**Project Selection:**

Black Friday sales analysis

550069 observations, 12 variables

**Analysis questions**

1. What effects do the masked variable have on the Purchase amount? Can we predict Purchase amount with the masked variables?(Occupation and Product categories)
2. What product categories influence the purchase amount the most? Or none?
3. What gender and age groups like to purchase more on Black Friday?
4. What variables are strongest in influencing purchasing amount?
5. Is it possible for us to narrow most purchases to a certain group of people? Such as women which are not married, men in a certain age range, etc.… E.g. multivariate analysis

**Study Framework:**

Y variables:

Purchase amount

X variables:

Product\_Category (1,2,3) (Masked)

Gender

Age

Occupation (Masked)

City\_Category

Marital\_Status

Stay\_In\_Current\_City\_Years

**Descriptive analysis**

-Distribution of Key Variables

-Summary of Key Attributes (inputs and outcomes)

**Exploratory analysis**

-Outlier Detection/ Missing Values

-Visualization

-Prediction: Regression – Prediction of the purchase amount

**Conclusion/Recommendation**

