

## RESOURCES ON CASES AND DATA SETS

11-6-15

### INFORMS RESOURCES

#### INFORMS TRANSACTIONS ON EDUCATION (ITE) CASE STUDIES

<http://pubsonline.informs.org/page/ited/cases>

The INFORMS Transactions on Education (ITE) journal currently provides more than 20 case studies. All cases are accompanied by case articles in which the authors comment on uses of the case, teaching suggestions, classroom experience, and other helpful pointers. Some cases also offer teaching notes and instructor materials; access to these materials is restricted. Cases cover a wide range of O.R and analytics areas, including decision analysis, forecasting, integer programming, linear programming, optimization, probability, revenue management, scheduling, simulation, stochastics, and others. Site users have immediate and free access to all information provided, including PDF versions of all case studies. By clicking on keywords within each article, users gain additional access to related articles published in INFORMS journals that list the same keyword. However, most of these articles are not case studies.

#### INFORMS JOURNALS

<http://pubsonline.informs.org/action/doSearch?AllField=%22Data%2C+as+supplemental+material%2C+are+available+%22>

Four INFORMS journals currently offer more than 100 published articles that are accompanied by data sets. To access these articles you can search “data, as supplemental material, are available” on the INFORMS PubsOnline homepage. These articles include an abstract, PDF of the full article, data sets, and references. They include nearly 100 different topics such as analytical modeling, auctions, behavioral operations, data mining, decision analysis, decision-making, forecasting, healthcare, heuristics, inventory management, optimization, risk management, scheduling, stochastic programming, supply chain management, transportation and others.

For those with a journal subscription, access to all information provided is free of charge. For those without a subscription, access to the abstract, references, and supplemental material is of no charge, but you must pay for access to the case study information. INFORMS offers the option of 2 days of access to that publication and all relevant information for \$30. This \$30 fee must be paid for each article accessed without a journal subscription.

#### INFORMS MARKETING SCIENCE SOCIETY-ISMS RESEARCH DATASETS

<https://www.informs.org/Community/ISMS/ISMS-Research-Datasets>

INFORMS Society for Marketing Science (ISMS) offers two raw data sets in relation to retail—a panel database and a field test database. The first data set spans over 6 years in a retail location, recording

nearly 175,000 transactions of 20,000 households. The second data set records the success of a holiday marketing campaign with a database of more than 175,000 customers. ISMS offers a detailed 25-page description of both data sets, research issues, and limitations of the data. The cost of each data set is \$200 for ISMS members, \$500 for non-members. Annual membership for ISMS currently ranges from \$39-\$158 depending on status as student, retired, or regular member.

## **SAS-INFORMS ANALYTICAL SCHOLAR COMPETITION**

<https://www.informs.org/Recognize-Excellence/Community-Prizes-and-Awards/Analytics-Section/SAS-and-INFORMS-Analytics-Section-Student-Analytical-Scholar-Competition/SAS-Analytics-Section-Student-Analytical-Scholar-Competition-Application-Process>

SAS and INFORMS sponsor an annual student competition that allows students studying business analytics and operations research (or a similar field) to practice the process of structuring and presenting a complete proposal for analytical work. Students are asked to produce a statement of work (SOW) for a case study based on a real-life project. SAS provides all initial documentation for the case study and acts as the 'customer' for the competition. In an online forum, SAS answers questions as the customer and encourages students to interact and explore the data provided.

The winner of the competition is selected based on cohesion, proper use of assumptions, demonstrated technical and presentation skill in the SOW, and practicality of the proposed approach to solve the business problem. The winner is invited to attend the INFORMS Annual Meeting at no charge and given the opportunity to learn how the case study was addressed in real-world practice.

Since the competition inception in 2012, INFORMS and SAS have made available all the case-study data online at no charge. Current case-studies include topics such as forecasting, optimization, inventory management, scheduling, and revenue management. The competition is open to all full-time students, at least 18 years old, enrolled in an accredited university, studying analytics, operations research, management science, or a related field, and who will still have their full-time student status at the time of the annual conference.

