LEAD SCORE CASE STUDY

Pramod Bhaskar

Karthik Ramadass

Vishwas Yadav

AGENDA

Topic one

Topic two

Topic three

Topic four

Business Objective

An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.

The company markets its courses on several websites and search engines like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals. Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.

Build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads. A higher score would mean that the lead is hot, i.e. is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted.

HIGH LEVEL APPROACH USING LOGISTIC REGRESSION ALGORITHM

A) Data Extraction and Cleansing
- ❑ Handle NaN, duplicate values
- ❑ Handle Outliers, Impute missing values

B) Data Preparation
- ❑ Convert Binary variable (Y/N) to 0's and 1's
- ❑ Create Dummy variable for multiple categorical values
- ❑ Dropping the extra dummy variables

C) Test and Train Split

D) Feature Scaling

E) Identify Correlation through heatmap and acting upon

F) Model Building

G) Feature selection using RFE
- ❑ Assess the Stats model
- ❑ Handle high p values and VIF (to handle multicollinearity)

H) Metrics on Confusion Matrix

I) Plotting the ROC curve

J) Finding Optimal cut-off Point

K) Making Prediction on test set

L) Validation of the Model for Train and Test data

EDA (to draft)
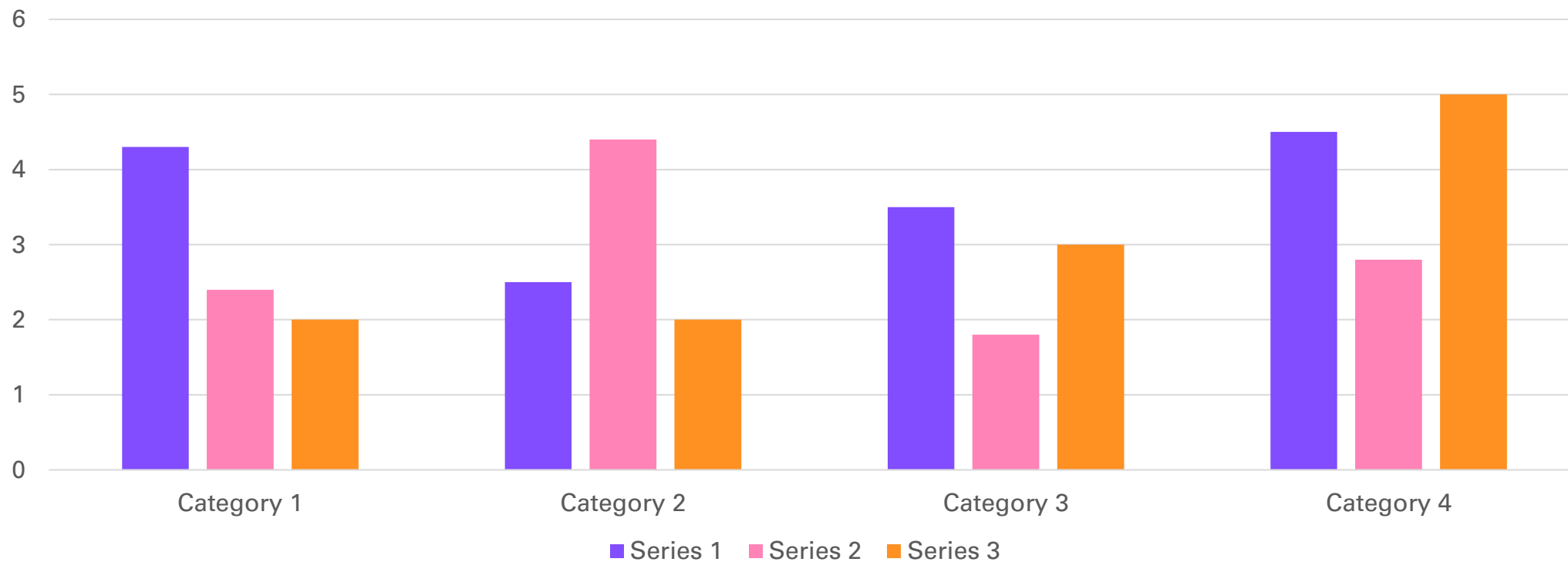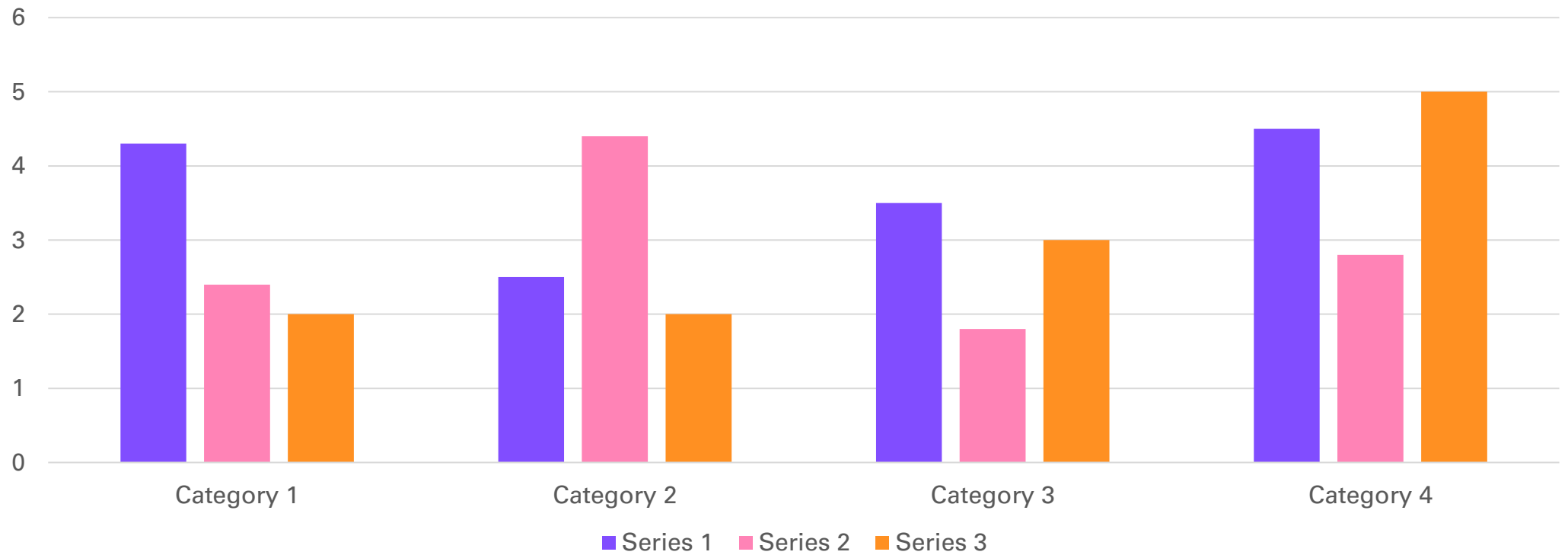
Model Building Summary

- To write up

Conclusion

- To write up

THANK YOU