



# MASA Hackathon 2022

S3S2'19


*Dexter, Dongheng, Jason, Ze Li, Ming Roong*



# Diverse STEM Backgrounds

---


## **Dexter Woo Teng Koon**

- The Chinese University of Hong Kong
-  BSc in Quantitative Finance and Risk Management Science, Minor in Mathematics and Statistics


## **Lee Dongheng**

- Nanyang Technological University, Singapore
-  BSc in Mathematical Sciences with a Minor in Finance

## **Lee Jason**

- Asia Pacific University of Technology & Innovation (APU)
-  BSc (Hons.) in Computer Science with a specialism in Data Analytics

## **Liew Ze Li**

- Monash University
-  BEng (Hons.) in Chemical Engineering

## **Ku Ming Roong**

- UOW Malaysia KDU University College
-  BSc (Hons.) in Computer Science







# Presentation Outline

---

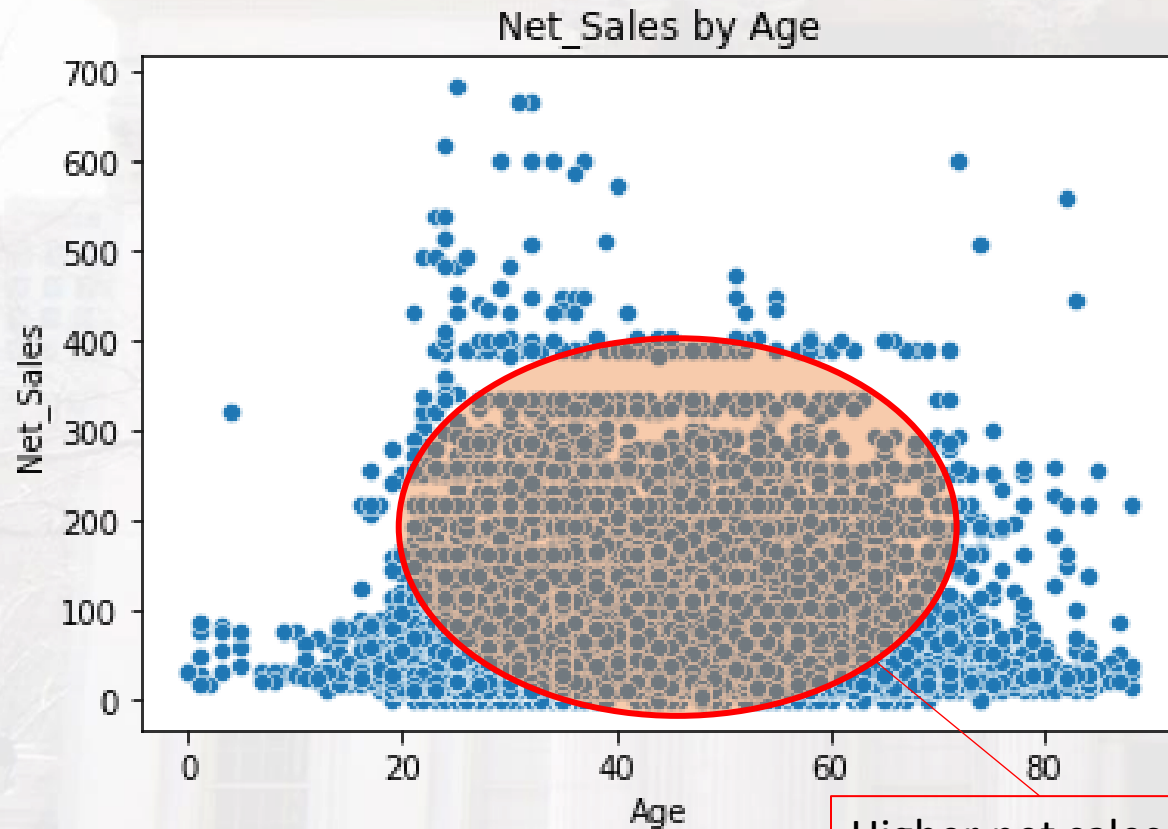
1. Sales Strategy
  2. Pricing Strategy
  3. Results
  4. Conclusion
-

