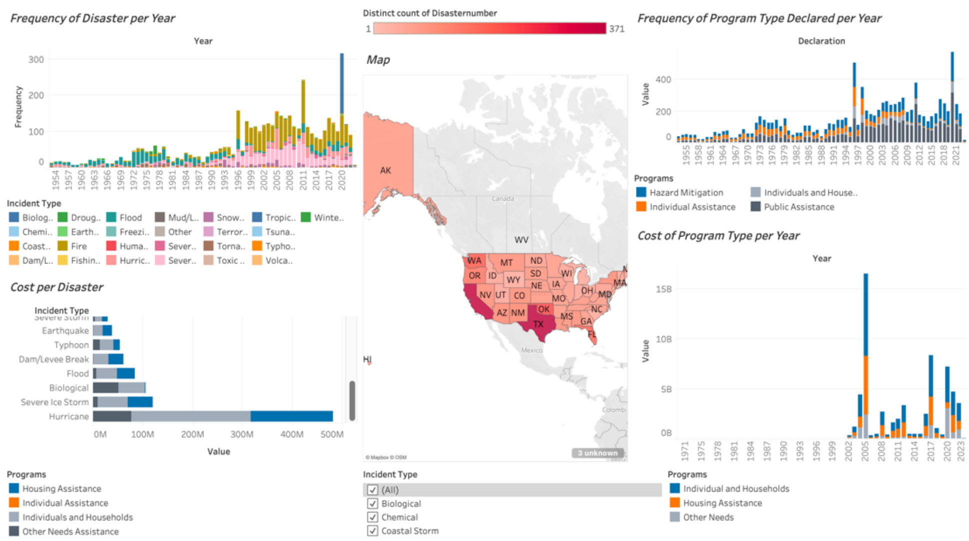

Figure 4: Dashboard

4.2 Design choices

The group had identified five crucial KPIs and intended to present them all on a dashboard to provide users with insights. Several decisions had to be made concerning the dashboard's layout and design. Firstly, the group decided that the map would be positioned at the center of the dashboard as it acts as a filter and provides the most informative insights when hovered over.

Additionally, the remaining four diagrams were related to either the frequency or cost of events, and so one pair was placed on each side of the map. This decision was made to provide a balanced look to the dashboard and allow users to easily compare and contrast the KPIs.

Furthermore, after conducting short research, the group determined that a color palette consisting of dark blue, blue, light blue, grey, and orange was the most optimal. However, to attract user attention, the group decided to color-code the map in red. This decision was made because the map was an essential component of the dashboard, and the red color would help it stand out amongst the other components.

Thirdly, the group added filters to the dashboard to provide users with an additional level of interactivity. However, this addition required some layout adjustments. As a result, the group had to modify the layout to accommodate the filters' placement and ensure that the dashboard's overall look remained balanced.

Lastly, the group decided to abbreviate the names of filters to increase readability. This decision was made to reduce clutter on the dashboard and ensure that the user could easily identify the different filters. To mitigate the potential for confusion, the group provided users with the option to view the full names of the filters by clicking on them. Additionally, users could view the diagram in full-screen mode, which would display the full-sized filters.

Overall, the group made several well-considered decisions regarding the dashboard's layout and design to ensure that it was informative, easy to use, and visually appealing to users.

4.3 Login credentials

Tableau requires login credentials to be able to connect to the database and thus view the dashboard in an interactive manner, which is why they will be provided below:

Username: user_ce

Password: uKoUgeILQXjNZswFzJX5qiVN-48PgKPZ

5 Value proposition

5.1 Overview

As briefly discussed within the KPI motivation section of this report, the insurance case requires to translate data provided by a government agency of the United States to a value proposition to insurance companies. This section will outline the thereby created value proposition to insurance companies in further detail.

