A Deep Dive into Insurance Data using PowerBI

Capstone Project Summary

Scope of the Project

The Great American Insurance Group lacked transparency in its operations. To address this, the company is digitalizing and implementing a comprehensive reporting solution. This project involves creating multiple dashboards and reports to present key statistics required for stakeholders to make informed business decisions. Additionally, this project explores the use of predictive analytics with PowerBI functions, utilizing time series forecasting and regression algorithms to make future predictions for target values.

Confidential Data and Sample Data Usage

After considerable discussion on how data would be made available to carry out the project as insurance data is very confidential and one of the risks initially mentioned in the project was regarding the data and how the scope can be achieved by performing the analysis using 1000 observations and later linking it with the original data warehouse after successfully deploying the reports onto the company's Integrated PowerBI service and switching credentials to the Azure data warehouse. Significant effort was dedicated to replicating real configurations and setting up the data accurately. This involved creating calculated columns and measures using various formulas to derive meaningful insights. Coordination with a data engineer ensured proper naming conventions. Once the data was prepared, it was connected to PowerBI for visualization of the multiple reports shown below:

Visualisations



As we can see, all these pages represent the statistics needed by the insurance company's team to make decisions and to understand the performance of the company. These visualizations not only display basic stats but also produce derived statistics which is of key importance.

The scope of this project was to increase transparency which has been achieved and also to explore the predictive features of PowerBI functions that use well trained models and provide target values and can be used to create regression models as well.

The next area that will be delved into as per the discussion is to include models that can predict what variables are more important in sales using regression. The variable importance feature will be used to understand what makes a customer choose GAIG over other things and what will be the required steps to boost customer satisfaction.

Conclusion

We can see what value we have added to the company and whether this project will be reproducible in the future and of good use. Since the actual data warehouse was connected and the target values on the dashboards made sense to the analytics team, it seems to have added value in the discussions and meetings at GAIG already and has been noted as a thing they want to dive into an improve over the course of the next few months at least.

The learning experience in this project helped us understand that when working with data, there will be various roadblocks on the way, transformations will be required, data might have to be molded into the right shape that can take a huge chunk of time provided for the project but this showed how there can be workarounds to each problem, how important it is for the data to be understood, the domain knowledge to be perfect to be able to understand what variables can be thought of by the client and every addition that can be useful in making decisions in meetings. I also learnt how the industrial mindset is and how we can use the real power of data to move in the right direction.