

Predicting Purchase Intentions

And Other Trends in User Shopping Sessions



Overview

- E-Commerce sales have jumped from \$1.3 trillion in 2014 to \$3.5 trillion in 2019.
- Share of e-commerce in retail has increased in the U.S. from 9.7% in 2014 to 16% in 2019
- The share of e-commerce in India is only 1.6%, but revenue from e-commerce is expected to grow at 51% per annum.
- Advanced analytics can ensure the company is well positioned to take advantage of the e-tail boom that will come.



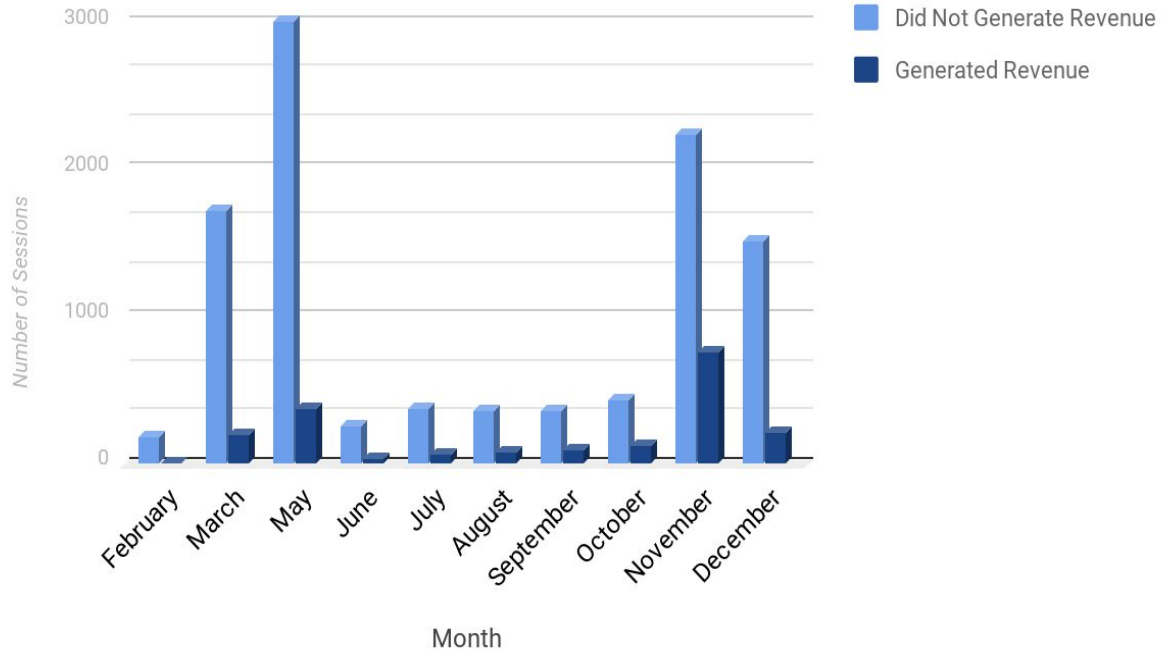
Objectives

- Discovering trends in online user shopping sessions, for better decision making.
- Predicting if the user will make a purchase, based on the characteristics of the session.

