Harvard College Data Analytics

Harvard Undergraduate Data Analytics Group

PREPARED FOR

The City of Philadelphia

PREPARED DATE

Aug 6, 2023

ENGAGEMENT TIMEFRAME

September – December 2023

Harvard College Data Analytics Group (HDAG) is a non-profit student organization at Harvard dedicated to helping organizations make smarter and more data-driven decisions. We assist companies in achieving their strategic goals by translating their data into meaningful and actionable information. We aim to pair teams of well-trained, highly-motivated Harvard students with our partners, specifically focusing associates and analysts in industries where they have experience or interest, in order to produce the highest quality of work possible. From data collection to strategy implementation, we want to be there every step of the way to help organizations make data their new superpower.

We competitively recruit undergraduate students at Harvard with demonstrated competence, dedication, and problem-solving skills, many of whom have prior experience working in top management consulting or data science teams. All our team leaders have experience working in or leading data science teams at Fortune 500 companies, and our board of technical advisors include members of the Harvard faculty. Each team, composed of around seven to eight Harvard students, commits over 600 hours to a case over the course of a 10 week span.

We enjoy different challenges and work with a diverse set of organizations and problems. Our clients range from local businesses to Fortune 500 companies to international non-profits. Using our capabilities in visualization, machine learning, and predictive analytics, among others, we help organizations diagnose problems and identify strategies across their sales, marketing, financial or operational functions. Client confidentiality is our utmost priority.

Harvard College Data Analytics

Team Capabilities

1. Data Analytics Consulting: deriving valuable insights from data

a. Case study 1 - Providing IT resource management analytics for a multinational Fortune 500 company in energy and automation: Through statistical analysis of over 100k anonymized employees, we identified help desk call volume and demographic trends to help inform executive decisions on employee satisfaction and IT resource allocation.

b. Case study 2 - Providing data processing service for a Wall Street fintech company: Through scraping the Securities and Exchange Commission (SEC) website and extracting relevant data en masse, we created well-formatted databases to advance the client's core digital offerings.

2. Machine Learning Algorithms: training and deploying predictive models

a. Case study 1 - Providing IT security service for a multinational Fortune 500 company in energy and automation: By building ML models, we enabled predictive analytics for the company's future spending on Indirect Procurements and introduced data integrity improvement design to the purchase request process.

b. Case study 2 - Providing AI algorithm advancements for a leading sports analytics company: Using “Big 5” European club leagues’ pre-game and in-game data, we created models that predict win, loss, and draw probability and provided an evaluation of the accuracy and probability calibration of the models.

3. Business Intelligence Visualizations: creating interactive visual dashboards

a. Case study: Providing visualization services for the World Health Organization Region for the Americas: We developed a web app to visualize models on COVID-19 outbreak to predict rate of transmission and epidemic curves; product delivered to WHO country offices in Latin America for projections of varying health intervention measures.

4. Whole-Set Solutions: providing comprehensive digitalization systems

a. Case study: Creating an HR and user management system for an educational foundation in China: We developed a system from scratch to help the management team keep track of employee’s progress and KPI and to help employees better manage student feedback.

Harvard
College
Data
Analytics

Proposal for the City of Philadelphia:

Broadly speaking, the project that HDAG will engage with the client would consist of two stages. Our offered skill sets include: Data Analytics: Exploring, cleaning, processing, and deriving valuable insights from data; Dashboarding: Creating interactive visualizations; Whole-Set Solutions: leveraging data and research to create technology-driven forecasts.

1) Cleaning, processing, and matching data sources. The HDAG team will start by analyzing the provided point of sale datasets on wholesaler transactions. The datasets will be processed, combined, and cleaned such that data common to several similar, but unique, datasets can be reliably cross-referenced and combined. This phase of data processing serves as the foundation for data analysis in the second stage of the project.
2) Exploratory data analysis, visualization, and insight generation. Guided by the client's specific analysis goals and insights, the team will analyze the processed datasets and explore meaningful trends, relationships, and patterns. The team will address specific hypotheses regarding the data provided by the client and investigate their occurrence in the data.

The team will thoroughly document their methodology and findings for all technical work throughout the duration of the engagement. A final presentation in slide format and a final writeup detailing the work, processes, motivations, and findings will be created and delivered to the client at the end of the engagement. The client will additionally receive all technical materials (i.e. code, cleaned and processed datasets) related to the case.

Contingent on the Case Team Lead's agreement, they will serve as the lead author and point of contact for the final report and its contents. Whether or not the Case Team Lead is accessible, the work remains property of the client, and the client is free to seek other parties to validate and present the findings.

Harvard College Data Analytics

Rough Engagement Timeline

Dates	Week	Tentative Schedule
9.18-9.24	0	Each HDAG Case Team Leader (CTL) will have a call with the respective Client liaison to better understand work expectations and align goals for this semester (in terms of research questions, final format of deliverables, etc.) After the meeting, CTL will consult with two associates of the HDAG case team and map out the weekly work plan for the semester.
9.25-10.1	1	CTL will introduce the project and the work plan to the rest of the case team and start delegating tasks to each individual. (In each team we have data scientists who are proficient in Python, R, SQL and other analytical tools as well as business analysts who have experience working in industry).
10.2-10.8	2	Every member of each Client Case Team will follow the work plan, start both the data analytics, which includes every aspect of the data pipeline: data transferring, cleaning, exploration, modeling, visualization etc. Every week, each CTL will update the Client liaison on the progress that the case team has made over the past week. There is also a weekly meeting between the case team where each member will discuss their work with the others, and
10.9-10.15	3	
10.16-10.22	4	

		the CTL will delegate work for next week.
10.23-10.29	5	Wrap up the work for the first half of semester, and prepare for the midway presentation to Client which can include both a technical product (algorithm, statistical model, web app) and a business presentation (slides).
10.30-11.5	6	Midway presentations with Client: each whole team will present their findings and recommendations from the first half of the semester to the Client team. Each HDAG case team will follow up with any questions the Client team might have during or after the presentation.
11.6-11.12	7	After the midway presentations, each CTL will integrate comments or suggestions from the Client team to the work plan. Each CTL will list out the remaining questions or technical tasks for the latter half of the semester and delegate them to each individual of the case team.
11.13-11.19	8	
11.20-11.26	9	
11.27-12.3	10	The case team will summarize their work for the entire semester and give a final presentation to Client. This will include both technical deliverables and the business presentation. HDAG team will follow up with any questions the Client business team might have during or after the presentation.
12.4-12.17	Post- Project	HDAG team will follow up with Client on the implementation of suggestions and deployment of analytical tools. We will ask for feedback on their work for the Fall of 2023.

Harvard
College
Data
Analytics

Pricing

- Engagement Timeline: 12 weeks, September - December, 2023
- Semester Case Fee: TBD