The effect of pandemic measures on different modalities in Amsterdam: A dashboard visualisation

Joris Bruil, Arno Garstman, Ivo Cornelis de Geus, Megiel Kerkhoven, Martijn Schendstok
Supervisor: Frank Nack
University of Amsterdam
Amsterdam, The Netherlands
In collaboration with:
City of Amsterdam
Amsterdam, The Netherlands

Figure 1. Empty Amsterdam, Reddit 2020.

Abstract

CCS Concepts: • **Human-centered computing** → *Geographic visualization*; *Visual analytics.*

Keywords: Data visualisation, COVID-19, Amsterdam, Transportation

1 Introduction

Through the ages, Amsterdam has faced several pandemics and epidemics (e.g., Spanish flu and smallpox). However, until the rise of the corona virus, modern generations were spared from such pandemics. In order to fight the negative consequences for human life, policy makers have attempted various public interventions to fight the virus. However, due to the novelty of the virus, little is known of the effectiveness of various interventions developed by policy makers on the behaviour of citizens. Therefore, the City of Amsterdam has requested an interactive visualisation which gives insight on the effectiveness of certain measures on public behaviour. In order to make such an interactive visualisation, the City of Amsterdam provided us with a dataset containing traffic data on various modalities (e.g., pedestrian, public transport and parking). Using this data, this project aimed to build an interactive visualisation that will give decision makers insight on the effectiveness of corona measures on the use of various traffic modalities.

With the visualisation we want to answer the following question; How can we effectively visualise the effect of corona measures on traffic modalities? To answer this question we look into the following sub-questions, the visualisation should be able to answer these question:

- How does the distribution of usage change over the day?
- How do measures impact transit crowdedness in different boroughs?
- How does the effect of measures change over time?
- How do corona measures influence foot traffic at key locations?
- How can we clearly present data of different spatial and temporal dimensions to analyse impact of measures?

2 Literature Review

2.1 Heuristic Evaluation

In order to optimise the usability with the user interface design of the dashboard, the Heuristics of Nielsen[10, 11] were used. Although not all the heuristics were equally applicable to every interface, the heuristics provide ten widely accepted rules of thumb to critically improve the interface of the dashboard. In this section, we will briefly describe each the ten heuristics.

Visibility of system status. This heuristic refers to the rule that the system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

Match between system and the real world. This heuristic refers to the rule that the dashboard should contain concepts that are familiar to the user.

User control and freedom. The second heuristic refers to control and freedom, which means that the users should be free to leave any unwanted state without having to to a extended dialogue. This is an important feature, as users often choose system functions by mistake.

Consistency and standards. The third heuristic refers to consistency and standards of the interface. This heuristic refers to the rule that users should never have to wonder whether different concepts (e.g., words, situations, or actions mean the same thing).

Error prevention. The fourth heuristic refers to error prevention. At all times, it should be avoided that users get error messages. One way to comply with this heuristic, is to eliminate error-prone conditions or check for them.

Recognition rather than recall. The fifth heuristic is known as "recognition rather than recall". This means that users should be able to recognise information in the dashboard, rather than recall it. This is achieved by making objects, actions and options visible and recognisable in the dashboard. In other words, the user should not have to remember information from on part of the dashboard to user another part of the dashboard.

Flexibility and efficiency of use. This heuristic refers to rule of creating additional options to sidestep the regular interaction process. In other words, this rule indicates that a satisfactory dashboard should have options to take shortcuts in the interactions.

Aesthetic and minimalist design. The last heuristic we have considered refers to aesthetic and minimalist design. Users should not be confronted with information which is irrelevant or rarely need. As every piece of information enlarges the cognitive load of a user, it is important that the information displayed on the dashboard is concise.

Help users recognise, diagnose, and recover from errors This heuristics refers to the rule that error messages should contain information that helps the user recover from certain errors. In other words, the error message should precisely indicate the problem, and constructively suggest a solution.

Help and documentation Although a good system should be used without documentation, this heuristic refers to the rule that a good systems sometimes also need external explanation

2.2 User Testing

Where heuristic evaluation provides the groundwork for usability optimisation, user testing suits itself for refining the usability. Using [9, Chapter: 'Usability Testing'] as inspiration a test setup could be constructed to efficiently gather more insight for usability optimisation. Lewis describes three steps to accomplish a well structured test.

- Determining appropriate measurements
- Formulating Goals
- Collecting Data & Formulating Goals

2.2.1 Determining appropriate measurements. Lewis examines several different frameworks for appropriate. measurements but ultimately concludes the following to be the most important:

- Successful task completion rates
- Mean task completion times
- Mean participant satisfaction ratings

2.2.2 Formulating Goals. The formulation of goals generally involves comparing results across different iterations of tests or different tests altogether. To this end the goals should be quantified. In this case however there is no historical test data to compare with and therefore the goals need to be formulated ad hoc.

2.2.3 Collecting Data & Verifying Goals. The testing is supposed to be executed till the aforementioned goals have been reached. This means that a certain quantity of tests needs to be conducted to subsequently conduct statistical tests. In this adjusted format however testing can be conducted till the ad hoc formulated goals have been reached.

3 Methodology

3.1 Data Sources

To create a useful dashboard, we were given access to internal data sources provided by the City of Amsterdam. These were accessed through a custom system and were internally distributed through Resilio Sync[1], a peer-to-peer file-synchronisation utility. In this section, we will further describe the different data sources which were looked into.

