“METRO” 17.03.2022
Janicki Jakub
Twardowski Hubert

Requirements specification for Monitoring Occupancy business 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, and possible current analytical problems.

The process looks as follows:
1. The passenger goes through a gate and scans a ticket
2. Ticket information with the entry gate and entry time is submitted to the system
3. Passengers use the metro and then leave the train
4. Passengers leave through a gate, scanning tickets again
5. Ticket information in the database is updated with the exit gate and exit time
6. Using entry and exit times and gates, we can measure occupancy of metro line at certain points in time with high accuracy
7. Then, we can figure out if metro has enough throughput for comfortable travel

b. Typical questions
What is the occupancy of the train at midnight?
Is the train overloaded during peak hours?
How many times does the train ride empty per week?
How many stops does one ride on average?
What is the average time of the ride?

c. Data

All data are kept in a database.
After one, scans a ticket at entry gate, ticket information is submitted to the database. After, one scans the ticket at exit gate, relevant entry is updated with exit information.

GOAL:

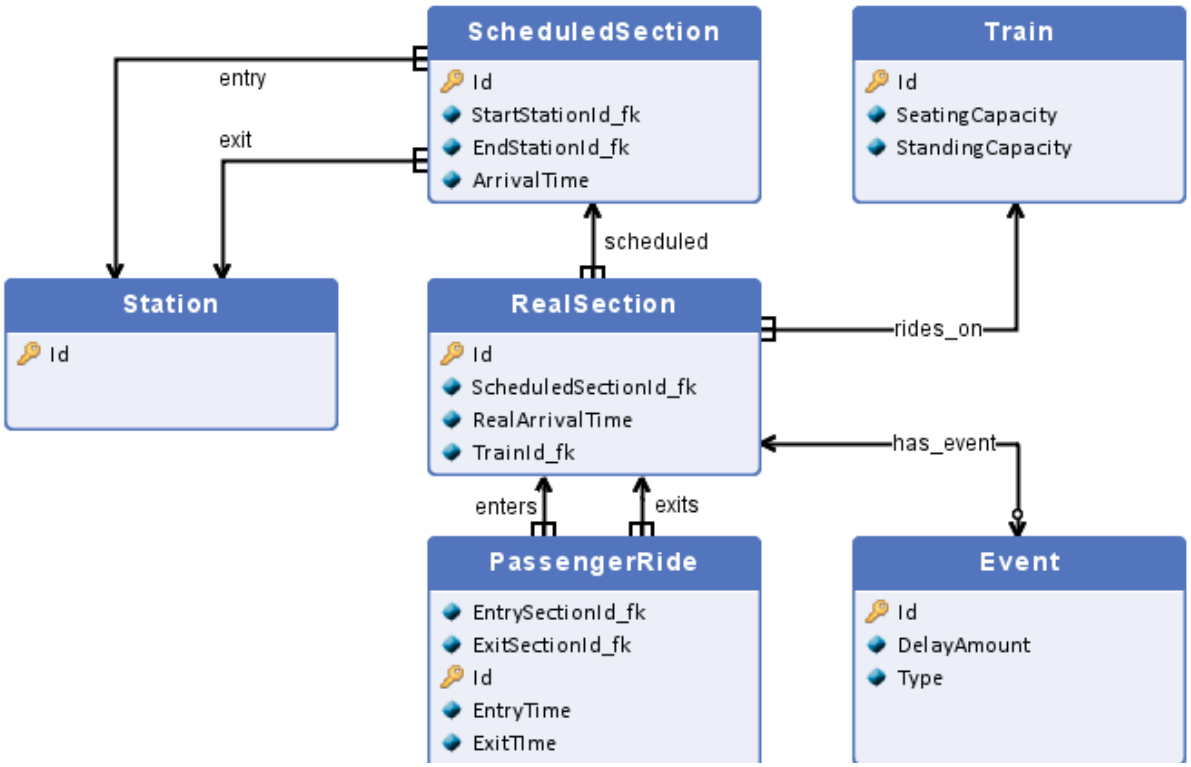
Less than 70 people per train should be forced to stand on 90% of rides in one year's time.

2. Data sources

TABLE NAME	ATTRIBUTE	TYPE	DESCRIPTION
PassengerRide	Contains information about a given train ride.		
	Id	int	PK
	EntrySectionId	int	The number of the entry on which the passenger starts subway journey.
	ExitSectionId	int	The number of the exit that the passenger finished subway journey.
	EntryTime	smalldatetime	The exact time and date passenger went through the entry gate
	ExitTime	smalldatetime	The exact time and date passenger went through the exit gate
Train	Train. Uses electricity.		

