STIB - ULB
Data mining project

Presentation business context

Justine BARET

12/10/2022

STIB – Some figures

- Leading urban public transport company in Belgium
- Operates an integrated network of almost **650 km**
  - > 4 metro lines
  - > 17 tram lines
  - > 55 bus lines
  - > 11 night bus lines
- Every day, it provides transport for more than **1,2 million** Brussels residents and commuters
- More than **10.000** collaborators

Our Mission

*To provide the most suitable travel solution with a **competitive travel time** at an **affordable price** in very good **comfort and safety** conditions*

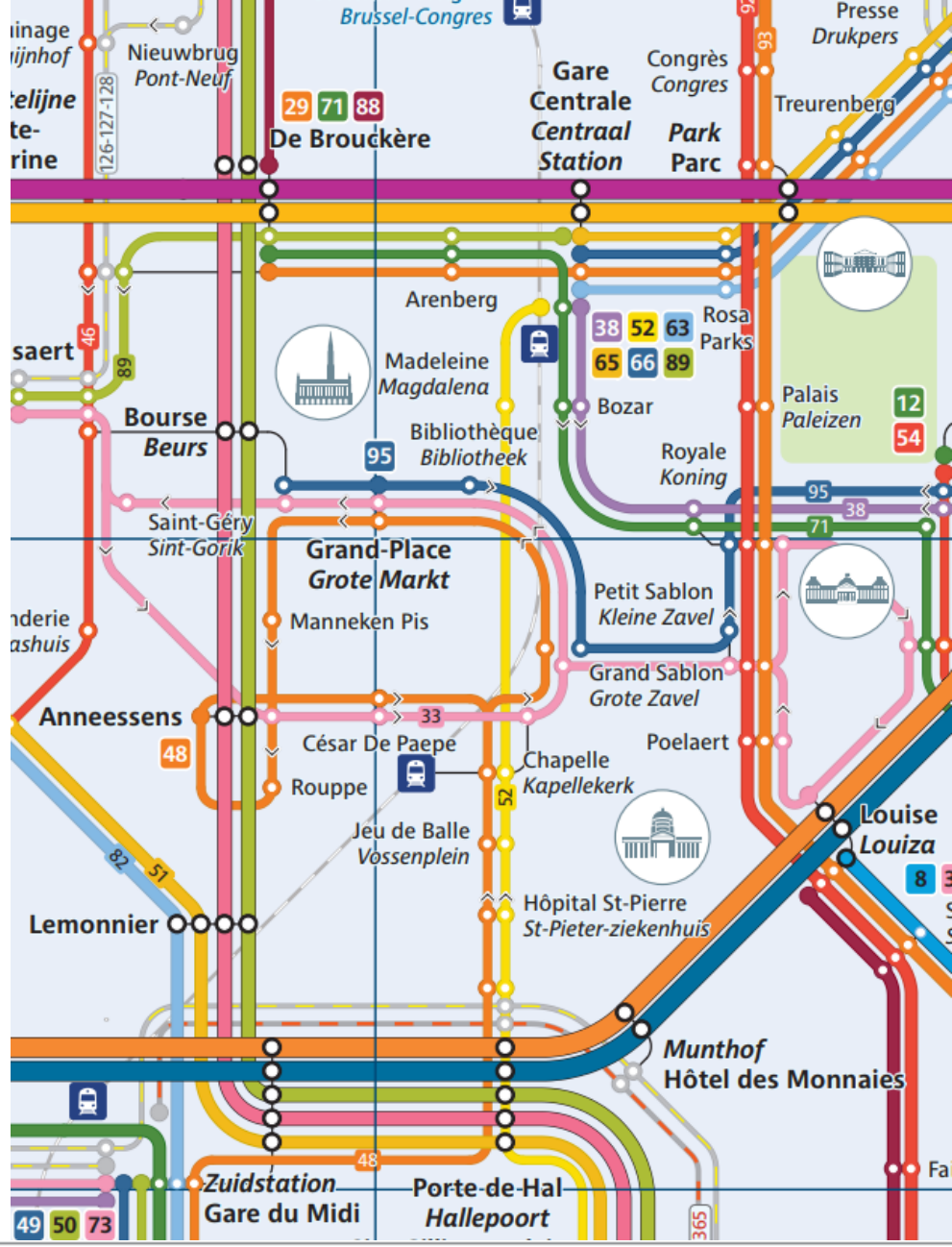
STIB - Network – Network Data Analytics

- **Network** is the division responsible for the design and planification of the network

- **Network Data Analytics** :
  - 5 Data Scientists
  - Management and analysis of ridership and travel time data

→ Understand all the phenomena that occur on the network and their consequences on the operation and quality of service for the customer

→ Need for a quantitative approach to obtain the best arbitrage decisions



Assessment of Quality of Service – Reliability

- **Service reliability = no excess waiting times. It can be considered either in terms of punctuality or regularity**

  - Punctuality : if the customer comes to the stop based on the scheduled timetable, a delay causes an increase of waiting time
  - Regularity : if the customer comes to the stop randomly, irregularity of headways will cause an increase in waiting time

→ It is typically assumed that travelers will arrive randomly as the headway is of 12min of less. If the headway is more than 12min they will adapt their arrival pattern to the schedule

→ As the headways on the network decrease the need for a regularity assessment increases

Regularity indicator : Excess Waiting Time (EWT)

- The additional waiting time that passengers experience due to irregular vehicle arrivals
- As to be calculated at an aggregated level (a period of homogenous scheduled headways) and not at passage level