3.1.1 CMSA. The Crowd Monitoring System Amsterdam (CMSA) data consists of pedestrian traffic captured using different devices. The time window for CMSA data used in this case was from the 30th of July 2020 till the 13th of December 2020. The devices used to capture movement were as follows: 3D sensors, counting cameras, or WiFi sensors. All of these devices aim to capture movements of pedestrians and these measurements are represented in intervals of 15 minutes in separate files. Each device was distinguished by a unique code but the location of the devices wasn't present in the data. The binary direction of movement, relative busyness, average movement speed, the amount of inputs for the 15 minute summation were also present but not used in this case.

[1]Resilio Sync: https://www.resilio.com/

3.1.2 GVB. The Gemeentevervoerbedrijf Amsterdam (GVB, Amsterdam Municipal Transport Company) data set consists of information on bus, tram, and metro journeys in Amsterdam. The time window for this data set is from the first of January 2019 to the 12th of December 2020. The information is split into different files with each a different perspectives on the same general data. The first perspective was journeys compared to trips. Journeys contain information on the start of a journey and trips are part of these journeys. Next there was a difference between files which mention the stop a journey or trip started compared to files which mention the destination of a journey or trip. Next there was a difference between the origin of a journey and the destination. Every contained either the origin or destination per trip or journey. For the purpose of this dashboard the information of the complete journeys was taken with the destination information. The reason being that when a journey is finished at a certain destination the person in question will impact the local crowdedness. Important to note is that metros were letter coded in contrast to trams and busses which were indistinguishably numerically coded. This lead to trams and busses being grouped in the dashboard while metros were possible to display separately. One more important note is that a lot of data was unusable in its current form. This was because the stop was identified as unknown ('Onbekend') or other ('Overig'). This incorrect labelling was mostly present in 2019 and didn't render the visualisations unusable as can be seen in Figure 2. This is because there was still a strong contrast to be found between 2019 and 2020. The final remark for the data is that the coordinates present for each stop were swapped. What should've been longitude was stated as latitude and the other way around.

3.1.3 Resono. The Resono data set consists of pedestrian traffic information across Amsterdam. The time window for the Resono data ranges from the 30th of September till the 14th of December. Each file contains the data of one day from all 423 locations with 15 minute intervals. Ultimately the data showed very large values for certain areas which rendered implausible results. An example is Amstelveen-Zuid at 7:00-7:59 which has 22.342.238 pedestrians registered. This a very implausible scenario considering Amstelveen has approximately 90.000 inhabitants. It was found that the specialists at the city of Amsterdam were aware of this erroneous data and attributed the cause to an incorrectly functioning normalisation algorithm.

3.1.4 Coronavirus Government Response Tracker. The Coronavirus Government Response Tracker is a project created at Oxford University to track government responses around the world[8]. Using an extensive team of researchers, it maintains 20 different fields, the stringency of a measure in a certain country. This can range from school closings and public events to fiscal measures and public testing. For this project we took five different measures from the Coronavirus Government Response Tracker project, including the calculated meta-index called Stringency Index, as shown in Table 1. The Stringency Index is a simple additive of seven ordinal indicators, changed to range from 0 to 100. The selected measures were taken as we considered them the most relevant for measures taken in the Netherlands and they gave more context to the displayed statistics.

Table 1. Measures used in Dashboard as defined by the Coronavirus Government Response Tracker[8].

Code	Description	Range
C1	Closings of schools	0-3
C2	Closings of workplaces	0-3
C4	Restrictions on gathering	0-4
C6	Stay at Home orders	0-2
–	Stringency Index (meta)	0-100

3.1.5 Covid Measures Calendar & News. An existing excel-sheet delivered from the City of Amsterdam. This resource had a daily description of changes in measures, both including national, regional and local changes. It also notes significant changes in policy or events, such as the first contagion in Amsterdam.

Besides the measures received from the City of Amsterdam, other potential resources were scraped as well. These include the news regarding the pandemic published by the city[5] and daily reports by a dedicated dutch website aggregating news related to the pandemic[12].

3.1.6 Maps Data. The City of Amsterdam has publicly available maps data[1–4], this data can be used for the visualisation of spatial elements of the data. The specific data sets used for spatial visualisation are the quarters[1], the city district[2], and the tram and metro lines 2020[4] data sets. The maps data of the City of Amsterdam includes GeoJSON-files, these files are easy to implement in Tableau and include shape files made out of coordinates, in this case for the areas of the quarters and city districts of Amsterdam and the tram and metro lines.

The last maps data used from the City of Amsterdam is the "Sensors Crowd Monitoring System Amsterdam"[3] data set. This data set includes information on the CMSA sensors, including the coordinates for each sensor.

3.2 Data Preparation

All the data workflows used in this process have been published online in a public git repository, found on https://github.com/idegeus/amsterdam-covid-mobility.

3.2.1 CMSA. In order to use the CMSA data the temporal dimension needed changing and the sensor required geographic coordinates to be mapped. First off the time intervals

