The Agile Crisp data science project methodology combines a general set of phases to do data science with a visual pipeline process management. The system’s highly visual nature (by leveraging a board) allows teams to communicate more easily on what work is being done and what needs to be done. It allows agility since it shows the team the work as moves across the board and knowledge from previous tasks can be used to help determine what should be started next.

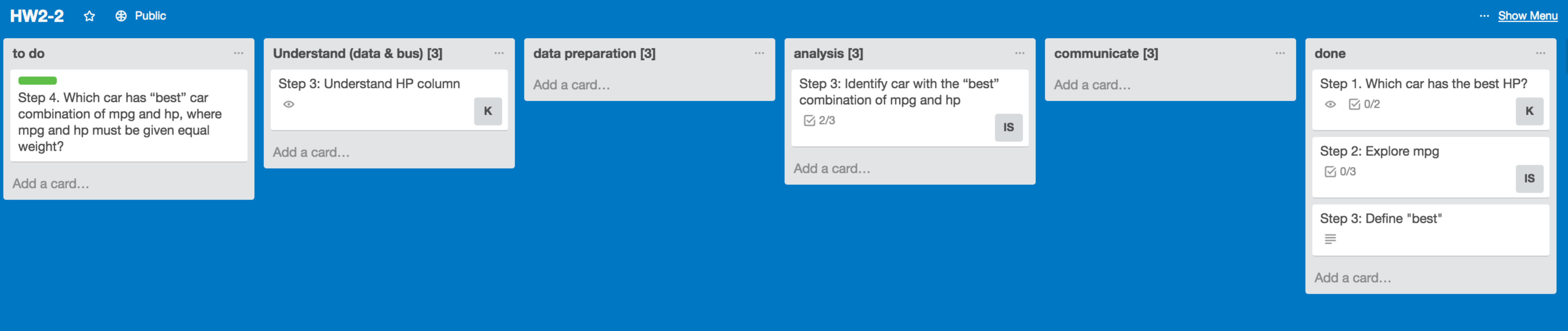
**Steps of the Agile Crisp Methodology**

From a task perspective, each team will define what to investigate (i.e., high level tasks, such as “link weather data to our previously collected data”). These ideas will all be listed (in a prioritized order) in a “to do” column. Then, as the team prioritizes the work, tasks flow to the next column (such as “understand”).

When the task is done, it gets moved to the next column across the board. Some of the tasks might just focus on preparation or analysis. That is fine; it just takes zero work to move through a phase when that task has no work in that phase. As the team deems appropriate, new tasks can be started.

The team collectively decides how large the “chunks of work” are – but the smaller / more detailed, the easier it will be for the team to understand potential bottlenecks.

A sample board is shown below. The appendix shows two possible boards (columns) to use. The team needs to collectively determine what are the best columns for the team to use.

****

**Summary**

In other words, the process to do data science can be thought of as a pipeline with requests entering one end and improved data insight coming out the other end. The team works through the project pipeline throughout the project with no defined schedule. The goal is to make sure, at the end; there was not a lot of time spent on an effort that did not complete (better to get a fewer number of tasks all the way through the pipeline).

**Trello: our Getting Things Done visual Tool**

Trello allows planning, prioritizing, and tracking anything you’re working on. It’s primary purposes is keeping a [GTD](https://en.wikipedia.org/wiki/Getting_Things_Done" \t "_blank) (Getting Things Done) framework. It allows cross-team communication by showing a visual workflow. It also has an activity history.

**Reviewing the Key Principles of the Methodology:**

Split the work into pieces; write each item on a “card” and put on the “wall” and using named columns to illustrate where each item is in the workflow. Making the work visible—typically leads to increased communication and collaboration.

**What is Required for Project Updates:**

For each project update, the team needs to submit a one page summary (powerpoint), that lists what has been completed since last update, what is your plan for next update, any issues/roadblocks and what is working well (see the example at the bottom of this page).

