Requirements specification for Tram’s course process.

1. General description of business process

a) A general description of the business process and a description of the performance metrics generated by this process, possible current analytical problems.

Each course starts with the driver signing the beginning of the course by scanning his work card in the system. The tram's course consists of visiting all the tram stops and checking the number of free seats on each of them. Each course generates the summary of itself that is automatically sent to the system. Course ends when the tram arrives at its base and the driver pulls out the card from the tram’s terminal.

The increase of the average number of free seats by 5% (from 7am to 10 am and from 3pm to 6pm) on the 4 main lines within 12 months, in comparison to the previous year.

Reduce, by 20%, the number of the courses that are with more than 90% of free seats on the 4 main lines within 12 months, in comparison to the previous year.

b) Typical questions

What are the most overpopulated courses?

Give the most populated tram among the 4 main lines.

Give the total amount of the passengers for each line for 7-10am and 3-6pm.

Find the most overpopulated tram stop.

Give the average amount of free seats for each course.

c) Data

All the data is measured by the tickets that were put into the ticket machines and scanned transport cards at the start and at the end of the passenger's route, by which we are able to track the current number of passengers inside the tram. The beginning, the end and each stop is noted in the system by the driver. The tram system sends the visited tram stop information with the amount of the passengers automatically to the system.

We gather data from the GVBTable, Workers table and from the ticket price chart.

2. Data sources structures

GVBTable

TABLE NAME	ATTRIBUTE	ATTR. TYPE	DESC.
Passenger	A unique person identified by his/her ticket number.		
	Ticket number	String	PK, a unique ticket number for each ticket sold.
	Start time	Datetime	The hour that the ticket/card was put into the machine while getting into the tram.
	End time	Datetime	The hour that the ticket/card was put into the machine while getting out of the tram.
	Ticket type	String	Defines a ticket type such as an hourly ticket or one course ticket.
	FK_coursenu	Int	FK pointing at the

	m		course. Implementation of the n to one relationship between Passenger and Course.
	FK_start_stop	Int	FK pointing at the stop when a passenger gets on the tram. Implementation of the n to one relationship between Passenger and Stop.
	FK_leave_stop	Int	FK pointing at the stop when a passenger leaves the tram. Implementation of the n to one relationship between Passenger and Stop.
Tram	A tram identified by its ID number.		
	ID	Int	PK, a unique ID of the tram.
	Number of repairs	Int	Number of repairs carried out on the specific tram.
	Year of production	Datetime	The year of production of the tram.
	FK_model	String	FK pointing at the the model. Implementation of the n to one relationship between Tram and Model.

Model	Table presenting specification of a tram model.		
	Name	String	Unique name of the model.
	Num. of seats	Int	Number of seats in the tram.
	Start of service	Datetime	Date that this model was introduced “on the rails”.
Line	Predefined route that the tram follows each day.		
	Number	Int	PK, number of the route.
	Num. of stops	Int	Number of the stops that are on the specific line.
Course	Each time a driver starts a course an object of this entity is created.		
	Course ID	Int	PK, ID of the course.
	Start time	Datetime	Hour that the driver started the course from the first stop.
	End time	Datetime	Hour that the driver ended the course at the last stop.
	FK_tram	Int	FK pointing at the tram. Implementation of the n to one relationship between Course and Tram.
	FK_line	Int	FK pointing at the line. Implementation of the n to one relationship

			between Course and Line.
Stopped_at	A unique stopped at a stop identified by his/her ID number.		
	ID	Int	PK, a unique ID of the stop at the entit.
	Num_of_pass	Int	Number of passengers in the tram while leaving the tram stop.
	Stop_time	Datetime	The hour of the stop.
	FK_course	Int	FK pointing at the course. Implementation of the n to one relationship between Stopped_at and Course.
	FK_stop	Int	FK pointing at the course. Implementation of the n to one relationship between Stopped_at and Stop.
Stop	A unique stop identified by his/her ID number.		
	ID	Int	PK, a unique ID of the stop.
	FK_line	Int	FK pointing at the line. Implementation of the n to one relationship between Stop and Line. Additional information whether the tram spot is cross-stop

