

## **Scenario**

Let's say you work at an online e-commerce company, abc.com, selling books in US, CA and Latin America. You are part of the marketing data analytics team. You are the Viz engineer in the team. Your teammates are John, who is a data engineer and Jyoti, a data analyst. Don, who is the marketing manager in the company, comes and tells the team that he needs some insights on conversion rates of people who visit the shopping card page.

He tells his objective is two fold:

- Improve conversion rates from shopping cart to converted sales
- Reduce the time taken to adding the item to the cart and completing the sale.
- 

John built a data model combining web log data and sales transaction data. Jyoti has been extremely busy preparing the following insights for Don.

- 1) Daily/Weekly/monthly trend of conversions from shopping cart to completed sales
- 2) Rate of folks dropping out at each stage of the shopping process
- 3) Any outliers on conversion rates for specific product types
- 4) Any correlations between promotions and the conversion rates
- 5) Average time taken to convert from shopping cart to sales
- 6) Average time spent in shopping cart page before leaving the page
- 7) Any outliers on the above 2 average time for different products or clusters of customers

The above insights are great, but Don now wants self service capabilities so that he can directly get the above insights without being dependent on Jyoti. All the above metrics can be sliced, diced and filtered by Product categories, product, calendar, visit country, visit state, visit zip, browser and OS.

Your role is to build an interactive data product that allows Don to get his insights on his own.

## **Assumptions**

1. abc.com has around 100 million visitors a day and has around \$1B in revenue per year. The data volume for conversion analytics is huge
2. The data model to calculate the above metrics are already available in a table format. The attributes and hierarchies are all readily available
3. Don is looking for long term opportunities to optimize and does not need a tool to monitor the conversion rate. There is no business necessity for real time monitoring.
4. Thresholds for metrics to identify outliers are already part of the data model.
5. A good desktop app that works on a browser is sufficient. No need for mobile enablement.

## **Things that we look forward to discussing with you**

We expect this exercise to give you a perception of the responsibilities for the role and have a brainstorming discussion on the solution. The following are some of the things that we would like to discuss.

- How would you architect a visualization app for the above insights?
- What JS libraries and frameworks would you recommend for the solution and why?
- Can you come up with creative ways of communicating the above insights through visualization?
- What would be the navigation flow?
- How do you tell a compelling story?
- How do you NOT overwhelm users with too much information, but still design rich?

## High level data model