In addition, the team should estimate how much of their time they spent in the different phases of data science (which is also shown below). Finally, in addition to the one page update, please provide a link to your Trello Board, with the following email accounts allowed to view the board ([jsaltz@syr.edu](mailto:jsaltz@syr.edu), [gekrudys@syr.edu](mailto:gekrudys@syr.edu), [esande01@syr.edu](mailto:esande01@syr.edu), and [ivanshamshurin@gmail.com](mailto:ivanshamshurin@gmail.com) )

|  |  |  |  |
| --- | --- | --- | --- |
| **Understand** | **Prepare** | **Analysis** | **Communicate** |
| X% of time | X% of time | X% of time | X% of time |

**What is working well**

* One
* Two
* Three

**Accomplishments this past period**

* One
* Two
* Three

**Plans for next update**

* One
* Two
* Three

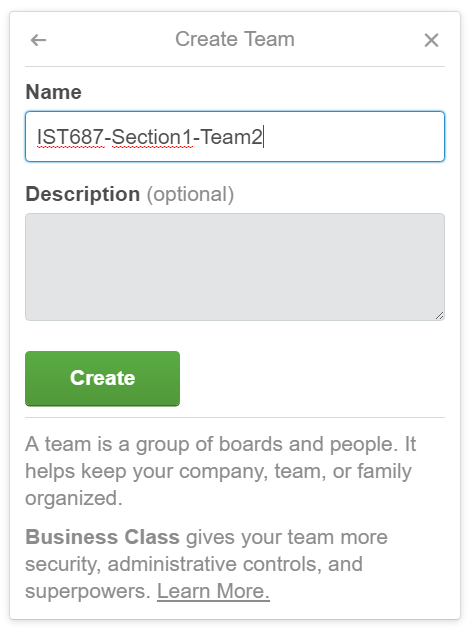
**Issues / What is not working well**

* One
* Two
* Three

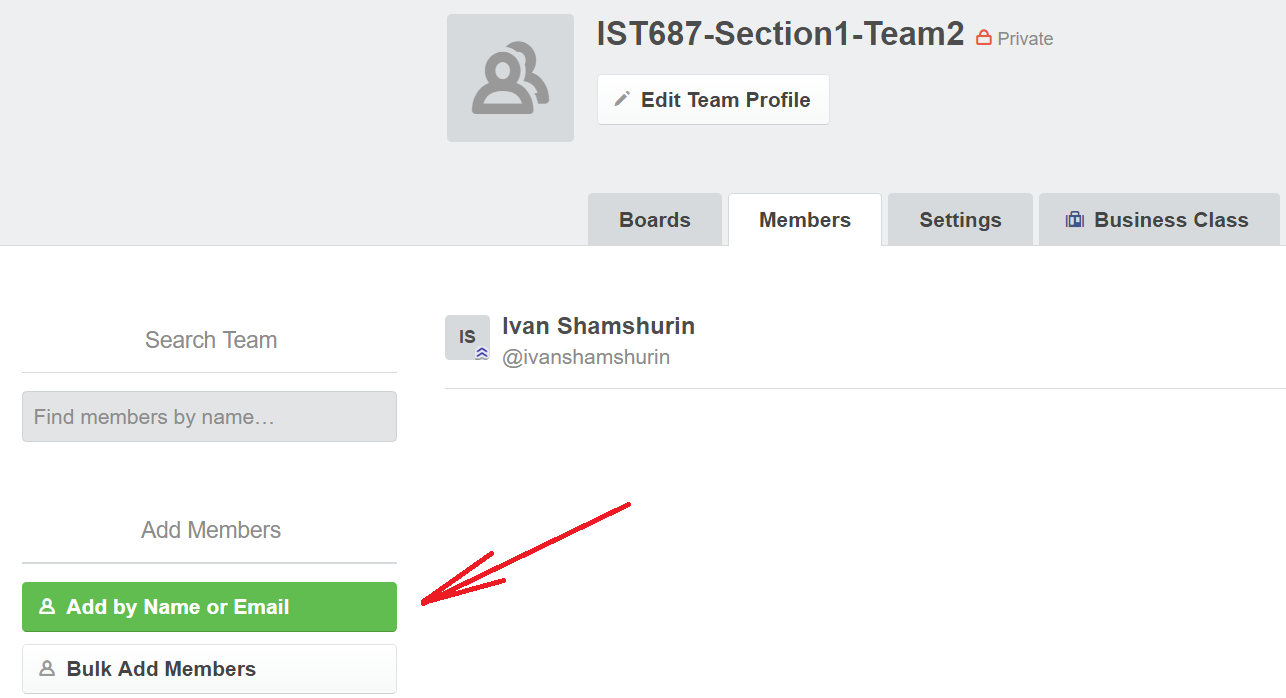
**Appendix 1: Using Trello**

Step 1. Create an account and log in to trello - <https://trello.com/>

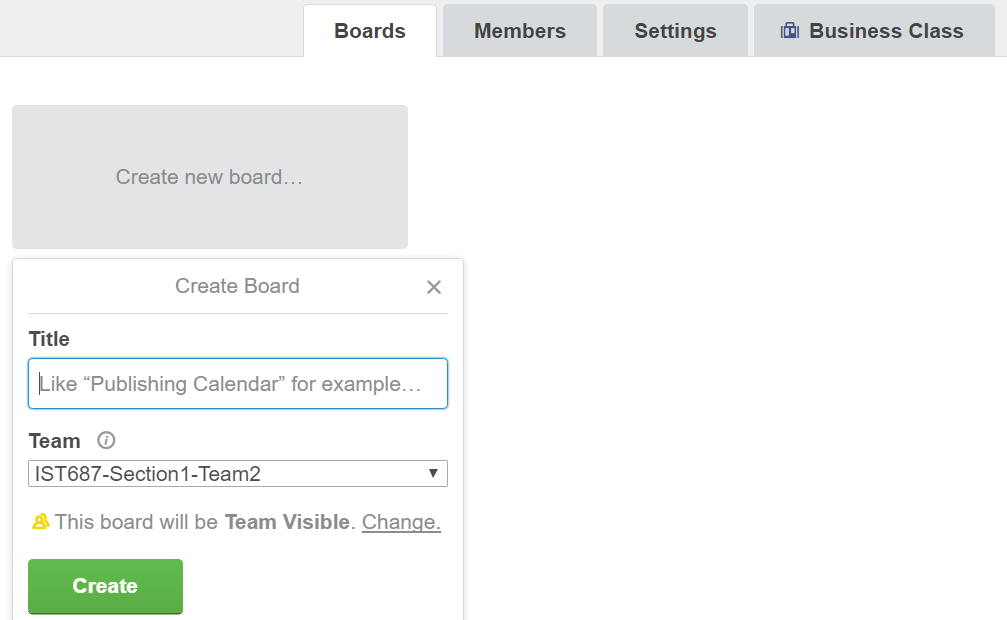
Step 2. Click on “Create a new team”



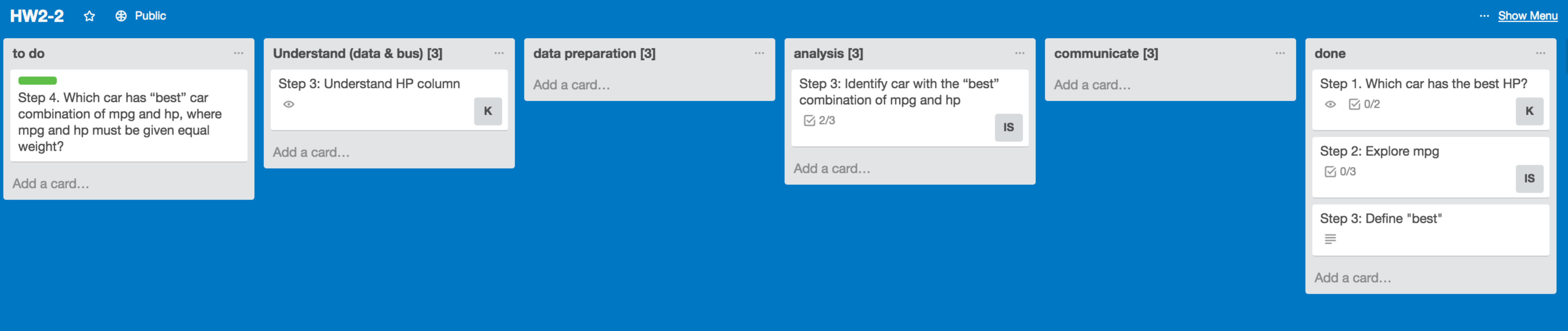
Step 3. Add members to your team by their email or Name



