**We have accomplished:**

1. The GUI
2. Upload .csv files
3. Use Spark parse the file and data
4. Get required data from Spark
5. Source data analysis
6. Test data file make

**Next:**

1. Let the user choose which city to make prediction.
2. Separate source data by each point.
3. Collect all data for each same point from different year.
4. Make prediction (machine learning) for each point.
5. Draw 2D graph of the PM2.5 concentration.

**I am thinking of:**

1. If time allows, try to make a 3D graph for all points.