The dashboard provides a holistic view on different characteristics and impacts of past natural disasters in the United States. It displays five KPI's expressing both a financial and a business process (product innovation) perspective. The interactive dashboard helps insurance companies identify a variety of things including but not limited to highly affected states for a certain disaster type, deriving where insurance for that should be more expensive, which part of the infrastructure is most affected by a disaster (decl. of program types hint at that) or what is the most "expensive" type of disaster. Furthermore, it helps identifying general trends in natural disasters like does the general number of disasters increase or does a certain type of disaster become more common. This can be considered on both a federal and state level.

5.2 Product Innovation

Furthermore, the insurance dataset provides additional data from a conducted survey asking 7.192 households how well they are prepared for a natural disaster. Only 44% said that they were prepared for a disaster and only 27% have documented & insured property.

The severity of the damages can be decreased significantly with proper preparation for a natural disaster. Eslamian et al. (2021) defines four priorities for risk reduction & resilience:

Priority 1 - Understanding disaster risk

Priority 2 - Strengthening disaster risk governance to manage disaster risk

Priority 3 - Investing in disaster risk reduction for resilience

Priority 4 - Enhancing disaster preparedness for effective response"

In this case, Priority 4 "Enhancing disaster preparedness for effective response" will be mostly considered. While the responsibility and liability of the insurance company can be argued, the insurance company has an interest in mitigating the financial impact of a disaster. Therefore, the group proposes two types of product innovations:

In states that are highly affected by a certain natural disaster (e.g Hurricanes in Florida), insurance companies could:

1. Bound the insurance offering to a certain level of preparedness.
2. Offer reduced premiums for well-prepared customers.

In order to realize this, preparedness would have to be standardized and a norm would have to be created to be able to measure if a certain customer is granted an insurance offer or reduced premiums. The benefit for the insurance company is a financial risk mitigation, if well-prepared the financial impact of a disaster can be reduced. (Onuma et al., 2017)

Because the customers of insurance companies come from various domains (individuals, businesses etc.) this would drive a generally higher level of preparedness for common disasters in a specific region. This could then be evaluated individually for each state, considering what the most common disaster is, what customer domain it impacts most and how "expensive" it is.

5.3 Outcome analysis

As described in a previous section of this report, the dashboard provides valuable insights for insurance companies into the geographical occurrence, cost and frequency of natural disasters in the United States. This section will outline some of the findings that can be of use for insurance companies.

Firstly, California and Texas are the states most affected by disasters in general, as the dark red coloring in the map, the center of the dashboard, suggests. Based on this, it may be wise to make insurance products in these highly affected states subject to a higher premium than in other states, because the likelihood of disasters occurring is significantly higher. Furthermore, Texas and California are the two most populated states in the US, which makes these states even more vulnerable to excessive damages.

The data also shows that the FEMA has released the most money in the Individual Assistance and Household programs, this shows that these areas are the most cost-intensive areas of impact of a disaster, this is particularly interesting to insurance companies as these are the categories a vast majority of their customers fall in. The trend moreover shows that the share of these programs are increasing over time. This can be seen in particular, in densely populated states like the aforementioned states California and Texas. Given this increase and large share, insurances have to be aware that they are liable for a considerable amount of damages caused by disasters, which will continue to grow over time.

The data also shows that fires are the most common type of disaster in the United States, as mentioned above insurance companies should consider binding premiums or offerings to a certain degree of preparedness or incentivize more robust and less vulnerable building structures. This can mitigate damages significantly and help reduce the overall payout of the insurance company in response to fires. Wood is a highly popular building material in the US but is more vulnerable to fires than other building material. Incentivizing the use of more resilient building material can help reduce overall damages caused by disasters and thus help reduce payouts in the long run. This is of even more significance, when considering that hurricanes are the most expensive type of disaster and more robust building structures can also help mitigate their impacts.

The state most affected by hurricanes is Florida, which is why this state would make a great candidate for pilot programs and trials for novel and innovative insurance products like this. Incentivizing more robust construction however goes beyond a financial incentive for insurance companies, as it can help spare people the suffering of losing their homes and save lives, which for the insurance companies ultimately results in more satisfied customers and a better overall image.