Trends in User Shopping Sessions



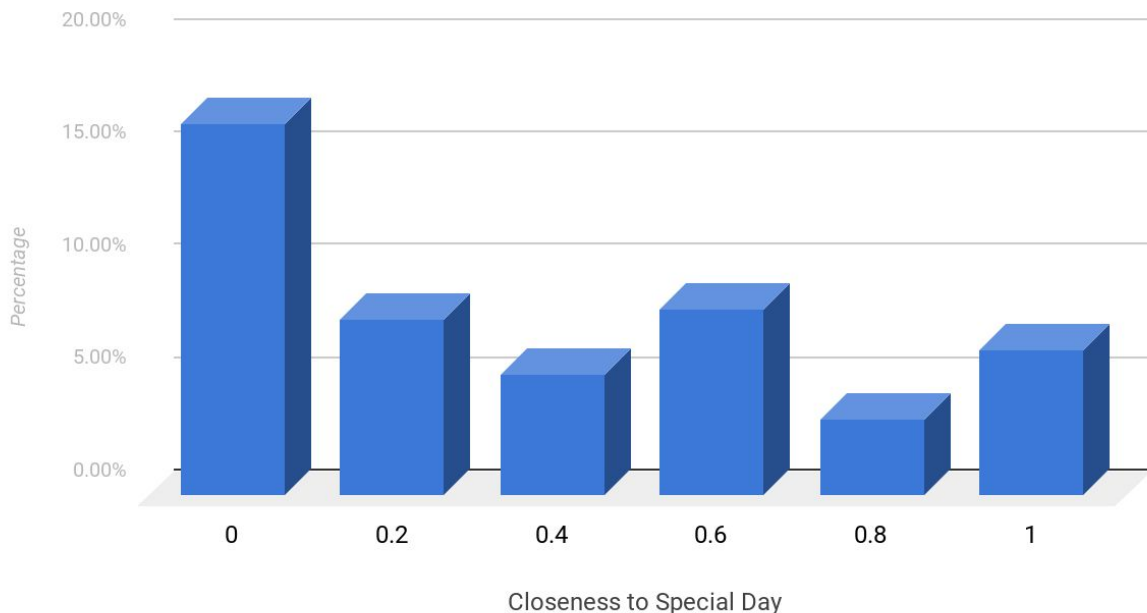
Conversions Per Month



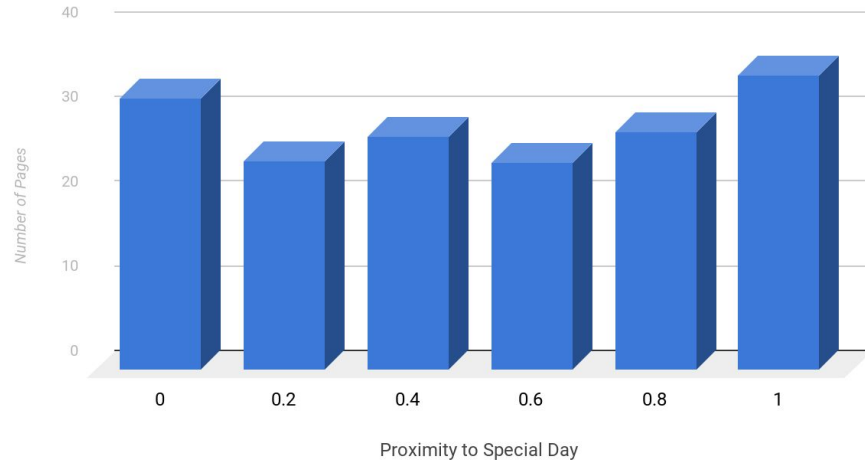
- November was the best month, with high conversion rates due to upcoming holiday period.
- May presents the best opportunity, with many visitors but few conversions.
- December was the outlier here. Sales for e-tailers during this month are usually very high, but that is not the case here. The site needs to sell more holiday related merchandise.

- Special days are days like holidays, Mother's Day etc.
- The value of the special day increases as the day nears, and gradually decreases back to 0 after it occurs.
- Sales should ideally increase near special days, but that was not the case here. Conversion rates decreased as a special day neared.
- Conversions just before and after were the lowest. A priority, one-day delivery can help change that.
- Themed products (eg. mother's day products near mother's day), and appropriate advertising can increase conversions and sales.

