* How did your team approach the initial stages of exploring a large, complex, and unfamiliar dataset?

It was great working with Khushi and Zakir during the data fest. Dataset was too huge with 3-4 different dataset files. We tried to breakdown the dataset reading codebook and understanding the connection between each dataset to create a story and analyze the dataset. The dataset was really great to play around. It was complex yet interesting dataset where we understood great information regarding leasing in United States.

* What was your strategy for dividing tasks and working collaboratively under time pressure?

Strategy as such was nothing firstly we started to understand what is data and how do we turn this into analysis with great outcomes. Time pressure was truly hard to manage with such large data but we tried to focus on certain variables rather than entire dataset. We collaboratively did all the task with each others understanding and insights. We could not try much due to less time constraint but we tried to focus on one outcome to provide some insight using data answering the challenge.

* What challenges did you encounter, and how did you overcome them?

One big challenge we faced was tones of missing data around 1.5lakh data was missing from one sheet so it was difficult to come with exact details as they were not available in dataset. We tried to focus on Top-15 market which had heavy impact post-covid or had most leasing space occupied where the missing data didn’t bother a lot. We also tried to get rid of missing data but it screwed up the entire dataset.

* What decisions did you make about what to analyze and how to present your findings?

Initially, we were confused looking at large dataset but after understanding codebook and data we tried to combine the links between some variables and find story to provide good outcomes. We started from Top-15 markets then we connected that to analyze where does the top-15 market lies whether it was Central Business District or Suburban market. Connecting to that post-pandemic effect and unemployment effect on leasing in those top-15 markets.

Moreover, we tried to connect all the analysis with one another and make a story to answer our research questions or the challenge.

* If you could do the challenge again, what would you do differently?

It would be great to take challenge again. I would stick to understanding what data is and how can it be used effectively to answer the challenge. Focusing on minor trends than major would give more insights regarding the dataset. I would practice more on presentation to understand the outcomes and ways to make it easier to understand for the audience. Also, including external data to explore more deeply would provide great outcomes.

* How did this experience compare to prior data challenges or hackathons that you have participated in?

If I had to compare to other data challenges or hackathons. It was pretty similar experience but this data was quite large and complex at the same time with lots of missing data and moreover important data to work on. It was great working with this with focused and different perspectives to look over the given challenge. It was fun and a great learning experience.