Requirements specification for Car Sharing 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, possible current analytical problems.

The process of renting a car is as follows: the customer must register in the application and pass the verification process contained therein. Next, he/she selects the car he/she wants to rent, then goes to a third-party payment provider to plug in his/her means of payment to pay for the service. After proper verification, the car is unlocked and available for driving. Once the ride is complete, the user confirms completion in the app and a bill is issued, which is sent to him/her via email.

An increase in quarterly income of no less than 3% compared to the previous quarter.

An increase of at least 10 000 new users per month.

a. Typical questions

Which locations have the most car rentals?
What car models are rented the most?
Do users prefer to use an automatic transmission or a manual transmission?
Does engine power translate into willingness to rent a particular car?
Give the total profit for the last year.
Give the total profit of each car model for the last year.
Determine during which periods the most cars are rented.
What is the average duration and length of customer rides.

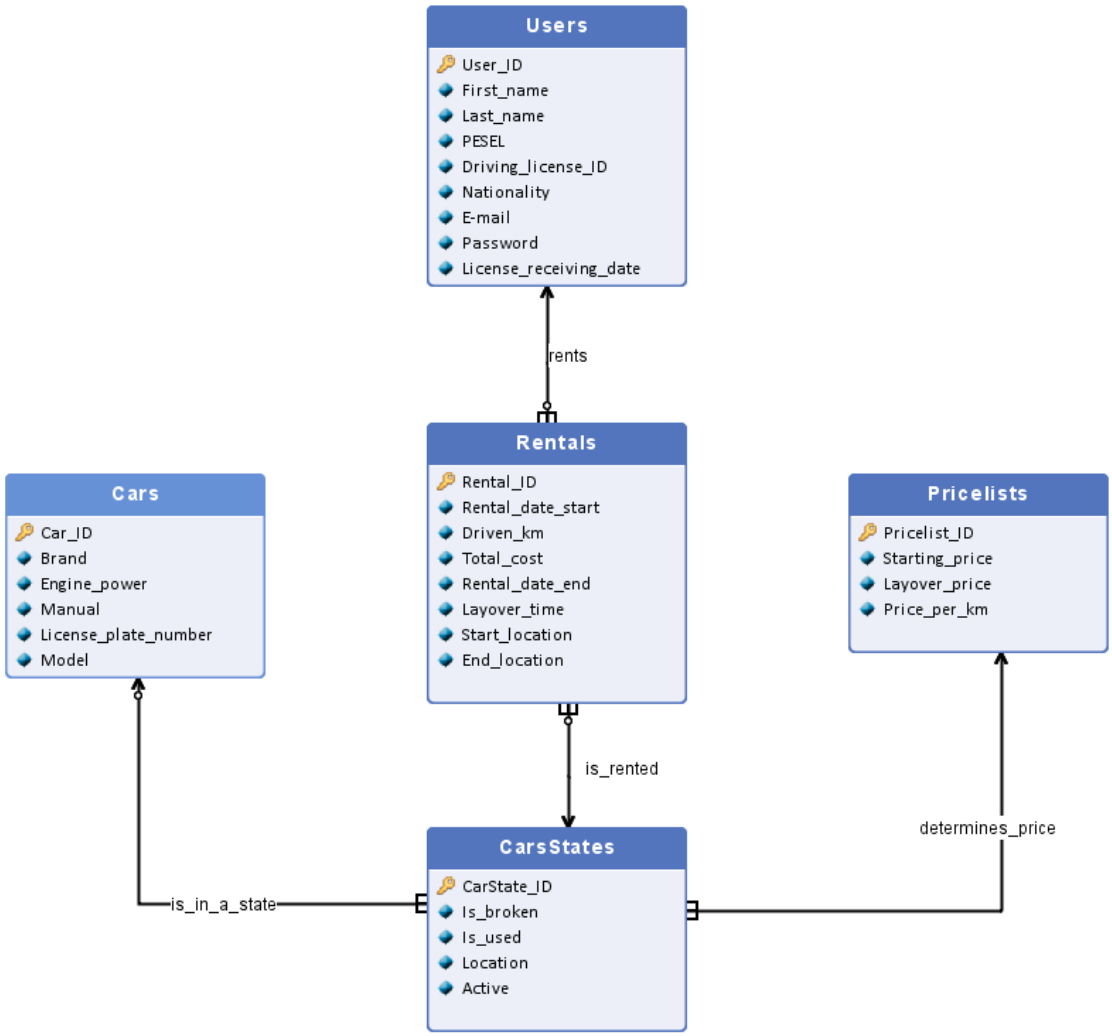
b. Data

All data is curated from the company's application database system. The system stores information about all rentals, the most important about each car including its cost of rent and all user data.
Furthermore, there is a separate excel sheet with all the information about damage done by users.