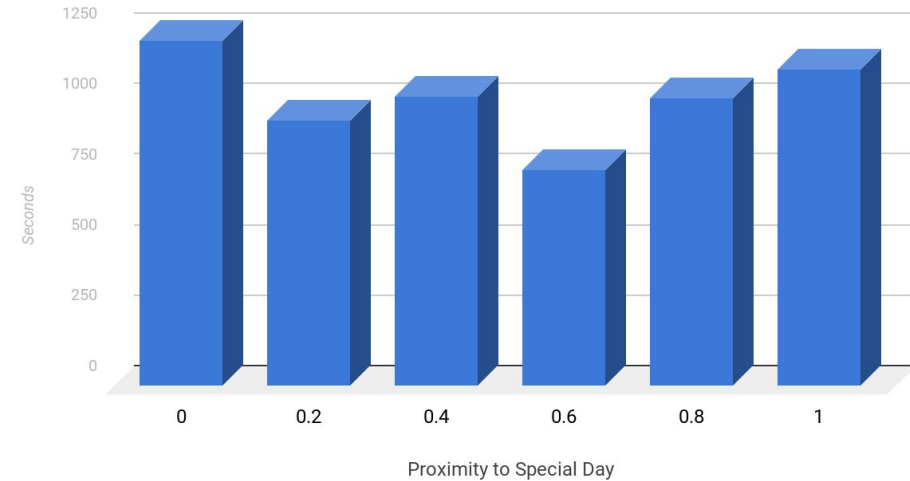
Conversion Rates - Special Days



Product Pages Visited by Special Days

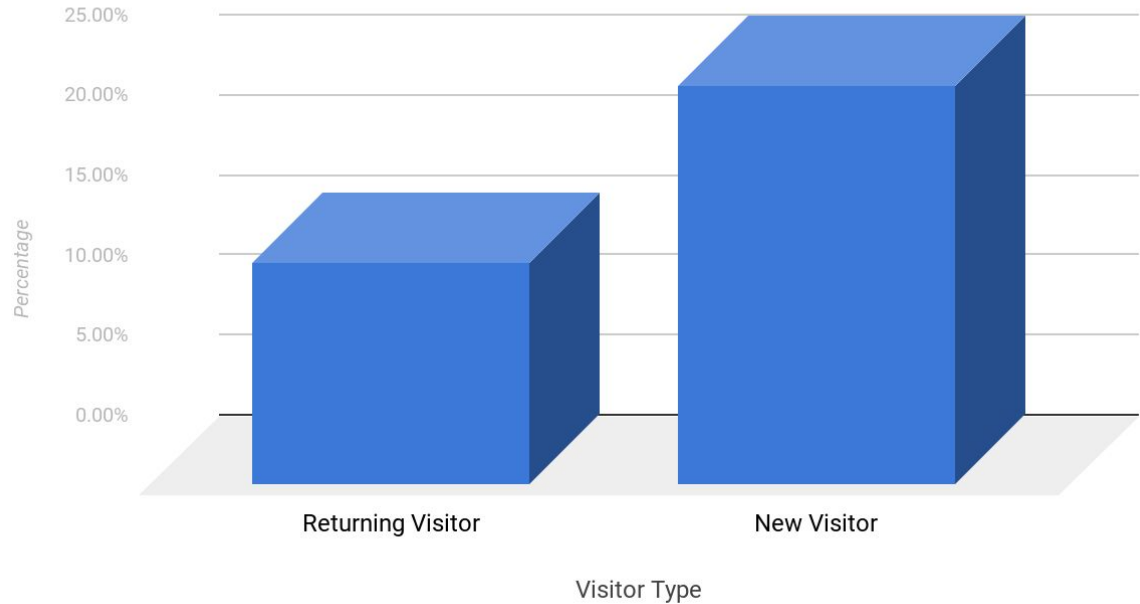


Time Spent on Product Pages (in seconds)



- Low conversion rates on special days need to be further analyzed. The chart on the left shows the number of product related pages visited, on average, per special day. The chart on the right shows the average time spent on these pages.
- The low conversion rates on special days were not due to less time spent on the site. Even on or around special days, users were still spending time on the website and visiting quite a few pages. Transactions completed, however, were low.
- Ideally, the number of pages visited and the time spent on these pages should increase as the special day nears. Marketing strategies focused on maximizing special day sales need to be created and implemented.

