EASE – THE – ERROR 3.0

PROBLEM STATEMENT : Create a machine learning model that can analyze customer behavior and predict their likelihood to purchase a product or service

TEAM NAME : INFINITY HACKERS

ASHWIN M (TEAM LEADER)
HARINI J
HARINI S
JOHN JEHIEL E

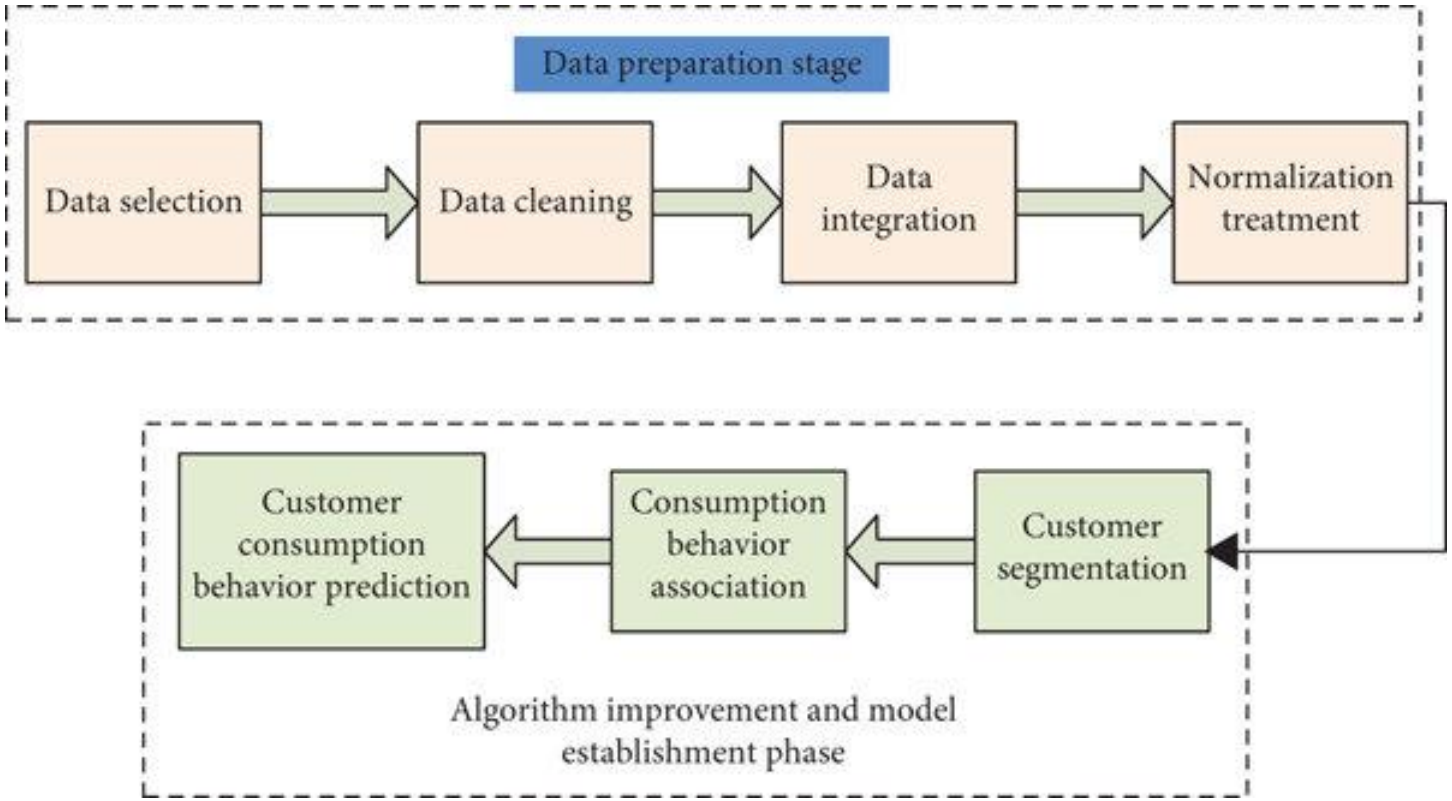
SOLUTION

- To focus on developing a model that can understand customer behavior patterns and use that information to predict whether a customer is likely to make a purchase or not.
- The goal is to utilize historical data on customer behavior and other relevant features to build a predictive model
- The model will be trained on this data and will learn patterns and relationships between customer behavior and their purchase decisions
- Once trained, the model can be used to predict the likelihood of a new customer making a purchase based on their observed behavior.