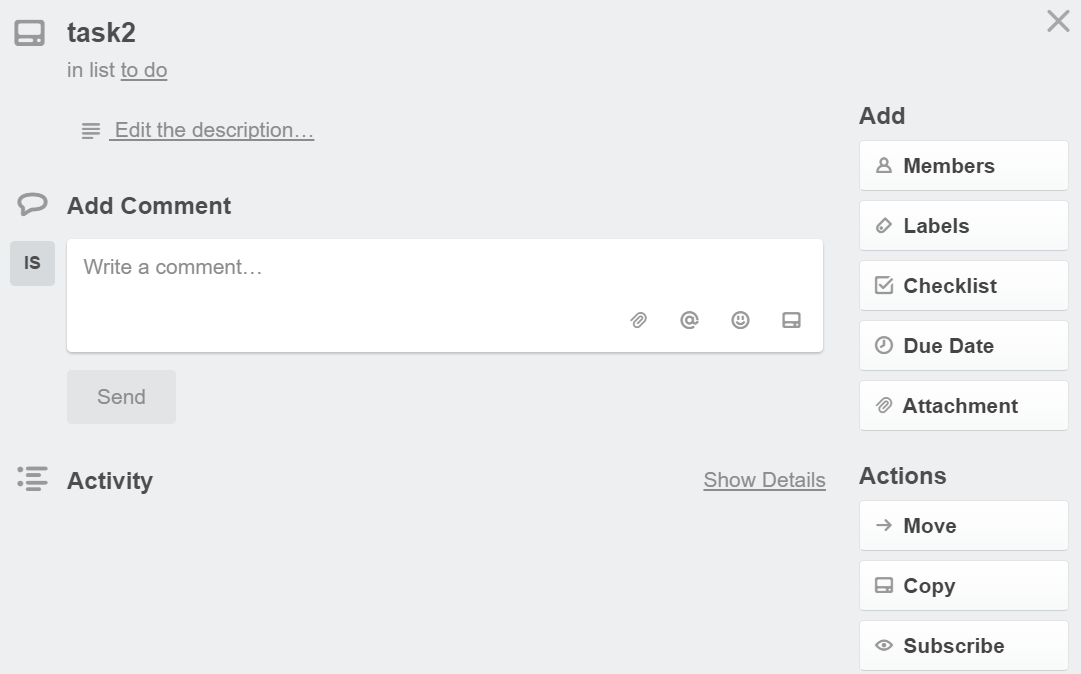
Step 4. Create a new board within the team and put the title for the board.



Step 5. Create columns (phases of Kanban methodology) and tasks (cards).

****

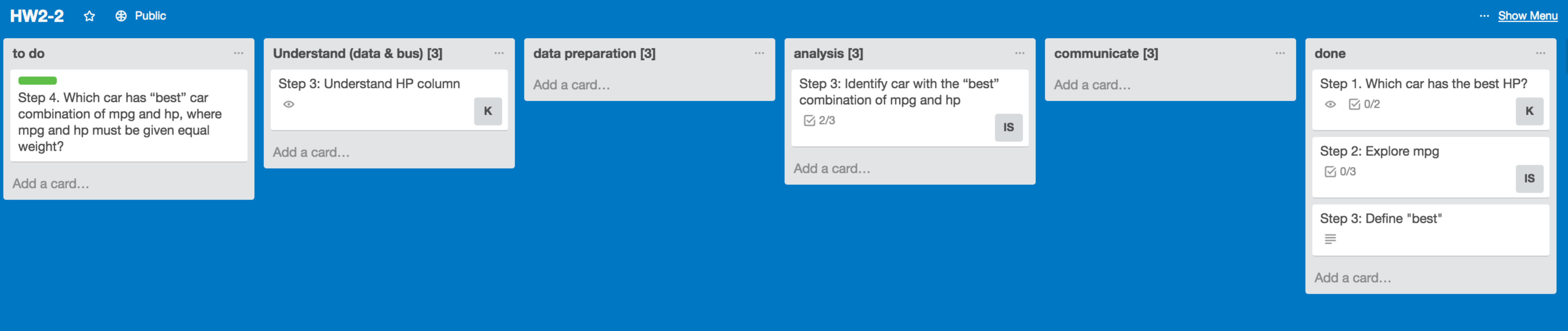
Step 6. When creating a task, it is possible to change/add its label (color), due date, attachments and most importantly ***teammates responsible for this task***.

