# Input analysis passenger arrivals

Moddeling assumption: we use Poisson distribution

**Using the 12a and 12b data sheets**

We want to compute the averge passengers/minute value (the rate)

* Per stop
* Per interval of 15 minutes

Way of working:

* Each arrival is mapped to a bin of 15 minutes (starting at 6:00)
* Each arrival is mapped to the bin that corresponds with *departure time* of the bus + the amount of *traveltime* required to reach the given stop.
  + These times were collected using the 9292OV App.
* An arrival has the value of the amount of passengers entering divided by the amount of time that passed between the current bus and the previous bus (so: the average amount of passengers arrived per minute, AKA “rate”)
* We then take, per interval, the average of all the rates that were mapped to our bin per stop.
  + **This is our initial rate schedule**
* We paste the values of the very last bin to up to 2 bins after the last bus is supposed to leave, because we want to be able to simulate extreme delays in our program. The assumption is that the rate of passengers arriving after 21:00 is at a constant, very low value. We therefore decide to copy these values for later stops.

*Translating the initial rates of busses to the passenger prognosis of the tramline*

* We are interested in the average amount of passengers that arrive at **tram stops** (not bus stops) per interval
* We assume that the passenger prognosis at the tram stops is a result of a statistical experiment or an educated guess, and might therefore not be an accurate representation.
* We therefore give higher priority to the predicted passenger arrivals using the rates deducted from the bus schedule rather than the predicted passengers in the prognosis data sheet
  + For example: prognosis predicts 99 passengers to enter at Padualaan on the CS tram
  + Our sheet indicates that the expected amount of passengers during that interval, using the rates of the bus, is only 14.
  + Overall, the total amount of passengers using the tram line over the bus line only increased by a maximum of 3.
  + The new average amount of arrivals we use for the padualaan will therefore be around a maximum of 45, and not around the 99 which was indicated by the passenger prognosis.
* Our assumption in all this is the following: the exact same type of people that use the bus line 12 will be using the new tram line to travel. Only, this group has grown in size because of the increased resources at the Uithof. This group using the tram line has therefore grown even faster. That is the reason we aim to multiply the existing rates by a “growth-parameter”, depending on the time of the interval.

*Way of working*

* We calculate the ratio of total passengers predicted during the peak by the passenger prognosis and by the bus rates. We calculate this ratio for 3 intervals:
* We do this for the first and second interval, and scale the bus rates using that ratio to match the prediction of the bus rates with the total tram passenger prognosis
* We then calculate the ratio for all the remaining intervals combined, and use this to scale these intervals individually to match the grand total of passenger predictions per day of the bus rates and the tram prognosis.
* Used parameters:
  + We multiply the bus rates during the first peak with 2,8
  + We multiply the bus rates during the second peak with 2,692
  + We multiply the bus rates during the off-peak interval with 2,0088
* These are the rates we assume are the most accurate for predicting the amount of passengers that will be using the new tram line.

*Handling the new route schema near the Uithof and near the CS area*

At the end of the Uithof line, one stop (AZU) has been turned into 3 stops (UMC, WKZ, and P&R)

The same buildings and facilities are present at this end of the Uithof, therefore, we assume that the same group of people using the bus will now use the tram line. The only exception is that the large P&R facility is now present, which was not present during the collection of bus data.

We use the passenger prognosis to make an estimate of the comparison between the rates of entering passengers at these stops, expressed as a fraction of the scaled amount of people arriving at AZU. These were determined to be 2/3 and 1/(3,2) for UMC and WKZ respectively. The amount of passengers arriving at the P&R is then estimated to be the remainder of the passengers previously arriving at the AZU. This, in combination with the scaling of the size of the group of arriving passenger in total, creates a good fit to the passenger prognosis sheet at the stops UMC, WKZ and P&R.

Three stops have disappeared when comparing the bus and tram tracks, and one is added:

* Rubenslaan, Sterrenwijk and Bleekstraat are gone
* Vaartsche Rijn is added

Judging from the geographical stationing of the different stops, we assume that:

* Passengers previously boarding from Rubenslaan will now board at Stadion Galgenwaard
* Passengers previously boarding from Sterrenwijk and Bleekstraat will now board at Station Vaartsche Rijn.

This, in combination with the scaling of the size of the group of arriving passenger in total, creates a decent fit to the passenger prognosis sheet at the stops Rubenslaan, Sterrenwijk , Bleekstraat and Vaartsche Rijn