EXPLORE AI

INSURANCE CLAIM PREDICTION

BY TEAM 15

EXPLORE AI ACADEMY

AGENDA

01 Introduction

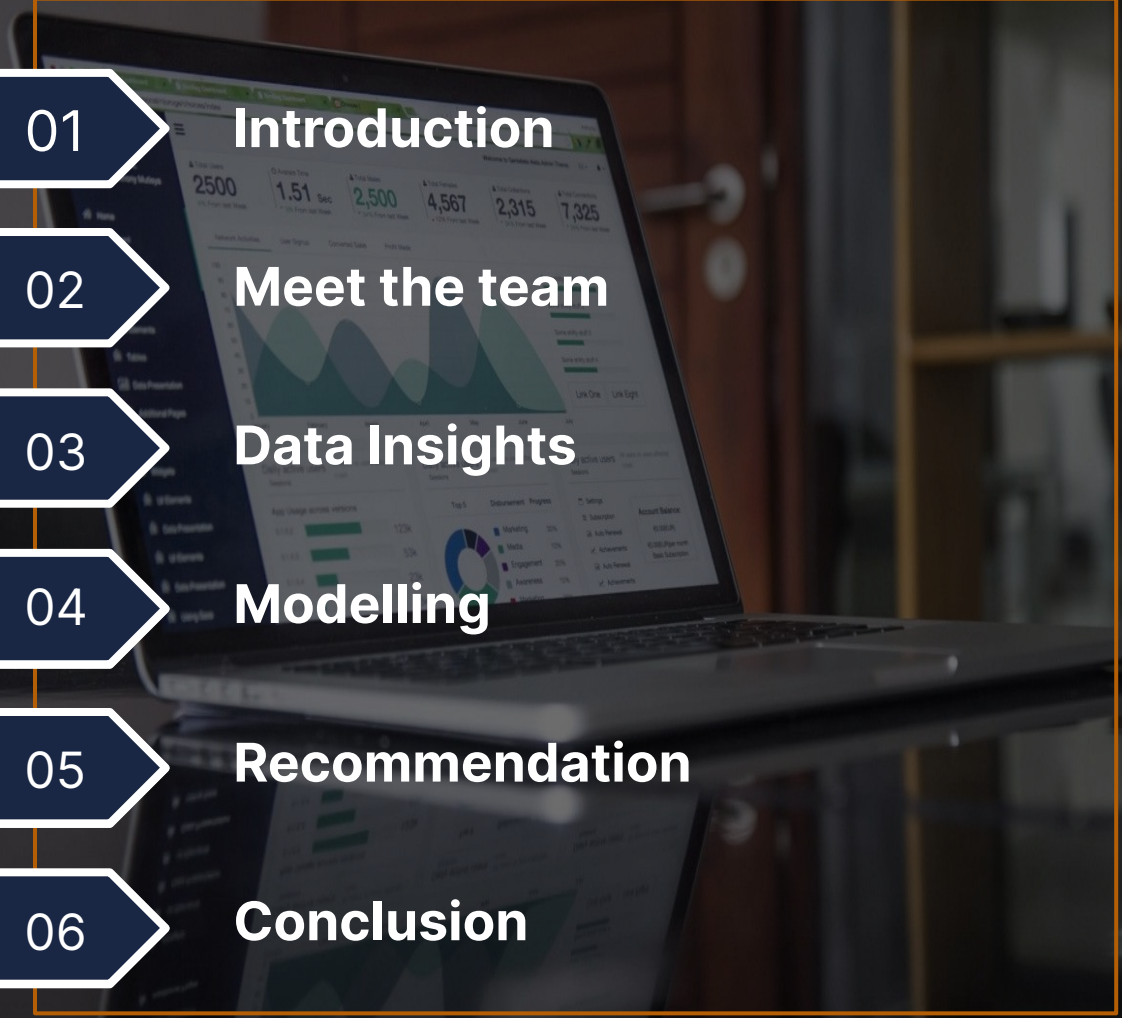
02 Meet the team

03 Data Insights

04 Modelling

05 Recommendation

06 Conclusion



EXPLORE AI ACADEMY

BY TEAM 15

Introduction

01 Project definition

02 Details of the data set

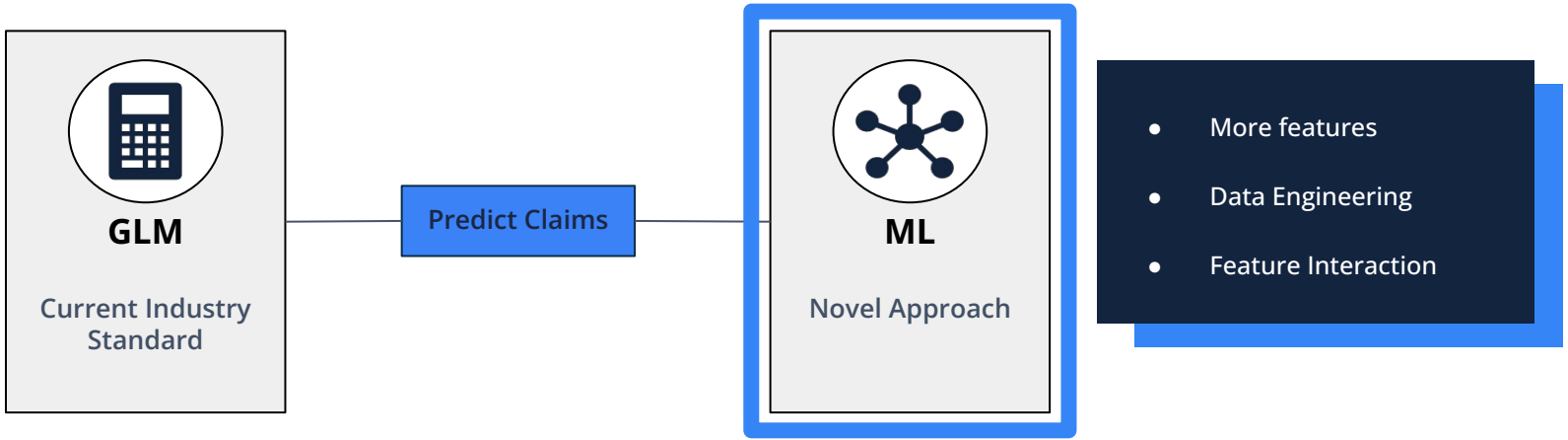
03 Summary of the data set

04 Observations

Project **Overview**

Aim

- Machine learning model **to Predict the claims** over the next year for a motor insurance client
- Using **personal** and **vehicle** information
- With a **non-conservative approach (Machine Learning models)**
- Instead of **simple Generalised Linear Models**
- Used to predict **claim severity of potential policyholders**

