|  |  |
| --- | --- |
| **Titel of your Project** |  |
| Mobility in Hamburg - Seasonal effects and geographic distribution of car/bike traffic | |

|  |  |
| --- | --- |
| **Short summary (4-5 sentences)** |  |
| The city of Hamburg installed a large amount of infrared sensors around the city where the number of cars and bikes passing by are counted. The up-to-date results of certain time intervals are published and constantly updated online (Car: <https://metaver.de/trefferanzeige?docuuid=2936465E-C045-4F5D-8614-24C3FBB522E2>; Bike: <https://metaver.de/trefferanzeige?docuuid=9072E37F-8505-41F0-9332-B80C02C7E802> )  Our main interest was to visualize and compare the traffic for bike and car data during the last year, on the one hand visualizing the overall traffic volume measured over the course of the year and the changing behavior of traffic over the different seasons. Also, geographically analyzing the traffic distribution over Hamburg. | |

**- The first two parts will also appear on your Digital Shaper certificate!**

|  |  |
| --- | --- |
| **Main section** |  |
| **Introduction:**  Since Mobility is a relevant topic in every big city we wanted to dig deeper in understanding how traffic in Hamburg behaves. Therefore we found an analysis of geo data from digital counting points for bikes and cars a very interesting topic to look into.  **Methodology:**  A very big challenge we had right in the beginning was to extract a “usable” dataset from the initial data source: The traffic data was provided in JSON format we had to access via an API request (Application Programming Interface). To receive data we had to make a request to the server, through which we could ask for the data in 15 min / hourly/ daily/ weekly intervals. We chose the daily version. No one of us has worked yet with Python and loops , so data extraction at the beginning was a major challenge, as various parts of the data we needed was hidden behind links, but we needed those in our final data set for analysis, so it took a long time until we had looped through the various “levels” and created a data set to start cleaning/describing/analysing.  After we got all the data the first step was to modify and clean the data so it was easy to handle. One thing we added, for example, was the weekday of each date in the dataset. Therefore we needed a for-loop and the knowledge of how to split up strings. Also we enriched the dataset with the weekly mean for each counting point, so we could visualize the weekly numbers over the course of the last year.  After we cleaned and enriched the dataset we started visualizing in terms of the seasonal effects. First we started visualizing all counting points for every day with their respective results, colored differently according to the season (spring, summer, autumn, winter):  Car data: Bike data:    but for a better overview we ended up using seaborn barplots (sns.barplot) to look at the mean values for all measured vehicles in weekly intervals (in the pictures week 2020 - 44 to week 2021 – 32 are shown):  Car data: Bike data:    Also we compared bike and car traffic over the course of the previous year:    Secondly, we tried to figure out how to visualize the geo data in form of a map. Here we didn’t have any knowledge from our course but our mentor could help us out in this case and suggested some tools we tried out. During the online learning course we got to know some ways how to look up problems online on our own. That was really helpful during the whole project phase. We learned the usage of plotly express for geographically plotting the Zählstelle locations on the map. This helped to pinpoint the locations of the source of the data. Secondly, we used pydeck.gl, in order to visualize the vehicle traffic distribution of throughout the year within Hamburg. Alternatively, we have used Kepler.gl as well to get more geo visualization like clustering of the data points and understanding of the traffic distribution better. We have observed that most of the traffic is focused on the center of Hamburg and spread evenly across major entry point to the city.    **Results of the project:**  The data showed us (as expected) that there is far more car traffic than bike counts. Both sort of vehicles have weeks with far less traffic than usual, for example between Christmas and during the holidays. Nevertheless, for bikes there is a greater difference between summer and winter weeks, while for cars this difference is less pronounced. We also observed that, even though the number of bikes is lower as compared to the cars, the traffic behavior tends to show similarities (e.g., most of the traffic is concentrated at the central city irrespective of type of vehicle).    Unfortunately, the daily data is only provided for maximum the previous year. As this was a year with the pandemic also influencing traffic streams due to home office, lock-down, etc. it would have been interesting to have comparable data from the pre-corona years, so our analysis unfortunately couldn´t find out whether and what influence the pandemic had on the traffic. In general, further projects could also use data from this platform to analyze 15-minute, - and hourly intervals to get for example insights into rush hours, etc.  **Your learning journey:**  In the beginning it was unfamiliar to just work remotely. We learned that a good and constant communication is really important. Sometimes tasks weren’t split up clearly so two people worked on the same problem and sometimes it wasn´t communicated to the group that someone was having issues with his task, so the project got stuck. Also sadly 2 of our members left the project on the way, so we ended up with 3 members instead of 5. In the end all remaining group members learned a lot but of course it did cost us some time. Overall we can say that coding is a lot of fun but you have to keep at it and it is no shame to ask really simple questions because it is always easier to work on difficulties together. | |

|  |  |
| --- | --- |
| **Further details** |  |
| * Names + learning track of all active team members * Maren Hagemeister: Data Science * Anshul Tomar: Data Science * Patricia Sewing: Data Science * Names of your mentor * Philipp * Optional: linking LinkedIn/XING profiles, linking your GitHub repo, roles/tasks distribution within the team. | |