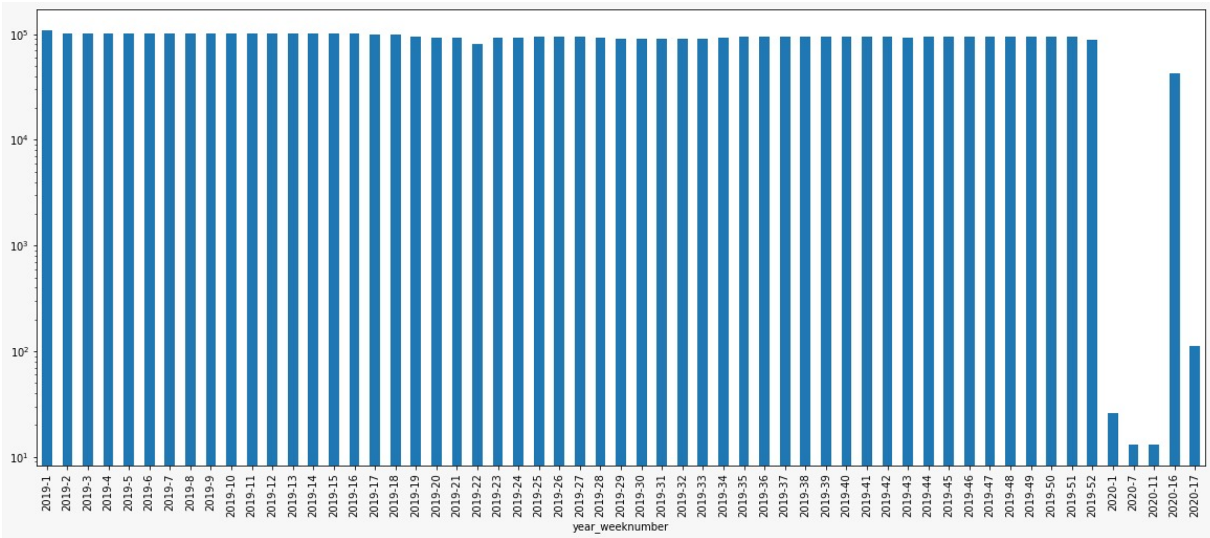

Figure 2. Missing GVB data with a Logarithmic scale on the y axis

of 15 minutes were aggregated to a 1 hour interval to compress the data. Next the codes of the devices were extracted and the coordinates were added using a data set from the City of Amsterdam which used the same device codes[3]. Using these addresses the GeoPy library[6] could be applied to find the corresponding geographical coordinates. Finally all files are combined into one containing the device code, the time with the aforementioned 1 hour interval, coordinates, and amount of movements. In regards of erroneous data there were only two files from the 8th of August and the 9th of August which were empty and therefore not present in the final result.

3.2.2 GVB. The GVB data required little preparation. Only the modality, borough, and neighbourhood needed specification. For the modality this could be done based on the difference between letter codes for metros and numerical codes for trams and busses. The borough and neighbourhood could be extracted using an extra source from the city [1, 2]. This data set contained coordinates for every neighbourhood and borough which made it possible to label each stop.

3.2.3 Resono. The Resono data needed, similarly to the CMSA data set, to be changed from 15 minute intervals to 1 hour intervals. Afterwards the geographic names were accompanied by geographic coordinates by matching the locations to the locations in the same data set used in for processing the GVB locations.

3.2.4 Oxford Stringency Index. This resource was offered as a work-in-progress excel-file, where each sheet was a measure, each row was a country and each column was a day. This resource was scraped using Python Notebooks and transformed into a CSV-timeseries, which was then aggregated on a weekly basis. These quantitative values per week aligned with the resolution of the other data values, and could therefore be shown in a colour-bar with a change in hue, from grey-yellow to dark-red, depending on the stringency of a measure (see Figure **??**).

3.2.5 Covid Measure News. This resource too was scraped from the provided excel-calendar, and aggregated on a weekly basis in a CSV-timeseries format. In this way, the dashboard can show significant measures and events when a specific week is selected (see Figure 5).

3.3 System description

The goals of this project was to give oversight into -and possibly find links between- the effects of corona-measures on different modalities and modal split. Next to that, the project also aimed to find out if these effects were different per borough.

The system is built in a way that aims to immediately see the links between different measures and the crowdedness in the Amsterdam metro, bus and tram in different geographical parts of Amsterdam. To bring all this different information together in a logical way a data pipeline was built and fed into tableau. The first part of the pipeline is the data collection. The Netherlands uses a contactless smart card system for all public transport in the Netherlands. Users are required to 'check-in' at their station of departure and 'check-out' at their station of arrival. Data about both check-in's and check-out's is administered and saved in data sets that either contain check-outs, check-ins or both. Data is also both saved per hour and per day.

After the distribution of data across team members, some data pre-processing and cleaning steps are taken that are described to further detail in the data section. After the first cleaning and pre-processing part is done, a second script uses the coordinates of the different stations to add boroughs and more detailed neighbourhoods ('buurtcombinaties') to each row. This data is then linked in tableau to CMSA-data, a data set that describes the intensity of corona measures (Oxford), some manual textual descriptions of measures taken per week. Next to that a two GeoJSON-files are added to create the boroughs and public transport lines on the maps. All the data except the GeoJSON-files are linked to the public transport data by the date of each entry. The GeoJSON-files are linked to the public transport date by modality (lines) and by area (boroughs).

In tableau, various graphs are created. The main bar chart, as seen in Figure 3, is created by plotting the sum of trips per week number of 2019 and 2020 on the same graph. As 2020 is our year of interest and 2019 is merely used as a reference point, specific choices are made to highlight 2020. Bars are not stacked and 2020 is brought to the front, 2020 is coloured blue and 2019 is coloured grey. The whole graph changes when subsets are chosen (i.e. a specific set of hours, a specific set of stations or a specific modality).

It can be quite confusing to remember which different measures were taken against the spread of COVID-19, and to which extent over which specific period they were used. Therefore, we have looked for data that tracks different levels of intensity of a specific measure (for instance schools closing) over time. The university of Oxford already had such a file and also keeps track of each country's more general stringency around COVID-19. This data is therefore plotted on as a row of squares ordered by date as seen in Figure 4. The squares have a colour according to that measure's level of stringency. On the order of stringency (most stringent first) they colour red, then yellow and finally grey if not applicable or at level zero. Only the Stringency index is a scalar variable, the other measures are tracked as indicators with 2-4 levels.