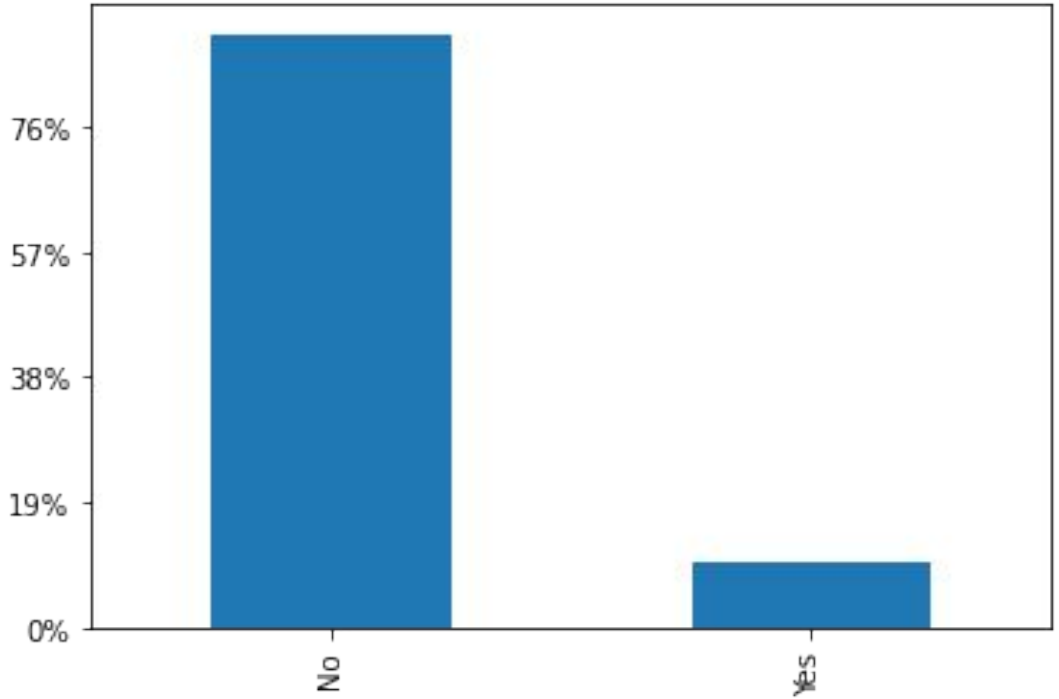


Exploratory Data **Analysis Overview**

Things we looked at

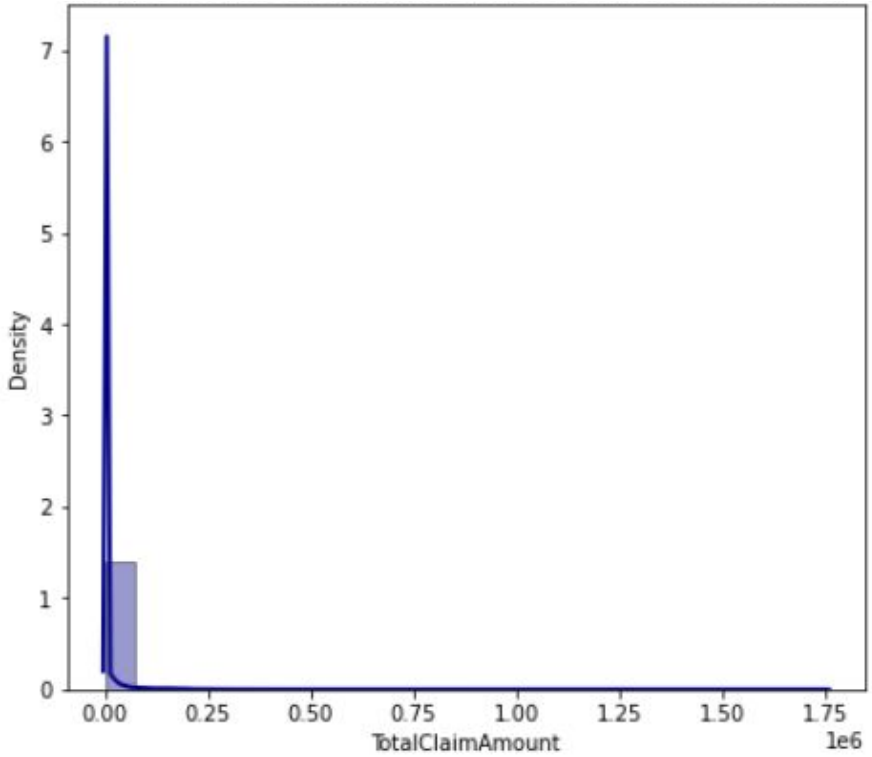
- Distribution of the data - claim counts and claim amounts
- Claim amounts for different categorical features
- Claim count percentages for different categorical features
- Relationships of numerical values with claim amount

Who made claims?

