Harvard
College
Data
Analytics

Harvard Undergraduate Data Analytics Group

PREPARED FOR

Moderna

PREPARED DATE

Aug 12, 2022

ENGAGEMENT TIMEFRAME

Sep - Dec 2022

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 10-12 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 Moderna:

The goal of this project is to develop an ML use case (project code name: Maverick) based on User Data located in ServiceNow, which is a help desk service to manage tickets & satisfaction and organize onboarding / offboarding tasks, and OKTA, a device & software authentication service. The ultimate goal of this engagement is to help accelerate the digitalization of Moderna.

There are a few potential use cases outlined within the following verticals: “Perceive”, “Predict”, “Recommend”, and “Categorize”. After spending some time evaluating and discussing the feasibility of each use case, the team will choose the use case(s) which have the most value to Moderna. The use cases are briefly described below,

- **Perceive:** Detecting “Unusual Incident Trends” to alert stakeholders proactively (NLP to ServiceNow ticket description). Detect “Recurring Incident” to prioritize development of knowledge base and trainings
- **Predict:** Forecast “Incident Load Evolution” to adjust service availability and bandwidth (based on in incident quantity historical data)
- **Recommend:** Generate “Urgency Score” (likelihood to escalate or impact senior levels) to recommend what incident to solve in priority. Generate “At Risk Score” (likelihood of a slow resolution time) to trigger extra attention and set expectations. Generate “Response Quality Score” (speed/detailed/satisfaction/tone).
- **Categorize:** Auto-assign tickets to the right ServiceNow category.

Below is a list of the deliverables that will be provided by the end of the engagement. The HDAG team will deliver all models/scores developed and insights generated for the chosen use cases.

1) Presentations detailing the specific use cases which have the highest value proposition and the findings/insights generated from the baseline models/scores developed (plan to be delivered by Midpoint).
2) Technical deliverables, including code repository, curated data sets, and model-based findings. The case team will report on the scalability of the use case and methods by which it can be deployed onto a larger scale.

Harvard College Data Analytics

Rough Engagement Timeline

Dates	Week	Tentative Schedule
9.5-9.18	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 the 1-2 associates of the HDAG case team and map out a more detailed weekly work plan for the semester: from both the perspective of technical execution and business analysis.
9.19-9.25	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). At the end of Week 2, the team will have chosen which of the potential use cases to work on out of those provided in the "Perceive", "Predict", "Recommend", and "Categorize" verticals by determining which use case will have the highest value proposition to Moderna. Additionally, exploratory data analysis will have started on the most promising verticals.
9.26-10.2	2	
10.3-10.9	3	Every member of each Client Case Team will continue to follow the work plan, mainly focusing on the 1-2 use cases that we decided in the first 2 weeks. Exploratory data analysis will continue in each of the verticals that we've narrowed down to have the highest potential value after which baseline models will be created for each of the use cases. Every week, each CTL will update the Client liaison on the
10.10-10.16	4	
10.17-10.23	5	

		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 the CTL will delegate work for next week. By the end of Week 5, exploratory data analysis and baseline model generation will have been finished for each of the use cases and we will begin creating a presentation to display our findings so far. This will include the reasoning behind why we determined a specific case to have the highest value proposition and the findings we have generated from the baseline models so far.
10.24-10.30	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.
10.31-11.6	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. Additionally, more complicated models will continue to be developed by the case team in order to implement the chosen use cases. The performance of these models will be compared to the baseline models that were created prior to the midpoint presentation. In parallel, if desired by the client, the case team will be exploring the scalability of the use case and methods by which it can be deployed onto a larger scale.
11.7-11.13	8	
11.14-11.20	9	
11.21-11.27	10	
11.28 - 12.4	11	The case team will summarize their work for the entire

		semester and give a final presentation to Client. This will include both technical deliverables (e.g. code repository, curated data sets) and the business presentation (e.g. use-case summaries, model-based findings). The HDAG team will follow up with any questions the Client business team might have during or after the presentation.
12.5-12.11	Post-Project	The 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 2022. This extra week is a time that is mainly kept in the case of any delays in finishing deliverables when final presentations can still occur.

Pricing

- Engagement Timeline: 12 weeks, September - December, 2022
- Semester Case Fee: $45,000