It is worth noting that the coronavirus pandemic disrupted the spending pattern of the FEMA, making 2020 the year with the highest total of aids paid in its history. With globalization and interconnectedness on the rise, pandemics pose a new dimension of risks and relating insurance products have been rather vague and insufficient. This provides an opportunity to offer clear and structured insurance products tailored towards biological disasters. After the devastating coronavirus, demand for these types of products is given and worth taking a look at for insurance companies.

This is just touching the surface of what can be derived with the dashboard and information at hand. The KPI's and visualizations should be interpreted and considered on a regional level, as the frequency, devastation and occurrence of different disaster types highly varies for different states. Taking this into account can help insurance companies make more data-driven decisions, reduce payout costs and innovate new products.

6 Data mining

To gain a deeper insight into the FEMA data sets, the group used rapid miner a sophisticated data mining tool. A main area of interest was using the predict functionality of rapid miner to predict types of disasters in combination with a time of year and an area. While this approach yielded some results, it turned out to be a much too complex endeavour. The FEMA data simply cannot be used to predict future disasters, rather it can be used to identify general trends and high-risk areas. There are many companies and government agencies devoted to providing disaster predictions, that use way more comprehensive data including the development of CO2 emissions, atmospheric pressure, ocean currents as well as general temperature changes around the world. Even with a combination of these data markers predicting natural disasters remains difficult and is only reliably possible in the short term.

While rapid miner offers many other functionalities, we could not find any approaches to gain deeper insight into the FEMA data sets. Rather, we used rapid miner to create charts that sum up relevant data for insurance companies to get a solid overview of most common disasters and declaration types.

4 Distinct Values:

Value	Count	Percentage
Major Disaster	2,684	56.73%
Fire Management	993	20.99%
Emergency	591	12.49%
Fire Suppression	463	9.79%

Figure 5: Declaration types

Figure 5 shows that there are four declaration types, namely Major Disaster, Fire Management, Emergency and Fire Suppression. It should be mentioned that Fire Suppression and Fire Management fall under very similar categories. The most common of these declaration types is the Major Disaster declaration with almost 57% of all declarations or 2684 declarations since 1953 (Appendix 1). This declaration type is particularly interesting for insurance companies, as the amount of anticipated assistance contributed by FEMA to support a disaster region is reduced by the "amount of insurance coverage that is in force or should have been in force as required by law and regulation at the time of disaster" (https://www.fema.gov/disaster/how-declared). This means FEMA will not cover damages that at the time of disaster were covered by an insurance company.

Value	Count	Percentage
Fire	1,550	32.85%
Severe Storm(s)	1,023	21.58%
Flood	859	18.20%
Hurricane	430	9.11%
Tornado	177	3.75%
Snowstorm	170	3.60%
Biological	167	3.54%
Severe Ice Storm	74	1.57%

Figure 6: Most common incidents

Next, the group considered the most common incident types, where fires made up 32.85% of declared incidents (1550 individual incidents), followed by severe storms with 21.58% (1023 individual incidents) and then floods with 18.20% (859 individual incidents) since 1953(Appendix 2). While fires are treated as a separate category when it comes to FEMA disaster declaration types, it can be seen that for 1550 declared incidents there were a combined 1456 Fire Management and Fire Suppression disaster declarations(Appendix 1 & 2). For FEMA to grant funds to a disaster location a fire cost threshold must be met, meaning the damages must cross a specified threshold that varies strongly between states. For example, in California this threshold was at $9,667,096 in 2022, while for states such as Alaska the threshold was at $500,000(https://www.fema.gov/assistance/public/fire-management-assistance/fire-cost-thresholds). This is the cumulative threshold, meaning that total damages must amount to it, while individuals in California can be granted $3,222,365 and only $100,000 in Alaska, once the cumulative threshold is crossed. This means any amounts larger than the individual fire cost threshold would then be covered by insurance companies.