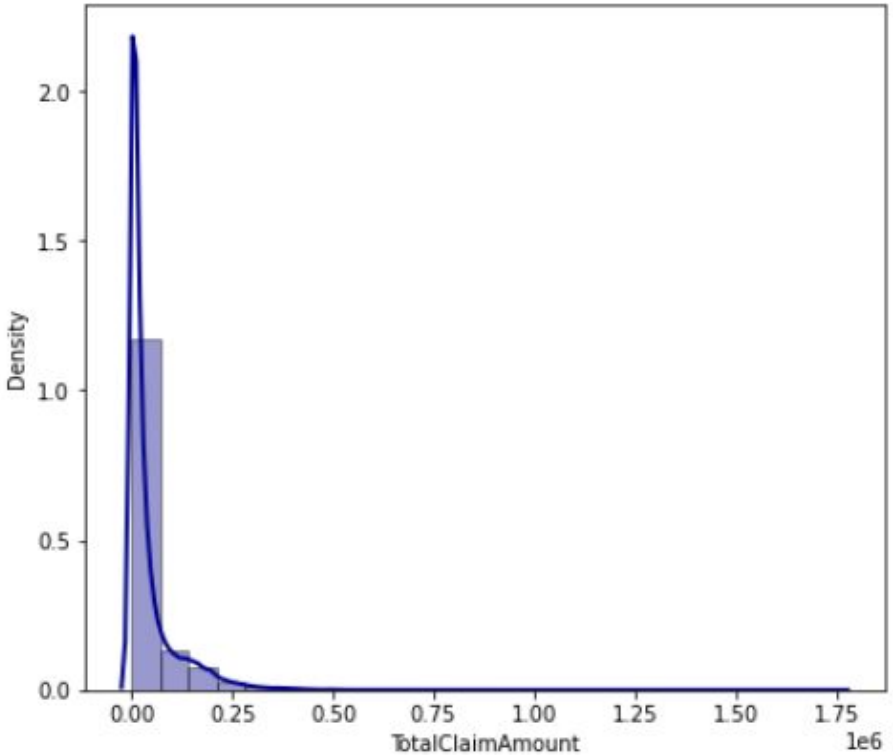


Claim amount **distribution**

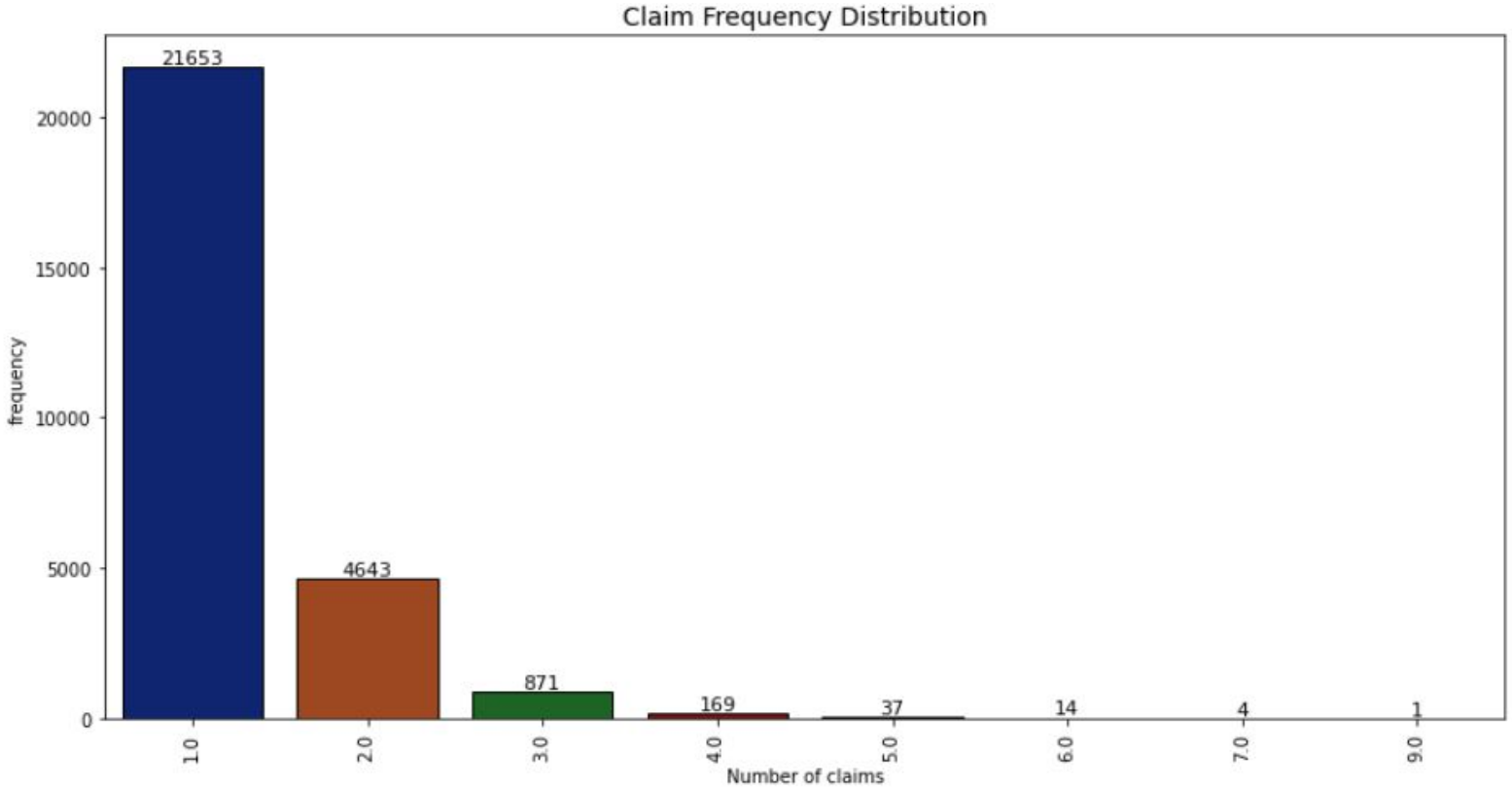
Claim amount (including zero claims)



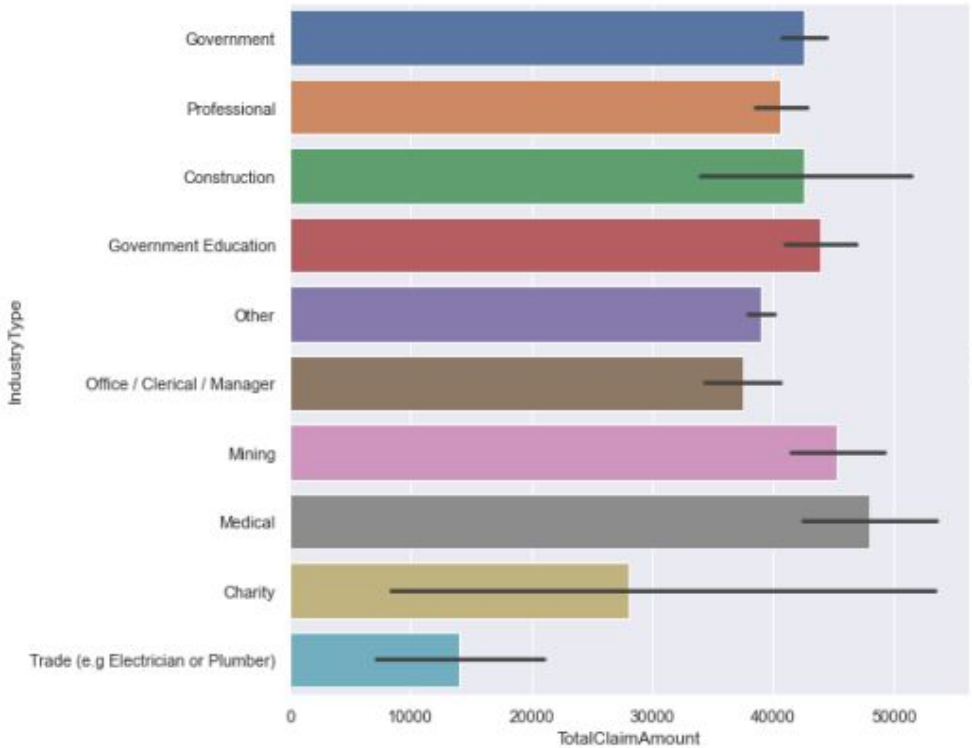
Claim amount (excluding zero claims)