			or single stop.
Driver	A unique person identified by his/her ID number.		
	ID	Int	PK, a unique ID of the driver.
	Name	String	Name of the driver
	Surname	String	Surname of the driver.
	FK_tran	Int	FK pointing at the tram. Implementation of the n to one relationship between Driver and Tram.
Hangar	A unique hangar represented by the hangar ID.		
	Hang. num.	Int	PK, a unique hangar number.
	Number of total plots.	Int	Number of the repair plots in the hangar.
Service_in	Implementation of many to many relationship between Tram and Hangar. Identified by foreign keys to tables Tram and Hangar.		
	FK_tram	Int	KF, part of PK
	FK_hangar	Int	KF, part of PK
Workers team	Team of workers led by a manager.		
	Team ID	Int	PK, ID of the team.
	Num. of workers	Int	Number of the workers in a team.
	Manager	String	Name and Surname of the team manager.

	FK_hangar	Int	FK pointing at the hangar. Implementation of the n to one relationship between Workers team and Hangar.
Repair	Entity containing repair details.		
	Repair ID	Int	PK, ID of the repair
	Start date	Datetime	Start date of the repair.
	End date	Datetime	End date of the repair.
	Complexity	Int	Complexity in scale 1-10
	FK_tram	Int	FK pointing at the tram. Implementation of the n to one relationship between Repair and Tram.
	FK_team	Int	FK pointing at the team. Implementation of the n to one relationship between Repair and Workers team.

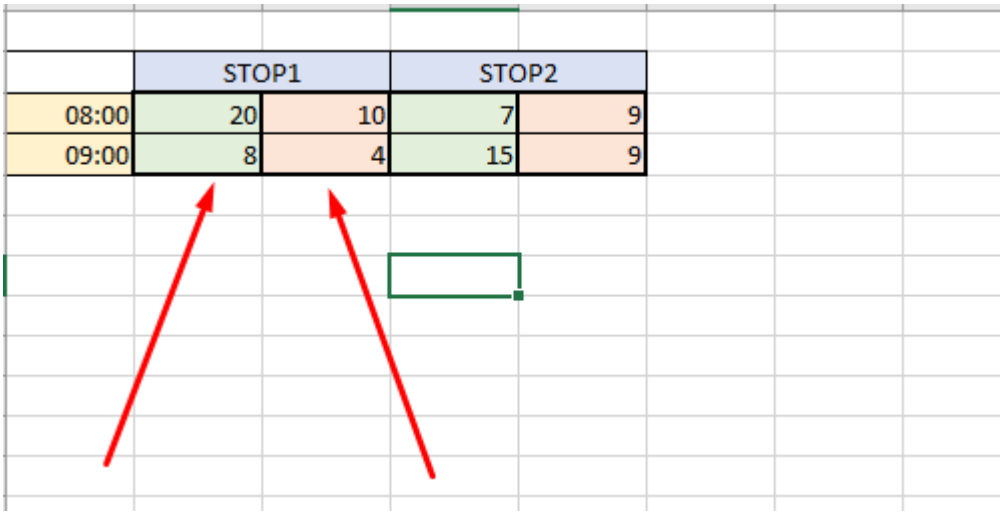
Workers table

Sheet 1-4 (Informations provided by the drivers containing the amount of the people that left and hop onto the trams on each stop on the course, sheet - 1 for line - 1, sheet - 2 for line - 2 etc.)

For each line,

Each row represents a course in which there is noted how many people entered and left on each stop on the course.

(Image) For example, we can see a course that takes place on the line with just 2 stops (STOP1, STOP2) and takes 2 courses at 8:00 and 9:00.



	STOP1		STOP2	
08:00	20	10	7	9
09:00	8	4	15	9

People entered People left

Ticket prices

3. Scenarios of analytical problems

Why was there an increase/decrease in the passengers number after changing the distribution of the trams?

1. Find and compare the average time of the course.
2. Find and compare the average free seats for the course.
3. What is the most populated course of each line before and after the distribution change.
4. How many, on average, people are getting into the tram on the stops of the 4 main lines in the rush hours, before and after the distribution change?
5. What part (section between 4 stops) of the most populated course (from question number 3) of each of the main lines is the most populated, before and after the distribution change?
6. Which of the courses (each day) has the least number of passengers?