The usage per hour graph (Figure 6) is made in a way that it most resembles the main bar chart of crowdedness over the year. Colours and stacking is done in a similar fashion. The only difference is that the bins here contain a total of trips taken in a certain hour. This is shown for all weeks when the dashboard opens but can be toggled to a certain subset in a similar fashion as sub-setting is executed in other parts of the dashboard.

The foot traffic data (CMSA) is grouped by 4 area's in Amsterdam. These are displayed as bars (figure 7) that change according to temporal subsets (weeks and hours). It does not respond to the maps as they do not contain information about stations.

To show how trips are distributed across Amsterdam we have created 2 maps (figure 8). One with dots for each station that grow bigger if there are more trips at that station. These dots are pie charts if a station has both bus and tram (orange) and metro (blue). The other map is a heatmap per one of the eight boroughs of Amsterdam. These are mapped on a sequential colour scale of grey (least busy) to orange (average) to red (busiest).

Lastly we have created a text-box in which textual descriptions of weekly measure changes against COVID-19 are described (see Figure 5). These are from a calendar that is kept by the City of Amsterdam. Therefore, these texts are unfortunately in Dutch. As a lot of people that work for the City still speak Dutch, we thought it was still a nice addition to the dashboard. It shows descriptions per week if a week is selected. In the descriptions are some national as well as local measures specific for Amsterdam.

3.4 Evaluation

3.4.1 Heuristics. As discussed in the method section, we have applied several heuristics of Nielsen[10]. In particular we focused on seven heuristics that were ought to be applicable to our case: "match between system and the real world", "user control and freedom", "consistency and standards", "error prevention", "recognition rather than recall", "aesthetic and minimalist design" and "help and documentation".

Match between system and the real world. This heuristic refers to the rule that the dashboard should contain concepts that are familiar to the user. We have implemented this heuristic by showing the map of Amsterdam, accompanied by the metro and tram lines (see Figure 8) As these concepts are ought to be very familiar to the users, we consider this as a design choice that helps comply with this first heuristic.

User control and freedom. This heuristic is applied by creating an option which enables the users to set the setting of the dashboard to the default settings (see Figure 9). This way users is able to restore its actions without having to go through an extended dialogue.

Consistency and standards. To comply with this heuristic, the visualisation styles and functionalities were made consistent across the different visualisations in the dashboard. The first example of such consistency is visible in the hover functionality for graphs which gives more insight into the specific part of the visualisation in every case. Secondly clicking part of a visualisation selects this specific time period or area and alters the related visualisations.

Error prevention. To comply with this heuristic we aimed to eliminate all elements that would be error-prone. After the minimal viable product of the dashboard was done, we directly notices that the dashboard was a computationally expensive application. As a result, the dashboard freezed when selecting multiple filters. As we could not prevent this from happening, we strongly advise the users to only run the dashboard on an external server that meets the computational requirements. We did this by prompting the user with a message that informed them with the requirements, when starting the dashboard.

Recognition rather than recall. To comply with this heuristic, we ensured that the user is not forced to recall certain concepts from one part of the dashboard to the use another