Generally, it makes sense to connect the data of the most common disasters with their location, which has been done in the interactive Dashboard. This is useful to gain insight into where certain disasters occur most often and how insurance companies can adjust their insurance policies in certain areas to account for reoccurring natural disasters such as wildfires in California or floods in Florida.

7 Contribution

7.1 Overview

As a team, we have been working on the Business Intelligence and Databases project for the past few weeks. Through our collective efforts, we have made significant progress, and everybody has contributed to the success of the project.
The team held multiple meetings a week, these meetings always had a leader, that was responsible for creating an agenda, and making sure it was executed during the meeting. This resulted in lively discussions on the diagrams and their contents, in which everybody actively participated. This really helped to enhance the profoundness of our understanding of the diagrams and the contents of this project.
Even though the workload was sometimes overwhelming, we were able to manage it effectively by dividing tasks equally and fairly among team members. This allowed us to make steady progress on our projects and avoid overwhelming any individual team member. Overall, our team's ability to effectively divide and manage the workload was a key factor in our success.

In conclusion, our team has worked hard and made great progress on the Business Intelligence and Databases project. We have had productive meetings, divided tasks fairly, provided support and guidance to each other, and consistently sought out ways to improve the project, especially by considering external feedback. Through our collective efforts, we have successfully completed the project and are more than content with the outcome.

7.2 Personal contribution & reflection

7.2.1 Tom Hansult

Drawing upon my prior experience in data visualization and user interface design, I contributed to developing a visually engaging and user-friendly dashboard that convincingly conveyed the defined KPIs to insurance companies. Over the course of the project we were able to design a dashboard that makes use of color-coding and interactive elements to enable viewers to explore the data in greater detail.
One aspect of the project that I particularly enjoyed was working collaboratively with my group. Everyone brought their own strengths and perspectives to the table everybody was able to contribute to the projects final success.
Looking back on the project, I am proud of the contributions I made to the team, as well as the skills and knowledge that I gained throughout the process.
Working on this project allowed me to expand on my dashboard design and data visualization skills. Overall, I am thankful for the insights I gained and look forward to applying the skills I acquired to future projects.

7.2.2 *Gracjan Chmielnicki*

Drawing on my background in research and data analysis, I played a key role in providing insights that helped shape the direction of our team's project. I worked closely with my colleagues to develop hypotheses, design experiments, and analyze data, ultimately contributing to the project's success by providing meaningful and actionable insights. One aspect of the project that I found particularly enjoyable was the opportunity to collaborate with experts from a range of disciplines. Working alongside individuals with diverse skill sets and backgrounds helped me to broaden my perspective and approach problems from different angles. Looking back on the project, I am proud of the contributions that I made to the team and the ways in which my expertise helped to move the project forward. This experience allowed me to further develop my analytical and problem-solving skills and reinforced the importance of effective communication in interdisciplinary teams. Moving forward, I am excited to apply the knowledge and skills I gained from this experience to future projects and to continue learning and growing as a professional. Overall, I am grateful for the opportunity to have worked on such a challenging and rewarding project with such a talented team.

7.2.3 *Felix van Delden*

Looking back at the entirety of this project, it was a great opportunity to learn how to work with data and gain useful insights into a data set, which at first glance just seems overwhelming. The different programs used such as DBeaver, Tableau and Rapid Miner are all tools which offer a much wider scope of usefulness than just this project. This leads me to conclude that the structure and content of the project was useful even for future endeavor in the Business Intelligence and Databases Landscape. Working with a great team made the entire project even more fulfilling as it offered the opportunity to help each other in many forms and have interesting discussions about the content of the project. It also taught me how to work in a group efficiently, finding tasks that can be divided but working together when decisions had to be made that affect the entire group. Overall, I felt like the project was well organized and offered great learning opportunities in the BI&DB landscape.