- This can provide valuable insights for businesses to understand and cater to customer preferences, optimize marketing strategies, and improve customer acquisition and retention efforts.

BUSINESS MODEL

- Improved Customer Targeting
- Increased Conversion Rates
- Enhanced Customer Experience
- Optimized Product Development
- Risk Management and Fraud Detection
- Data-Driven Decision Making

WORKFLOW


Data preparation stage
Data selection
Data cleaning
Data integration
Normalization treatment
Customer consumption behavior prediction
Consumption behavior association
Customer segmentation
Algorithm improvement and model establishment phase

EXPLORATORY DATA ANALYSIS (EDA)

SPENDING SCORE

The spending score is typically used to measure a customer's purchasing behavior and their potential value t o a business.

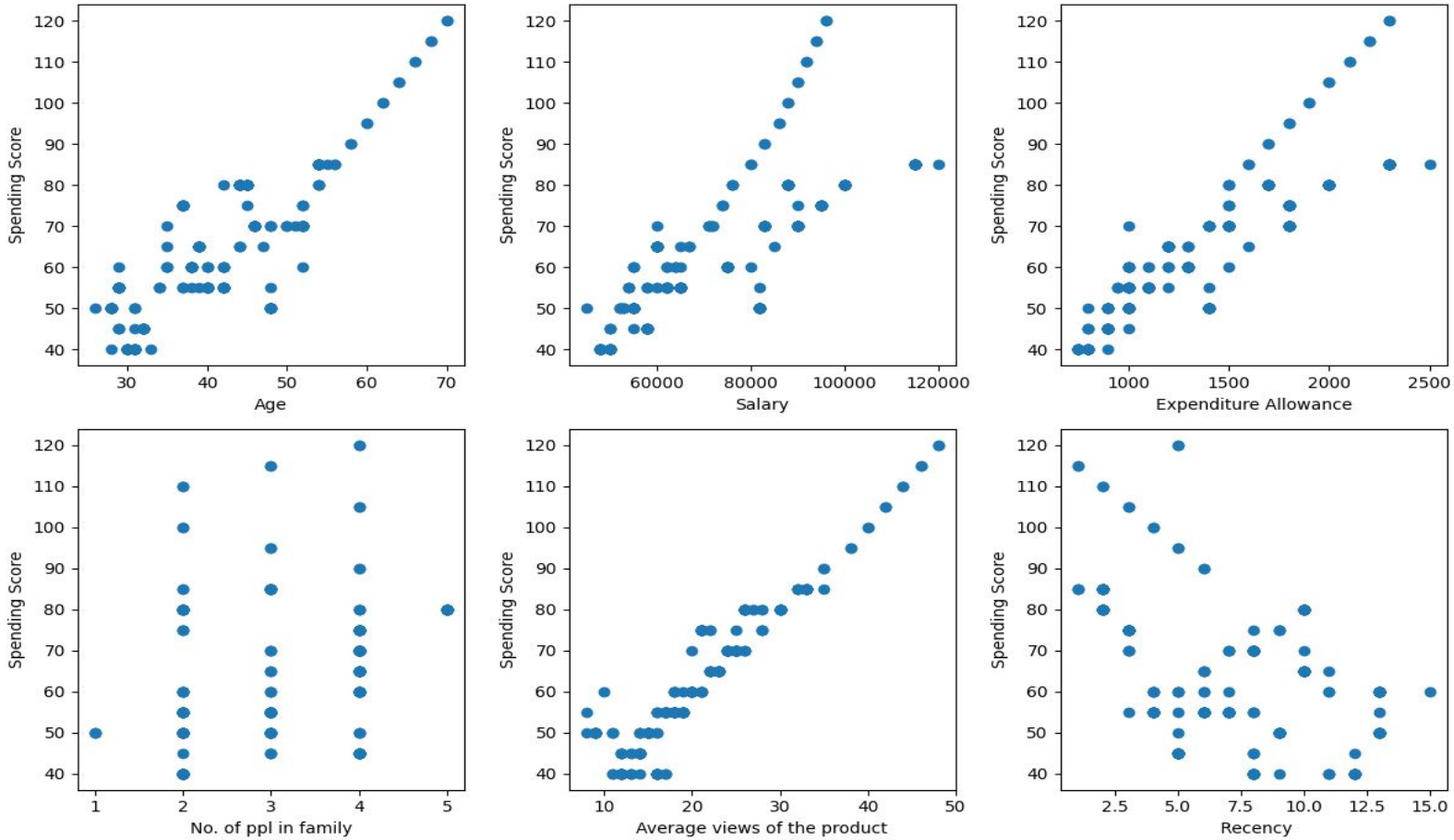
The commonly used method to calculate spending score is the RFM model, which stands for **Recency, Frequency, and Monetary Value**

