Friedrich-Alexander-Universität
Erlangen-Nürnberg

FAU

Analyzing the Impact of Weather Conditions on Insurance Claims in Germany

Presented By

Mohammad Naim

MSc. in Artificial Intelligence

Introduction

Overview of the Project and its Goal

The impact of weather on our daily lives extends to various aspects, notably influencing insurance claims. Germany's diverse climate presents an intriguing opportunity to delve into the correlation between weather conditions and insurance claims. This study seeks to examine this relationship, exploring how various weather patterns may affect the frequency and severity of insurance claims in Germany. Unraveling these connections could provide valuable insights for insurers and policymakers in effectively managing risks associated with weather-related claims in the country.

Project Structure

Overview of the Project Structure

- The project stucture consists of multiple files and directories.
- All the necessary variables are present in *project/config/config_var.py*
- The main entry file of the project is *project/main.py*

```
project/
├── config/
│   ├── __init__.py
│   ├── config_var.py                   # Main configuration file
│   ├── config_var.example.py           # Dummy configuration file to duplicate
│   └── source_info.json                # Data sources file
├── data/
│   ├── fau_made_project_ws23.sqlite    # Sqlite Database
├── main.py                             # Main entry point to run the pipeline
├── pipeline.py                         # Data Pipeline
├── pipeline.sh                         # Bash script of running pipeline
├── project-plan.md                     # Project Plan
├── report.ipynb                        # Notebook of final project report
├── tests.py                            # Unit testing file
└── tests.sh                            # Bash script for running tests
```

Data Sources

Overview of the Data Sources

Datasource 1: Weather Data Source

- Source: **Deutscher Wetterdienst - DWD**
- Data Type: CSV
- Authentication: not required
- License Type: OpenData License
- This dataset offers comprehensive weather-related data covering Germany and its individual states.

Datasource 2: Insurance claims Data Source

- Source: **GENESIS**
- Data Type: CSV
- Authentication: Required
- License Type: OpenData License
- This dataset provides detailed information on insurance claims across various categories within Germany.