7.2.4 *Arsen Ordokov*

Working on a project with a great team and the opportunity to work with datasets and build a dashboard was a valuable experience. I am grateful for the knowledge gained through this project and for the support of my team members. Collaborating with a good team makes all the difference in the success of a project. Each team member brings their unique skills and perspectives, and working together allows for better problem-solving and a more efficient workflow. I am thankful for the open communication, constructive feedback, and teamwork displayed by my colleagues. Working with datasets and building a dashboard taught me how to organize and manipulate data effectively. It also helped me to visualize complex data and communicate insights more clearly. This skillset will be useful in my future work and studies. Overall, I am grateful for this project and the opportunity to work with such a fantastic team. It has been a fulfilling and educational experience that will benefit me in the future.

8 Summary

8.1 Conclusion

In conclusion, this project aimed at helping insurance companies in their decision-making process by transforming and using data provided by FEMA on natural disasters. By designing a database and constructing a balanced scorecard, we were able to define relevant KPIs for insurance companies and develop a dashboard that provides an overview of important natural disaster data. Through the dashboard, insurance companies can identify trends in damages, types of disasters, and disaster frequency, aiding in product innovation and premium calculation. Additionally, the dashboard provides insight into the demands of customers in various domains. Overall, this project provides a valuable tool to insurance companies to enable more data driven decision-making and improve their risk assessment in the field of natural disasters.

8.2 Feedback reflection

The group received peer-feedback from two other groups and a TA, about halfway through the project. By that time, the project was still in progress and the submission did not include all the screenshots and was missing a proper introduction and case description. The first peer feedback was not helpful as the group that did the feedback did not find any weaknesses in the project and only praised the work. The second peer feedback was more useful because the group highlighted what is missing in our project. The TA's feedback was the most helpful because the teacher was not shy about criticizing and finding mistakes in the project, basically the TA found the key issues of the project. The feedback clearly pointed that out, so the group reflected on it in a meaningful way.

After summing up all the feedback, it was mainly concerning the missing screenshots of the ER-diagrams and the problem statement of the insurance case. The feedback on the submitted parts of the project was highly positive, bringing confidence in what has already been done and providing a good foundation for the remainder of the project.

Getting another view on the work is always important and useful. All suggestions from the three sources were accepted with respect and were implemented immediately. Modifications were made to the case description and ER-diagram, and the team adjusted the report according to the feedback, implementing all relevant screenshots and a proper introduction and problem statement.

8.3 References

Eslamian, S., & Eslamian, F. A. (2021). Handbook of disaster risk reduction for resilience : new frameworks for building resilience to disasters. Springer. https://doi.org/10.1007/978-3-030-61278-8

Onuma, H., Shin, K. J., & Managi, S. (2017). *Household preparedness for natural disasters: Impact of disaster experience and implications for future disaster risks in Japan.* International Journal of Disaster Risk Reduction, 21, 148–158. https://doi.org/10.1016/j.ijdrr.2016.11.004

9 Appendix: SQL Queries

The team faced an issue during the optimization of the database when modifying attribute types, resulting in the loss of some essential functionalities in the dashboard. The team acknowledges that alternative modifications could have been implemented, prompting the decision to incorporate two versions of CREATE queries: the current version and an improved version.

9.1 CREATE

```
CREATE TABLE public.disaster_declarations (
        disasternumber int NOT NULL,
        declarationdate date,
        incidentbegindate date,
        incidentenddate date,
        declarationtype varchar(50),
        incidenttype varchar(50),
        entrydate date,
        updatedate date,
        closeoutdate date,
        ihprogramdeclared int,
        iaprogramdeclared int,
        paprogramdeclared int,
        hmprogramdeclared int,
        CONSTRAINT disaster_declarations_pk PRIMARY KEY (disasternumber)
);
```