	Id	tinyint	PK
	SeatingCapacity	tinyint	Number of seats available on the train.
	StandingCapacity	tinyint	Number of standing places available on the train.
ScheduledSection	Contains information about the scheduled time station on which the train should stop.		
	Id	int	PK
	StartStationId	int	Fk, Id of start station of section
	EndStationId	int	Fk, Id of end station of section
	ArrivalTime	smalldatetime	Timestamp of scheduled arrival at the start of the section
Event	Unexpected situation which involves a route and the train.		
	Id	int	PK
	Type	char (256)	There can appear some unexpected events during the route: including delays, technical problems of train, accidents, etc.
	DelayAmount	time	Specific amount of time which presents the delay of train.

Station	A place on a subway line where trains stop so the passengers can get on or off.		
	Id	tinyint	PK
RealSection	Contains information about station on which the train stops.		
	Id	bigint	PK
	ScheduledSection Id	int	FK, Id of scheduled time station
	RealArrivalTime	smallda tetime	Timestamp of scheduled arrival at the start of the section.
	TrainId	tinyint	FK, Id of train assigned to specific section



Excel with train timetables, respectively for two routes:

TimetableId	0																
StationId	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
TravelTime	00:00	00:01	00:03	00:02	00:01	00:03	00:01	00:02	00:02	00:01	00:03	00:01	00:02	00:03	00:01	00:03	00:01
0	00:00	00:02	00:05	00:07	00:08	00:11	00:12	00:14	00:16	00:17	00:20	00:21	00:23	00:26	00:27	00:30	00:31
1	00:30	00:31	00:34	00:36	00:37	00:40	00:41	00:43	00:45	00:46	00:49	00:50	00:52	00:55	00:56	00:59	01:00
2	01:00	01:01	01:04	01:06	01:07	01:10	01:11	01:13	01:15	01:16	01:19	01:20	01:22	01:25	01:26	01:29	01:30
3	01:30	01:31	01:34	01:36	01:37	01:40	01:41	01:43	01:45	01:46	01:49	01:50	01:52	01:55	01:56	01:59	02:00
4	02:00	02:01	02:04	02:06	02:07	02:10	02:11	02:13	02:15	02:16	02:19	02:20	02:22	02:25	02:26	02:29	02:30
5	02:30	02:31	02:34	02:36	02:37	02:40	02:41	02:43	02:45	02:46	02:49	02:50	02:52	02:55	02:56	02:59	03:00
6	03:00	03:01	03:04	03:06	03:07	03:10	03:11	03:13	03:15	03:16	03:19	03:20	03:22	03:25	03:26	03:29	03:30
7	03:30	03:31	03:34	03:36	03:37	03:40	03:41	03:43	03:45	03:46	03:49	03:50	03:52	03:55	03:56	03:59	04:00
8	04:00	04:01	04:04	04:06	04:07	04:10	04:11	04:13	04:15	04:16	04:19	04:20	04:22	04:25	04:26	04:29	04:30
9	04:30	04:31	04:34	04:36	04:37	04:40	04:41	04:43	04:45	04:46	04:49	04:50	04:52	04:55	04:56	04:59	05:00
10	05:00	05:01	05:04	05:06	05:07	05:10	05:11	05:13	05:15	05:16	05:19	05:20	05:22	05:25	05:26	05:29	05:30

TimetableId	1															
StationId	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
TravelTime	00:00	00:01	00:03	00:01	00:03	00:02	00:01	00:03	00:01	00:02	00:02	00:01	00:03	00:01	00:02	00:03	00:01
0	00:00	00:01	00:04	00:06	00:07	00:10	00:11	00:13	00:15	00:16	00:19	00:20	00:22	00:25	00:26	00:29	00:30
1	00:30	00:31	00:34	00:36	00:37	00:40	00:41	00:43	00:45	00:46	00:49	00:50	00:52	00:55	00:56	00:59	01:00
2	01:00	01:01	01:04	01:06	01:07	01:10	01:11	01:13	01:15	01:16	01:19	01:20	01:22	01:25	01:26	01:29	01:30
3	01:30	01:31	01:34	01:36	01:37	01:40	01:41	01:43	01:45	01:46	01:49	01:50	01:52	01:55	01:56	01:59	02:00
4	02:00	02:01	02:04	02:06	02:07	02:10	02:11	02:13	02:15	02:16	02:19	02:20	02:22	02:25	02:26	02:29	02:30
5	02:30	02:31	02:34	02:36	02:37	02:40	02:41	02:43	02:45	02:46	02:49	02:50	02:52	02:55	02:56	02:59	03:00
6	03:00	03:01	03:04	03:06	03:07	03:10	03:11	03:13	03:15	03:16	03:19	03:20	03:22	03:25	03:26	03:29	03:30
7	03:30	03:31	03:34	03:36	03:37	03:40	03:41	03:43	03:45	03:46	03:49	03:50	03:52	03:55	03:56	03:59	04:00
8	04:00	04:01	04:04	04:06	04:07	04:10	04:11	04:13	04:15	04:16	04:19	04:20	04:22	04:25	04:26	04:29	04:30
9	04:30	04:31	04:34	04:36	04:37	04:40	04:41	04:43	04:45	04:46	04:49	04:50	04:52	04:55	04:56	04:59	05:00
10	05:00	05:01	05:04	05:06	05:07	05:10	05:11	05:13	05:15	05:16	05:19	05:20	05:22	05:25	05:26	05:29	05:30