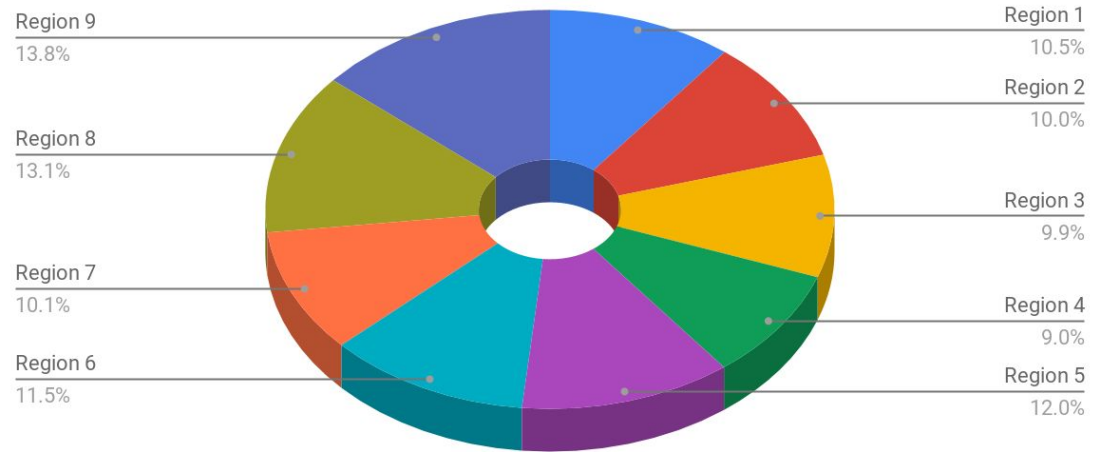
Conversion Rates - New vs Returning Visitors



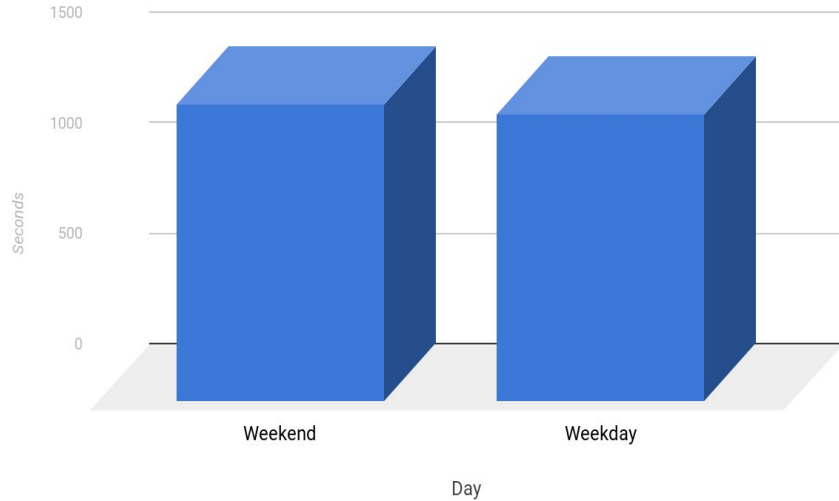
- Conversion rates of first time visitors were much higher than those of returning visitors. The main reason was that many users who signed up for the first time only did so because they already had a product to buy in mind.
- A sign-up bonus, or discount, can help improve the conversion rates of new visitors even more.

- To attract new visitors, we first need to know where they come from.
- New visitors come from almost all the regions equally. Region 4 accounts for the least share of new visitors (only 9%), while Region 9 (13.8%) accounts for the most.
- Targeted marketing in these regions, both offline and online, can help increase user base.

