

# Report: Iteration 2

DS 5110: Introduction to Data Management and Processing

Section 4

Krishnan Narayanan

Neha Reddy Palnati

Reghunaath Ajith Kumar Ahila

[GitHub](#)

## Contents

1	PROJECT KICKOFF	1
2	TEAM DISCUSSIONS	2
3	SKILLS AND TOOLS ASSESSMENT	2
4	INITIAL SETUP	2
5	PROGRESS REVIEW AND PLAN REVISION	2

## 1 PROJECT KICKOFF

As international students unfamiliar with the housing market in Boston, we faced significant challenges in choosing the right neighbourhoods, understanding the average rental costs in various areas, and navigating the overall housing landscape. A solution like what we are planning to build would have been incredibly helpful, providing us with the necessary insights and guidance to make informed decisions. To avoid scope creep, we have clearly documented and prioritized the requirements upfront. Any new requirements will only be considered once all existing requirements have been fully completed.

### Deliverables and Milestones

Phase-1: Collect the data and determine the platform on which the interactive dashboard should be created, between Streamlit and Plotly.

Phase 2: The data will be cleaned, and the interface will be developed in the platform that has been agreed upon.

Phase-3: Review the dashboard for any inconsistencies and validate that the functionalities that were listed are present.

Phase 4: After the dashboard has been finalized, host it as directed. Additionally, discuss potential future enhancements and updates.

The team's capabilities align with the objectives that have been established, as all members have prior experience in data mining, applying a variety of visualization tools and techniques, and developing a dashboard, etc.

We are currently in the process of extracting data from a variety of online sources in order to utilize it for the project, as the current relevant datasets that are available online do not meet the requirements.

## 2 TEAM DISCUSSIONS

### Skillset and Tools

All the team members; Neha, Reghunaath, and Krishnan are experienced in working with Python and basics of Exploratory Data Analysis. Apart from the common skillset that we have, each one of us has some unique skillset that each of us brings to the table. Here are some specific skills that each of us have:

- Neha: StreamLit, Data Pre-Processing, Git, Visualization.
- Reghunaath: Data Mining, Domain knowledge, Dashboard.
- Krishnan: Plotly and Interactive plots, Visualization, Dashboard.

Based on the project objectives and the team's expertise, we have decided to use the following programming languages and platforms for the development phase:

- Python
- Plotly.Dash/StreamLit
- GitHub

## 3 SKILLS AND TOOLS ASSESSMENT

The required tools, frameworks, and libraries are as follows:

- Pandas
- NumPy
- Plotly.Express
- Seaborn
- BeautifulSoup
- SciPy
- Scrapy
- Plotly.dash/StreamLit

The Frameworks were determined after a discussion of the team members' skillsets and expertise to guarantee that all members are at comfort with the tools. The tasks have been allotted to each member in accordance with their strengths and weaknesses.

## 4 INITIAL SETUP

We have decided to go with PyCharm, Jupyter, and command line as our main working environment.

We have installed and configured all the required software, libraries and tools.

We can run Integration testing to ensure the functionality of the environment.

For troubleshooting, we can create system logs so that it will be easier to track down the issue, identify the problem, test the components of the code and run a network diagnostic test to see the problem and take an action accordingly.

## 5 PROGRESS REVIEW AND PLAN REVISION

We have completed the initial setup and finished setting up a GitHub repository and started the data collection process.

We haven't encountered any issues so far, everything is right on track.

Yes, each team member contributing and understands their role as how we discussed earlier.

So far, we are on track with the timeline and milestones, but we might have to adjust it soon. in accordance with the intensity of the coursework, assignments and the upcoming exams.

We have setup a group in Microsoft teams and discord to establish efficient communication within the team and we have prepared an excel sheet to track the progress with respect to the overall project goals and checkpoints.