SPECIFIC REGULATIONS FOR THE POSTGRADUATE DIPLOMA PROGRAM IN DATA SCIENCE IN HEALTH AND CLIMATE CHANGE FOR SOCIAL IMPACT IN COLLABORATION WITH DATA.ORG, JPAL-SOUTH ASIA AND ARTPARK

General

a. This document presents the Regulations specifically applicable to the academic program, Postgraduate Diploma in data science in health and climate change for social impact (hereafter referred to as DSHCS) conducted and offered by IIIT-Delhi in collaboration with data.org, JPAL-South Asia and ARTPARK.

b. This is an academic program that runs under a collaboration with data.org, the responsibilities of each of the collaborators remain as defined in the MoU signed between IIIT-Delhi and data.org and ARTPARK.

c. The program being offered has already been approved by the Senate and therefore, any recommendation towards a change in the program shall be first approved by the Academic Affairs Committee (AAC), and thereafter, by the Senate of the Institute.

d. Both IIIT-Delhi and data.org will individually nominate a coordinator before the announcement of the start of a DSHCS. These coordinators shall collectively be the point of contact for the implementation of the joint program as per these Regulations.

e. The expenses towards any legal issue related to the program, if any arising at any stage due to these Regulations, will be met from the budget of the program.

Academic sessions and calendars

a. The Academic Session of each DSHCS batch (cohort) will run for about 37 weeks. At a given time, in steady state, only one batch will run.

b. The program will be delivered in a hybrid model, i.e., classes will be conducted offline and online candidates will attend those classes through online platforms like zoom. Classes will mainly be conducted by IIIT-Delhi faculty with support from ARTPARK and J-PAL South Asia.

c. The start date of the program will be announced with the due approval of the Senate on the recommendations of the AAC to the proposal of the coordinators of the DSHCS program. The grant available at present is to support the program for two years. The batch will comprise around 50 students each year.

d. The academic calendar and course schedule for every batch will be announced at least one month before the commencement of the batch, with the prior approval of the Senate.

Eligibility for Admissions and Admission Process

a. **Number of seats:** The intake will be ~ 50 seats per batch, with the actual number depending upon the quality of applications

b. **Reservations:** The reservations will be followed as stated in the MoU with data.org, or otherwise as approved by the Board of Governors of IIIT-Delhi.

c. **Eligibility:** The DSHCS program is targeted at both non-technical and technical professionals working in diverse sectors ranging from healthcare, including clinical and nursing professionals, climate scientists and data analysts, program implementers in government agencies and non-governmental organizations, and technically trained students. The candidate should satisfy the below criteria to be considered for the selection process.

   - Graduation (BTech/BE in any discipline, BSc/BCA in CS/Mathematics/Statistics/IT/Electronics or related streams OR BBA OR MBBS OR related bachelors in CS or IT or Economics or Bioinformatics or other areas involving a quantitative component) with at least 50% marks or equivalent CGPA.
   - Candidates who have appeared in the final semester examination and awaiting results may also apply. However, they have to submit proof of passing all semester examinations/final degree at the time of completion of the course. Otherwise, provisional admission shall stand cancelled.

d. **Shortlisting/ Selection process:** A selection committee will be shortlisting the candidates based on their academic and professional credentials. The shortlisting criteria as approved by the Senate, on the recommendations of the AAC to the mutually agreed proposal of the coordinators of IIIT-Delhi and data.org, will be followed.

e. **Enrolment:** Enrolment/ admission to the program by a selected student requires timely payment of fees by the given date and verification of the original documents in accordance with the procedure set out in the prospectus.

f. **Foreign students:** The program is also open for foreign students. However, this may be revisited subsequently by the Senate, on the recommendations of the AAC to the mutually agreed proposal of the coordinators of IIIT-Delhi and data.org.

g. **Cancellation of admission:** The admission of a candidate to the program may be cancelled if any material information provided is found to be incorrect, or s/he has not paid the fee by the given date. The admission of a student may be cancelled by the Senate even at a later date if it is found that the student had supplied some false information or suppressed any relevant information while seeking admission.