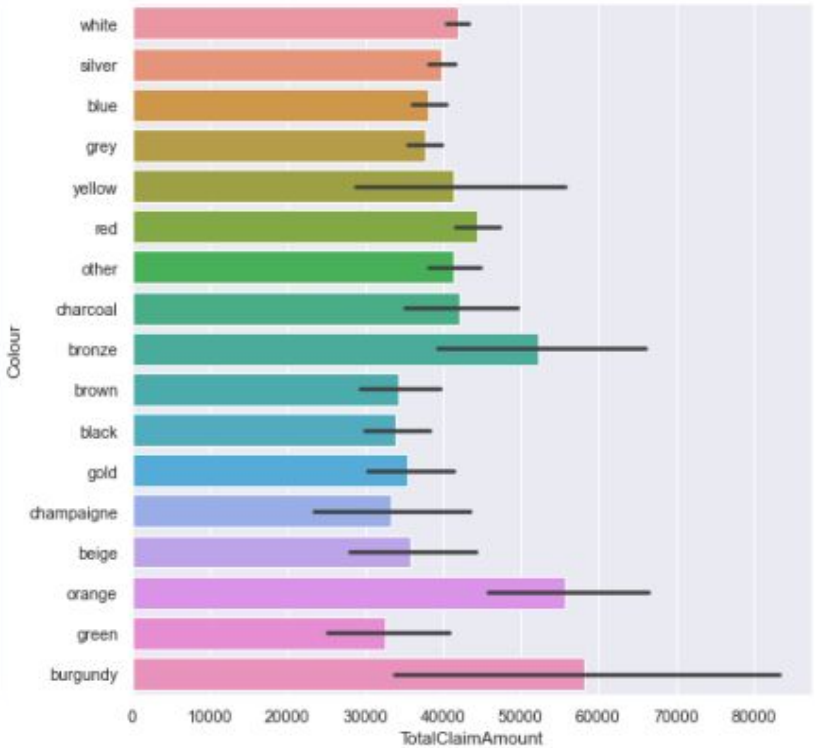


Claim count **distribution**

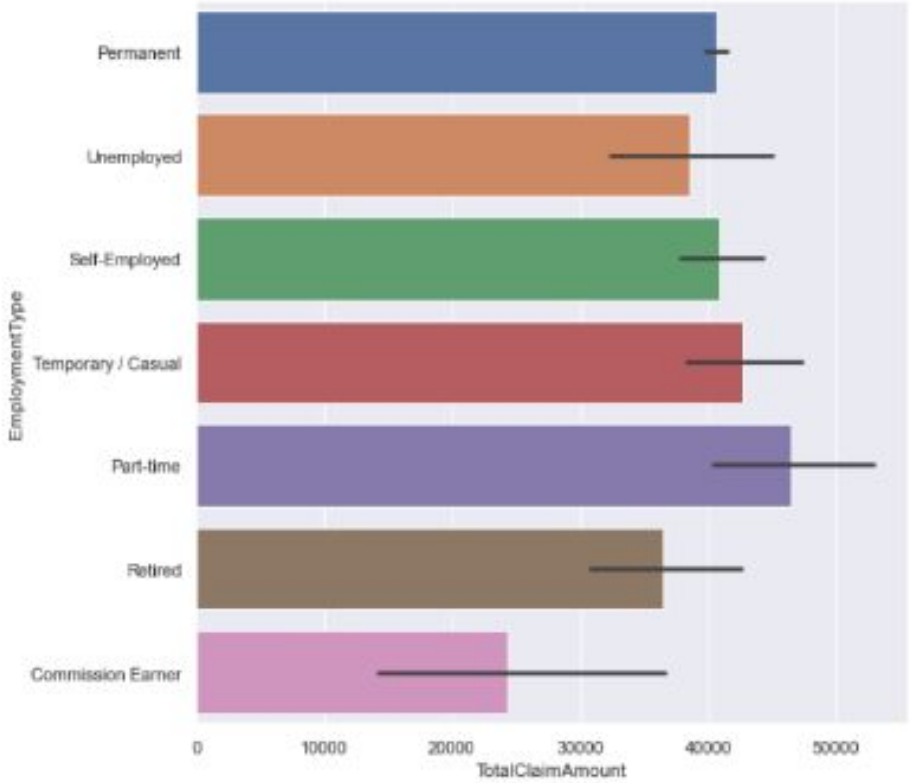


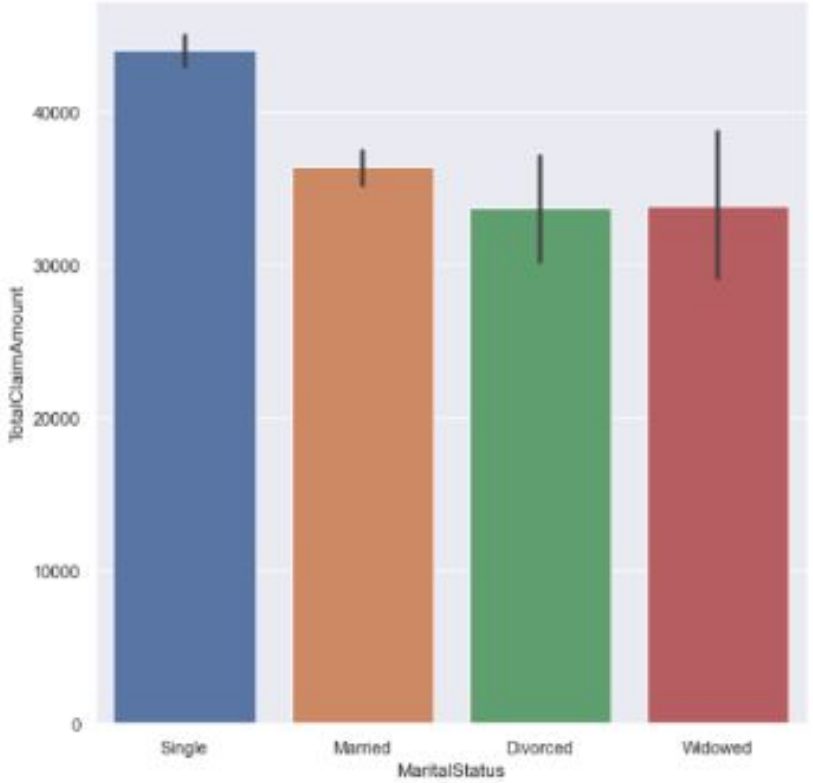
Claim amount **per category**





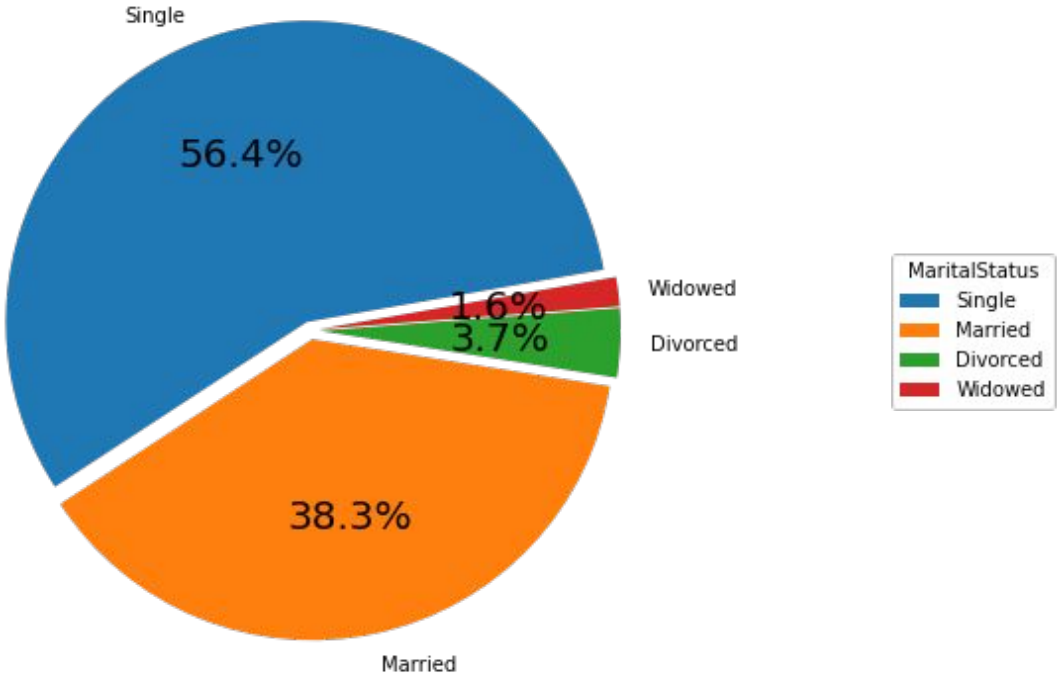
Claim amount **per category**





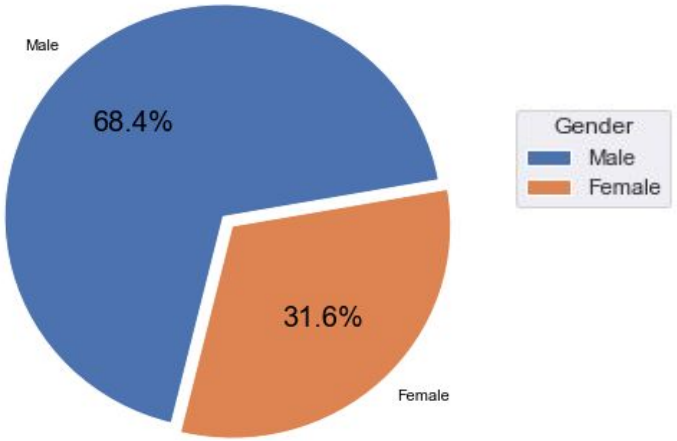
Categorical values **claim count differences**