Figure 7: Create disaster_declaration table (improved)

```
CREATE TABLE public.disaster_declarations (
        disasternumber int4 NOT NULL,
        declarationdate varchar(50) NULL,
        incidentbegindate varchar(50) NULL,
        incidentenddate varchar(50) NULL,
        declarationtype varchar(50) NULL,
        incidenttype varchar(50) NULL,
        entrydate varchar(50) NULL,
        updatedate varchar(50) NULL,
        closeoutdate varchar(50) NULL,
        ihprogramdeclared int4 NULL,
        iaprogramdeclared int4 NULL,
        paprogramdeclared int4 NULL,
        hmprogramdeclared int4 NULL,
        CONSTRAINT disaster_declarations_pk PRIMARY KEY (disasternumber)
);
```

Figure 8: Create disaster_declaration table

```
CREATE TABLE public.disaster_location (
        disasternumber int NOT NULL,
        statename varchar(50),
        CONSTRAINT disaster_location_pk PRIMARY KEY (disasternumber),
        CONSTRAINT disaster_location_fk FOREIGN KEY (disasternumber) REFERENCES
public.disaster_declarations(disasternumber),
        CONSTRAINT disaster_location_fk_1 FOREIGN KEY (statename) REFERENCES public.states(statename)
);
```

Figure 9: Create disaster_location table (improved)

```
CREATE TABLE public.disaster_location (
        disasternumber int4 NOT NULL,
        statename varchar(50) NULL,
        CONSTRAINT disaster_location_pk PRIMARY KEY (disasternumber),
        CONSTRAINT disaster_location_fk FOREIGN KEY (disasternumber) REFERENCES
public.disaster_declarations(disasternumber),
        CONSTRAINT disaster_location_fk_1 FOREIGN KEY (statename) REFERENCES public.states(statename)
);
```

Figure 10: Create disasterlocation05 table

```
CREATE TABLE public.states (
        statecode varchar(10) NOT NULL,
        statename varchar(50),
        CONSTRAINT states_pk PRIMARY KEY (statename)
);
```

Figure 11: Create states table (improved)

```
CREATE TABLE public.states (
        statecode varchar(250) NULL,
        statename varchar(250) NOT NULL,
        CONSTRAINT states_pk PRIMARY KEY (statename)
);
```

Figure 12: Create states table

```
CREATE TABLE public.disaster_summaries_amount (
        disasternumber int NOT NULL,
        totalnumberiaapproved int,
        totalamountihpapproved int,
        totalamounthaapproved int,
        totalamountonaapproved int,
        totalobligatedamountpa int,
        totalobligatedamountcatab int,
        totalobligatedamountcatc2g int,
        totalobligatedamounthmgp int
);
```

Figure 13: Create disaster_summaries_amount table (improved)

```
CREATE TABLE public.disaster_summaries_amount (
        disasternumber int4 NULL,
        totalnumberiaapproved int4 NULL,
        totalamountihpapproved float4 NULL,
        totalamounthaapproved float4 NULL,
        totalamountonaapproved float4 NULL,
        totalobligatedamountpa float4 NULL,
        totalobligatedamountcatab float4 NULL,
        totalobligatedamountcatc2g float4 NULL,
        totalobligatedamounthmgp float4 NULL
);
```

Figure 14: Create disaster_summaries_amount table

9.2 SELECT

```
SELECT dd.disasternumber, dl.statename, dsa.totalnumberiaapproved
FROM disaster_declarations dd
JOIN disaster_location dl ON dd.disasternumber = dl.disasternumber
JOIN disaster_summaries_amount dsa ON dd.disasternumber = dsa.disasternumber
WHERE dl.statename = 'California' AND dsa.totalnumberiaapproved > 1000;
```

9.3 ALTER

```
ALTER TABLE public.disaster_location ADD CONSTRAINT disaster_location_fk FOREIGN KEY (disasternumber) REFERENCES public.disaster_declarations(disasternumber);
```

9.4 DROP

```
ALTER TABLE disaster_summaries_amount
DROP COLUMN last_refresh,
DROP COLUMN hash,
DROP COLUMN id;
```