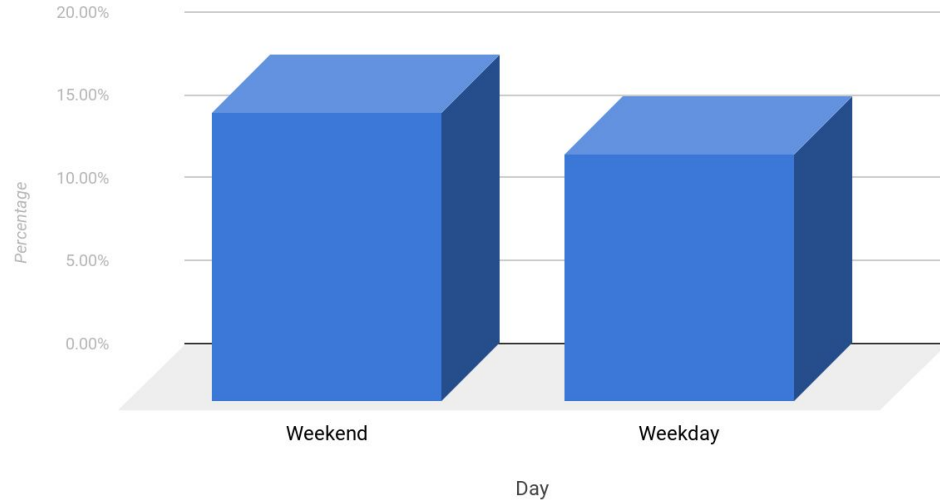
New Visitor Percentage per Region



Visit Duration - Weekends vs Weekdays



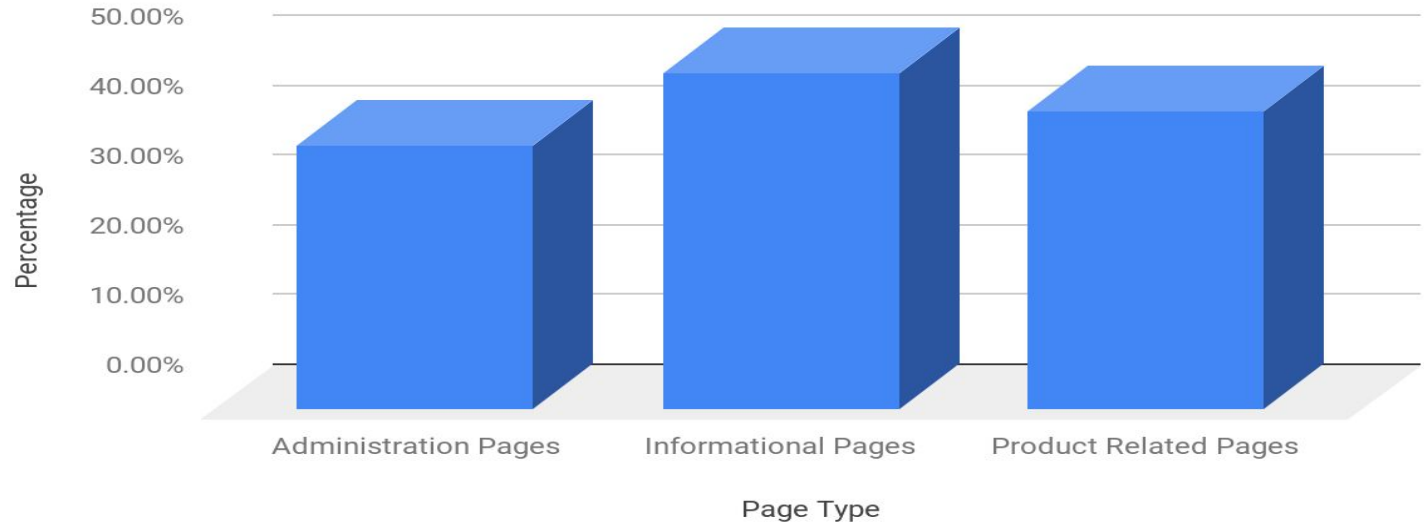
Conversion Rates - Weekday vs Weekend



- The chart on the left shows that weekday conversion rates were much lower than weekend conversion rates.
- However, users spent almost the same amount of time on the website during weekdays, so the low conversions were due to other factors. One reason could be that users usually put off decisions until weekends, when they have more time to think the purchase through.

