

Title: Web Analytics Dashboard

Author: AGGOR, Senanu and LU, Jamie

Date: February 20, 2016

INSEAD - Big Data & Analytics: Spanish Retailer Project

Introduction: Our client is a Portuguese retailer looking to expand its footprint in Spain, which it entered in 2008. As of 2015, there were 50 brick and mortar stores in Spain primarily in or around Barcelona and Madrid and along the Mediterranean coast. Our client is currently focused on rebranding in Spain by creating a new image, remodeling its stores and creating an omni-channel approach. A couple months ago, the company launched a website in Spain that offers all the products from stores in addition to enhanced product and order tracking abilities.

Goal: Given that the website is new (about 3 months) and considering the limited Google Analytics data we have been provided, our goal is to provide observations and insights about user website behavior activity across various devices: desktop, tablet, and mobile via an interactive dashboard.

Data:

Our data set consists of the following categories (10,000 rows):

- **medium:** the general category of the source
- **deviceCategory:** type of device used (desktop, mobile, tablet)
- **regionID:** the region ISO code of users, derived from IP addresses or Geographical IDs
- **users:** # of users per region
- **sessions:** # of sessions, group of interactions that take place on the website within a given time frame
- **pageviews:** # of total pageviews. A page view is recorded every time a page is viewed
- **transactions:** # of transactions made
- **transactionRevenue:** total revenue from all transactions

Note that each row in our data set corresponds to a specific regionID not to an individual user. Next we cleaned our data set in order to obtain our final data set for the analyses:

- **RegionID:** Matched the region ID with the appropriate country by creating a new column "country"; removed those that resulted in #N/A
- **RegionID:** Matched the region ID (region ISO code of users) with the appropriate region by creating a new column "region"
- **Regions:** Omitted regions labeled "(not set)"
- **Medium:** for those labeled "(none)", replaced with "direct"
- **Viewspersession:** Created a new column "viewspersession" by dividing "pageviews" by "sessions"
- **Revenuepertransaction:** Created a new column "revenuepertransaction" by dividing "transactionrevenue" by "transactions"