

# Readme

Mathews Philip Venattu

02/07/2021

## Readme File for this Project Directory

In this Directory there are 6 Folders: Data, Figures, Literature Surveys, Presentation, Tasks Allotted and Thesis Report Files. NB: You can use the tab below to read more about each folder and the files in it.

### Data

**Data Folder** contains all the datasets that I used and created for this project. Inside this **Data Folder** there are 3 Folders (in github there are only 2 because of storage constraints).

**data\_version\_0** In **dataset\_version\_0** Folder it contains the raw data that I got for this project. Files and its Explanations

**Cass\_data\_ksm2005.rds**: Contains the Magnetic Field measurements of Cassini Space Craft and its Positions in KSM Coordinate System

**Cass\_data2005.rds**: Contains the Magnetic Field measurements of Cassini Space Craft and its Positions.

**Full\_Cassini\_Master\_MP\_BS\_CMJ\_revised2005.rds** : Contains the data of the Bow Shock, Magneto Pause Events (when they occurred)

**joined\_data** : It is a dataset created by me by joining the **Full\_Cassini\_Master\_MP\_BS\_CMJ\_revised2005.rds** and **Cass\_data2005.rds**

**bow\_shock\_only\_Data.csv** and **Bow\_Shock\_Data.csv** : Contains only the bow Shock Data

**15minutes\_interval\_boeshockData.csv** : Contains all the Data that is 15 minutes before and after a bow shock occurs

**15minData\_Wider.csv** : Contains all the Data that is 15 minutes before and after a bow shock occurs as columns

**data\_version\_1** Inside **data\_version\_1** Folder there are three files which are the slightly modified versions of **bow\_shock\_only\_Data.csv**, **15minutes\_interval\_boeshockData.csv** and **15minData\_Wider.csv** files in **data\_version\_0** Folder. Files and its Explanations

**bow\_shock\_only\_Data.csv** : Contains only the bow Shock Data

**15minutes\_interval\_boeshockData.csv** : Contains all the Data that is 15 minutes before and after a bow shock occurs

**15minData\_Wider.csv** : Contains all the Data that is 15 minutes before and after a bow shock occurs as columns

**data\_version\_2 (Not Available in GitHub)** Inside **data\_version\_2** Folder there are six files which are created for training the different Models I created. NB: Not pushed to the github repo because of Large Size. Files and its Explanations

**Average\_Data.rds** : Contains - average value of magnetic properties 15 minutes before each event and average value of magnetic properties 15 minutes after each event

**Average\_SD\_Data.rds** : Contains - average value of magnetic properties 15 minutes before each event, average value of magnetic properties 15 minutes after each event, standard deviation of magnetic properties 15 minutes before each event and standard deviation of magnetic properties 15 minutes after each event

**Full\_Data\_Wider.csv** and **Full\_Data\_Wider.rds** : Contains all the Data points (about 49000) and 15 minutes before and after each points as columns (dimensions are 49000,210)

**input\_dataset.rds** : Contains all the Data in **Full\_Data\_Wider.csv** and **Full\_Data\_Wider.rds** but imputed the NA values with required values and also it doesn't contains the variable **Timestamp**.

**input\_dataset.rds** : Contains all the Data in **input\_dataset.rds** but doesn't contains the variable **dirn\_cross**

## Figures

Inside this folder it contains some of the sample graphs I saved during Exploratory Data Analysis. Files and its Explanations All files are either in **png** or **pdf** format.

## Literature Surveys

Inside **Literature Surveys Folder** it contains some of the papers that i refered during my research work for this project and it also contains a documents which i created to prepare notes. Files and its Explanations

**.docx / .txt** : Files with this extensions are my notes

**.pdf** : Files with this extensions are the papers I refered.

## Presentation

Inside **Presentation Folder** it contains only the pptx file which i presented in the project review meeting. Files and its Explanations

**.pptx** : This is the presentation file I used. It contains a brief description of the project so far

## Tasks Alloted

This project was divided into different tasks to make it more understandable. Each task has its own folder and files. NB:You can use the tab below to read more about each folders and the files in it.