## **CASE PUBLISHERS**

### **DARDEN BUSINESS PUBLISHING-CASES**

<http://store.darden.virginia.edu/>

Darden Business Publishing provides case-based classroom content for management education at the collegiate level. There are more than 3,000 cases, technical notes, exercises, books, and simulations throughout the site, with approximately 150 new products added annually. Currently under the 'Management Science' tab there are about 80 classroom resources, including books, cases, classroom exercises, simulations, videos, and technical notes. Topics include auctions and bidding, corporate aviation, decision analysis, forecasting, games and competition, multi-attribute utility, optimization, probability and statistics, real options, risk analysis, simulation, spreadsheet modeling, stochastic processes, system dynamics and others. All resources are available for immediate PDF download or

hard-copy and are \$3-\$7 per resource for students or faculty. While site access is free, users must create an account in order to purchase any resource from Darden.

#### **DARDEN BUSINESS PUBLISHING-COURSE SYLLABI**

<http://store.darden.virginia.edu/syllabus#>

Darden Business Publishing provides current Darden course syllabi for the university's graduate curriculum. The list covers the Darden MBA core, electives, executive education, and case method and pedagogy courses. The 34 courses provided include decision analysis, data analysis and optimization, managerial quantitative analysis, and marketing analysis. Each course syllabus lists teaching resources being utilized and where to find them. Typically the resource is found and can be purchased directly through Darden's website for immediate download or shipment of a hard copy.

While access to the syllabi and course information is free, the resources and additional information must be purchased for \$3-\$7 depending on what format the user prefers. The user must create a website account, which is free, to be able to place any orders.

#### **IVEY PUBLISHING-CASES**

<https://www.iveycases.com/>

Ivey Publishing currently has a library of more than 30,000 resources with supplemental data such as data sets and additional case information. They currently offer resources in the categories of accounting, entrepreneurship, finance, general management /strategy, information systems, international, introductory business, management science, marketing, operations management, and organizational behavior/leadership. Ivey Publishing provides resources for all levels of collegiate study from introductory undergraduate courses to advanced graduate studies. The current library consists of about 400 cases in management science, 2,000 cases in operations management, and nearly 2,500 in marketing. Once within any category, users can refine their search for any number of resources based on intended audience, product type, industry, collections, region, product length, publication date, and available translations. When clicking on any resource, users are provided with a general overview of the case, learning objectives, and intended course audience.

Each resource costs \$3-\$6 depending on chosen format (digital download vs hard copy) and can be distributed to students via electronic or hard copy with prior permission directly from Ivey. While access to the site is free of charge, users must create an account in order to place any orders. Ivey will also help users find an appropriate list of resources for their course if provided with information such as level of study and main course subject.

#### **IVEY PUBLISHING-TEACHING TOOLS**

<https://www.iveycases.com/TeachingAuthoringTools.aspx>

Ivey Publishing provides insight to instructors about using cases and data sets for those who are either new to teaching with cases or who are looking for ways to improve their current methodology. The case writing process is described, including why cases are used, the value for participating companies, copyright and ownership issues, etc. Each year Ivey Publishing also provides case method workshops for those looking to hone their skills in writing cases. Information is also provided to individuals who are interested in having their cases published.

## **HARVARD BUSINESS PUBLISHING**

<https://cb.hbsp.harvard.edu/cbmp/pages/home>

Harvard Business Publishing (HBP) is the leading provider of teaching materials for management education. They offer more than 8,000 case studies, exercises, briefs, and industry notes. These materials can be used to add dynamic and real-life perspectives to undergraduate, MBA, and executive education programs worldwide. To assist educators, HBP has grouped a large amount of their most widely used cases based upon level of study. In each level of study (undergraduate, MBA, and executive education) they've broken down the topics into accounting, business and government relations, entrepreneurship, finance, general management, human resources management, information technology, international business, marketing, negotiation, operations management, organizational behavior, service management, social enterprise, and strategy.

HBP also hosts 1-2 day seminars several times per year, both in the US and internationally, to help instructors fine-tune their approach to case method teaching. The seminars are intended for all instructors, regardless of their level of experience with case method teaching.

