

Project proposal for Capstone 1: Predict Black Friday Sales by User behavior

The dataset here is a sample of the transactions made in a retail store.

The store wants to know better the customer purchase behavior against different products. Specifically, here the problem is a regression problem where we are trying to predict the dependent variable (the amount of purchase) with the help of the information contained in the other variables.

In the quest to do that, answers to following question can be sought:

1. What are the characteristics of my top revenue generators?
2. Who buys the top products?
3. To what accuracy can we predict purchase amount?

Note: The dataset featured on a Kaggle competition that was made available by a user at below location:

<https://www.kaggle.com/mehdidag/black-friday>.