**Task 1 (Merging the Data)** Inside **Task 1 (Merging the Data) Folder** it contains three files related to merging the **Cass\_data\_ksm2005.rds** dataset and **Full\_Cassini\_Master\_MP\_BS\_CMJ\_revised2005.rds** dataset by using **Timestamp** variable as the key. Files and its Explanations

**Task.R** :This a draft file that contains the code i wrote for merging the two datasets.

**Task.Rmd** : This Rmd File contains all the information including a brief description of each chunk of code that i wrote in **Task.R**

**Task.html** : HTML knitted version of **Task.Rmd**

**Task 2 (Data Creation 15 minutes)** Inside **Task 2 (Data Creation 15 minutes) Folder** it contains nine files related to creation of different datasets for training the models and also did some basic exploratory analysis on created Data(not a detailed one) Files and its Explanations

**Task 2 Draft.R** :This a draft file that contains the code i wrote for creating the files below.

**bow\_shock\_only\_Data.csv** : Contains only the bow Shock Data

`15minutes_interval_boeshockData.csv` : Contains all the Data that is 15 minutes before and after a bow shock occurs

`Exploratory Analysis on Bow Shock Data.R` : This a draft file that contains the code i wrote to do exploratory data Analysis on the Datasets i Created.

`Exploratory Data Analysis.Rmd` : This Rmd File contains all the information including a brief description of each chunk of code that i wrote in `Exploratory Analysis on Bow Shock Data.R`

`Data Creation (15 min duration).Rmd`, `Data-Creation--15-min-duration-.Rmd` and `Data Creation -15 min duration.Rmd` : This Rmd Files contains all the information including a brief description of each chunk of code that i wrote in `Task 2 Draft.R`

`Exploratory-Data-Analysis.html` : HTML knitted version of `Exploratory Data Analysis.Rmd`

`Data-Creation--15-min-duration.html` and `Data-Creation--15-min-duration-.html` : HTML knitted version of `'Data-Creation--15-min-duration-.Rmd` and `Data Creation -15 min duration.Rmd`

**Task 2 (Making Wider Data)** Inside `Task 2 (Making Wider Data)` Folder it contains five files related to creation of 30 minute window for every data points (about 49000 rows). NB: It takes a good amount of time to process the complete dataset. There is a completed Dataset in this folder Files and its Explanations

`Task 2(Wider Data) Draft.R` : This a draft file that contains the code i wrote for the creation of new columns (30 minute window).

`15minData_Wider.csv` : This is the csv formatted data that has BowShock with 30 minute window columns.

`Task 2(Wider Data).Rmd` : This Rmd File contains all the information including a brief description of each chunk of code that i wrote in `Task 2(Wider Data) Draft.R`

`Task-2-Wider-Data-.html` : HTML knitted version of `Task 2(Wider Data).Rmd`

**Task 3 (Exploratory Data Analysis)** Inside `Task 3 (Exploratory Data Analysis)` Folder it contains four files that is related to EDA of the newly created wider dataset. Files and its Explanations

`Draft (Analysis).R` : This a draft file that contains the code i wrote for EDA of the new 15 minute window wider dataset (`Full_Data_Wider.rds`).

`Data Analysis.Rmd` : This Rmd File contains all the information including a brief description of each chunk of code that i wrote in `Draft (Analysis).R`

`Data-Analysis` : HTML knitted version of `Data Analysis.Rmd`

**Task 4 (Model Creation)** Inside `Task 3 (Exploratory Data Analysis)` Folder it contains four files that is related to EDA of the newly created wider dataset. Files and its Explanations

`Draft (Analysis).R` : This a draft file that contains the code i wrote for EDA of the new 15 minute window wider dataset (`Full_Data_Wider.rds`).

`Data Analysis.Rmd` : This Rmd File contains all the information including a brief description of each chunk of code that i wrote in `Draft (Analysis).R`

`Data-Analysis` : HTML knitted version of `Data Analysis.Rmd`

**Task 5 (Complete ToDo List)** s

**Thesis Report Files**

Hi there