

MA-Project Data

Here are some data that you could consider for your final project.

Title	Data Description	Data File
Shopper Card Data	Data Dictionary from dunnhumby	Carbo Loading Complete Journey
Online Retail Service Quality	Online retail prices, service quality ratings, merchant website traffic, and merchant descriptors from Bizrate.com	Online Retail Price and Service Quality Rating Data
Pharmaceutical Data	Monthly physician drug prescriptions, monthly sales calls allocated per physician, and survey responses on sales person and drug quality for a single therapeutic category. Data is available for over 48 months.	Pharma Data
Brand Marketing Mix Data	Annual unit sales, price, and advertising investments for all vodka brands in the U.S. for more than 15 years.	Vodka Data
Mobile Data	Mobile Coupon Distribution and Redemption Data from Cardagin.	Cardagin Data
Scanner Data	Store level, monthly brand (&sku) sales, price, promotion, feature/display for 1000+ brands in over 30 categories from 2000 until 2006.	Marketing Science article: "The IRI Marketing Data Set" describing the data This is a very large data set that does not open in excel. Contact me for help with accessing this data from UVA's cloud cluster. You will need R or SAS to work with this data.
Emerging Market Sales Data	Sales and marketing mix data by sku for all the brands in the beverage category across retail formats spanning mom and pop, small supermarket, and large supermarket in an emerging economy.	Beverage Data Mom and Pop Beverage Data Small Super Market Beverage Data Large Super Market
Movie Preference Data	Movie preference ratings for 2000+ consumers.	Movie Preference Data ReadMe Genres Data Size Summary

		Movie Information Occupation Information User Information Masterdata Sample Train Data Sample Test Data
Mobile Keyword Data	Data that links keyword searches, and Xbox game purchases for best buy.com. Information on the data files is provided in the following link: https://www.kaggle.com/c/acm-sf-chapter-hackathon-big/data	Best Buy Training Data Best Buy Test Data Best Buy Sku Description
Rotten Tomatoes Dataset	<p>Classify the sentiment of sentences from the Rotten Tomatoes dataset.</p> <p>The Rotten Tomatoes movie review dataset is a corpus of movie reviews used for sentiment analysis.</p> <p>This data presents a chance to benchmark your sentiment-analysis ideas on the Rotten Tomatoes dataset.</p> <p>You are asked to label phrases on a scale of five values: negative, somewhat negative, neutral, somewhat positive, or positive. Obstacles like sentence negation, sarcasm, terseness, language ambiguity, and many others make this task very challenging.</p>	Train Data Test Data
Multi-channel Marketing	Data is from a multichannel company with sales of several hundred million dollars per year. This nationally known company has a network of retail stores, a well-established traditional catalog channel, and a website. The majority of its sales are food products purchased as gifts during the Christmas season. This data set includes over 100,000 customer	Demo Codes Data Description Data <p>This is a very large data set that does not open in excel. Contact me for help with accessing this data from UVA's cloud</p>

	records and over 3.3 million marketing contact records.	cluster. You will need R or SAS to work with this data.
--	---	---