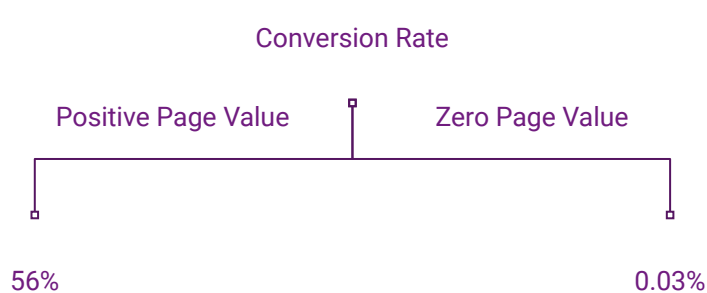
Difference in Time Spent for Revenue Generating Sessions

Extra Time Spent (in %) in Sessions Where Revenue was Generated

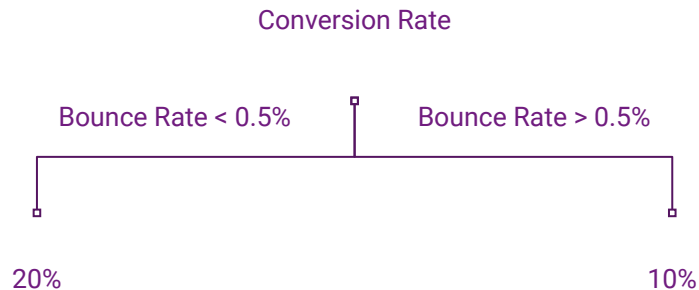


- Users who completed their transaction spent, on average, 35% more time on administration related pages, 48% more time on informational pages and 42% more time on product related pages.
- It shows that while the product pages are important, it is just as important to focus on and improve the non-product related pages. Users who complete their transactions spend more time on all supporting pages, not just the main product page.

Conversion Rates - by Page Value and Bounce Rates



- There were 9600 sessions when the user did not visit any page with a positive page value.
- Only 2730 sessions visited pages with a positive page value.
- Streamlining the website and removing redundant pages can improve the average page value and hence conversions.