EWT = AWT – SWT

$$SWT = \frac{\sum scheduled\ headways^2}{2*\sum scheduled}$$

$$AWT = \frac{\sum real\ headways^2}{2*\sum real\ headway}$$

Schedule

Average Headway = 6min

Scheduled Waiting Time = 3 min

Real situation 1

Average Headway = 6min36

Actual Waiting Time = $\frac{7^2+6^2+5^2+7^2+8^2}{2*(7+6+5+7+8)} = \frac{223}{2*33} = 3{,}38 = 3min23sec$

Excess Waiting Time = **23 sec**

Real situation 2

Fréquence Moyenne = 6min36

Actual Waiting Time = $\frac{7^2+6^2+2^2+8^2+10^2}{2*(7+6+2+8+10)} = \frac{253}{2*33} = 3{,}83 = 3min50sec$

Excess Waiting Time = **50 sec**

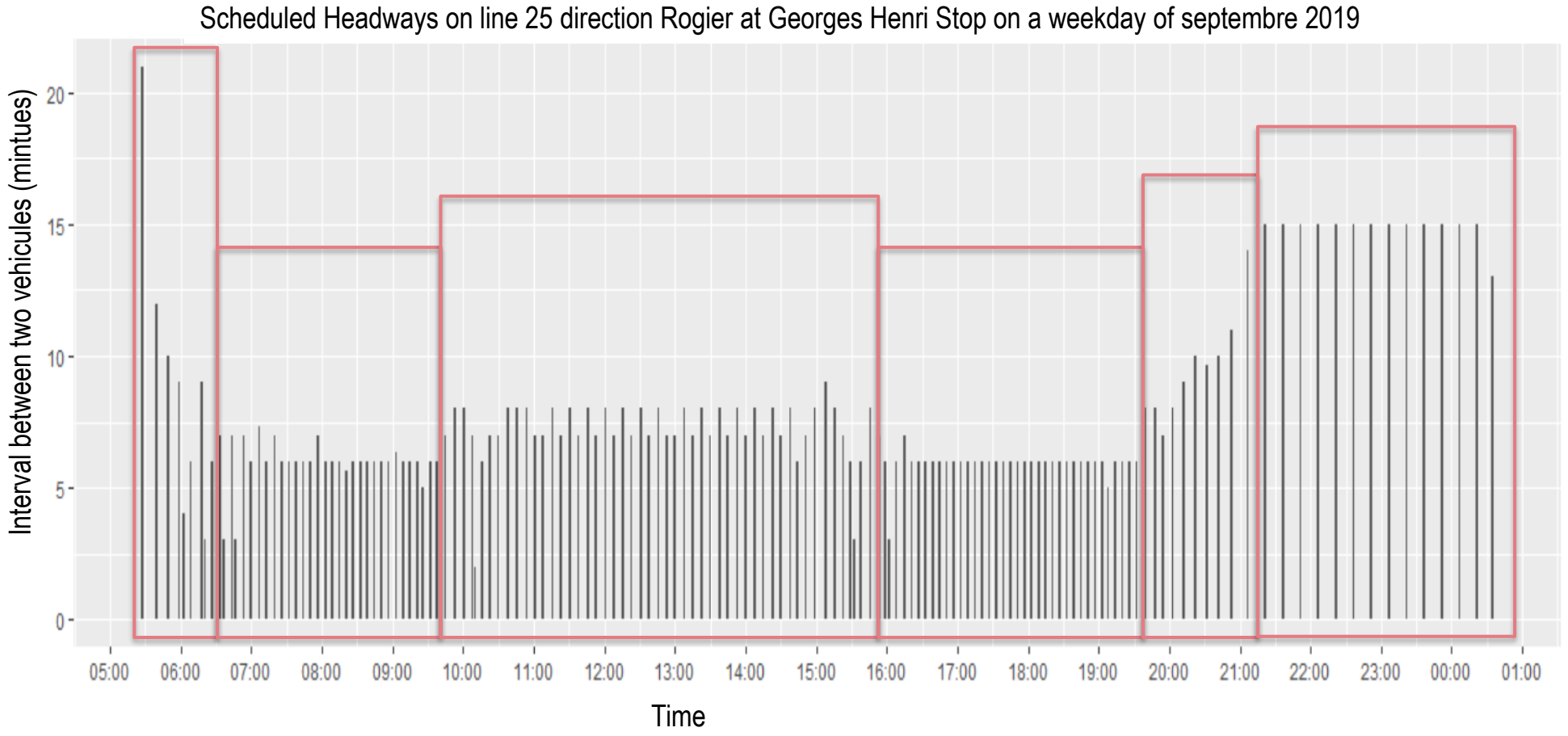
The indicator translates the **customer's perception of irregularity** into an easily communicable and comparable measure

Methodology

Step 1 : Based on the schedule, define the time groups

A time group =

- Either a time period where there is no pattern in the headway
- Either a time period of homogeneous headways



Note that the schedule is different a weekday, a Saturday and a Sunday

Beyond the assessment – analysis of the results

- **Assess the quality of service** (optimum per line / stop / timegroup / calendar)
- **Visualise the results**
- **Beyond assessment find valuable insights. Eg : where are the difficult spots**
  - What lines have issues in terms of service reliability ? When ?
  - Does the service deteriorate progressively from stop to stop or is there a rupture at some point ? Can you identify where the rupture happens (what is the difficult segment of the line – the root cause of the unreliability that then propagates trough the line) ?
- **All other insights from the data would be appreciated ☺**