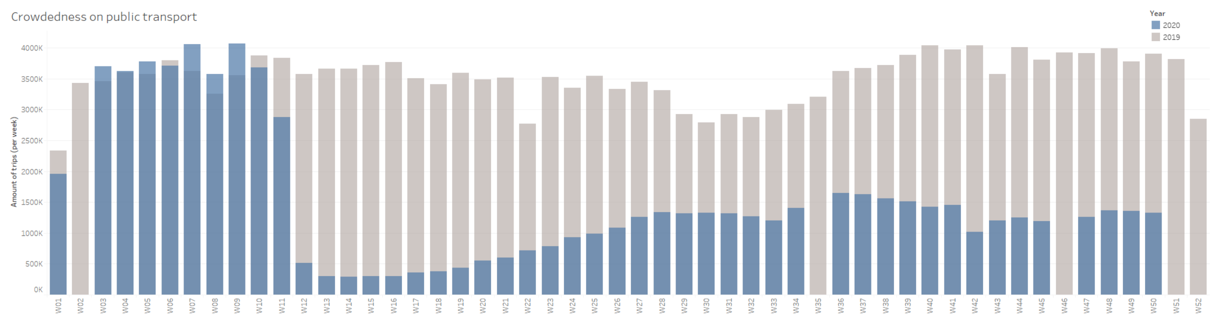

Figure 3. Crowdedness on public transport per week

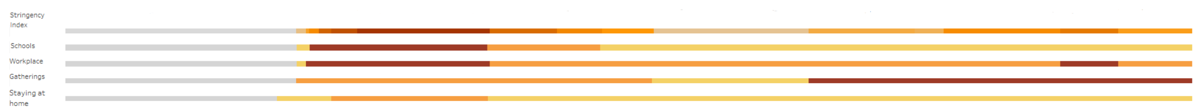

Figure 4. Measures taken against COVID-19 over 2020

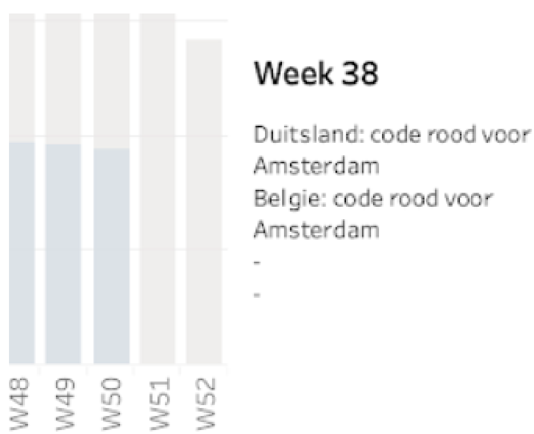

Figure 5. News tracker per week as provided by the City of Amsterdam

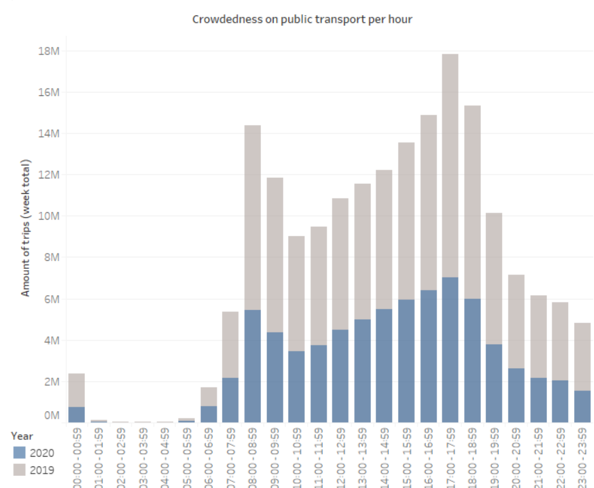

Figure 6. Crowdedness on public transport per hour of the day

part of the dashboard. Therefore, we aimed that the user should recognise the concepts immediately when interacting with the dashboard. We ensured that each selection from one part of the dashboard, also applies to other parts of the dashboard. As a result, the users adequately sees what one selections means to other area's of interest, without having to recall certain previous selections.

Aesthetic and minimalist design. To comply with this last heuristic, we aimed to only exclude information that is irrelevant or rarely needed. Because the dashboard contained a lot of information, we aimed to group all the information by its respective dimension (i.e., time vs. space). For instance, we visualised the time dimension (i.e., the amount of check-ins per day) in the upper part of the dashboard, and the space dimension (i.e., amount of check-ins per borough) in the lower part of the dashboard.

Help and documentation. In order to help a novel user with his initial dashboard use, we have created a video explaining how to answer to most relevant questions using the dashboard (see [7]).

3.4.2 User Testing. After using Nielsen's heuristics as core design principles all that remained was a validation of the usability of the dashboard. The user testing was guided by the three steps described in 2.3.

- Determining Appropriate measurements
- Formulating goals

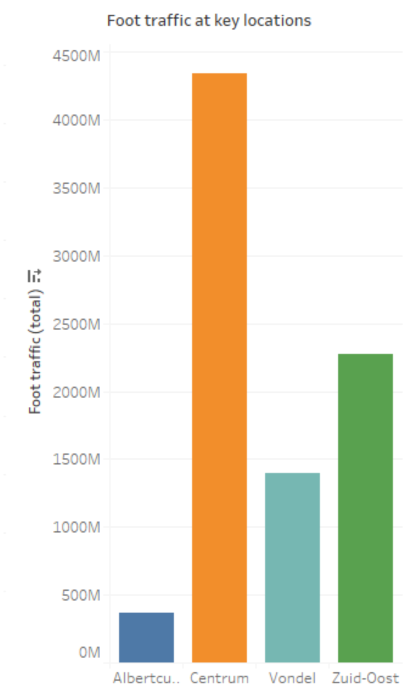

Figure 7. Foot traffic measured by sensor data

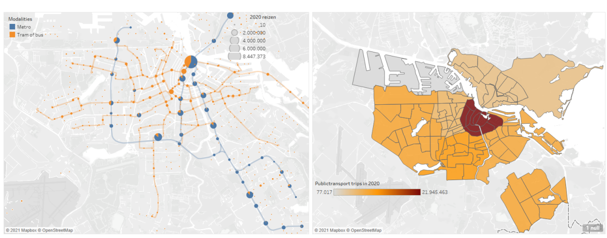

Figure 8. Maps of public transport traffic in Amsterdam

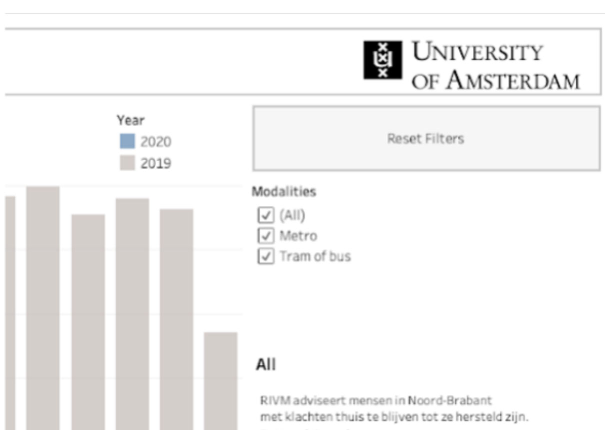

Figure 9. Reset filter option

- Collecting Data & verifying Goals

For the measurements it was decided to measure the task completion success rate by creating tasks that required the tester to answer a question. Besides the task success the tester was asked to rate the questions from the respective subject to formulate a satisfaction rate as well. The formulated goal was to optimise the usability beyond the initial design based on Nielsen's heuristics. The goal verification was done by deriving concrete points of improvement from the feedback of the testing. Besides this improvement of the design the satisfactory rating could commend the existing design in case of high satisfaction rating. Ultimately the test consisted of 8 questions (see Appendix A, p.11) spread over five subjects that correspond to the five research questions.

4 Results

As we did not perform any statistical analysis on the visible trends in the dashboard it is impossible to say what effects can be seen of each measure. However, The dashboard shows some trends that a user could spot. These will be described in this section.

