LINEAR REGRESSION MODEL FOR AN AUTO-INSURANCE

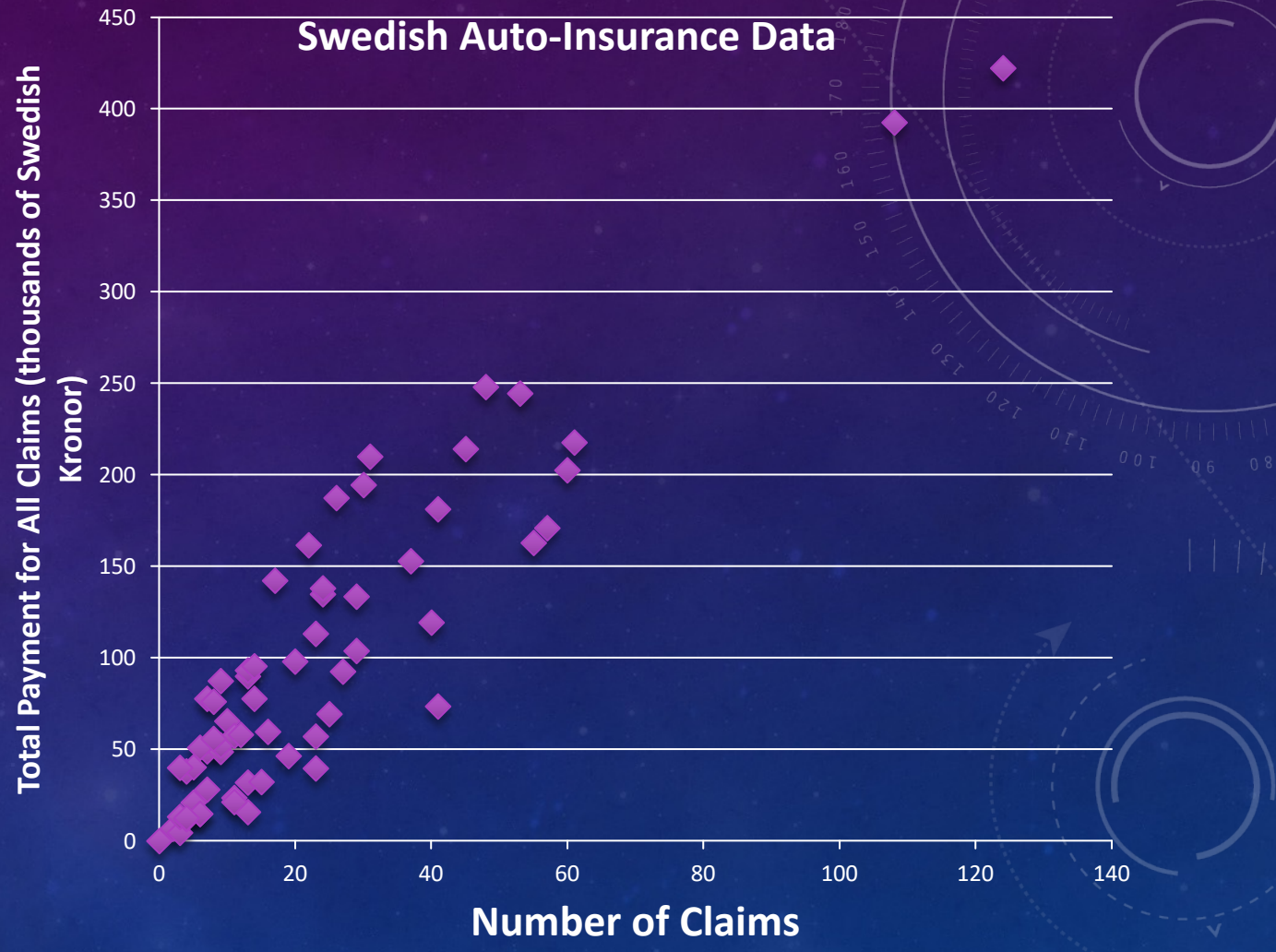
BY: DIEGO A. RODRIGUEZ MORENO

CONTENTS

- About The Data
- Methodology
- Linear Regression Model
- Results
- Understanding the results
- Conclusions, recommendations, possible errors

SWEDISH AUTO-INSURANCE DATA

- The Swedish Auto Insurance Dataset involves predicting the total payment for all claims in thousands of Swedish Kronor, given the total number of claims.
- It is comprised of 63 observations with 1 input variable and 1 output variable. The variable names are as follows:
  1. Number of claims.
  2. Total payment for all claims in thousands of Swedish Kronor.