****

**Appendix 2: Examples of working with Trello**

Below is an example of using the board for a homework assignment. There are two different boards, showing two alternatives for how to define the columns on the board.

[https://trello.com/b/jb87nnWb](https://trello.com/b/jb87nnWb" \t "_blank)

****

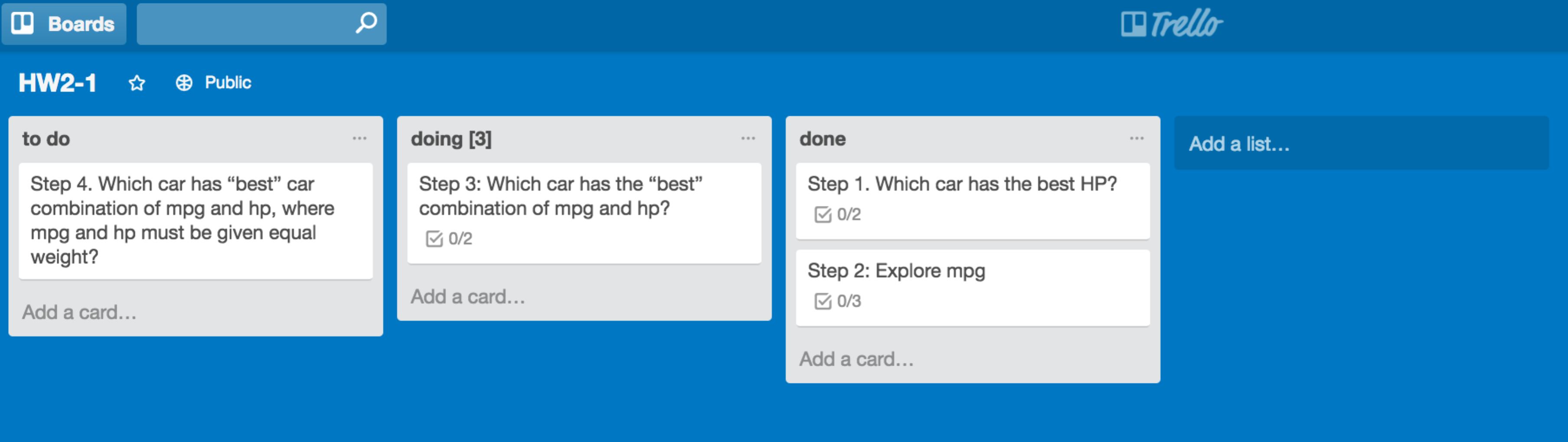
Each of the phases is briefly described below:

* **To do** - create list of possible tasks (put tasks into the “to do” bucket)
* **Understand** (Business & Data) - Team understands the business value of the task and also understands data to use (sources, meaning)
* **Preparation**: This phase includes data capture and data extraction.
* **Analysis**: This phase includes Modeling the data (ex. creating a linear model) as well as visualizing the data.
* **Communicate**: This final phase focuses on trying to ensure that the team has performed a valid analysis, and so, includes any testing / validation required by the team, as well as sharing the results (ex. writing up the results of the analysis in an easy to understand manner).
* **Done** – showing all the tasks that have been completed

However, the team is free to define a different board, if the team feels it will be easier to use. One simplified version is shown below, where the board just shows work to be done, work in progress and work that is finished.

*A simpler version:* to do, doing, done

[https://trello.com/b/5avpFOL5](https://trello.com/b/5avpFOL5" \t "_blank)

****