4.1 Effects of measures over time

First of all it is clearly visible how the first package of measures and warnings (March) drastically reduced the crowdedness in the Amsterdam public transport. Week 10 (starting March 2), still had 3.6 million trips taken. Over the following days, some early measures are taken. On the 16th of March, Prime-Minister Mark Rutte delivers his infamous speech from his presidential office, announcing that most Dutch people will be infected with COVID-19 and that there is no 'fast way out' of this pandemic. Week 13 (starting March 23) contains a bit more than 500 thousand trips. This is a weekly reduction of traffic of 85.8%.

Such a drastic reduction was not achieved again later in the year. Traffic came to a minimum in week 14 (292.365 trips) but in the weeks following rose back to a plateau of 1.35 million trips per week (37.9% of traffic during the same period in 2019). These levels did not drastically change in the rest of the year. The only real drop that is visible in the second half of 2020 is halfway through October. On the 13th of October, Mark Rutte gave a press conference in which he introduced a package of measures that he dubbed as a "partial lockdown". Week 42 (starting on October the 12th) saw a reduction of 30% compared to the week before. This is a reduction that is a lot smaller than the one in mid-March and might even be a coincidence rather than an effect from the measures taken on the 12th of October. One might say that no measure had as much effect as the first package mid-March. However, it is hard to conclude such statements as we have no idea what the levels in the second half of 2020 would be if there were no measures taken. We can only see that the amount of trips stayed around this level in the second half of 2020.

4.2 Usage per hours of the day

The usage over the different hours of the day changes proportionally similar as before the pandemic. After the first package of measures, a reduction is visible in rush hour traffic (defined as traffic between 7 AM - 10 AM and between 4 PM - 6 PM) relative to other hours' traffic. However, in the similar period in 2019, a similar reduction is also visible. This means that no difference is visible in this metric to the extent of work-related traffic and other traffic. Over the rest of the year, the proportion of rush hour traffic to other hours' traffic stays quite similar to similar periods in 2019. Commuting traffic thus followed the same line as normal traffic in responding to the different measures taken in 2020.

4.3 Crowdedness in different boroughs

Measures seem to have a minor difference in how they impact transit crowdedness in different boroughs. The boroughs that are on the outskirts of Amsterdam (Nieuw-West, Noord and Zuid-Oost) contain more residential area's than the one's closer to the city centre (West, Zuid, Oost and Centrum). The mid-March drop in residential zones is relatively a bit lower than the drop is in the area's that also contain a lot of shops and offices. The borough with the most extreme response to the first package of measures is Westpoort. This is an industrial zone with only 5 total stops, most traffic is between 6 and 10 AM and 98.4% of traffic happens at one stop, the Koivistokade. It is thus expected that this work-heavy borough has the most extreme of responses to the measures. One could conclude from the dashboard that the more offices or other non-residential buildings there are in an area, the bigger the response is to the measures. In the three residential-heavy boroughs, traffic also comes closer to its original crowdedness in the second half of 2020 compared to the other boroughs.

4.4 Foot traffic at key locations

The question of how corona measures influence foot traffic at key locations is harder to answer from this data. The circumstances of data-collection changed over the year. Next to that, data is only available starting in July 2020. Therefore it is hard to compare this data with situations before the corona pandemic. The data also contains a lot of spikes that lead to questioning the validity in which this data was brought about. In week 40 for instance, the sensors register a record high of 1,2 billion people walking past in the centre only. This are rather unbelievable numbers that make it hard to compare anything in this data set.

4.5 Displaying information of different dimensions

Our dashboard contains different scales or dimensions. The main bar chart shows 2 years of public transport usage per week (as a bin) on a one year axis. The measure intensity bars show 2020 and are aligned to the main bar chart. The crowdedness per hour bar chart, the foot traffic bar chart and the textual change show data for a full week. The maps however, show the geographical distribution of usage over the whole city of Amsterdam. These are all merely dimensions of the data that provide the user with a choice of sub setting the data. One user might be interested in the change per hour over different boroughs whilst another might be specifically interested in the development of rush hour traffic over 2020 and compare it to the intensity of corona measures. This dashboard does not select any specific combinations of such settings by itself or guide the user to a certain conclusion or insight that could be taken from the data. Whatever combination or subset of the different dimensions a user might be interested in, this dashboard can show it. This also has its downsides. When everything is possible, how does a user who doesn't know where to look find something of interest? Adding all these elements in one dashboard always opens up the risk to cognitive overload. We have decided to go for this lay-out nonetheless. We think that the drop in public transport usage in 2020 immediately catches the eye to a user that opens the dashboard for the first time. Therefore, the main message of the dashboard is not lost in the multiple possibilities to further analyse specific facets of the impact of pandemic measures.

4.6 User Testing

The user testing was done with six individuals which were in some way related to the authors. On average 7.71 of the 8 questions were answered correctly. The questions grouped by topic were also rated on user satisfaction which led to a satisfaction rating of 3.78 out of 5. Beyond these two metrics concrete feedback was provided to suggest improvements to the dashboard. This resulted in the following two concrete improvements:

- Consistent application of graph titles and legends
- Additional labelling

Besides these improvements the user testing also served as positive affirmation of the usability in the existing dashboard. This is suggested by the high rate of correctly answered questions as well as the satisfaction rating to a lesser extent.

5 Conclusion

In summary, our general aim in this project was to generate useful insights into the effects of taken measures in the pandemic. By providing a clear oversight of the measures, the change in transport occupancy and street business, we aimed to proved a useful tool to generate insights. We subsequently created tests to to make sure these insights were actually usable for stakeholders.

Regarding the effectiveness of the measures, it's quite difficult to make hard conclusions about them. The amount of variables that have an influence on street business and transit occupancy is large, which makes it difficult to isolate the effect of one measure. The aim of the dashboard is mostly to allow a first exploratory data analysis, which can show where further adjustments to measures have to be made.