Data Exploration

Overview of Weather Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 156 entries, 0 to 155
Data columns (total 20 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Year                          156 non-null    int64
 1   Month                         156 non-null    object
 2   Date                          156 non-null    object
 3   Brandenburg/Berlin            156 non-null    float64
 4   Brandenburg                   156 non-null    float64
 5   Baden-Wuerttemberg            156 non-null    float64
 6   Bayern                        156 non-null    float64
 7   Hessen                        156 non-null    float64
 8   Mecklenburg-Vorpommern        156 non-null    float64
 9   Niedersachsen                 156 non-null    float64
 10  Niedersachsen/Hamburg/Bremen  156 non-null    float64
 11  Nordrhein-Westfalen           156 non-null    float64
 12  Rheinland-Pfalz               156 non-null    float64
 13  Schleswig-Holstein            156 non-null    float64
 14  Saarland                      156 non-null    float64
 15  Sachsen                       156 non-null    float64
 16  Sachsen-Anhalt                156 non-null    float64
 17  Thueringen/Sachsen-Anhalt     156 non-null    float64
 18  Thueringen                    156 non-null    float64
 19  Germany                       156 non-null    float64
dtypes: float64(17), int64(1), object(2)
memory usage: 24.5+ KB
```

	Year	Month	Date	Brandenburg/Berlin	Brandenburg	Baden-Wuerttemberg	Bayern	Hessen	Mecklenburg-Vorpommern	Niedersach
0	2010	January	January-2010	-5.14	-5.15	-2.76	-3.77	-3.36	-4.64	-3
1	2010	February	February-2010	-0.61	-0.63	0.13	-1.12	-0.23	-1.03	-(
2	2010	March	March-2010	4.57	4.56	3.81	3.28	4.34	3.96	4
3	2010	April	April-2010	8.95	8.93	8.79	8.21	8.97	8.01	8
4	2010	May	May-2010	11.13	11.12	10.65	10.69	10.32	9.96	1(
...	...	...	...	...	...	...	...	...	...	
151	2022	August	August-2022	20.92	20.90	20.26	19.57	20.84	20.02	2(
152	2022	September	September-2022	13.56	13.55	13.29	12.56	13.39	13.43	13
153	2022	October	October-2022	12.52	12.51	12.92	11.92	12.15	12.40	12
154	2022	November	November-2022	5.66	5.65	6.37	5.31	6.72	6.21	7
155	2022	December	December-2022	1.56	1.55	1.95	1.10	1.85	1.20	2

Data Exploration

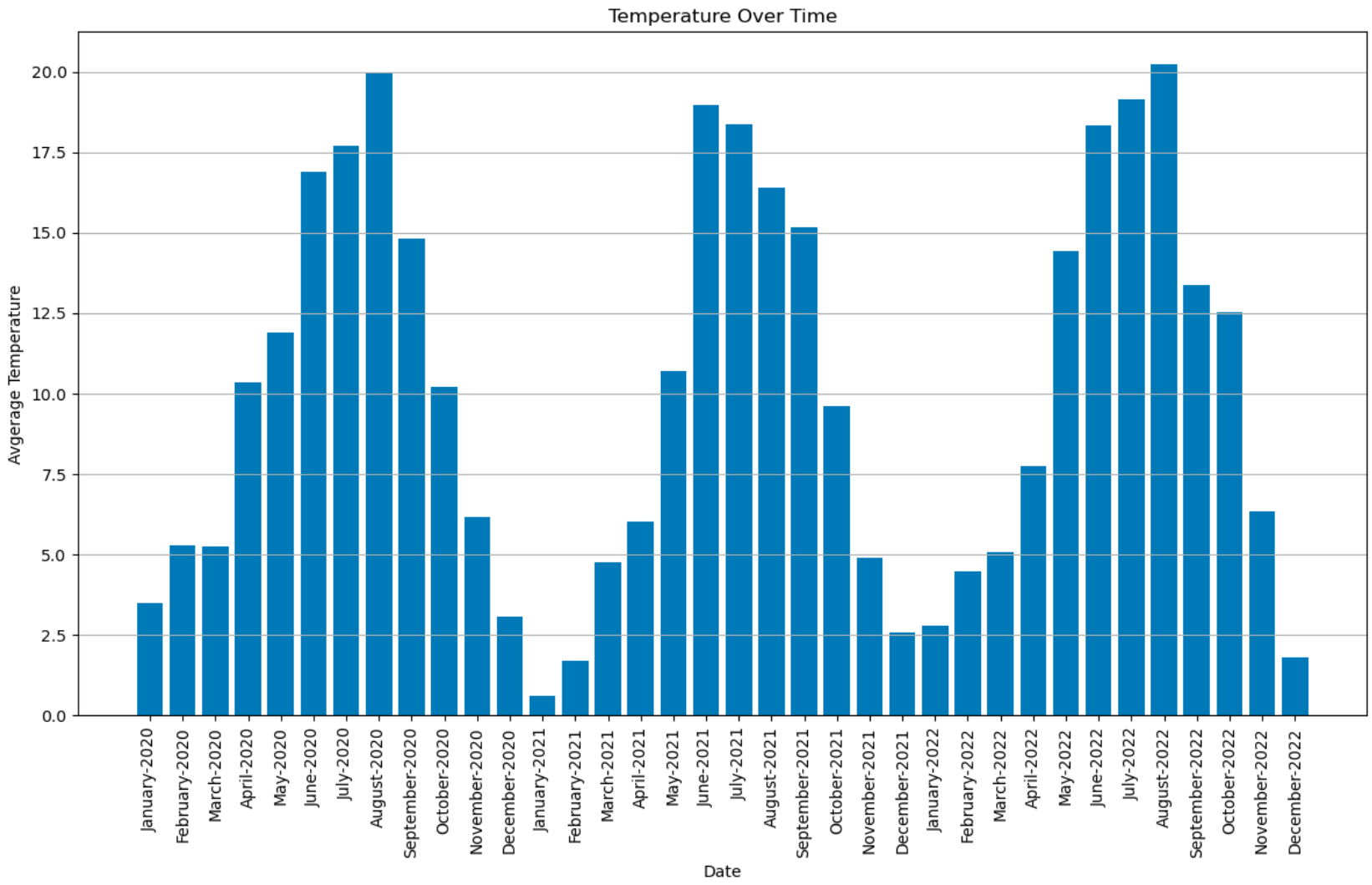
Overview of Insurance Claims Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 624 entries, 0 to 623
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype 
---  ------         --------------  ----- 
 0   Year           624 non-null    int64 
 1   Month          624 non-null    object
 2   Date           624 non-null    object
 3   Category_Type  624 non-null    object
 4   Category_Name  624 non-null    object
 5   Total_Claim    624 non-null    int64 
dtypes: int64(2), object(4)
memory usage: 29.4+ KB
```