2. Data sources structures

Traficar's app database

ERD Diagram



Description of the entity set

Users			
An entity set containing data on all individuals. It defines their basic data, such as first name, last name, PESEL, driving license number and nationality. Number: 500 000, annual growth of people at about 10 000 Each entity is distinguished by its User ID, which is unique for each user. The entity is added to the collection for each new user and is deleted after the user request			
Attribute	**Primary key**	**Attribute type**	**Description**
User_ID	**Yes**	Numerical	A unique identifier for each user, created on the basis of internal arrangements
First_name	No	String characters - 15	User first name
Last_name	No	String characters - 15	User last name
PESEL	No	String characters - 11	The PESEL number is an eleven-digit numeric symbol that contains the date of birth, sequence number, gender designation and a check number
Driving_license_ID	No	String characters - 8	A unique identifier for each driver's licenses
License_receiving_date	No	Date	Date of receiving driver license by the user
Nationality	No	String characters - 20	Nationality of the user
E-mail	No	String characters - 30	User's email
Password	No	String characters - 40	User's password stored as hash

Cars			
An entity set containing data of all rental cars. It defines their basic data, such as brand, engine power, type of car transmission and license plate number. Number: 2000, annual growth of cars at about 200 Each entity is distinguished by its Car ID, which is unique for each car. The entity is added to the collection for each new car and is never deleted.			
Attribute	**Primary key**	**Attribute type**	**Description**
Car_ID	**Yes**	Numerical	A unique identifier for each car, created on the basis of internal arrangements
Brand	No	String characters - 15	Brand of rental car
Model	No	String characters – 15	Model of rental car
Engine_power	No	Numerical	Engine power of rental car
Manual	No	Boolean	Determines whether the car's transmission is manual or automatic
License_plate_number	No	String characters - 10	License plate number of the car

Rentals
An entity set containing data of all rentals. It defines their basic data, such as start and end date of rental, kilometers traveled during car rental and total cost of car rental. Number: 7 500 000, annual growth of rentals at about 1 000 000 Each entity is distinguished by its Rental ID, which is unique for each rental. The entity is added to the collection for each new rental and is never deleted.

Attribute	Primary key	Attribute type	Description
Rental_ID	**Yes**	Numerical	A unique identifier for each rental, created on the basis of internal arrangements
Rental_date_start	No	Datetime	Start date and hour of car rental
Driven_km	No	Float	Kilometers traveled during car rental
Total_cost	No	Float	Total cost of car rental
Rental_date_end	No	Datetime	End date and hour of car rental
Layover_time	No	Float	Time that the user spent on layover
Start_location	No	String characters - 40	Location where the user started the ride given in coordinates
End_location	No	String characters - 40	Location where the user ended the ride given in coordinates

CarsStates

An entity set containing data of cars states. It defines the location in which the car currently resides and determines if the car is ready for use or whether the state is current for the car.

Number: 10 000, annual growth of cars at about 1000

Each entity is distinguished by its CarState_ID, which is unique for each car-state. A new entity is added when there is a new car or a change of state for an existing car and is never deleted.

The values of is_broken, is_used are updated on an ongoing basis, while location is updated when a car's journey is completed.