6 Discussion

During this project, we found several factors that could make this project even stronger if integrated in the future. One of the options with the most potential is a larger data set with less noise. This data set could potentially show traffic patterns in past years, which could grant a perspective in the exceptional circumstances.

Regarding the noise in data, especially seen in the CMSA- and Resono-data, this could provide more insight in more local changes in street business. These data-sources have a big potential, but had some problems in this project. For both, only 3 to 5 months of data was found, with spikes in the data sets which made it unreliable at times.

Out of the three initially scraped resources on qualitative news and context from the pandemic, only one was incorporated in the end: the measures from the City of Amsterdam. The other two resources (newspage of City of Amsterdam and another external website) were found out to be too verbose to give any real context in the dashboard. This ran the risk of possibly overloading the user with information. It was decided that they were not to be included. These data was considered to be used for a sentiment analysis but this was found to be out of scope for the project.

The data quality of GVB was relatively high. We noticed that stops with less than 10 travellers per hour were removed. We suspect that this is done for anonymization purposes. As there is no insight into the scale of data being deleted, it is hard to say how representative the data set is to the real world. Some descriptive statistics about the data which is not in this data set could really help to perform some checks on whether or not the deleted data is substantially different from the data that is available. Without such information, it is impossible to guarantee the external validity of the information that is shown in the dashboard.

Regarding the usability of the dashboard, the current version is interactive but shows some points for improvement in improving usability and recall. The current version does not aid the user in comparing the values of different weeks which, according to Nielson's heuristics, leads to usability points. For example, finding the decrease in tram usage per week between two different districts now requires the user to maintain certain numbers in memory, which can be optimised.

A points that might further increase the usefulness of the dashboard is a better division between different modes. Currently, bus and tram are combined in the same usage statistics, which makes it difficult to say whether there is a difference in usage between those two. This data is available and might be included by incorporating data from TransLink, the integrated fare system in Amsterdam.

7 Future works

The current version of the dashboard is static, caused by the nature of the resources provided. Data was accessed through a file-system which was only accessible through a user interface and by using a VPN. To improve this, a data-pipeline can be set up that automatically fetches delta changes every week, transforms them and incorporates them in the existing dashboard. If this is done properly, the dashboard can truly become a crowdedness dashboard that shows how crowded Amsterdam public transport is and how it is developing. This would not have as much a focus on corona-measures as the current dashboard has but it would be designed in a way that new data is automatically shown in an insightful manner and could thus suffice as a method to ensure data-driven decisions. This is purely a technical problem and as all methods used to change and reshape data is available in our public GitHub repository.

If the CMSA and Resono data were complete, they could provide two extra layers of measuring crowdedness. If this data was made available, the dashboard could really be a tool that measures different dimensions of crowdedness. If the data is combined this could also be used to cross-reference certain findings in one dataset by checking if that effect is also visible in other dimensions of crowdedness. For instance; somebody finds out that a certain measurement had a higher effect in a certain borough by the amount of trips taken in public transport. It could be really interesting to then automatically analyze if that reduction in trips also has an effect on the crowdedness on the street.

The current version of the dashboard has some usability limitations, which impend easy understanding and usage. We were able to solve some requests and problems we found while user testing. Examples of improvement points are the visibility of modal split and change per station over time in the map. Points which were not immediately clear to first-time users, such as the meaning of the overlapping weekly usage, have been adapted in the final version of the dashboard. Other points of improvements are a clear way to see all applied filters and points on which is currently being dug down on and a lack of options in changing resolution. While we have tried to provide the user with hour-to-hour data, this is very costly in data processing power, and did not work very well on local computers. Some of these, especially the last items, are fixable by using the professional version of Tableau (Tableau Server) which provides a different interface for the dashboard online.

Regarding the context-data such as news items and Oxford-stringency-index, the item with most potential is a translation of the measures. Currently, qualitative measures in the dashboard are displayed in Dutch. This is, according to Nielsen's heuristics, a discrepancy in user experience. Besides translating this, it might be interesting to look into more measures that could have had a big effect on the change in travel behaviour.

References

[1] City of Amsterdam. 2016. Maps Data - City districts. https://maps.amsterdam.nl/open_geodata/?k=202 Accessed: 07-01-2021.

[2] City of Amsterdam. 2016. Maps Data - Quarters. https://maps.amsterdam.nl/open_geodata/?k=200 Accessed: 07-01-2021.

[3] City of Amsterdam. 2020. Maps Data - Sensors Crowd Monitoring System Amsterdam. https://maps.amsterdam.nl/open_geodata/?k=264 Accessed: 08-01-2021.

[4] City of Amsterdam. 2020. Maps Data - Tram and metro lines 2020. https://maps.amsterdam.nl/open_geodata/?k=220 Accessed: 08-01-2021.

[5] City of Amsterdam. 2021. Nieuws over het coronavirus. https://www.amsterdam.nl/nieuws/coronavirus/nieuwsoverzicht-coronavirus/ Accessed: 14-01-2021.

[6] Esmukov et al. 2020. GeoPy. https://github.com/geopy/geopy

[7] Arno Garstman. 2021. *Data Systems Project - Corona dashboard demo: measures vs. public transport (UvA)*. https://www.youtube.com/watch?v=8B8qSIwmhaA

[8] Thomas Hale, Samuel Webster, Helen Tatlow, Toby Phillips, Anna Petherik, Saptarshi Mujamdar, Beatriz Kira, Laura Hallas, Emily Cameron-Blake, Noam Angrist, and et al. 2020. Coronavirus Government Response Tracker. https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker

[9] James Lewis. 2006. *'Handbook of Human Factors and Ergonomics'*. Hoboken, NJ: John Wiley. http://sistemas-humano-computacionais.wdfiles.com/local--files/capitulo%3Amodelagem-e-simulacao-de-sistemas-humano-computacio/usabilitytesting-ral.pdf

[10] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 152–158.