2. Analytical problems

What is the cause of overcrowded* metro*?

1. Compare the amount of people entering the metro each hour with throughput.
2. What is the increase above median on amount of people in each train, caused by a delayed train.
3. What is the average delay of trains on Mondays?
4. How many trains does metro need to run each hour not to reach heavy congestion?
5. What is the average wait time of a passenger on the platform, before the train arrives?

***How do events influence* metro occupancy?**

1. Which stop is the greatest source of delays on Fridays?
2. How significant is the increase in demand caused by an event in calendar spring in comparison to mean demand?
3. What is the median delay caused by an event in calendar winter?

4. How many trains are delayed on Fridays on average?
5. What type of events has the greatest impact on metro occupancy on weekends?

3. Data needed for analytical problems

Analytical problem: "***What is the cause of overcrowded* metro*?"***

1. Compare the amount of people entering the metro each hour with throughput.
    - ❖ PassengerRide – count number of passengers each hour at each station – demand for throughput
    - ❖ RealSection– count total train arrivals and total their capacity each hour - throughput
2. What is the increase above median on amount of people in each train, caused by a delayed train.
    - ❖ PassengerRide– count number of passengers each hour at each station
    - ❖ Excel timetable – calculate optimal throughput from timetable
    - ❖ RealSection– calculate actual throughput
3. What is the average delay of trains on Mondays?
    - ❖ Excel timetable - taking planned time for particular StationId
    - ❖ RealSection - retrieving actual arrival time and date (extract Mondays) for specific StationId and then calculating the average delay of trains across stations
4. How many trains does metro need to run each hour not to reach heavy congestion?
    - ❖ PassengerRide– count number of passengers each hour at each station – demand for throughput
    - ❖ Train – get capacity of a train, calculate how many trains per unit of time need to be on the track to manage congestion.
5. What is the average wait time of a passenger on the platform, before the train arrives?
    - ❖ PassengerRide– check entry time
    - ❖ RealSection– check arrival time

Analytical problem: "***How do events influence* metro occupancy?*"***

1. Which stop is the greatest source of delays on Fridays?
    - ❖ RealSection– taking actual arrival time and date (extract data for fridays) for specific StationId and then extracting the greatest number of delays
    - ❖ Excel Timetable - taking planned time for particular StationId
2. How significant is the increase in demand caused by an event in calendar spring in comparison to mean demand in spring?
    - ❖ PassengerRide– count number of passengers each hour at each station
    - ❖ Excel timetable – calculate optimal throughput from timetable
    - ❖ RealSection– calculate actual throughput
3. What is the median delay caused by an event in calendar winter?
    - ❖ Excel timetable - taking planned time for StationId
    - ❖ RealSection- retrieving actual arrival time and date (extract data for calendar winter by date) for specific StationId and then calculating the average delay of trains across stations
4. Is the congestion on Mondays heavy only on parts of the route?
    - ❖ PassengerRide– count number of passengers each hour at each station – demand for throughput,
      either EntryTime or ExitTime can be used to extract data for mondays only
    - ❖ Train – get capacity of a train, calculate how many trains per unit of time need to be on the track to manage congestion.
5. What type of events has the greatest impact on metro occupancy on weekends?
    - ❖ PassengerRide– count number of passengers each hour at each station – demand for throughput,
      either EntryTime or ExitTime can be used to extract data for saturdays and sundays only
    - ❖ Train – get capacity of a train, calculate how many trains per unit of time need to be on the track to manage congestion.

4. Additional “Special” Queries

1. query which demands additional data sources, but does not demand changing the business process

What is the dependence of congestion on weather conditions?

2. query which demands additional data, which can be gathered only by changes in the business processes.

Do people entering metro with bicycles or trolleys decrease comfort of other passengers?

5. Definitions

Throughput – amount of trains * train capacity / 1 hour

Congestion – amount of people in each metro carriage

- Light: number of persons in a carriage <= carriage capacity
- Medium: number of persons in a carriage - carriage capacity <= 70
- Heavy: number of persons in a carriage - carriage capacity > 70