Percentage claims per marital status

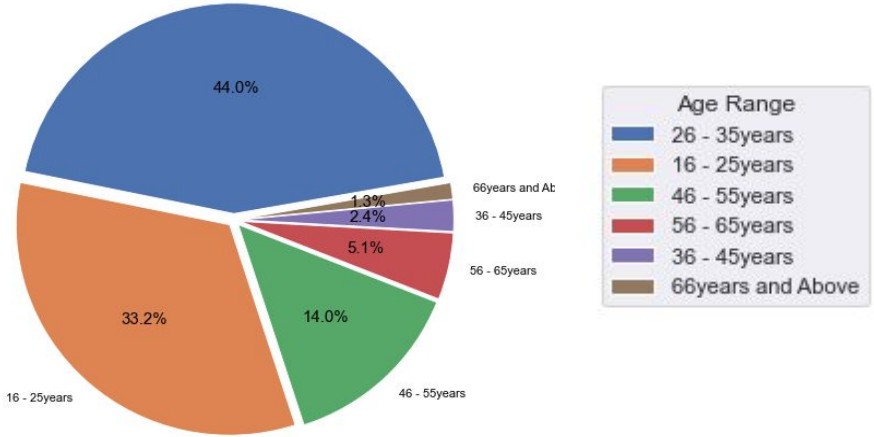


Claim count differences

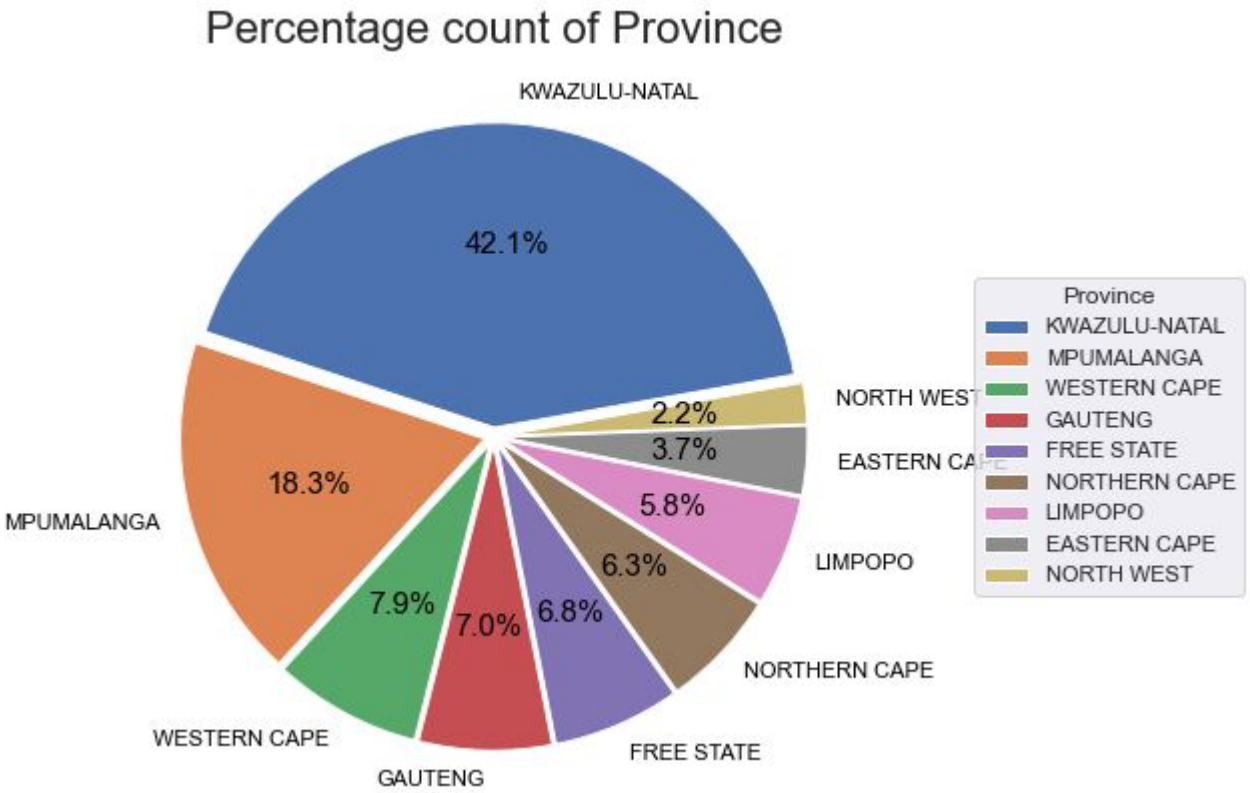
Percentage claims per Gender



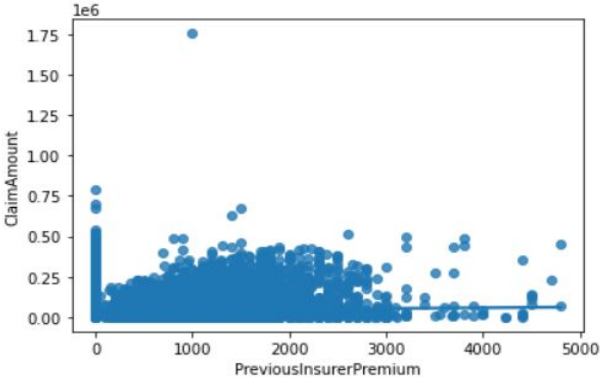
Percentage claims per Age Range

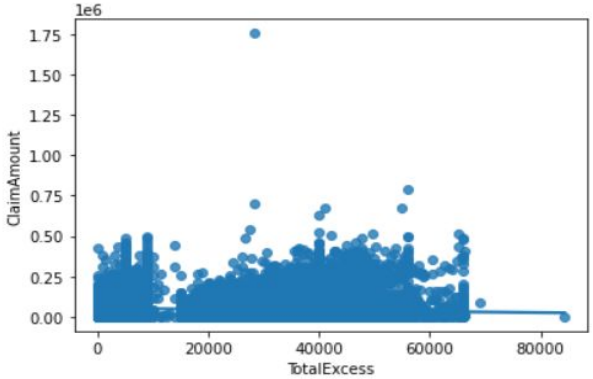


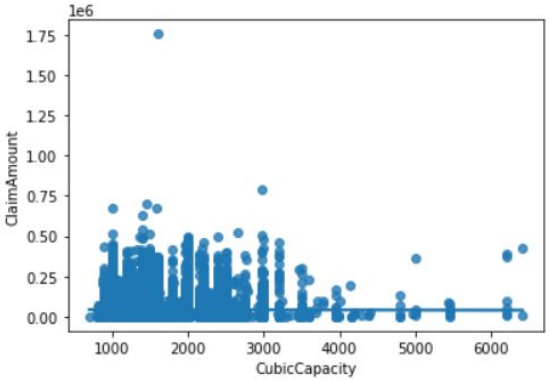
Claim count differences

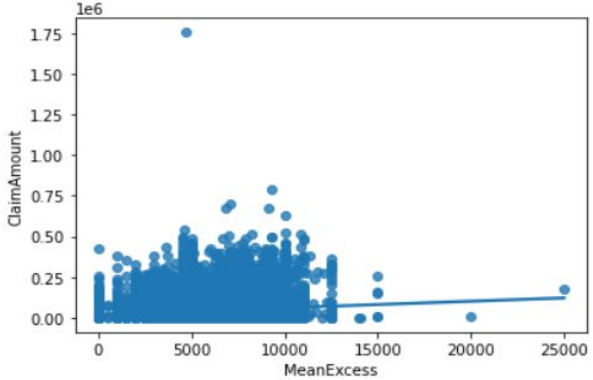


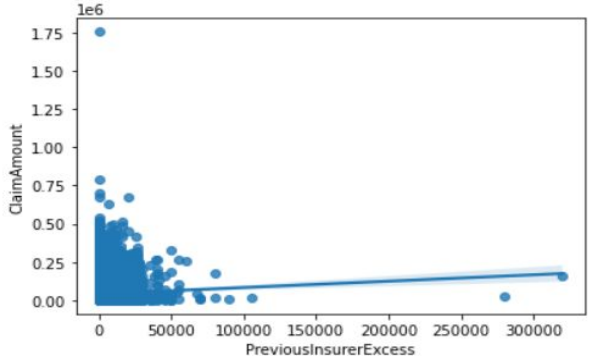
Numerical values claim **amount relationship**