# 1. Sales Strategy

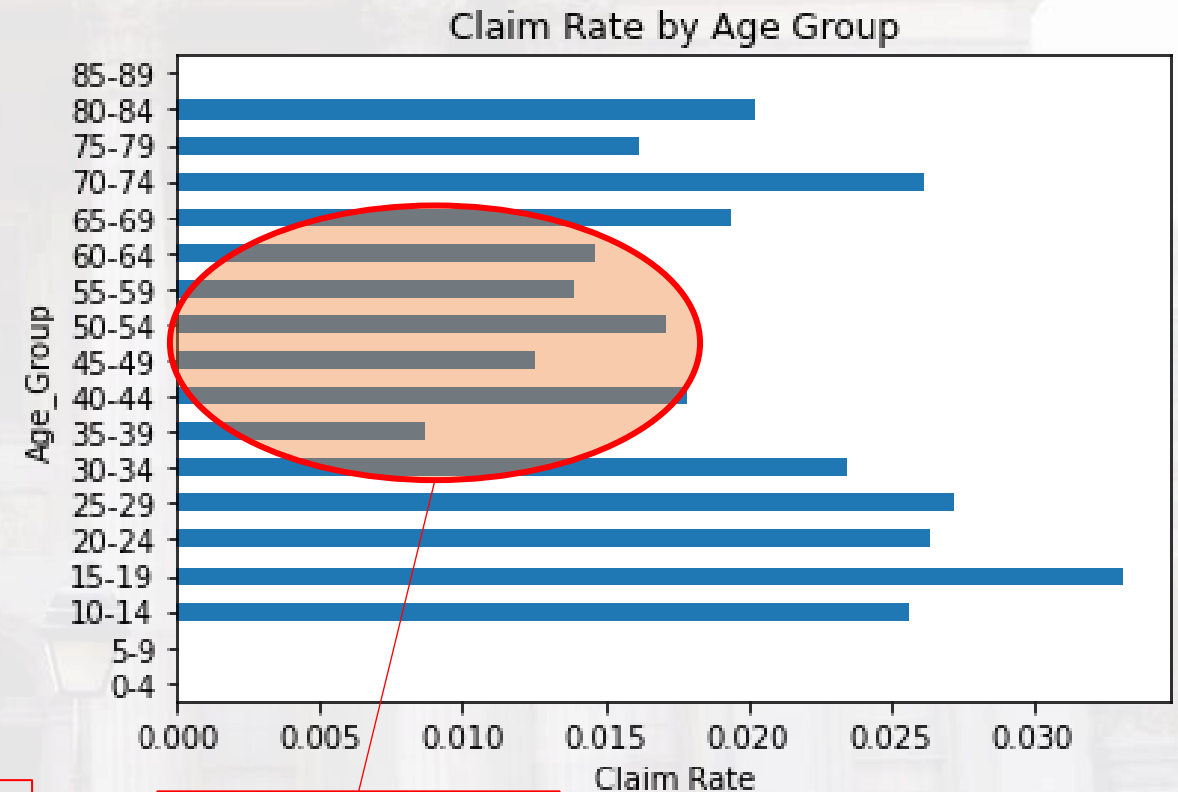
---

- **Want:** High  average profit, high  profit margin and low  claim risk
- **Where:** Online 
- **Who:** Middle-aged group  (30-70 years old)
  - Overpriced  plans on the market, despite the lowest claim rate among all age groups

# 1. Sales Strategy (con't)



Higher net sales in  
ages 30 to 70



Lower claim rates  
in ages 30 to 70

## 2. Pricing Strategy

---

- High Claim Amount  $\Rightarrow$  High Premium
- High Risk 🦴  $\Rightarrow$  High Premium