Courses, credits, evaluation and graduation

a. **Courses**

The course is divided into ten modules spanned over two semesters. The details of the ten modules as given below:

S. No.	Module	Topics	Description
1	Introduction to Data	Introduction to Statistics and Data Science	Overview of the Program: Unique Opportunities and Challenges
			Overview of Data Science Approaches: How
			Types of Data
			Open Sources of Data, Introduction to the Key Datasets for the Curriculum
			Generating Your Own Data: Study Design
2	Working with Data	Introduction to Excel, R and Python	Introduction to Excel, R & Python
			Reading and Writing Data using R & Python
			Basic Exploration of Data
			Data Quality (Seeing and Analyzing)
			Remediation Approaches (Filtering, Transformations, Imputation)
			Real World Example Summarizing the module
3	Inferential Statistics and Comparative Data Sciences	Basics of Biostatistics	Getting Data into Shape (Long and Wide Data)
			Random variable and probability
			Measures of Centrality and Deviation
			Probability Distributions: Discrete and Continuous
			Moments of distributions (mean, variance, skewness, kurtosis)
			Probability Density Function and Cumulative Density Function
			CLT and Normal Distribution
			Sampling distributions (e.g., sampling distribution of the mean)
			Central Limit Theorem

			Histograms, Density Plots, Boxplots, Violin Plots
		Beginning Data Analytics	Relationship between distributions (e.g., normal approximation to binomial)
			Quantiles and Rank Statistics
			Testing for Normality and Data Transformations
			Statistical Inference: Parametric Tests
			Statistical Inference: Non-parametric Tests
			Real World Example Summarizing the module
4	Modeling and Basic Machine Learning	Linear Algebra and Regression Models	Introductory Linear Algebra
			Supervised and Unsupervised Learning
			Regression: Linear Model General Linear Model
			Classification: Logistic Regression and Multinomial Model
			Variable Selection: Stepwise approaches
		Machine Learning	Introduction to Machine Learning & Bayes Theorem
			High Dimensional Data: Principal Component Analysis and t-SNE
			Machine Learning Approaches: Decision Tree, Neural Networks. SVM, Random Forest, Bayesian Networks
			Overfitting and Generalization
			Regularization: Lasso, Ridge and Elastic Net Regression
			Unsupervised Learning: Clustering
			Real World Example Summarizing the module
5	Data Visualization	Advanced Visualization Techniques in Data Sciences	Introduction to ggplot (R) and seaborn (python)
			Advanced Visualization Strategies in Data Sciences e.g., heatmaps, sankey plots, radar charts, sunbursts, word clouds, waterplots, icicle plots
			Interactive Data Visualization (using plotly, networkD3)

			3D data visualization
			Creating Dashboard
		Introduction to Geographical Data, Spatial Statistics	(A) Introduction to Spatial Data and Spatial Models: Geostatistical data; Lattice sata; and Point data. (B) Characterising Spatial Autocorrelation (Metrics) and Relevant Issues for Classical Regression Analysis;
			(C) Exploratory Spatial Data Analysis and Stationarity of Spatial Random Processes
			(D) Measuring Spatial Dependence and Spatial Heterogeneity
			(E) Environmental Pollution and Economic Growth application with hands-on exercises on ArcGIS and R.
6	Advanced Data Sciences	Artificial Intelligence and Deep Neural Networks: Models	Introduction and Multi-layer perceptron
			Introduction to TensorFlow and Feedforward Neural Networks
			Activation Functions and Training Neural Networks
			Learning from images: Convolutional Neural Networks (CNNs)
			Learning from sequences: Recurrent Neural Networks (RNNs)
			Applications of AI and Deep Neural Networks
			Interpretability
		Nature Language Processing	Parts of Speech, Parsing, Ontologies, WordSense
			Introduction to Modeling with Text: Real World Examples
			Modeling Text: Hidden Markov Model, Recurrent Neural Networks, LSTMs
			Word Embeddings & word2Vec
			Transformer
			BERT and its Applications
			Social Media, Emotion Modeling, Chatbots & ChatGPT
7	Data Science in Health Care	Data Ethics	Introduction to Data Ethics