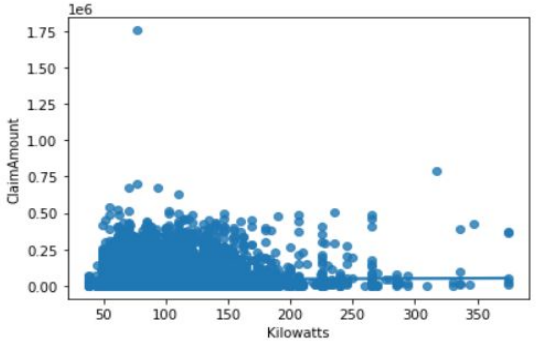




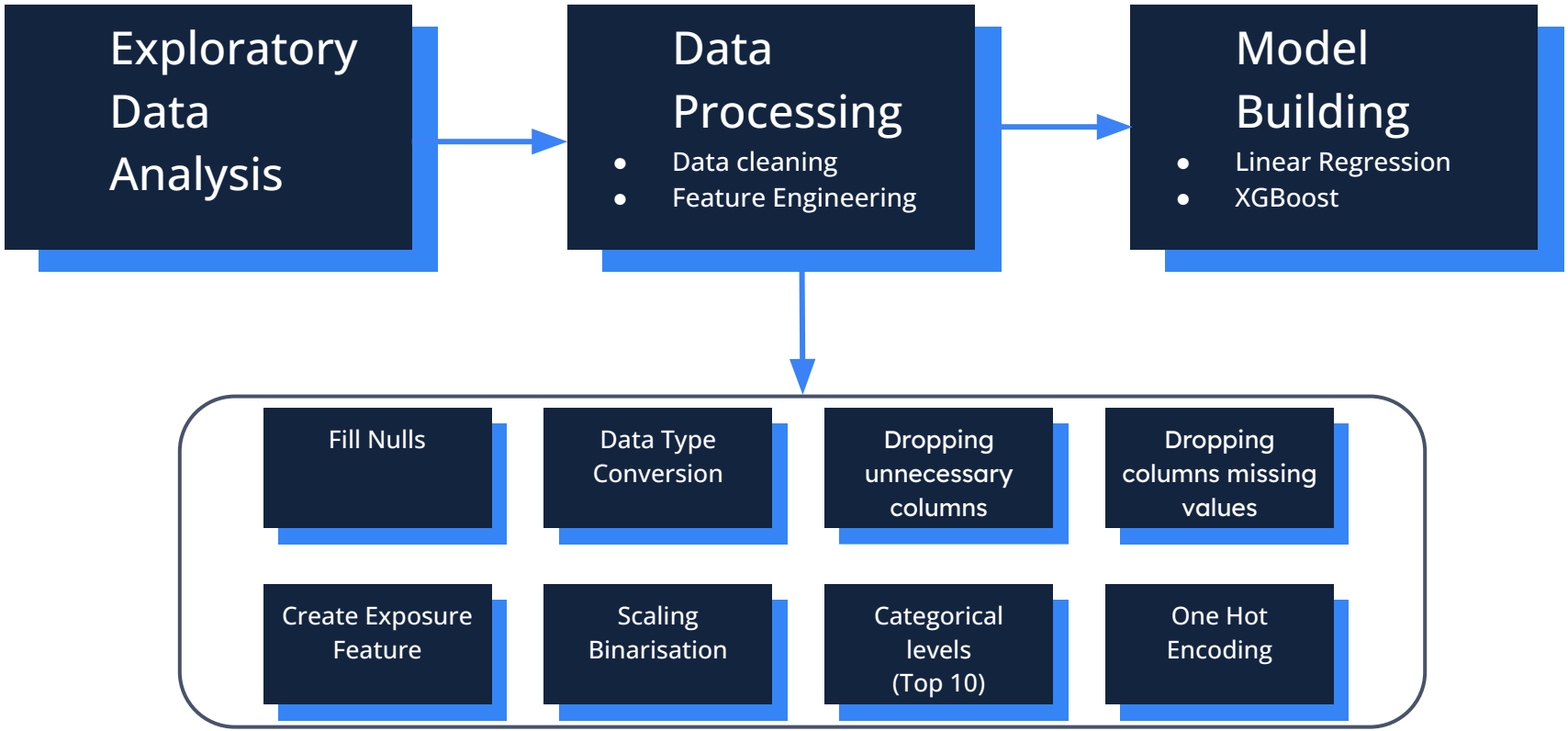








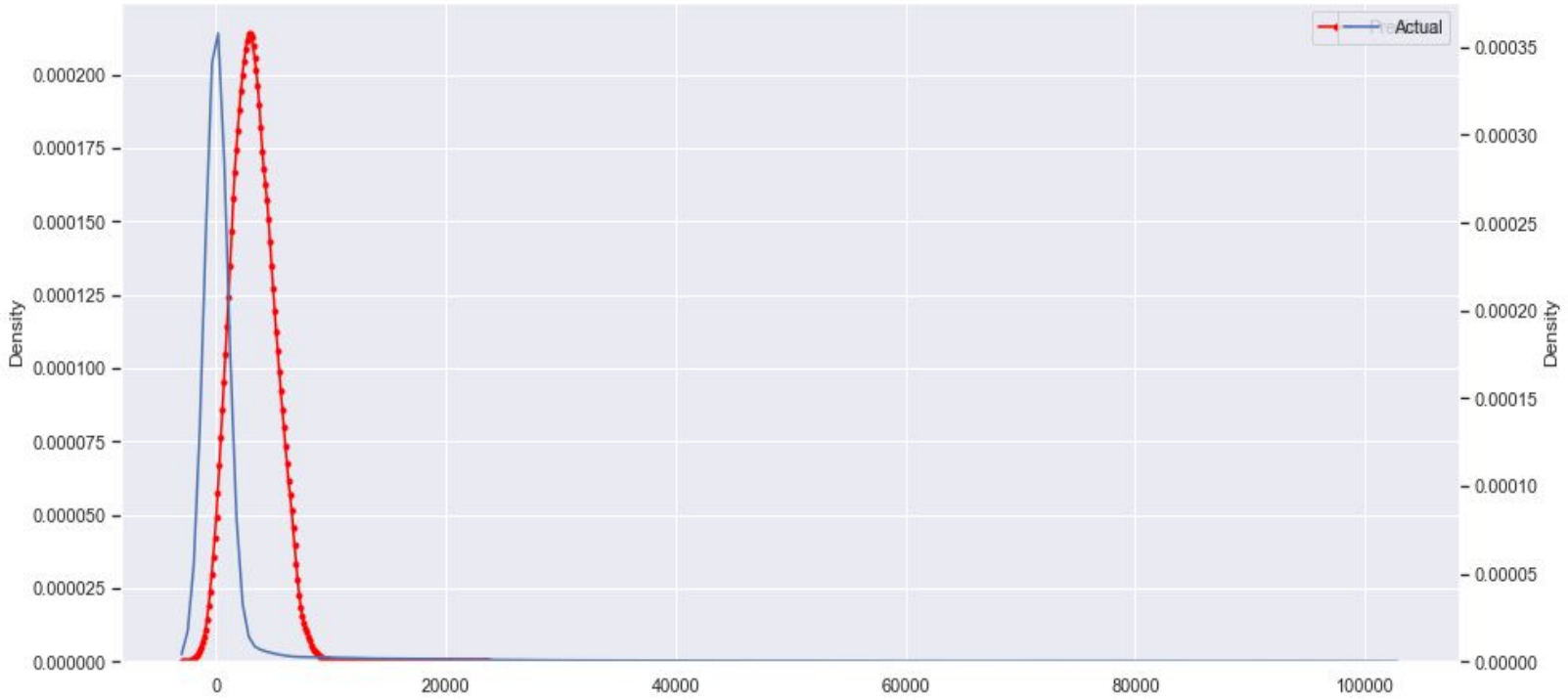
High Level Overview of **Modeling Process**