### Risk Evaluation 📊

Features investigated:

*Agency\_Type, Distribution\_Channel, Product, Duration, Destination, Age\_Group, Claims\_Amount*

---

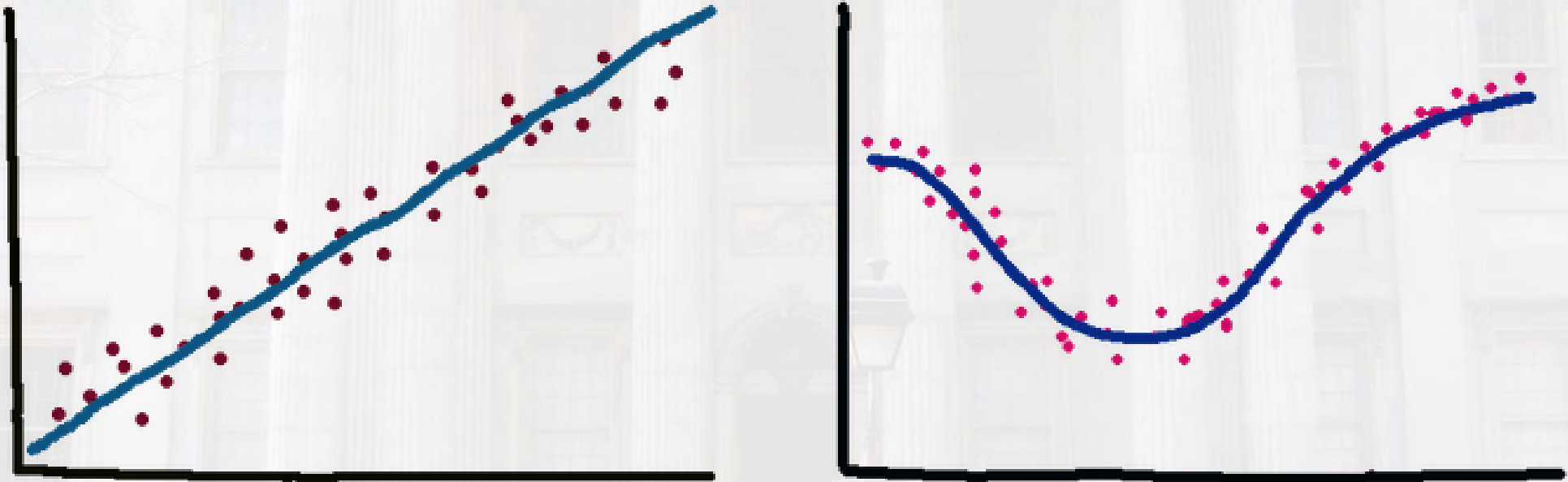


# Risk Quantification with Generalised Additive Model (GAM)

---

## What is Generalised Additive Model (GAM)?

A linear model where the target variable is the sum of a non-linear combination of the predictor variables



# Risk Quantification with Generalised Additive Model (GAM) (con't)

---

How?

**Predictor Variables:** *Agency\_Type, Distribution\_Channel, Product, Duration, Destination, Age\_Group, Claims\_Amount*

