MINUTES OF MEETING 2

|  |  |  |  |
| --- | --- | --- | --- |
| **Meeting/Project Name:** | Estimating & Comparing Public Transport Emissions using gtfs2emis | | |
| **Date of Meeting:**  (MM/DD/YYYY) | 23/02/2023 | **Time:** | **11:00 A.M** |
| **Meeting Facilitator:** | Nishi & Hemant | **Project Status** |  |

**The meeting was called to discuss the implementation of General Transit Feed Specification (GTFS) to estimate emissions from public transportation systems.**

1. Discussed on the analysis of the Detroit data and how to proceed with other countries data and performing EDA on it.
2. Discussion on the Machine learning models which will be helpful.
3. Discussion on the Team Presentation.
4. Discussion on the how to compare all the Cities data.

**Meeting Objective**

|  |  |  |
| --- | --- | --- |
| **Attendees** | | |
| **Name** | **Student Id** | **E-mail** |
| Hemant Chowdary | 0788804 | [W0788804@myscc.ca](mailto:W0788804@myscc.ca) |
| Sai Krishna | 0789428 | [W0789428@myscc.ca](mailto:W0788804@myscc.ca) |
| Eswar Kiran Pathuri | 0788366 | [W0788366@myscc.ca](mailto:W0788801@myscc.ca) |
| Nishi Shrivastava | 0770047 | [W0770047@myscc.ca](mailto:W0770047@myscc.ca) |
| Prayas Baliyan | 0790447 | [W0790447@myscc.ca](mailto:W0790447@myscc.ca) |

|  |  |
| --- | --- |
| **Project Daily Activities for this week** | |
| **Topics** | **Owner** |
| 1. After combining the data from the MOVES and EMFAC models, an analysis was conducted to review the EDA findings. 2. The analysis revealed variations in pollutant emission levels when the data from both models were combined. 3. During the analysis, an inconsistency in the pollutant units within the MOVES data was identified, leading to the creation of an issue in the GTFS2EMIS Github repository. Collaborators responded to the issue, and the pollutant units were confirmed to be in grams per kilometer. 4. Helped the team in data assessment 2 by performing data Science Ethics Checklist 5. Ticket raised and closed <https://github.com/ipeaGIT/gtfs2emis/issues/96> | Prayas |
| * I used the R programming language to combine the USA EMFAC and MOVES datasets, removing unmatched columns, and visualized the observed emission values of different pollutants in the combined dataset using facet wrap. * Helped the Team for the Data Assessment 2 by performing the Data Cleaning & research Questions. | Sai Krishna |
| * Understanding in detail regarding R package of the gtfs2emis package. * Worked on the Data Assessment 2 feedback and find out the scaling needs to be done for the **Emission** column as EC pollutant values were very high. * Worked on the Different Scaling techniques to see how the values show after performing like **Log Transformation & Box- Cox transformation.** * Finally, we found out log transformation is better by adding the **constant Value** to all the ef values to handle the negative values. * Working on the Windsor data and trying to get the emission data out of it * Performed EDA on the Combined data and scaling data both the EMFAC and moves data. * Helped the team with Data Assessment 2 on Bivariant Analysis * Understanding the existing research paper <https://osf.io/8m2cy/> | Hemant |
| * Worked on the scaling technique **square root** transformation. * Worked on the data assessment 2 feedback points and helping with EDA part with Hemant. * Commenced preparations for the team presentation and consolidated key points. * Working on the Machine learning model for the Detroit data * Performed EDA on the Combined data and scaling data both the EMFAC and moves data. * Working on the Research Questions * Helped the team with Data assessment 2 by performing the Multivariant Analysis and Data Quality part. * Working on the Toronto Data to get the emission data out of it. * Study of the different emission estimation models available emission\_model() and their suitability for estimating emissions from public transportation systems. | Nishi Shrivastava |
| * Helped the team with data assessment 2 by performing the Univariant Analysis. * Working on the Machine learning model for the Detroit data * Working on the Ireland data and performing the EDA | Eswar Kiran Pathuri |

|  |
| --- |
| **Challenges** |
| 1. How to import the other country data to the R package as the package expects data in a particular format. 2. Facing Challenges scaling techniques as most of the techniques generating the negative values. 3. Facing the challenge on how to compare the different countries as they have different columns. 4. Understanding the different emissions standards as different countries use different methodology and rate them accordingly. |

|  |
| --- |
| **Planned Activities for Next Week** |
| 1. All the Team Members will be working on the Brazil and Ireland as we have a good understanding of the data. 2. All the team members will be working on the EDA part of Brazil and Ireland. 3. Preparation of the Project Presentation and team presentation 4. Team will be working in the Toronto data and Windsor try to get the emission data out of it. |