	Year	Month	Date	Category_Type	Category_Name	Total_Claim
0	2010	January	January-2010	WZ08-G	Wholesale, retail trade, repair of motor vehicles	496065
1	2010	January	January-2010	WZ08-H	Transportation and storage	89613
2	2010	January	January-2010	WZ08-I	Accommodation and food service activities	33444
3	2010	January	January-2010	WZ08-Q	Human health and social work activities	14137
4	2010	February	February-2010	WZ08-G	Wholesale, retail trade, repair of motor vehicles	390388
...	...	...	...	...	...	...
619	2022	November	November-2022	WZ08-Q	Human health and social work activities	6892
620	2022	December	December-2022	WZ08-G	Wholesale, retail trade, repair of motor vehicles	230722
621	2022	December	December-2022	WZ08-H	Transportation and storage	22437
622	2022	December	December-2022	WZ08-I	Accommodation and food service activities	21600
623	2022	December	December-2022	WZ08-Q	Human health and social work activities	56256

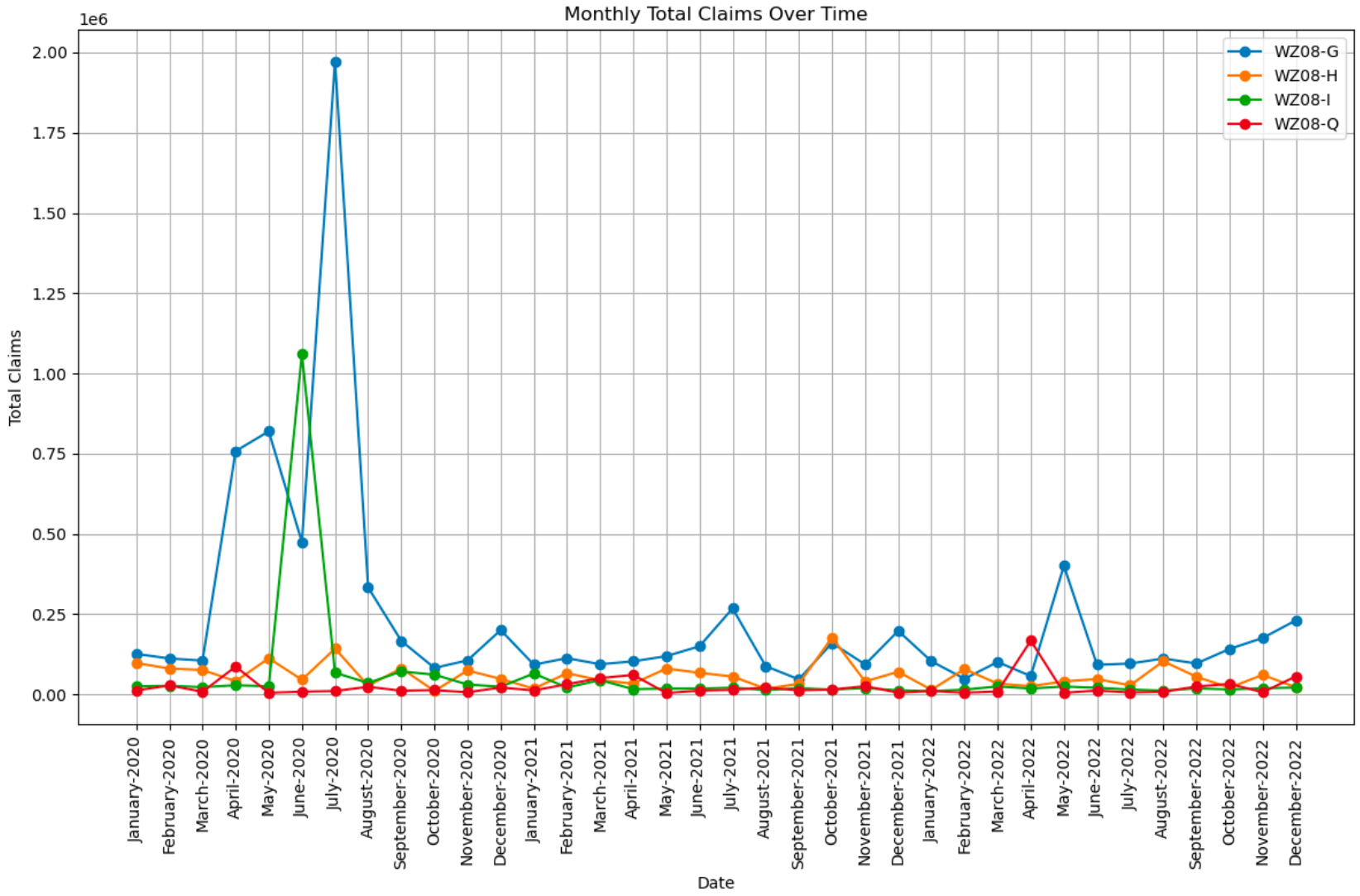
Temperature Pattern

Temperature pattern in Germany from 2020 to 2022



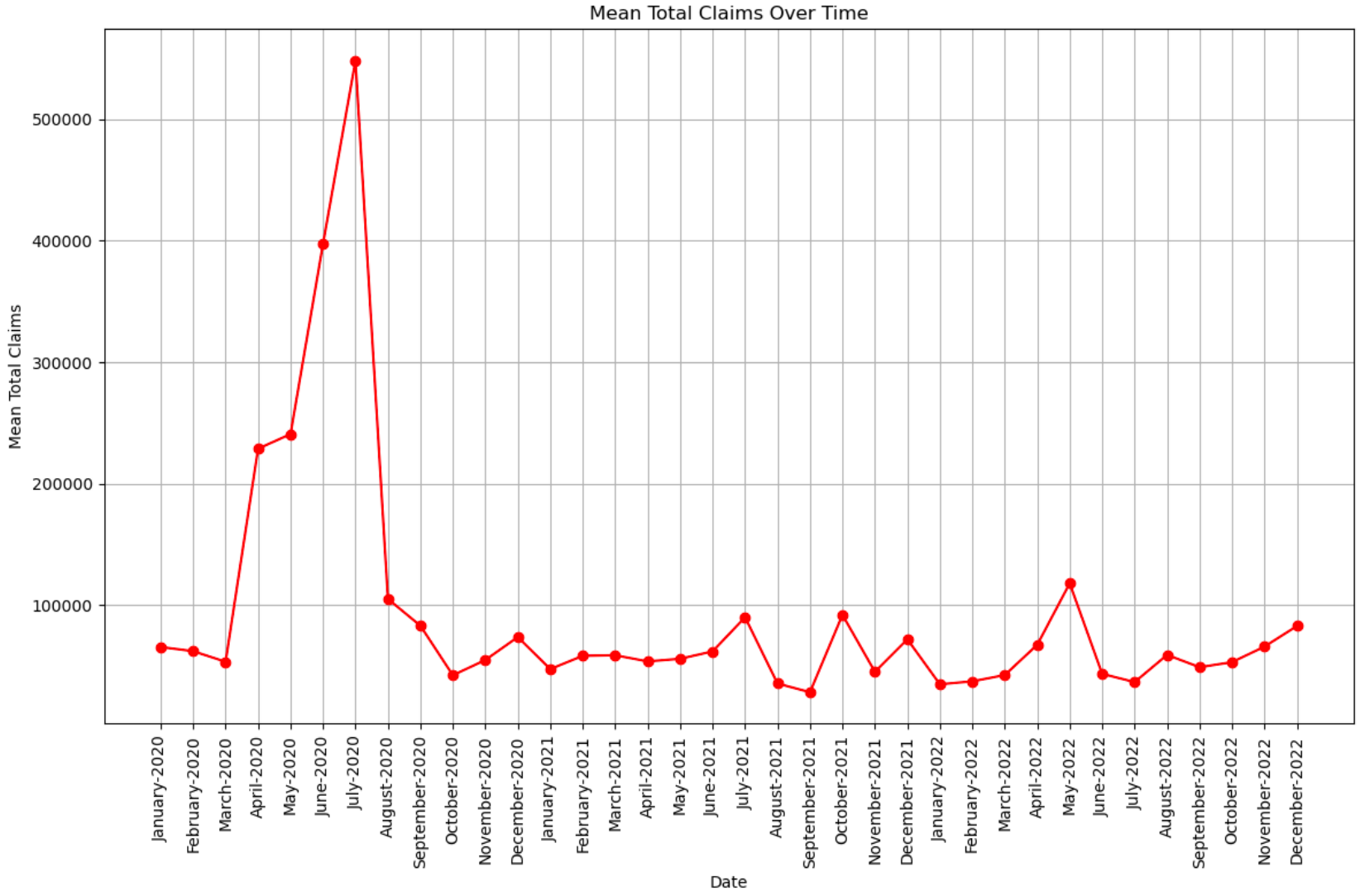
Insurance Claims Pattern

Insurance claims pattern across various categories in Germany from 2020 to 2022



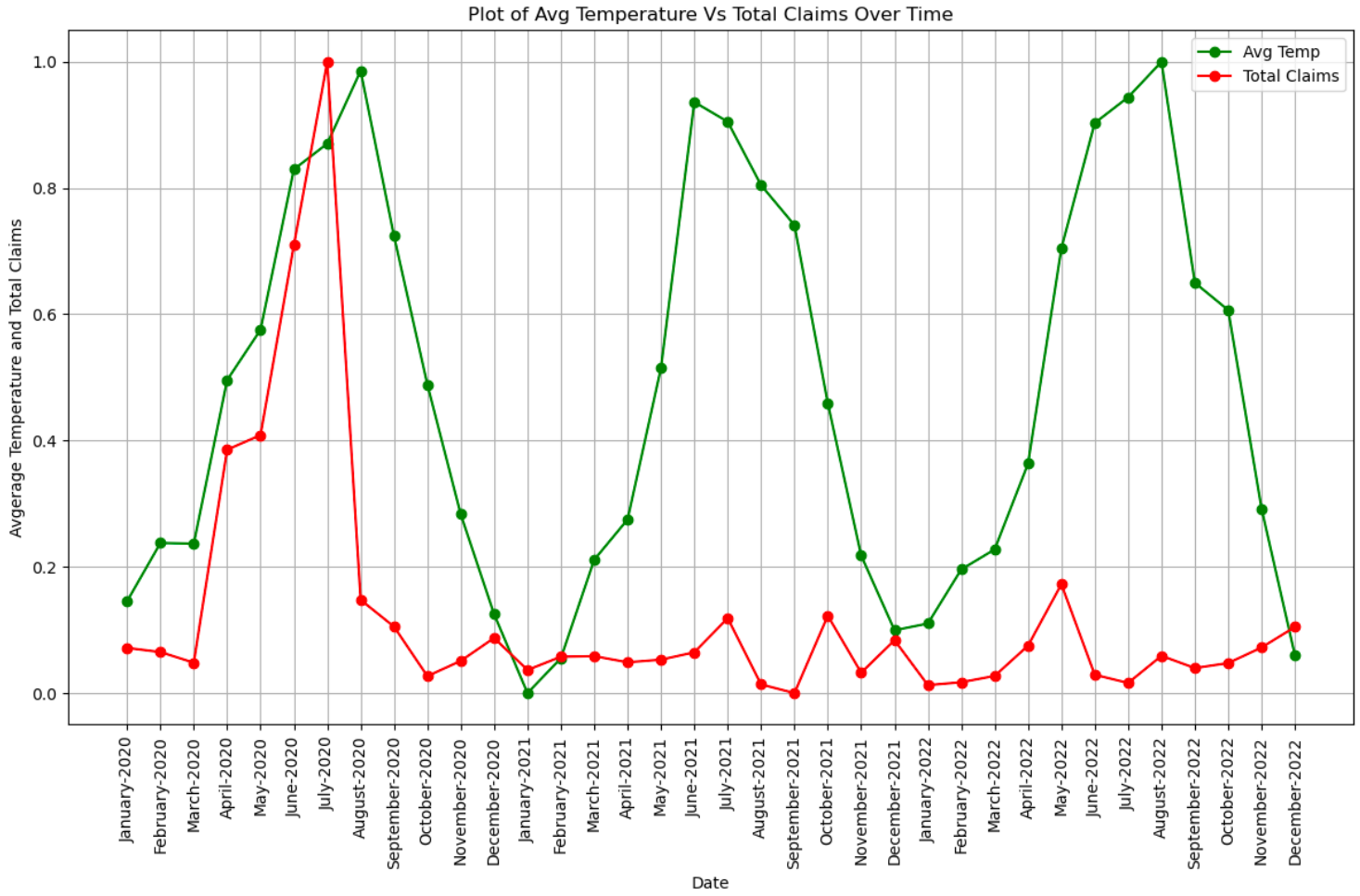
Mean Insurance Claims Pattern

Mean Insurance claims pattern in Germany from 2020 to 2022



Comparison & Analysis

Detailed comparison, analysis and findings

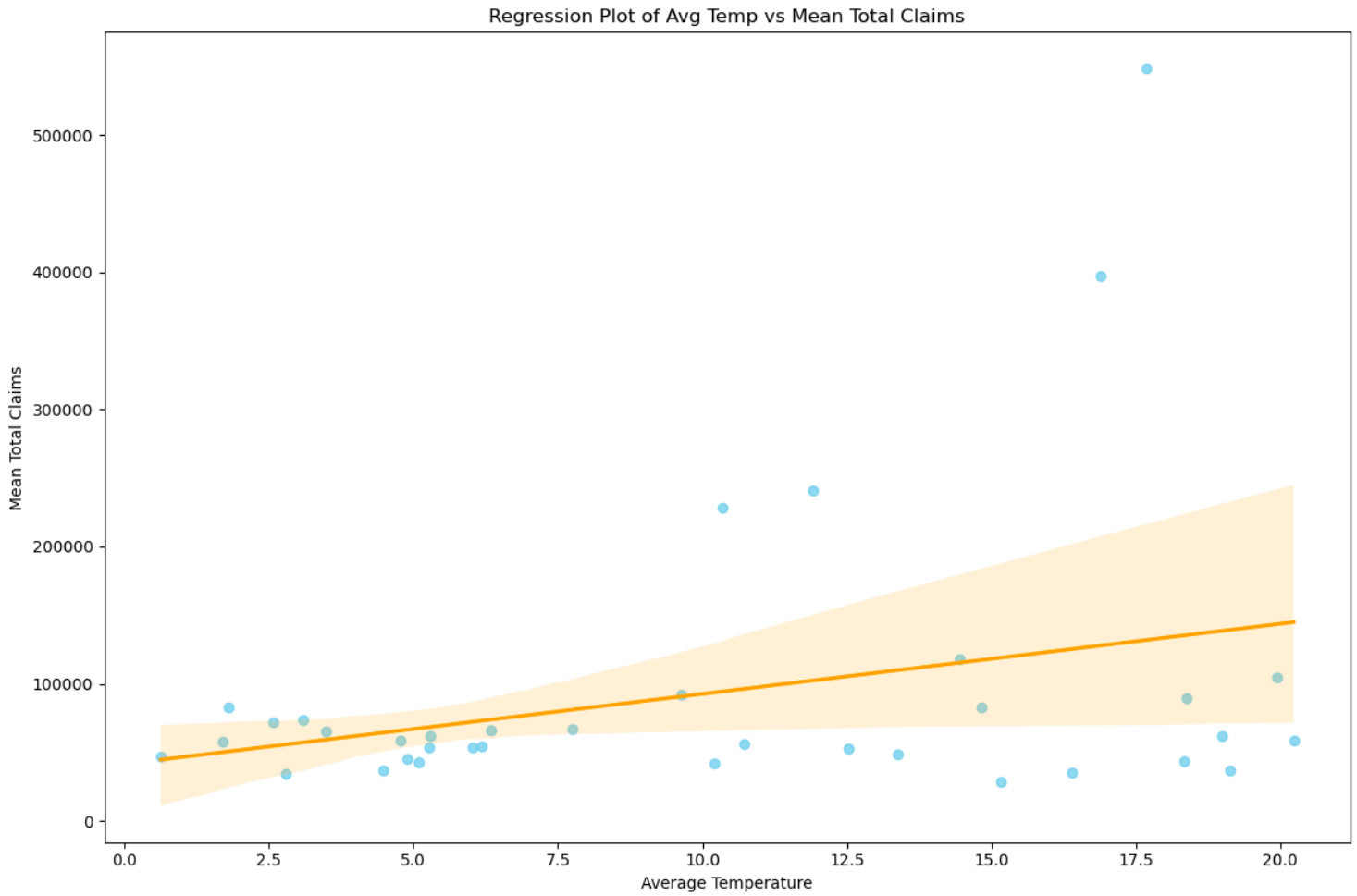


Analysis with Normalization

- Here, we have first normalized the average temperature and the total claims to plot them in the same graph.
- The average temperature peaked in July 2022, and then declined in the following months. The total number of claims peaked in August 2022, and then declined in the following months.
