

Google Analytics Customer Revenue Prediction

Kaggle Project: Predict how much GStore customers will spend (<https://www.kaggle.com/c/ga-customer-revenue-prediction/data>)

This data set is included in a Kaggle competition wherein you're challenged to analyze a Google Merchandise Store (also known as GStore, where Google swag is sold) customer dataset to predict revenue per customer. Hopefully, the outcome will be more actionable operational changes and a better use of marketing budgets for those companies who choose to use data analysis on top of GA data. For BUAN 6320 class project, you are not required to participate in the competition. No extra credit will be given for participating in the competition.

test.csv contains the columns listed under Data Fields. Each row in the dataset is one visit to the store.

Data Fields

- **fullVisitorId** - A unique identifier for each user of the Google Merchandise Store.
- **channelGrouping** - The channel via which the user came to the Store.
- **date** - The date on which the user visited the Store.
- **device** - The specifications for the device used to access the Store.
- **geoNetwork** - This section contains information about the geography of the user.
- **sessionId** - A unique identifier for this visit to the store.
- **socialEngagementType** - Engagement type, either "Socially Engaged" or "Not Socially Engaged".
- **totals** - This section contains aggregate values across the session.
- **trafficSource** - This section contains information about the Traffic Source from which the session originated.
- **visitId** - An identifier for this session. This is part of the value usually stored as the _utmb cookie. This is only unique to the user. For a completely unique ID, you should use a combination of fullVisitorId and visitId.
- **visitNumber** - The session number for this user. If this is the first session, then this is set to 1.
- **visitStartTime** - The timestamp (expressed as POSIX time).
- **IMPORTANT:** Due to the formatting of *fullVisitorId* you **must** load the Id's as strings in order for all Id's to be properly unique!