1. **Recency (R):** Determine the most recent transaction date for each customer.

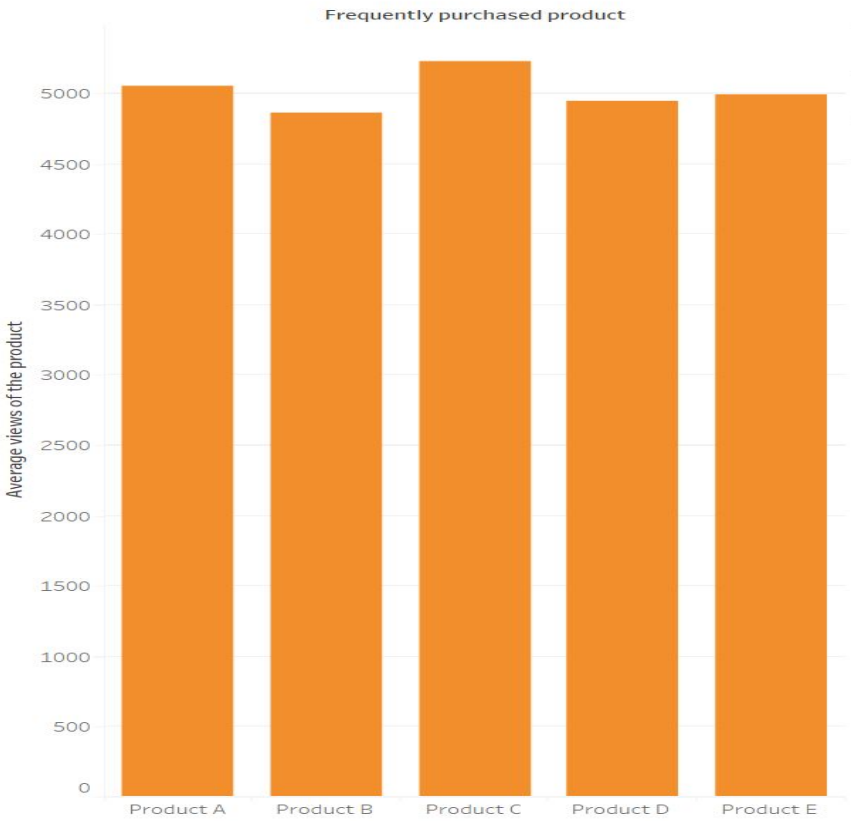
2. **Frequency (F):** Count the total number of transactions for each customer within a specific time period.