**Target Variable:** *Claim*

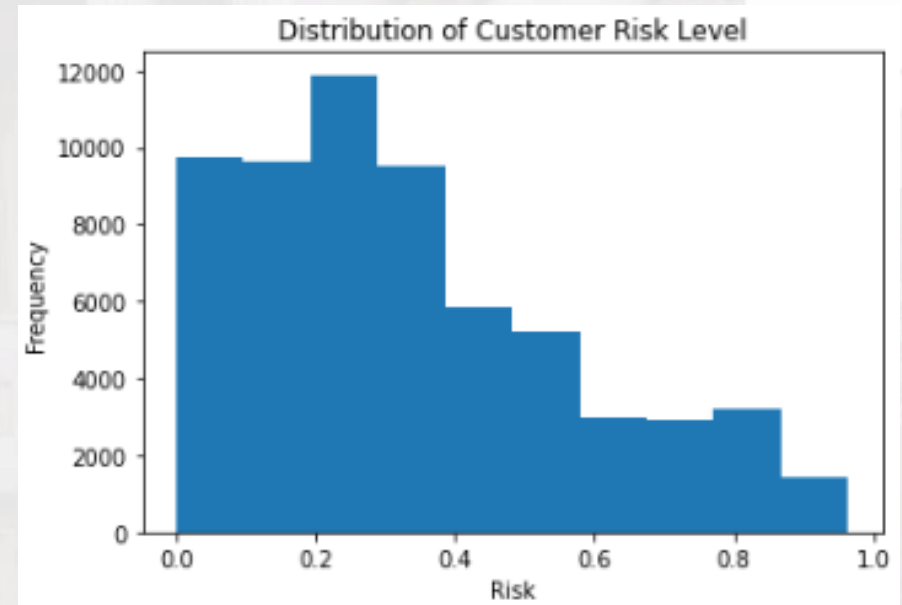
Why?

Given [Bronze Plan, Travel to Malaysia, 30 years old, ...]

Predicted *Claim* = 0.23

Outcome

Introduce a new column *Risk* to store the predicted *Claim* for each customer





## 2. Pricing Strategy (con't)

---

The premium is mainly determined by the *Claims\_Amount* and *Claim\_Rate*, and customers with higher *Risk* pay more.

$$\text{Fair premium} = \text{Claims\_Amount} \times \text{Claim\_Rate}$$

Hence, the formula for the adjusted premium (inclusive of administrative cost) is

$$\text{Premium} = \text{Claims\_Amount} \times \max\{0.5 \times \text{Risk}, \text{Claim\_Rate}\}$$

### Note:

- Only products with *Claim\_Rate*  $\leq 0.5$  are sold
  - For *Risk*  $\leq 1$ , **maximum premium** payable by client is  $0.5 \times \text{Claims\_Amount}$  of their policy
-

### 3. Results

Calculating average profit and profit margin:

- Profit 📊

If *Claim* is “Yes 👍”:

$$\text{Profit} = \text{Premium} - \text{Claims Amount}$$

If *Claim* is “No 👎”:

$$\text{Profit} = \text{Premium}$$

- Profit Margin 🚀

$$\text{Profit Margin} = \frac{\text{Profit}}{\text{Revenue}}$$


Product	Agency_Type	Profit Margin	Average Profit
1 way Comprehensive Plan	Travel Agency	0.972980	6017.143672
2 way Comprehensive Plan	Travel Agency	0.909339	15255.171915
24 Protect	Airlines	1.000000	877.112245
Annual Gold Plan	Airlines	0.735439	3343.714927
Annual Silver Plan	Airlines	0.744974	4243.900840
Annual Travel Protect Gold	Travel Agency	0.647657	10403.947956
Annual Travel Protect Platinum	Travel Agency	0.794040	12565.449313
Annual Travel Protect Silver	Travel Agency	0.782831	8846.106376
Basic Plan	Airlines	0.985785	4870.785090
Bronze Plan	Airlines	0.837696	5043.318691
Cancellation Plan	Travel Agency	0.972523	16662.386531
Child Comprehensive Plan	Travel Agency	1.000000	394.809337
Comprehensive Plan	Travel Agency	0.844689	243.099956
Gold Plan	Airlines	0.806787	2587.701623
Individual Comprehensive Plan	Travel Agency	0.892480	691.754099
Premier Plan	Airlines	1.000000	3543.763998
	Travel Agency	0.865628	1094.830209
Rental Vehicle Excess Insurance	Travel Agency	0.935724	8725.429921
Silver Plan	Airlines	0.788439	3645.750807
Single Trip Travel Protect Gold	Travel Agency	0.435871	792.174699
Single Trip Travel Protect Platinum	Travel Agency	0.056984	101.959571
Single Trip Travel Protect Silver	Travel Agency	0.718487	1548.938415
Spouse or Parents Comprehensive Plan	Travel Agency	0.677483	516.190264
Ticket Protector	Airlines	0.969898	1258.567217
Travel Cruise Protect	Travel Agency	0.979073	285.154350
Value Plan	Airlines	0.970962	4681.478001
	Travel Agency	0.850785	1361.282696

# Example

---

Case: [*Claims\_Amount* = \$4000, Travel to Malaysia, *Claim* = “Yes”, Cancellation Plan, ...]

