

Assignment 2 – Uncovering Marketing insights

Digital Marketing Analytics

Spring 2020

Analyzing Digital Marketing datasets

Goals

To work with datasets using Pandas, xsv(xcsv) and Trifacta(<https://community.trifacta.com/s/lesson-1-introduction-to-trifacta-wrangler>) tools

- To be able to analyze marketing data using Salesforce Einstein analytics studio
- Derive insights from the datasets
- Crisply communicate and document your findings

Case

Marketa analytics has hired you to as an Algorithmic marketing analyst. Marketa is a consulting organization specializing in Marketing analytical solutions. Your client (see allocations by team number below) has provided you a sample dataset and asked you to analyze and build an analytical dashboard as a **Proof-of-concept** to illustrate the value of data driven analytics. The themes to be considered include:

- Pricing
- Promotion
- Search
- Recommendations etc.

Marketa wants you to analyze the data using tools (Pandas, xsv, Trifacta) and build a dashboard using Einstein analytics. They also want you to build a codelabs document to crisply illustrate the value analytical solutions would

bring to the company. You are also asked to discuss what additional datasets and methodologies could be used. The company has a challenge using large scale datasets and are considering using Pandas, Trifacta and xsv as data tools to work with. You are expected to illustrate how you would:

- Use the tools for joining datasets
- Filtering
- Aggregating
- Missing value handling
- Deriving additional columns from existing datasets
- Cleaning (for example removing blank spaces, formatting dates, Capitalizing etc.)

In order to do that you are asked to illustrate the strengths and weakness of each tool/package

Dashboards:

Once you clean the data, import the data into Salesforce and illustrate how to use the Einstein analytical dashboard to illustrate various aspects of analysis. Questions to consider:

- Which columns are dimensions, which columns are measures?
- How would you generate new dimensions? What will you do to summarize measures?
- Who would use this dashboard?
- What value would generated using this dashboard ?

Deliverables:

- How to work with the Pandas, xsv and Trifacta
- Analytics Dashboard using Salesforce Einstein Analytics
- A Google Codelabs document summarizing the insights

Team allocations:

Team number	Link
1	https://www.kaggle.com/jpallard/google-store-ecommerce-data-fake-retail-data#Online.csv
2	https://www.kaggle.com/mkechinov/ecommerce-behavior-data-from-multi-category-store
3	https://www.kaggle.com/carrie1/ecommerce-data
4	https://www.kaggle.com/nicapotato/womens-ecommerce-clothing-reviews
5	https://www.kaggle.com/olistbr/marketing-funnel-olist
6	https://www.kaggle.com/AppleEcommerceInfo/ecommerce-information#category.txt
7	https://www.kaggle.com/artakusuma/laptopecomercee#1_jdid_22-3-19.csv
8	https://www.kaggle.com/rmdw127/cms-open-payments-data-2017#OP_DTL_RSRCH_PGYR2017_P01182019.csv
9	https://www.kaggle.com/vasudeva009/predicting-coupon-redemption
10	https://www.kaggle.com/retailrocket/ecommerce-dataset#item_properties_part1.csv
11	https://www.kaggle.com/vasudeva009/coupon-redemption-smote-feature-selection#final_train.csv