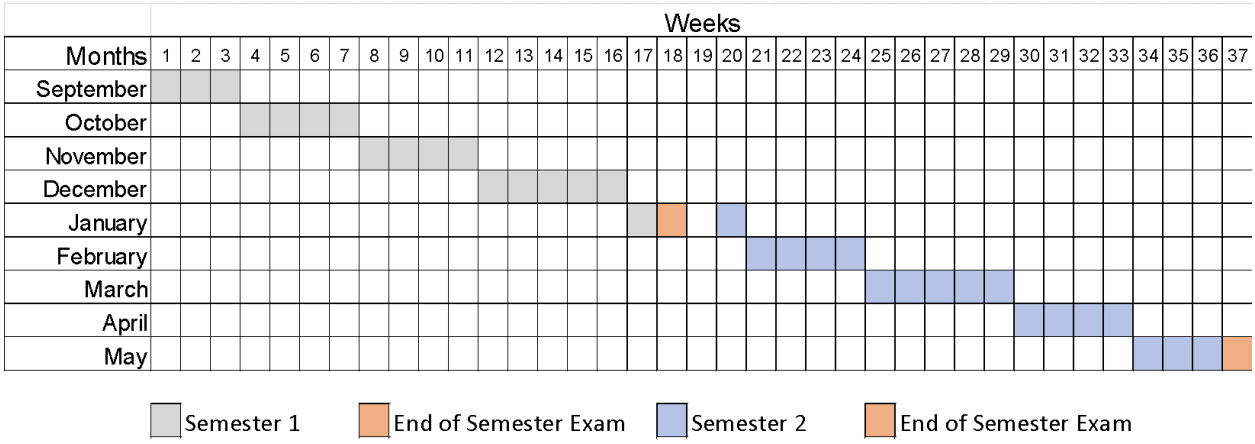
		Data Driven Health Sciences with Social Impact Case Studies	Each participating faculty is expected to demonstrate an actual health - related case study with demonstrated social impact. Preference will be on standard publicly available data sets that can be reused.
8	Data Science in Climate Emergency	Data and Climate Science	Different types of publicly available data products for studying climate change and emergency. Data driven models for climate predictions and projections.
		Data Driven Climate Sciences with Social Impact Case Studies	The module will cover data-driven case studies investigating the impact of climate change on rural communities, urban development, health and nutrition as well as studies linking climate change impact and public policy designs.
9	Data Science: Intersectionality and Representativeness of Data	Intersectionality and representativeness of data science	Importance of Intersectionality and representativeness of data science and its implications on policy making
		Case Studies	Participating faculty is expected to demonstrate an actual case-study involving intersectionality and data science and its impact on policy making
10	Futuristic Perspectives	Future of Data Science	Futuristic developments in Data Science: chatGPT etc

b. Credits and Teaching Hours

S. No.	Module	Teaching Hours		Credits
		Live Lectures	Asynchronous Learning	
1	Introduction to Data	12	8	2
2	Working with Data	13	10	2
3	Inferential Statistics and Comparative Data Sciences	25	15	4
4	Modelling and Basic Machine Learning	34	6	4
5	Data Visualization	28	12	4
6	Advanced Data Sciences	30	9	4
7	Data Science in Health Care	28	12	4
8	Data Science in Climate Emergency	22	6	3
9	Data Science: Intersectionality and Representativeness of Data	14	4	2
10	Futuristic Perspectives	4	4	1
Total		210	86	30

A cohort of ~ 50 students will go through the program together. Institute faculty along with some experts from ARTPARK and J-PAL South Asia will conduct lectures and required tutorials. The program will span approximately 37 weeks (9 months) with an overall 296 hours of teaching, of which 210 hours will be as live teaching sessions and 86 hours will be covered as asynchronous learning. The course will be divided into two semesters, with a total credit of 30 (the module-specific break-up of these credits is provided above).