What factors affect the population on the tram stops?

1. What is the change in the average passenger number after the increase/decrease in the ticket price?
2. What is the average number of the people on the “cross-stops” compared to the “single-stops”?
3. What is the change in the proportion between bus passengers and tram passengers on the similar trip after the distribution change?
4. Above what distance (0-500m, 500m-1km, 1km-2km, >2km) from house to the tram stop do passengers ride trams, on average, less than twice a week.
5. What is the change in the number of passengers between weekdays and weekends?
6. Which of the tram stops has the least number of passengers leaving and getting into the tram (cumulative value)?

4. Data needed for analytical problems

Analytical problem: “Why was there an increase/decrease in the passengers number after changing the distribution of the trams?”

1. Find and compare the average time of the course.
    - **Start time** - GVBTable, Course entity, column Start time
    - **End time** - GVBTable, Course entity, column End time
2. Find and compare the average free seats for the course.
    - **Number of the seats in the model** - GVBTable, Model entity, column num. of seats
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
3. What is the most populated course of each line before and after the distribution change.
    - **Course** - GVPTable, Course entity, column ID
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
4. How many, on average, people are getting into the tram on the stops of the 4 main lines in the rush hours, before and after the distribution change?
    - **Line** - GVPTable, Line entity, column Number
    - **Course** - GVPTable, Course entity, column ID
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
    - **Time** - GVBTable, Course entity, Stopped_at entity, column Stop_time / or from Excel

5. What part (section between 4 stops) of the most populated course (from question number 3) of each of the main lines is the most populated, before and after the distribution change?
    - **Course** - information for each Line from question number 3
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
6. Which of the courses (each day) has the least number of passengers?
    - **Course** - GVPTable, Course entity, column ID
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
    - **Day** - GVPTable, Course entity, column star_time

Analytical problem: “What factors affect the population on the tram stops?”

1. What is the change in the average passenger number after the increase/decrease in the ticket price?
    - **Ticket prices** - external chart of the ticket prices over the last 5 years.
    - **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel.
2. What is the average number of the people on the “cross-stops” compared to the “single-stops”?
    - **Number of lines crossing the stop** - GVBTable, Stop entity, column FK_line.

- **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel.

3. What is the change in the proportion between bus passengers and tram passengers on the similar trip after the distribution change?
    - **We get the information about the people riding buses from an external source. Weekly amount of bus passengers** - there is no such information available in our both data sources**.**
4. Above what distance (0-500m, 500m-1km, 1km-2km, >2km) from house to the tram stop do passengers ride trams, on average, less than twice a week.
    - **We get the information about the distance from the external source, which will be surveys** - there is no such information available in our both data sources**.**
    - **Distance** - Survey, chosen distance to the nearest tram stop, amount of tram rides per week.
    - **Nearest stop** - Survey, chosen ID of nearest stop from the list for each line
    - **Amount of weekly rides** - Survey, given amount of weekly rides.
5. What is the change in the number of passengers between weekdays and weekends?
    - **Average number of passengers on the stops each day** - GVBTable, Stopped_at entity, column stop_time.
    - **Number of passengers** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel
6. Which of the tram stops has the least population?

- **Number of the passengers at the stop** - GVBTable, Course entity, Stopped_at entity, column passengers / or from Excel.

1.It is not possible to measure people's amount of tram rides depending on the distance to the tram stop. We suggest introducing a survey system. Questionnaire might contain few questions:

- Choose the nearest tram stop (list containing all the tram stops)
- Choose the estimated distance to the nearest tram top - (0-500m, 500m-1km, 1-2km, >2km)
- Give the amount of tram rides each week.

Questionnaires are automatically generated in the GVB application after leaving the tram. Results are being uploaded to the survey system (into excel sheet). Structure of the sheet:

- Column A - number of the questionnaire
- Column B - Nearest tram stop name.
- Column C - Distance to the tram stop.
- Column D - Amount of tram rides each week.

In the rows we have each questionnaire with results.

2. External information of weekly bus passengers.