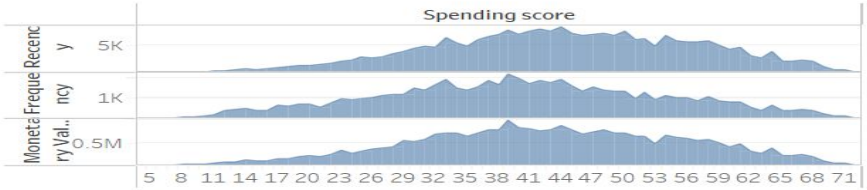
3. **Monetary Value (M):** Calculate the total monetary value of all transactions for each customer within a specific time period.

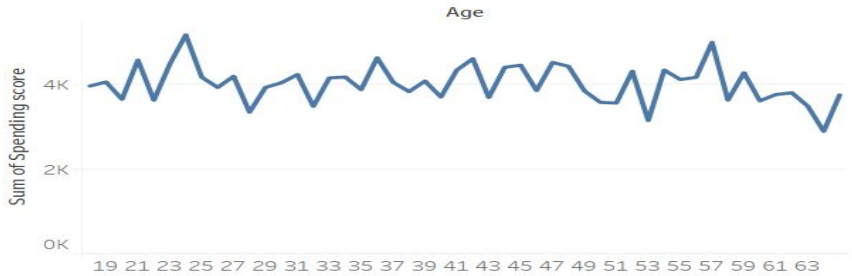
FEATURE SELECTION

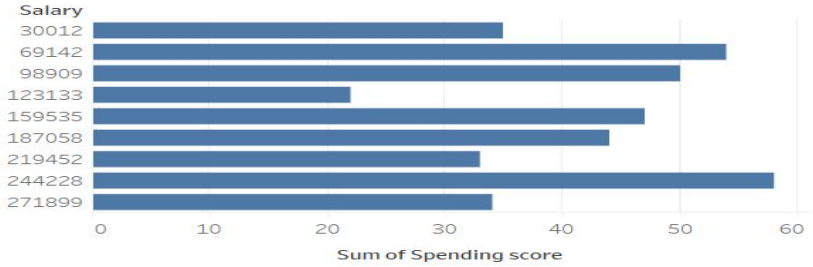


DASHBOARD USING TABLEAU









INFERENCE

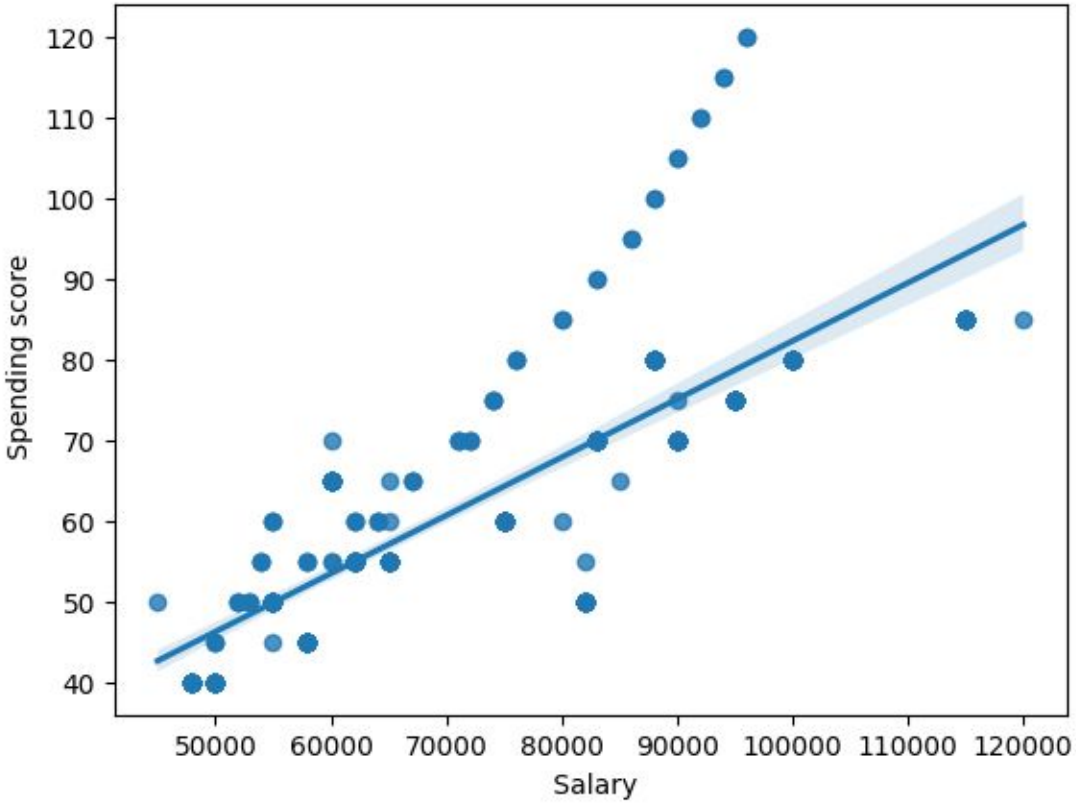
- We have analysed the various features of the dataset
- We have found that Age, Salary, Expenditure Allowance and No.of views per product are linearly dependent on the target label “Spending Score”
- Whereas Number of people in the family have no effect on the spending score

