Course: CSP571 Data Preparation and Analysis
Mind Matters - Mental Health in Tech Industry
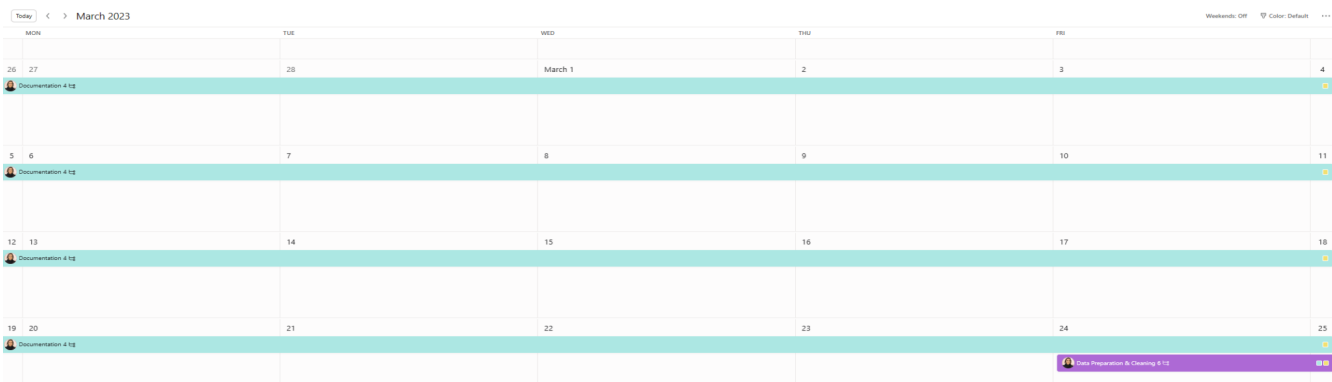
Project Plan and Detail

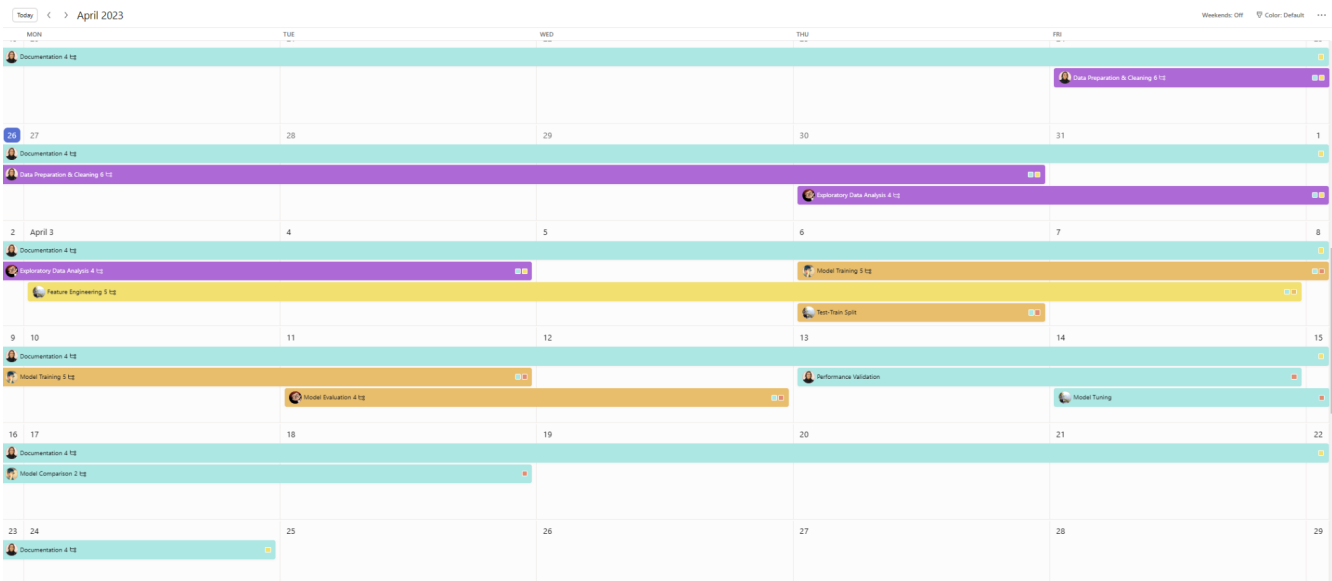### Team Members

Zainab Hasnain (Group Leader)	A20516879
Aditya Sharma	A20516668
Muhammad Ans Qadri	A20521573
Abhishek Jaiswal	A20380004

Detailed project breakdown can be found on this [Asana project link](#).