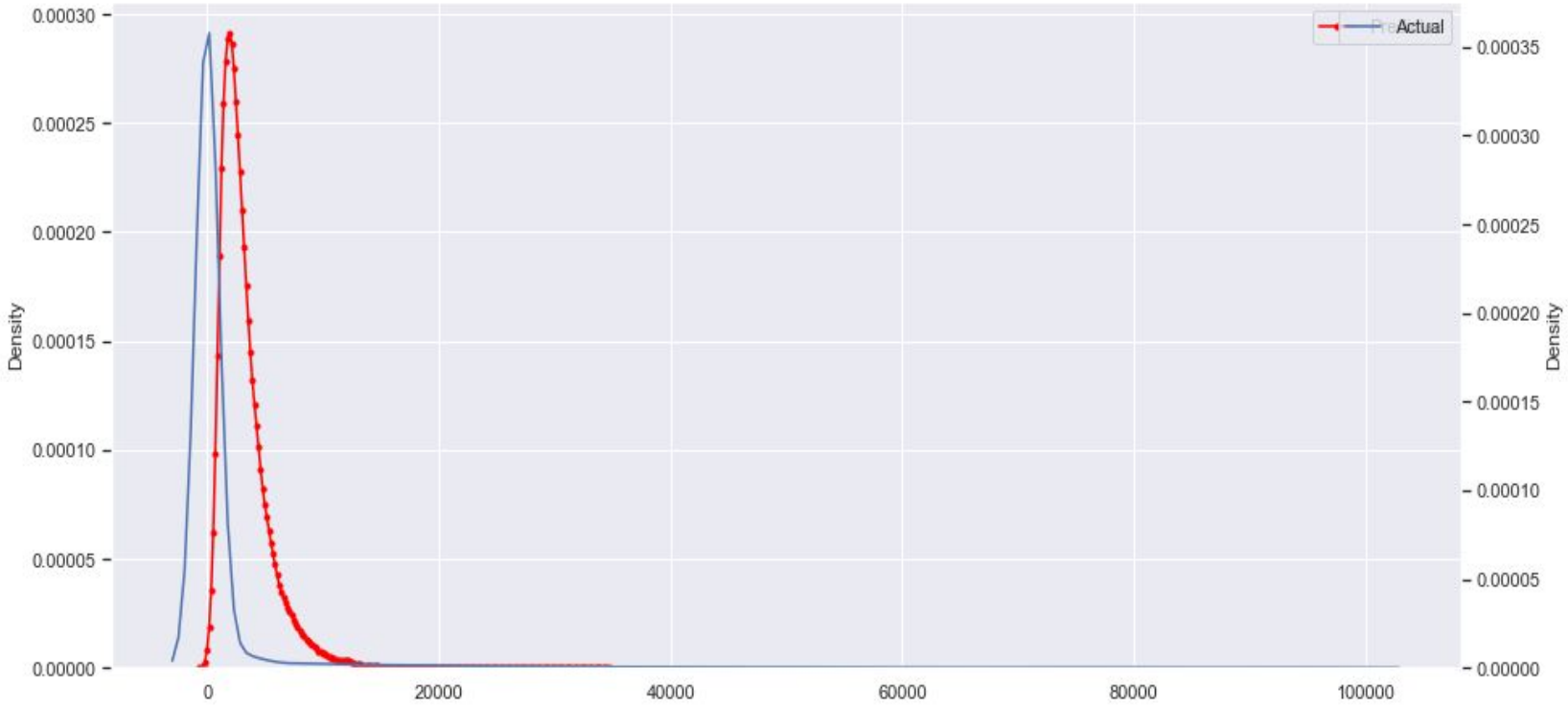
High Level Overview of **Modeling Metrics**

	Test MAE	Test RMSE
Linear Regression	5938	20035
XGBoost Regressor	5919	20003
FS Team	7400	24000

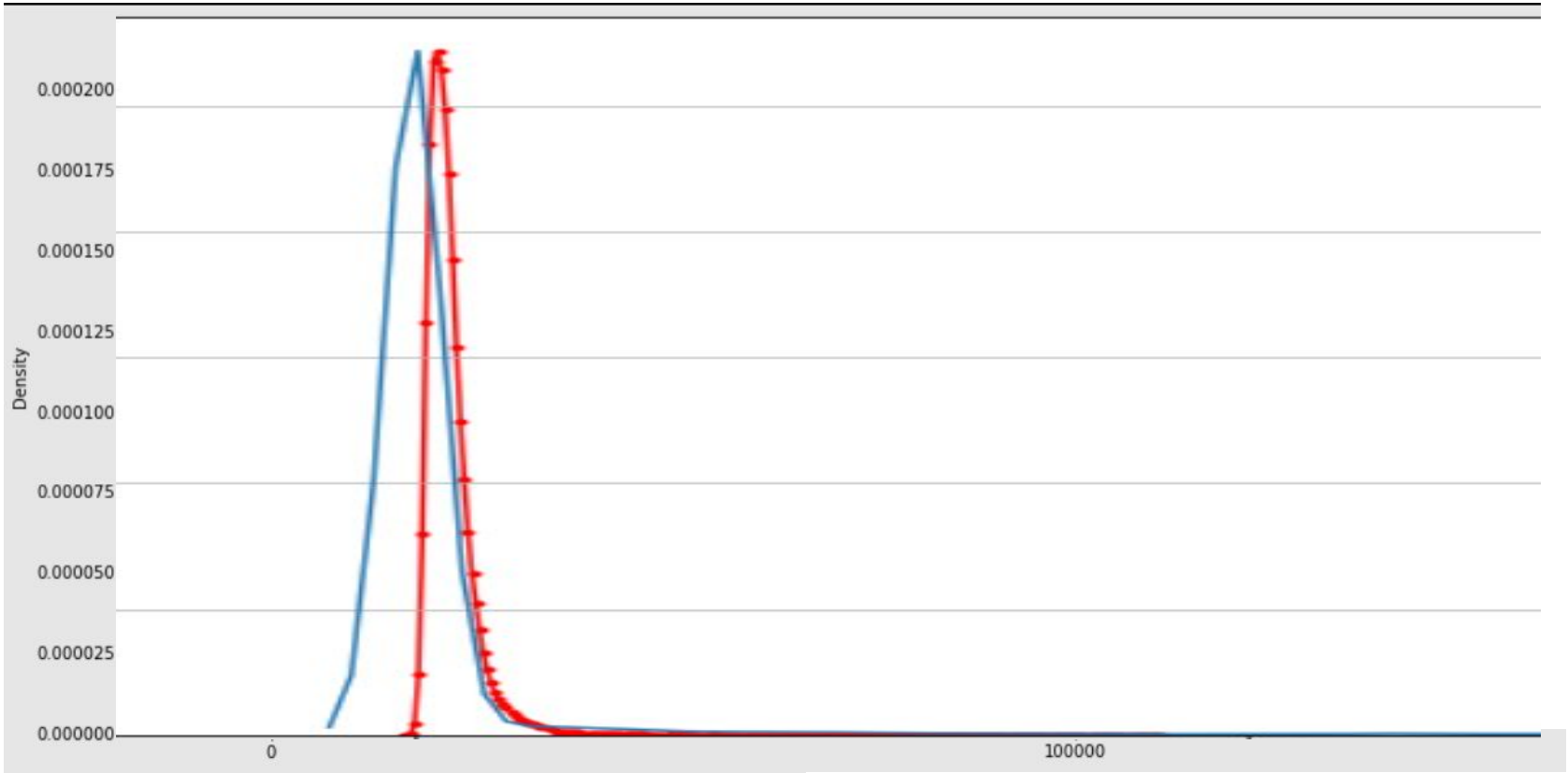
Linear Model **Predicted (Blue) vs Actual (Red)**