PRELIMINARY PREDICTION

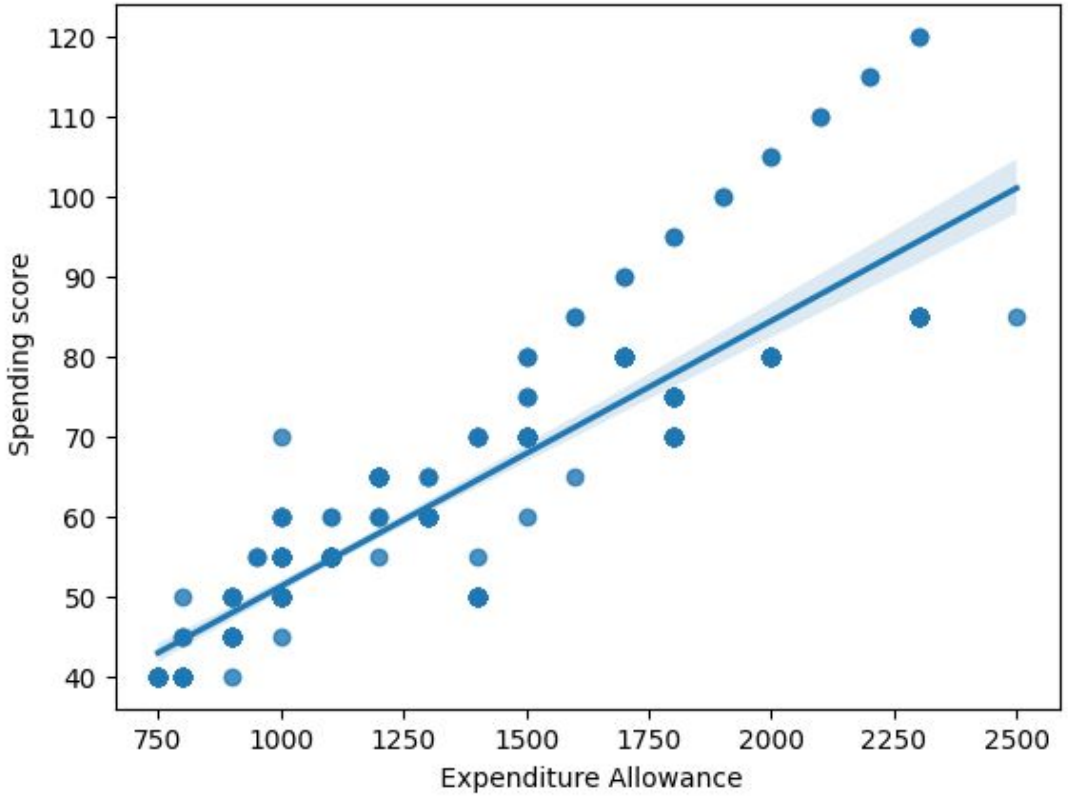
MULTI LINEAR REGRESSION

From the inference, we see that, there are 4 attributes which are linearly related to “Spending Score”. Hence we created a machine learning model using Multi Linear Regression algorithm.