Using GAM, predicted *Risk* = 0.8

Based on past data , *Claim\_Rate* for the product is 0.002362

$$\text{Premium} = 4000 \times \max\{0.5 \times 0.8, 0.002362\} = 4000 \times 0.4 = \text{\textcolor{green}{\$1600}} \text{💰}$$

$$\text{Profit} = 1600 - 4000 = \text{\textcolor{red}{-\$2400}} \text{💸}$$



# Discussion

We will prioritise products with

- High Average Profit
- High Profit Margin

Sort by average profit

		Average Profit	Profit Margin
Product	Agency_Type		
Cancellation Plan	Travel Agency	16662.386531	0.972523
2 way Comprehensive Plan	Travel Agency	15255.171915	0.909339
Annual Travel Protect Platinum	Travel Agency	12565.449313	0.794040
Annual Travel Protect Gold	Travel Agency	10403.947956	0.647657
Annual Travel Protect Silver	Travel Agency	8846.106376	0.782831
Rental Vehicle Excess Insurance	Travel Agency	8725.429921	0.935724
1 way Comprehensive Plan	Travel Agency	6017.143672	0.972980
Bronze Plan	Airlines	5043.318691	0.837696
Basic Plan	Airlines	4870.785090	0.985785
Value Plan	Airlines	4681.478001	0.970962
Annual Silver Plan	Airlines	4243.900840	0.744974
Silver Plan	Airlines	3645.750807	0.788439
Premier Plan	Airlines	3543.763998	1.000000
Annual Gold Plan	Airlines	3343.714927	0.735439
Gold Plan	Airlines	2587.701623	0.806787
Single Trip Travel Protect Silver	Travel Agency	1548.938415	0.718487
Value Plan	Travel Agency	1361.282696	0.850785
Ticket Protector	Airlines	1258.567217	0.969898
Premier Plan	Travel Agency	1094.830209	0.865628
24 Protect	Airlines	877.112245	1.000000
Single Trip Travel Protect Gold	Travel Agency	792.174699	0.435871
Individual Comprehensive Plan	Travel Agency	691.754099	0.892480
Spouse or Parents Comprehensive Plan	Travel Agency	516.190264	0.677483
Child Comprehensive Plan	Travel Agency	394.809337	1.000000
Travel Cruise Protect	Travel Agency	285.154350	0.979073
Comprehensive Plan	Travel Agency	243.099956	0.844689
Single Trip Travel Protect Platinum	Travel Agency	101.959571	0.056984