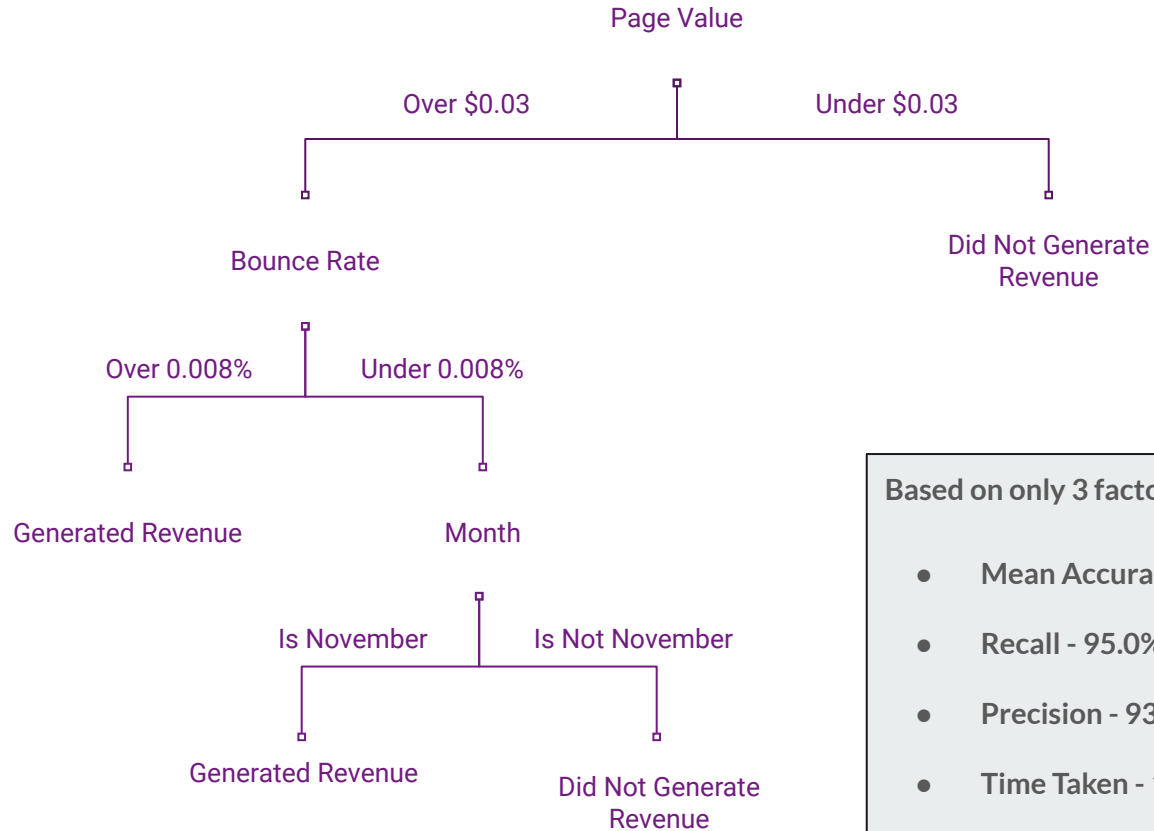


- 5555 visitors had a bounce rate of over 0.5%, with the average conversion rate for this group only 10%.
- Bounce rates can be lowered by showing *relevant* recommended products on product pages.

Using Machine Learning to Predict Shopper's Intention to Purchase



Decision Tree Model



Based on only 3 factors. Simple and easy to understand

- Mean Accuracy - 90.1%
- Recall - 95.0%
- Precision - 93.4%
- Time Taken - 1.72 seconds

Other Models Used - Logistic Regression

- Model Parameters:
 - Weights: 1:1
 - Threshold: 50%
- Advantages:
 - Fast
 - Easy to Understand
 - Can be used for factor analysis
- Model Performance
 - Mean Accuracy - 88.25%
 - Recall - 97.25%
 - Precision - 89%
 - Time Taken - 5.4 seconds

Other Models Used - Gradient Boosted Machines

- Model Parameters:
 - Learning Rate: 0.05
 - Number of Runs: 200
 - Max Depth of Individual Trees: 4
- Advantages:
 - Extremely accurate, gives excellent results
- Model Performance:
 - Mean Accuracy - 90.8%
 - Recall - 95.5%
 - Precision - 93.75%
 - Time Taken - 80 seconds

Other Models Used - Random Forest

- Model Parameters:
 - Number of Trees: 300
 - Max Depth of Individual Trees: No Limit
- Advantages:
 - Extremely accurate
 - Excellent for noisy data
- Model Performance
 - Mean Accuracy - 90.5%
 - Recall - 95.34%
 - Precision - 93.56%
 - Time Taken - 252 seconds

Conclusion

Data science can empower you to make better data driven decisions, discover new trends, identify opportunities and get actionable insights on your target audience. With the amount of data generated by your e-commerce platform, not using this data would be akin to going backwards.

The golden age of e-commerce in India is about to start. The company needs data science to take make the most of it.



Thank you.