[11] Jakob Nielsen. 2005. Ten usability heuristics.

[12] Tijdlijn Coronavirus. 2021. Tijdlijn corona Nederland. https://tijdlijn-coronavirus.nl Accessed: 06-01-2021.

A Questionnaire User Testing

A.1 Questions

List of every sub-question and asked question to validate the usability and potential usefulness of the built up dashboard.

- *How does the distribution of usage change over the day?*
  - Which hour was the most busy of travellers on public transport in week 21 of 2020?
  - How many people travelled in this peak hour?
  - How well does finding this answer work?
  - Feedback or Points of Improvement
- *How do measures impact transit crowdedness in different boroughs?*
  - Which borough had least amount of traffic registered in week 21 of 2020?
  - Which measure was introduced or lifted in week 23 of 2020?
  - How well does finding this answer work?
  - Feedback or Points of Improvement
- *How does the effect of measures change over time?*
  - Which type of measure was mandated the least strict throughout the pandemic according to the dashboard?
  - How well does finding this answer work?
  - Feedback or Points of Improvement
- *How do corona measures influence foot traffic at key locations?*
  - How much pedestrian movement was registered at Vondelpark during week 34 of 2020?
  - How well does finding this answer work?
  - Feedback or Points of Improvement
- *How can we clearly present data of different spatial and temporal dimensions to analyse impact of measures?*
  - The visualisations of foot traffic and public transport use the exact same time period.
  - The visualisations of foot traffic and public transport use the exact same areas.
  - How well does finding this answer work?
  - Feedback or Points of Improvement

A.2 Screenshots

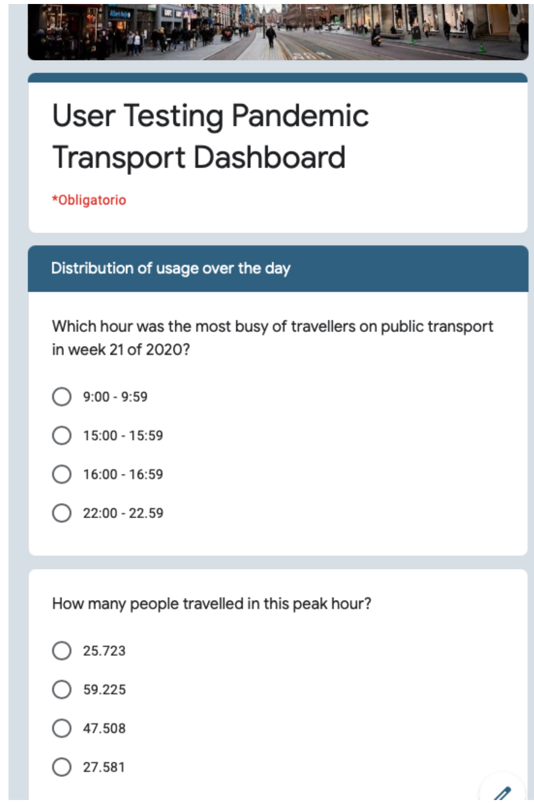

Figure 10. Screenshot of the first part of the questionnaire

B Screenshots Dashboard

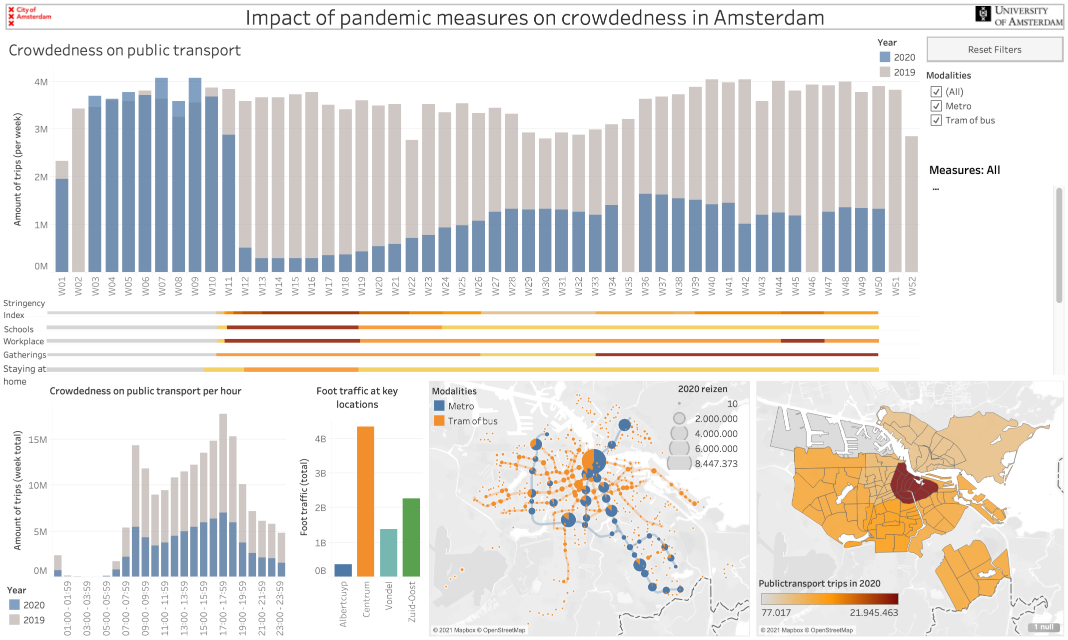

Figure 11. Corona measures dashboard