- The graph shows that there is a positive correlation between average temperature and total number of claims.

Comparison & Analysis

Detailed comparison, analysis and findings



Analysis with Regression

- In this graph, the points on the scatter plot show the average temperature and mean total claims for each month. The orange line is the regression line.
- The regression line is the best-fit line for the data.
- The regression line slopes upwards, which means that there is a positive correlation between average temperature and mean total claims.

Result

Overview of the Analysis & Findings

*Over the period spanning 2020 to 2022, an observable relationship emerges between the average temperature and insurance claims in Germany. This relationship exhibits **a moderate positive correlation** indicating that there is a tendency for both variables to move in sync, either increasing or decreasing together.*

Conclusion

Summary and Discussion of Result

- This correlation isn't absolute - there are instances when fluctuations in one factor don't align with changes in the other.
- This suggests that there are other factors besides temperature that are also influencing the number of insurance claims in Germany. For example, population increase.
- One possible explanation for the correlation between average temperature and insurance claims is that warmer weather leads to more outdoor activities, which can increase the risk of accidents and injuries.
- Another possibility is that warmer weather can lead to an increase in stress levels, which can in turn lead to more aggressive behavior and an increased risk of accidents.

Friedrich-Alexander-Universität
Erlangen-Nürnberg

FAU

Thank You.