LINEAR REGRESSION MODEL

For our regression model, we are only using the standard python library.

The model assumes a 1 to 1 function. That is: there is only one input variable and one output variable.

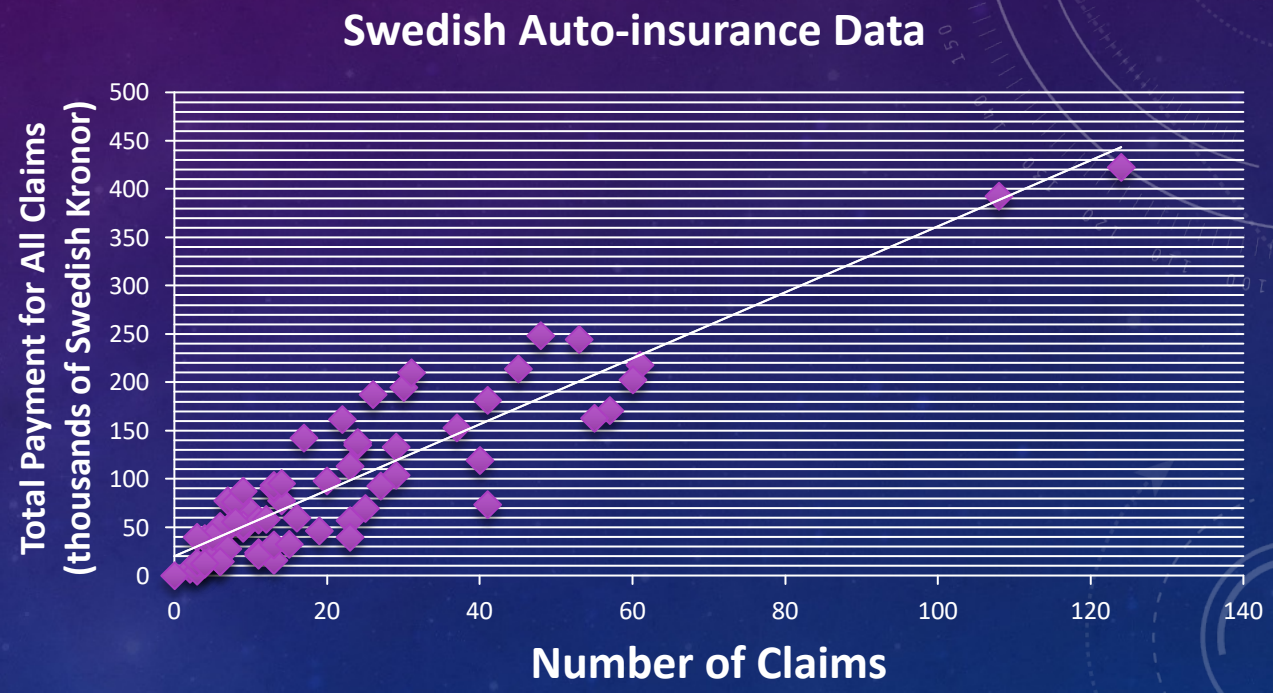
A training dataset of 60% of the data is used to prepare the model and predictions are made on the remaining 40%.

The model returns the coefficients for the linear regression model, the predicted output values, the mean of these values, and the root mean squared error from these values.

From excel, we can generate the line of best fit.

RESULTS FROM THE LINEAR MODEL

Coefficients	Mean value from the predictions	Root mean squared error
B0=28.229, B1=3.318	81.310	33.630



INTERPRETATION OF THE RESULTS

From our model, we obtain a mean value of 81 thousands of Swedish Kronor. That is the average total payment from all claims is 81 thousands of Kronor.

From our graph, we notice most of the outputs are below 80 thousands of Kronor. Therefore, we use the root mean squared error as out prediction.

Thus, we predict the average payment from all claims is 33 thousands of Kronor.

CONCLUSIONS AND ACKNOWLEDGEMENTS

The average value from our prediction is likely skewed by a few very large values.

The root mean squared error takes care of the outlier values.

The model is one of many predictions. Thus, we need to be open to different predictions.

RESOURCES

- **Auto Insurance in Sweden (small dataset). *Kaggle*, https://www.kaggle.com/datasets/sunmarkil/auto-insurance-in-sweden-small-dataset**

  Accessed by 16 Jul. 2022.

- Linear Regression Model. https://github.com/dbzdiego/autolinregmodel/blob/main/simple_linear_regression_insurance.py