An Update on Tram and Metro Statistics, including the Possibility of Short-Term Collection

Note by the secretariat

Summary

This document describes the work of the secretariat in continuing data collection on city-level tram and metro passenger statistics, including new dissemination methods. The document also describes data availability at the short-term level, and discusses automated collection techniques. Delegates are asked to reflect on the value of this data collection exercise, as well as the ideas floated of collecting these data automatically and/or on a short-term basis in the future.

I. Background

1. After a successful workshop on urban public transport statistics in 2019, the Working Party on Transport Statistics decided to send out a pilot questionnaire on tram and metro statistics afterwards, in order to ascertain data availability and collection methods (see ECE/TRANS/WP.6/2020/5). These data have not been collected at the international level before. Given the demonstrated good data availability as well as the relative simplicity of the data requested, the collection was continued in pilot form in 2020.

II. Annual Data Collection and Dissemination

2. Since the 2020 first publication of the tram and metro dataset, the secretariat has improved the accompanying map that shows data availability, using the Dash Plotly software run through Python. The resulting map can now be generated automatically, and in addition to showcasing data availability in a geospatial way, can also visualise passenger numbers and
passenger-km over time through an automatic line graph for each city. The user can toggle between passenger numbers and passenger-km.¹

Figure 1
Screenshot of the new interactive tram and metro data map

3. When requesting the data, the secretariat encouraged countries to share links to where these data may appear online, if possible. Given the simplicity of this dataset compared to others, if data were published through an Application Programme Interface (API) or other structured, machine-readable format, the secretariat offered to automate the data collection process, in order to remove any response burden at all. However, the only countries that responded positively to this proposal were Portugal and Spain. This was despite some countries indeed having these data available potentially automatically through their national databases. The secretariat would be interested to hear from these countries on why they have provided the data in a questionnaire despite this. It may be that statistical offices wish to ensure that collected data are verified manually before publication at the international level, or other reasons that delegates are encouraged to share.

4. Since the dataset was first published, data for France and Georgia have now been added, taking the number of cities or regions that are covered in this exercise to 146 over twenty-six member States.

III. Shorter-term Data

5. When the questionnaire was sent out in September of 2020, the secretariat also requested countries to provide any short-term data that they had; either quarterly or monthly passenger and/or passenger-km numbers. The main purpose of this was to ascertain data availability, and also to collect data that could better measure the impact of different lockdown policies across cities in the ECE region, given the ongoing COVID-19 crisis, without waiting for annual data that would in any case mask changes over the course of the year.

6. Out of 26 countries that had some form of annual data, seventeen countries provided quarterly data as well, or had quarterly data available through their website in an easy-to-manipulate format. Figure 2 shows an index of passenger numbers for just a few selected cities, highlighting the strong decline in the second quarter of 2020 in all of them, but that also there was a large degree of variation in the size of the decreases in the second quarter, reflecting the range of how COVID-19 affected each country. These short-term variations will be hidden when the annual data are released.

¹Publication of the map was pending at the time of document submitted but should be available at the time of the Working Party session.
6. The good availability of these short-term data show that a more regular collection exercise is viable, if there is sufficient interest in the data. The question remains, however, whether the usefulness of the data would warrant increasing the reporting burden on countries, in addition to the work imposed on the secretariat. This is why the possibility of importing these data manually remains attractive. If APIs or structured excel files are available, then data can be imported from multiple countries in a straightforward manner, with no additional reporting burden for countries and a minimal verification burden placed on the secretariat.

7. Of the seventeen countries who provided quarterly information, data were found on national websites on either a monthly or quarterly basis for nine countries (Belarus, Denmark, Hungary, Latvia, Portugal, Serbia, Spain, Ukraine and United States of America). Of these nine, data were in a machine-readable format for all except Serbia, where data were published as a PDF. However, data were not always broken down by city; Denmark, Latvia, Portugal, Spain and United States had the city breakdown, whereas data for Belarus, Hungary, Serbia and Ukraine were collated at the national level. While city-level data are indeed important to this exercise, national level short-term data would still have use as a way to quickly track changes in public transport use across countries.

IV. Value of these Data and Conclusions

8. On the annual level, the secretariat plans to continue collecting these data through a questionnaire, with the exception of those countries (namely Spain and Portugal) that have indicated they are happy for the data to be taken through their API directly to minimise their reporting burden. Given the importance of urban transport data to monitoring the Sustainable Development Goals, for example indicator 11.2.1 on access to public transport and 9.1.2 on passenger and freight volumes by mode of transport, the secretariat continuing to provide the collated dataset internationally is clearly of use to transport monitoring. These data could, for example, assist in calculating modal split patterns of specific cities or agglomerations, as well as countries. In addition, their trend over time show which cities have seen public transport grow or decline.
9. When the secretariat launched the pilot data collection, it kept the collection as simple as possible, asking for only passenger numbers and passenger-km of metro and tram systems. In order to put these numbers into context, additional information can sometimes be beneficial. For example, for those cities where the tram or metro system is new or has been expanded, the total line length each year would provide context into any sudden increases (or indeed decreases) in passenger numbers.

10. In addition to this, while searching for online links to the data the secretariat noticed that some countries were publishing either city bus or trolleybus information alongside the tram and metro data. If there were sufficient interest, these data could be collated with this dataset as well. While trams or metros are sometimes the principal public transport mode in cities, buses are more universal. Data availability of this is likely to be very variable across countries, given the vast array of different structures involved in public transport.

11. Finally, on short-term data, the secretariat encourages countries to consider publishing any available data in a machine-readable format, and broken down by city, if possible. This could allow a metro and tram index to be published for cities across the ECE region, allowing short-term urban transport patterns to be available quickly.