HBP has partnered with more than 35 other institutions such as Columbia, Darden, Ivey, MIT, McGraw Hill, and Princeton. The average cost for most materials from Harvard Business Publishing is about \$7 per resource but some resources such as full presentations can cost up to \$50. Clicking on any resource throughout the site provides you with a description, covered subjects, setting, and supplemental data varying on what has been provided. In order to purchase any materials from Harvard Business Publishing, users must create a free premium educator account. The account provides users with educator copies of materials, teaching notes, and discounts for students.

## **OTHER CASE & DATASET RESOURCES**

### **DATA.GOV**

<http://www.data.gov/>

Data.gov is the official portal for open data from the U.S. Government. Users can find federal, state, and local data, along with tools and resources to conduct research, build apps, design data visualizations, etc. There are more than 160,000 data sets, which can be filtered by topic area, category, data set type, tags, formats, location, organization types, organizations, and publishers. On average, there are more than 10,000 new data sets published on data.gov every month, spread throughout all of the different

departments listed. All federal data and the majority of state and local data provided is free and does not require registration in order to access it.

## **DATA SCIENCE CENTRAL**

<http://www.datasciencecentral.com/page/search?q=data+sets>

Data Science Central (DSC) is an online resource for big data practitioners that provides an editorial platform, social interaction, forum-based technical support, and information about technology, tools, industry trends, and job opportunities. When searching DSC for data sets, more than 2000 results come up with a wide variety of topics, including about 90 articles related to teaching. The results are largely forum-based (question and answer) and opinion pieces more specifically on methods of teaching different topics, tutorials, and interviews with big data scientists around the country.

Access to DSC is free but with a website membership you can also receive a weekly digest and full access to DSC's professional network, which includes access to other relevant information and the ability to post blogs, forum questions, and comment on both throughout the site. The weekly digest includes topics such as analytics, data science, operations research, big data, visualization, Hadoop, data integration, statistical science, analytics, and pure data science.

A more concentrated list of those data sets available throughout DSC's site, as well as some additional data sets not listed on their site, can be accessed at

<http://www.datasciencecentral.com/profiles/blogs/great-github-list-of-public-data-sets>. The data sets are organized by their study area, such as agriculture, biology, climate/weather, computer networks, economics, energy, finance, government, sports, transportation, etc., and are free to the public.

## **GOOGLE PUBLIC DATA EXPLORER**

<http://www.google.com/publicdata/directory>

Google Public Data currently offers nearly 170 different data sets (cases are not included). Google offers upfront filtering options of Eurostat, Destatis, Statistics Ireland, US Bureau of Labor Statistics, Central Statistics office of Ireland, and a generalized option of any data provider. The data sets can be manipulated to generate the desired results and can be specific or generalized depending on what they're being used to accomplish. Google Public Data Explorer also allows users to upload their own data sets for visualization and exploration. The data sets provided and the service of uploading your own data sets is free to users.

## **KAGGLE-KDD CUP (1997-2010)**

<http://www.sigkdd.org/kddcup/index.php>

KDD Cup is the annual data mining and knowledge discovery competition, now sponsored by Kaggle. The annual competition allows students to put their analytics and data mining skills to the test against students from all over the United States. Each year students are given two tasks, a varying number of

data sets, and other relevant information to complete the assignment. All of the information provided is free; however, users need to create a website account in order to access and download the data available. Below are links to recent competitions:

[1997-2010](#)

[2011](#)

[2012](#)

[2013](#)

[2014](#)

### **KAGGLE IN-CLASS INITIATIVE**

[\(http://inclass.kaggle.com/\)](http://inclass.kaggle.com/)

The Kaggle in Class initiative is a site devoted to helping instructors of analytics and data mining courses host competitions as part of their course work or outside projects. There are currently more than 500 student competitions listed on the website. Although most of the competitions listed have passed their submission dates, they do still have all of the data sets, case studies, and other relevant information available to the public. The services provided by Kaggle are free to hosts (course instructors) and users (students), but all hosts and users must create a free Kaggle account to access the site. Because the site is offered to users at no charge, the capabilities of the site are limited and Kaggle is unable to assist with issues that may arise.