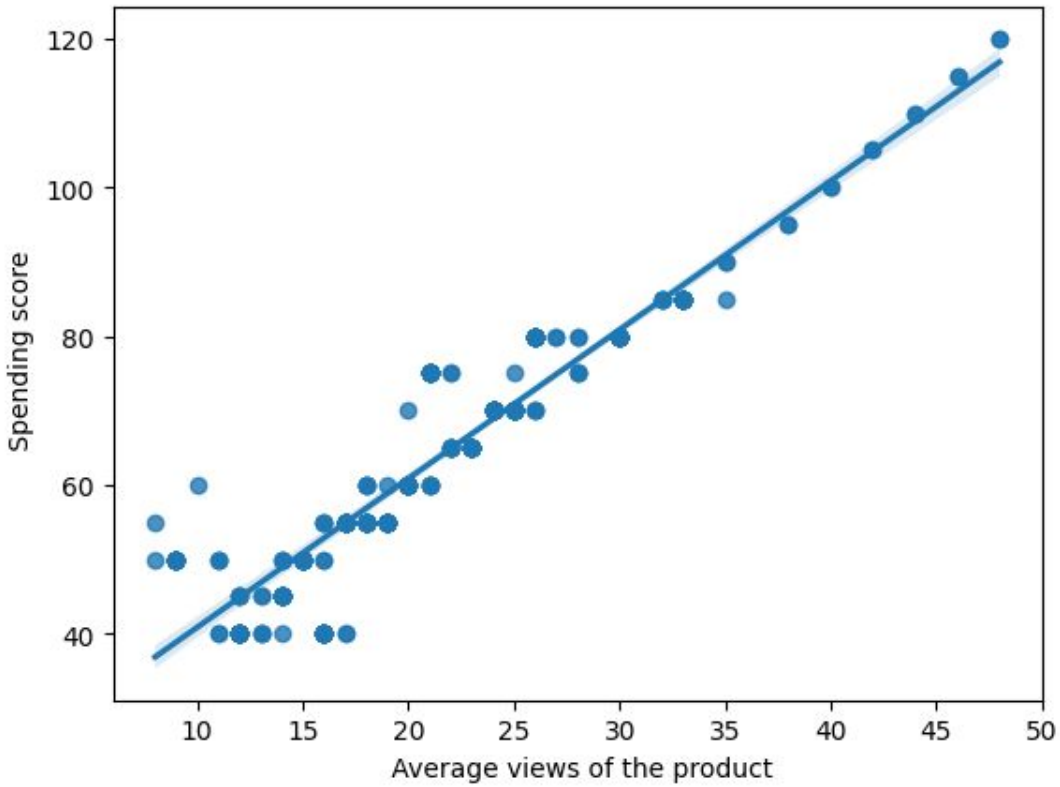
REGRESSION PLOTS



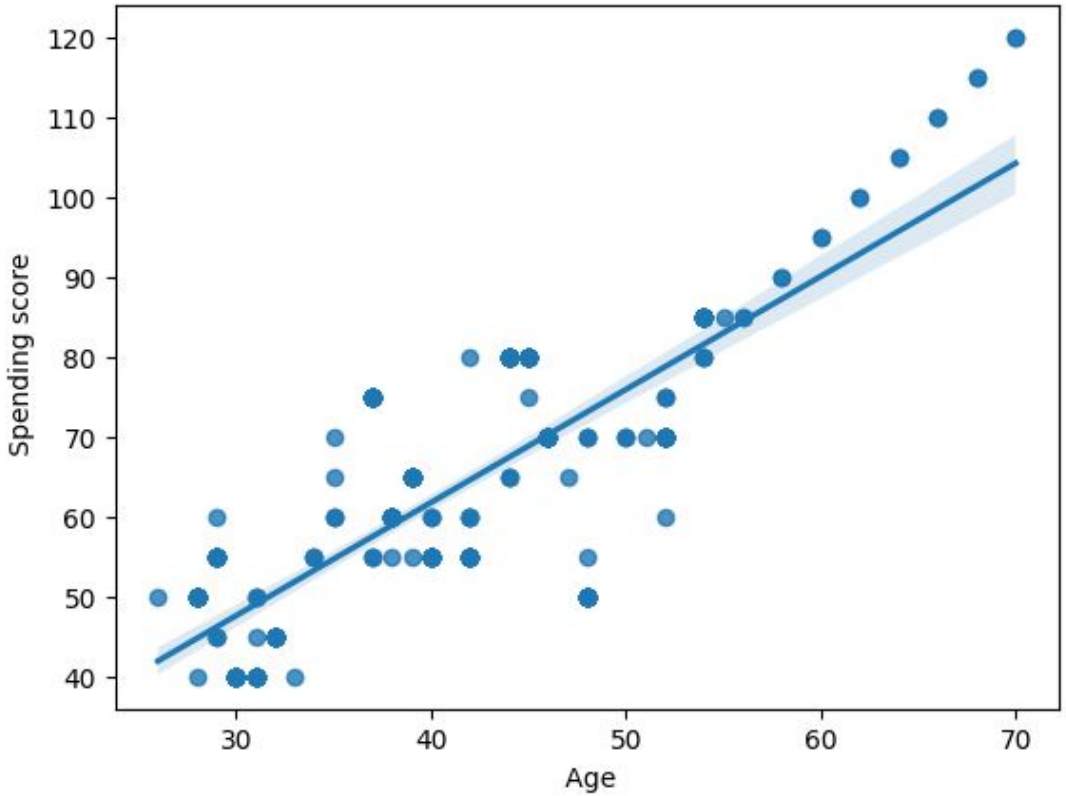
- The graph shows that the linear regression model predicted the linear relationship between Salary and Spending score similar to the previous graphs.
- Hence the choice of Salary as a prediction feature has helped to provide accuracy to the model.
- Hence, the chosen feature have proved to be useful for prediction of spending scores.



- The graph shows that the linear regression model predicted the linear relationship between Expenditure Allowance and Spending score similar to the previous graphs.
- Hence the choice of Expenditure Allowance as a prediction feature has helped to provide accuracy to the model.
- Hence, the chosen feature have proved to be useful for prediction o spending scores.



- The graph shows that the linear regression model predicted the linear relationship between No.of views per product and Spending score similar to the previous graph.
- Hence the choice of No.of views pe product as a prediction feature has helped to provide accuracy to the model.
- Hence, the chosen feature have proved to be useful for prediction spending scores.



- The graph shows that the linear regression model predicted the linear relationship between Age an Spending score similar to the previous graphs.
- Hence the choice of Age as a prediction feature has helped to provide accuracy to the model.
- Hence, the chosen feature have proved to be useful for prediction spending scores.

FAILURE OF LINEAR REGRESSION

- Accurate prediction is not possible when **large number of customers** are analysed as real world attributes are not necessarily in linear relationship.
- **Feature Engineering** is time-consuming because we need to select and engineer the features manually.
- It fails to extract **meaningful patterns** and representations from raw data while handling high-dimensional and big data.
- It does not contain activation functions, regularization techniques and **optimization algorithms.**

FINAL PREDICTION

NEURAL NETWORKS!

In order to implement a more efficient model to be used for non-linear mappings, complex relationships and scalable outputs, we use NEURAL NETWORKS with a total of 6 layers including the input and output layers.

THANK YOU