XGBoost Regressor **Predicted (Blue) vs Actual (Red)**



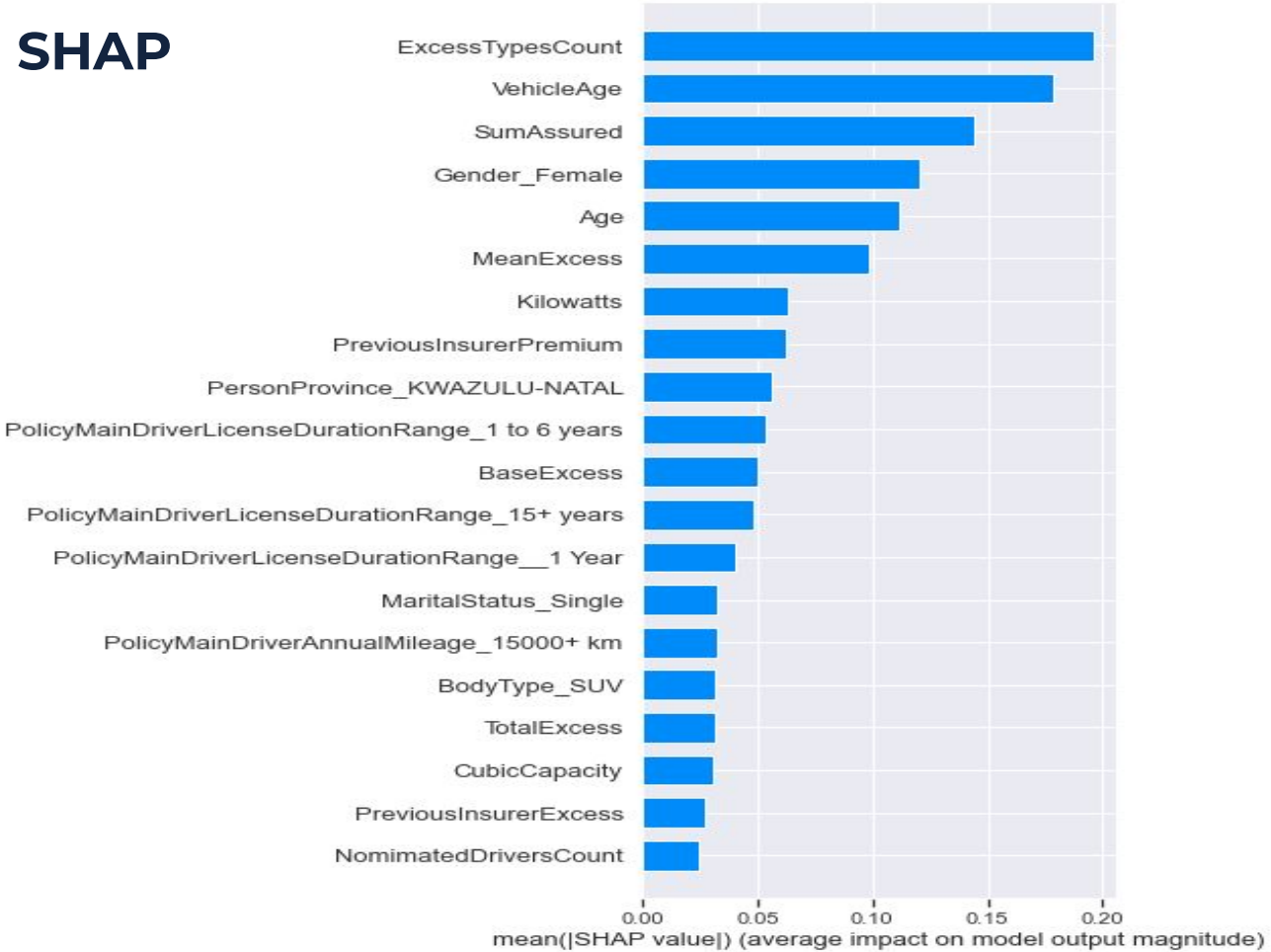
XGBoost Regressor **Predicted (Blue) vs Actual (Red)**



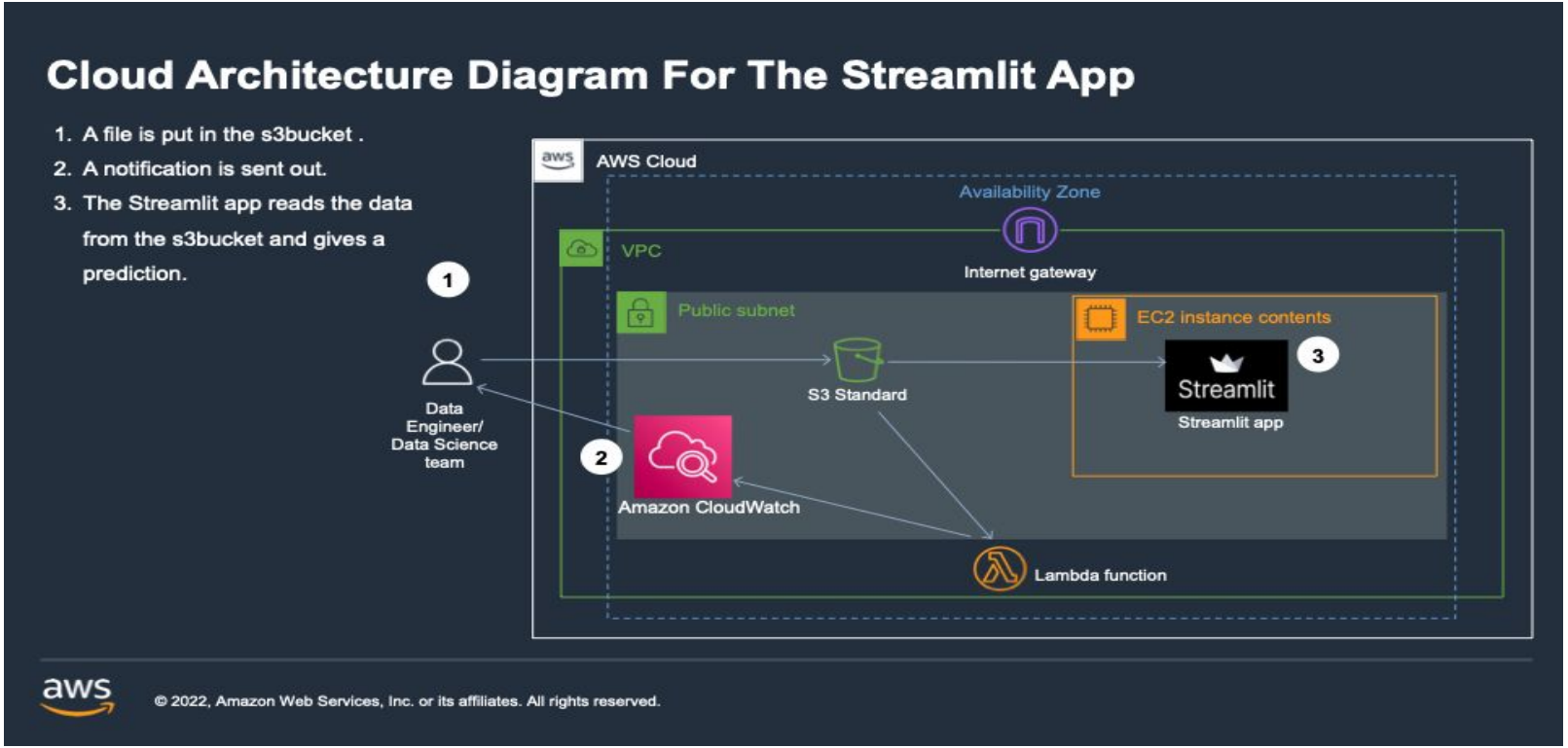
What has been **achieved by FS**

	Average Predicted Claim	Average Actual Claim	Average Difference
FS Team	**3.9K** Testing set	**4.5K** Testing set	**-600** Testing set
Team 15	**3.2K** Testing set	**3.4K** Testing set	**-200** Testing set

Best features **using SHAP**



AWS **Architecture**



Application **Demo**

Conclusion

01 Hypothesis Validation

02 Recommendations for Enhancement

EXPLORE AI

Thank you