### **KD NUGGETS**

<http://www.kdnuggets.com/datasets/index.html>

KD Nuggets is a website covering topics in business analytics, big data, data mining, and data science. The site currently has nearly 70 data repositories from various sources. The information included within each link varies from case studies to actual raw data sets and data sheets. While most of the data is provided at no cost to the user, the user often will need to create an account for the site providing the data in order to access it completely. The site does not offer a way to search through the data sets or filter the results, but does give you additional links to:

1. [Analytics, data science, and data mining competitions,](#)
2. [Preferred data APIs, hubs, marketplaces, platforms, and search engines, and](#)
3. [Government, state, city, local, and other public data.](#)

### **MARKETING EDGE**

<http://www.marketingedge.org/marketing-programs/data-set-library>

Marketing Edge is a data set library related to various forms of marketing. The information includes data in relation to retailers, non-profits, television, and healthcare. The site currently offers 15 data sets free to members or \$25/data set to non-members. The data sets are available to full time college faculty, PhD students, doctoral candidates, full-time researchers at colleges and universities, as well as website members. Included with a membership (\$100 basic or \$175 premium) is access to materials, seminars, data sets, course syllabi and various teaching materials. For those that choose to forego membership, requesting data sets requires sending an email which includes typical contact information (name, title, university, and contact number) and allowing the site to verify your academic status, as well as signing a non-disclosure agreement.

### **SAS-INSTRUCTOR RESOURCES**

<http://support.sas.com/learn/ap/tkit/list.html>

SAS currently offers teaching materials to professors at no charge. These materials include chapter instruction notes, PowerPoint presentations, course data sets, and practice exercises, all delivered electronically and ready for reproduction. To request this material, an email must be submitted to SAS with the desired materials and course information, a license is delivered to the professor, must be signed and returned, and materials are delivered electronically.

The list of available material covers a range of courses, both online and in-person, teaching fields such as programming, statistical analysis, forecasting and econometrics, data-mining, text analytics, optimization and simulation, Hadoop, visual analytics, data management, and risk management. These courses are designed for the beginner to advanced SAS user and can range from a single half-day course to multiple days depending on the speed of the program chosen.

SAS also offers workshops and professional development opportunities exclusively for university professors. They are offered at no charge (except travel expenses) but are by invitation only. They include topics such as cleaning up big data, advanced predictive modeling, applied mixed models, visual analytics, advanced business analytics, visual statistics, and data manipulation.

### **SAS-ANALYTICS SHOOTOUT**

<http://www.sas.com/events/analytics/us/contest.html>

The annual Analytics Shootout sponsored by SAS is a competition in which student competitors are given a collection of data sets and a business problem to solve using SAS software and advanced analytical methods. This competition gives students the opportunity to solve real-world advanced analytic problems, to test their skills and gain valuable experience. Students must work together to clean the provided data sets, comprise a modeling approach using a range of techniques, build the model, and submit a report to summarize their approach and solution to the problem.

The Analytics Shootout integrates relevant advanced analytics problems, use of real and complex data, multiple approaches to a solution with no clear-cut correct answer, and publication potential. The

competition is free to all participants and is set up to run concurrently with a typical university semester; registration opens in January, teams must be completed in May, submissions are due in June, winning teams are notified in August, and winners are recognized and present their solutions at the SAS conference in October. Access to previous years' problems and data sets has not been made available by SAS.

## **TERADATA RESOURCES**

<http://www.teradata.com/Templates/ResourceCenter/Landing.aspx?id=12884907670&LangType=1033&LangSelect=true>

Teradata, an analytics solutions company, offers various forms of analytics information, including a library of relevant reports, articles, case studies, briefs, infographics, podcasts, videos, white papers, and other materials. There are nearly 150 case studies relating to various areas of analytics, including 14 cases in finance, 3 cases related to government, 3 cases in relation to healthcare, and 6 cases in transportation and logistics, among other areas. The solutions included are in areas such as big data, marketing, analytics and operations, Hadoop, and unified data architecture, along with several other solutions. Teradata also offers about 50 data sheets which highlight successful usage of Teradata solutions (the full data sets are not provided). The information provided by Teradata, including case studies, is available for immediate and free download to anyone. Some of the case studies and other information are locked but can be accessed by registering for a free T-Pass membership.