Calendar View (March 2023): highlighting the tasks and subtasks

Calendar View (April 2023): highlighting the tasks and subtasks

Project Task and Subtasks including the Dates, Duration as well as the Assignee

Task name	Assignee	Due date	Priority	Status
▾ Completed				
▾ ✅ Data Gathering 2	Muhammad Ans Qadri	Feb 21	Medium	Done
◯ Gather responses of the OSMI survey conducted in 2014.	Muhammad Ans Qadri			
◯ Access the survey responses conducted in subsequent years.	Muhammad Ans Qadri			
Add task...				
▾ In Progress				
▾ ◯ Data Preparation & Cleaning 6	Zainab Hasnain	Mar 24 – 30	High	On track
◯ Create new datasets (pre-covid and post-covid)	Zainab Hasnain	Tuesday		
◯ Recognizable data type	Muhammad Ans Qadri	Wednesday		
◯ Data Imputation: handling null values	Aditya Sharma	Wednesday		
◯ Omitting any duplicate and invalid responses	Abhishek Jaiswal	Wednesday		
◯ Merging and quantifying the data in categorical columns	Aditya Sharma	Thursday		
◯ Converting categorical variables to numerical variables	Muhammad Ans Qadri	Thursday		
▾ ◯ Exploratory Data Analysis 4	Abhishek Jaiswal	Mar 30 – Apr 4	High	On track
◯ Correlation plots	Zainab Hasnain	Saturday		
◯ Colored histograms by label (seek treatment)	Muhammad Ans Qadri	Saturday		
◯ Visualizing our EDA results using plotting libraries like ggplot	Abhishek Jaiswal	Apr 2		
◯ Using box plots/histograms/heatmaps/bar charts wherever appropriate	Aditya Sharma	Apr 2		
▾ ◯ Feature Engineering 5	Muhammad Ans Qadri	Apr 3 – 7	Medium	On track
◯ Removing any redundant columns	Abhishek Jaiswal	Apr 3		
◯ Outlier detection and removal	Abhishek Jaiswal	Apr 4		
◯ Dimensionality Reduction	Aditya Sharma	Apr 5		
◯ Feature selection using Chi-Squared Test, PCA, Lasso (or Ridge) Regression	Abhishek Jaiswal	Apr 6		
◯ Feature scaling and normalization	Aditya Sharma	Apr 6		
▾ ◯ Documentation 4	Zainab Hasnain	Feb 10 – Apr 24	Low	On track
◯ Project Outline				
◯ Project plan and detail				
◯ Project presentation				
◯ Project report				
Add task...				
▾ To Do				
◯ Test-Train Split ⇅ Details ›	Muhammad Ans Qadri	Apr 6	Medium	To Do
▾ ◯ Model Training 5	Aditya Sharma	Apr 6 – 11	Medium	To Do
◯ Logistic Regression	Muhammad Ans Qadri	Apr 6		
◯ LDA	Zainab Hasnain	Apr 8		
◯ Lasso regression	Abhishek Jaiswal	Apr 9		
◯ Ridge regression	Abhishek Jaiswal	Apr 10		
◯ Decision Trees	Aditya Sharma	Apr 10		
▾ ◯ Model Evaluation 4	Abhishek Jaiswal	Apr 11 – 12	Medium	To Do
◯ Confusion Matrix.	Muhammad Ans Qadri	Apr 11		
◯ Precision, Recall, and F1-Score	Zainab Hasnain	Apr 11		
◯ The area under the curve and ROC.	Aditya Sharma	Apr 12		
◯ Cross validation	Abhishek Jaiswal	Apr 12		
◯ Performance Validation	Zainab Hasnain	Apr 13 – 14	Low	To Do
◯ Model Tuning	Muhammad Ans Qadri	Apr 14 – 15	Low	To Do
◯ Test Performance Evaluation	Abhishek Jaiswal	Apr 16	Low	To Do
▾ ◯ Model Comparison 2	Aditya Sharma	Apr 16 – 18	Low	To Do
◯ Comparison Analysis over the Training /Validation Set	Abhishek Jaiswal	Apr 16		
◯ Comparison Analysis over the Test Set	Aditya Sharma	Apr 17		
Add task...				