Sort by profit margin

		Average Profit	Profit Margin
Product	Agency_Type		
24 Protect	Airlines	877.112245	1.000000
Premier Plan	Airlines	3543.763998	1.000000
Child Comprehensive Plan	Travel Agency	394.809337	1.000000
Basic Plan	Airlines	4870.785090	0.985785
Travel Cruise Protect	Travel Agency	285.154350	0.979073
1 way Comprehensive Plan	Travel Agency	6017.143672	0.972980
Cancellation Plan	Travel Agency	16662.386531	0.972523
Value Plan	Airlines	4681.478001	0.970962
Ticket Protector	Airlines	1258.567217	0.969898
Rental Vehicle Excess Insurance	Travel Agency	8725.429921	0.935724
2 way Comprehensive Plan	Travel Agency	15255.171915	0.909339
Individual Comprehensive Plan	Travel Agency	691.754099	0.892480
Premier Plan	Travel Agency	1094.830209	0.865628
Value Plan	Travel Agency	1361.282696	0.850785
Comprehensive Plan	Travel Agency	243.099956	0.844689
Bronze Plan	Airlines	5043.318691	0.837696
Gold Plan	Airlines	2587.701623	0.806787
Annual Travel Protect Platinum	Travel Agency	12565.449313	0.794040
Silver Plan	Airlines	3645.750807	0.788439
Annual Travel Protect Silver	Travel Agency	8846.106376	0.782831
Annual Silver Plan	Airlines	4243.900840	0.744974
Annual Gold Plan	Airlines	3343.714927	0.735439
Single Trip Travel Protect Silver	Travel Agency	1548.938415	0.718487
Spouse or Parents Comprehensive Plan	Travel Agency	516.190264	0.677483
Annual Travel Protect Gold	Travel Agency	10403.947956	0.647657
Single Trip Travel Protect Gold	Travel Agency	792.174699	0.435871
Single Trip Travel Protect Platinum	Travel Agency	101.959571	0.056984

## 4. Conclusion

---

### Future Efforts

- Products that cater to a wider range of clients
  - Explore insurance plans in areas other than travel and airlines (e.g. health)
  - For plans similar to those offered by other agencies, strive to provide lower premiums and greater benefits
  - Collect more client data to improve our claim risk prediction algorithm
-