The duration and the time-line per batch is provided below:



b. **Changes in course structure:** Any changes in courses proposed will be done post approval by the AAC. Larger changes that include changing one or more courses will be done post approval by the AAC, and upon its recommendation, if required, upon the approval of the Senate.

c. **Assessment and Evaluation**

- Certification Body: PG Diploma in data science in health and climate change for social impact and course grade sheets and transcripts shall be issued by IIIT-Delhi.
- Grade sheets will be issued after the successful completion of each course as approved by the coordinators of the program for its compliance to the Senate approved norms.
- For successful completion of the program, the candidate will be required to successfully complete all the modules of the courses. Grading will be as per the IIITD norms (based on moderation).
- The PG Diploma certificate will be issued to a participant only post the successful completion of the program as described in the clause above.
- Course Repetition: If any of the participants is not able to successfully complete any of the modules, he or she will be given one more chance to successfully complete the respective failed modules without being charged any extra fee. In case, the participant is not able to successfully complete two or more courses even after having availed the

second chance, he or she would not be eligible to receive the Diploma and would not receive any fee refunds.

- Each course's assessment plan will be decided by the teaching faculties in consultation with the program coordinators.

d. **Change in grade:** A letter grade, once awarded, shall not be changed unless the request is made by the instructor of the course to the coordinators of the DSHCS and recommended by them to the AAC for approval. The Approval of the AAC will be reported to the Senate. Any such request for change of grade must, however, be made within two weeks of the declaration of the grade with all relevant records and justification.

e. **Attendance:** Candidates will be required to maintain at least 70% of their attendance to be eligible for the end-term exam.

f. **Graduation requirements:** Candidates will be required to successfully complete all the modules and achieve a minimum cumulative grade point average (CGPA) of 5.0 in order to graduate..

g. **Award of Diploma:**

- At the end of the offering for a batch, the CGPA will be computed as the average of the grade point values corresponding to the letter grades for all the modules and the term exams.
- Once the CGPA has been computed, the Senate may recommend the award of the DSHCS diploma certificate to the student. While awaiting the actual award of the diploma during a campus visit by the student on a day arrived at by IIIT-Delhi (AAC and Senate) and data.org, the DOAA may authorize the Registrar to issue a provisional certificate to a student who completes the requirements for graduation.
- Under extremely exceptional circumstances, where gross violation of the graduation requirements or use of dishonest means is detected at a later stage, the Senate may recommend to the Board of Governors to withdraw an already awarded diploma.

Fee payment

a. **Stratification:** IIIT-Delhi and data.org reserve the right to revisit the program fee, and can install a stratified merit-based fee payment regime.

b. All fee payment must be made by a student as per the declared timelines.

c. **Refund and withdrawal:** Fee once paid by the candidate will not be refunded back. Only in very rare and exceptional scenarios mostly on health or humanitarian grounds, the refund case may be referred to and assessed by the program coordinators at the institute and his or her decision on the case shall be final.

Leave Rules

The program doesn't have an explicit notion of the candidates taking a leave.

Placement/Internships

ARTPARK and J-PAL south Asia will be helping the candidates with internship opportunities. However, no commitment will be made for the placements.

Disciplinary Action

IIIT-Delhi maintains global academic standards in its own education system; this extends to and includes this DSHCS program. Therefore, it does not tolerate any form of indiscipline and/or academic dishonesty. Action will be taken against students found engaging in acts of indiscipline/ academic dishonesty by the Disciplinary Action Committee (DAC) constituted by the Director. The disciplinary action may result even in termination of the program in extreme cases. The rules and regulations in force at the Institute will be followed for the students of DSHCS as well.

If any candidate or candidates are found to be involved in any activity of plagiarism or copying, he or she or both will be awarded zero marks for the submitted assignment or in the assessment or exam and will be given a stern warning for the first time. Thereafter, if the same participant is found to be involved in similar activity again, he or she will be expelled from the program immediately without any fee refund.