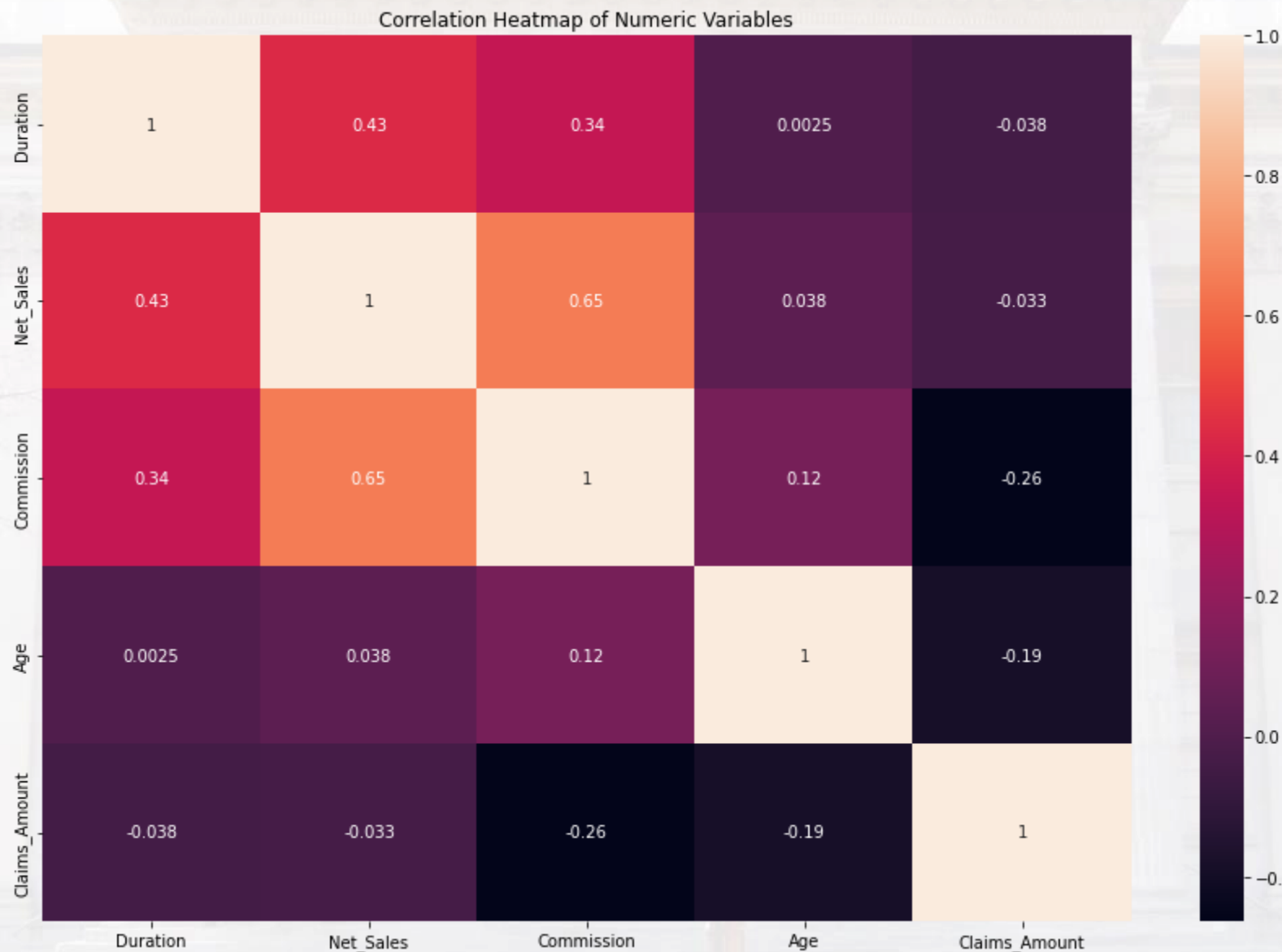




Thank you for listening!

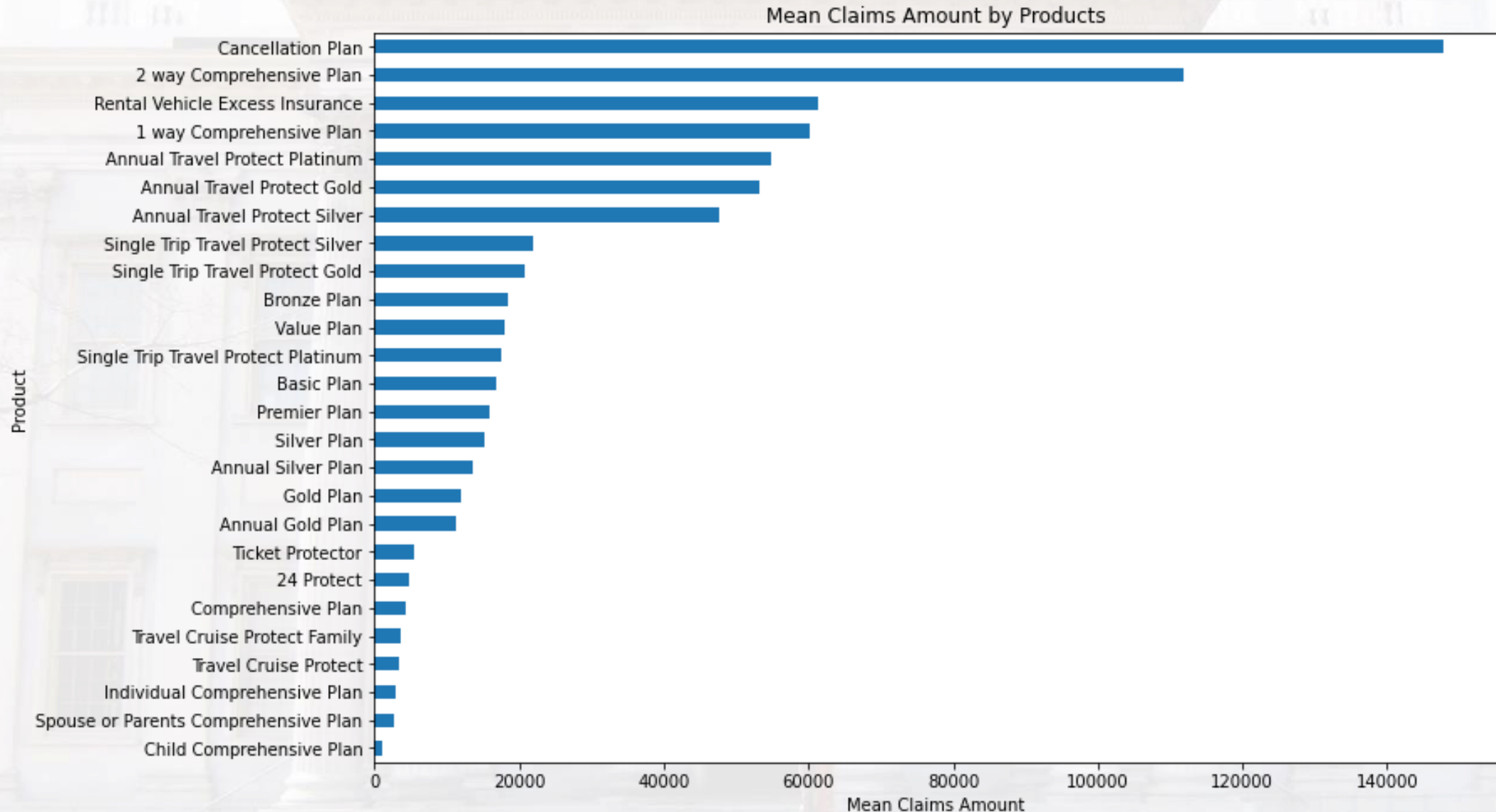


# Appendix: Data Visualisations



No significant correlations between the variables

# Appendix: Data Visualisations (con't)



# Appendix: Generalised Additive Model

---

```
from pygam import LogisticGAM
X3 = df11.drop(columns=['Claim', 'Net_Sales', 'Agency', 'Commission'], inplace=False)
y3 = df11['Claim']

# Encoding categorical data
le = LabelEncoder()
X3['Agency_Type'] = le.fit_transform(X3['Agency_Type'])
X3['Distribution_Channel'] = le.fit_transform(X3['Distribution_Channel'])
X3['Age_Group'] = le.fit_transform(X3['Age_Group'])
y3 = le.fit_transform(y3)
ohe = OneHotEncoder(drop=None, sparse=True, handle_unknown='error')
X3['Product'] = ohe.fit_transform(X3['Product'].values.reshape(-1, 1)).toarray()
X3['Destination'] = ohe.fit_transform(X3['Destination'].values.reshape(-1, 1)).toarray()

# Split the data into training and test sets
X3_train, X3_test, y3_train, y3_test = train_test_split(X3, y3, test_size=0.2, random_state=42)

# Oversampling imbalance data
smote = SMOTE(random_state=42)
X3_SMOTE, y3_SMOTE = smote.fit_resample(X3_train, y3_train)

# Train model
gam = LogisticGAM().fit(X3_SMOTE, y3_SMOTE)
```

---



## Appendix: Generalised Additive Model (con't)

LogisticGAM

```
=====
Distribution:          BinomialDist Effective DoF:          44.3325
Link Function:        LogitLink Log Likelihood:          -48904.5578
Number of Samples:    98224 AIC:          97897.7806
                        AICc:          97897.8234
                        UBRE:          2.997
                        Scale:          1.0
                        Pseudo R-Squared:          0.2817
=====
```

Feature Function	Lambda	Rank	EDoF	P > x	Sig. Code
s(0)	[0.6]	20	2.0	0.00e+00	***
s(1)	[0.6]	20	1.0	0.00e+00	***
s(2)	[0.6]	20	1.9	0.00e+00	***
s(3)	[0.6]	20	12.9	0.00e+00	***
s(4)	[0.6]	20	0.9	8.79e-02	.
s(5)	[0.6]	20	11.5	0.00e+00	***
s(6)	[0.6]	20	14.0	0.00e+00	***
intercept		1	0.0	9.81e-01	

```
=====
Significance codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
=====
```