Name	Primary key	Attribute type	Description
CarState_ID	**Yes**	Numerical	A unique identifier for each car-state created on the basis of internal arrangements
Is_broken	No	Boolean	Determines if the car is ready for use
Is_used	No	Boolean	Determines whether the car is currently rented

Location	No	String characters - 40	The location in which the car currently resides given in coordinates
Active	No	Boolean	Determines whether the state is current for the car

Pricelists			
An entity set containing data of pricelists for renting cars. It defines their basic data, such as starting cost, layover price or price per km. Number: 300, annual growth of cars at about 50 Each entity is distinguished by its Pricelist_ID, which is unique for each pricelist. The entity is added to the collection for each new pricelist and is never deleted.			
Name	**Primary key**	**Attribute type**	**Description**
Pricelist_ID	**Yes**	Numerical	A unique identifier for each pricelist, created on the basis of internal arrangements
Starting_price	No	Float	Determines the price of renting the car
Layover_price	No	Float	The price the user pays for layover
Price_per_km	No	Float	Determines the price per kilometer driven

Description of relationships

Relationship	**Entity set 1**	**Entity set 2**	**Cardinality of the relationship**	**Description**
is_rented	CarsStates	Rentals	1 : 0..n	The relationship assigns a car in the specific state to the rental in which it participates. A given rental may include only one car, but a car may have more rentals at all, while it may not

				participate in two or more rentals at the same time.
rents	Users	Rentals	1 : 0..n	The relationship assigns the user to the rental in which he participates. A given rental may include only one user, but a user may have more rentals at all, while a user may not participate in two or more rentals at the same time.
determines_price	Pricelists	CarsStates	1 : 1..n	The relationship assigns the appropriate price list for a given car state. A CarState can contain only one price list, while a price list can be assigned to multiple car states.
is_in_a_state	Cars	CarsStates	0..1 : 1..n	The relationship assigns a car to a given car state. A CarState can contain only one car, while a car can have multiple states at the same time of which a maximum of one can be active

Relational database schema

Users(**User_ID**, First_name, Last_name, PESEL, Driving_license_ID, License_receiving_date, Nationality, E-mail, Password)

Cars(**Car_ID**, Brand, Model, Engine_power, Manual, License_plate_number)

Pricelists(**Pricelist_ID**, Starting_price, Layover_price, Price_per_km)

CarsStates(**CarState_ID**, Is_broken, Is_used, Location, Active, Car_ID REF Cars, Pricelist_ID REF Pricelists)

Rentals(**Rental_ID**, Rental_date_start, Driven_km, Total_cost, Rental_date_end, Layover_time, Start_location, End_location, Car_State_ID REF CarsStates, User_ID REF Users)

Car damage information Excel sheet

Information about all car accidents or damage done by users, each line describes one incident, line 1 is header row:

Column A - ID of damaged car,

Column B – Damaged car brand,

Column C – Damaged car model,

Column D – Damage report date,

Column E – ID of user responsible for the damage,

Column F – ID of rental on which damage was made,

Column G – Damage repair cost,

Column H – Damage description.

3. Scenarios of analytical problems

1. Why are some cars used more frequently than others?

1. Compare popularity of available cars in terms of the number of rentals over the past year.
2. Which car models generate the highest revenue per vehicle?
3. Compare the earnings of cars by grouping them according to engine power and type of transmission.
4. Compare which cars were most frequently used in different city regions.
5. Create a ranking of the 5 car models with the highest total mileage driven by customers in the last year.

2. Is there a correlation between user age or driving experience and the likelihood of causing damage to cars?

1. Create a ranking of age groups that caused the most damage over the past year.
2. Create a ranking of driving experience levels that resulted in the highest number of damages over the past year.
3. What is the average total number of rides for users who caused damage compared to those who never caused damage?
4. Considering the number of rides taken and the number of reported damages, what is the theoretical probability of causing damage within each age group of users?
5. What is the average cost of repairs caused by users with different levels of driving experience?
6. What is the driving rating of users in different age groups?
7. Create a ranking of weather conditions in the context of the number of reported damages during rentals under selected weather conditions

4. Data needed for analytical problems

1) **Why are some cars used more frequently than others?**
    a) Compare popularity of available cars in terms of the number of rentals over the past year.

- **Number of rides of every specific car –** App database, table Rentals, attribute Car_State_ID
- **Car model –** App database, table Cars, attribute Model

b) Which car models generate the highest earnings per vehicle?
- **Earnings –** App database, table Rentals, attribute Total_cost
- **Every specific car rides –** App database, table Rentals, attribute Car_State_ID
- **Car model –** App database, table Cars, attribute Model

c) Compare the earnings of cars by grouping them according to engine power and type of transmission.
- **Earnings –** App database, table Rentals, attribute Total_cost
- **Every specific car rides –** App database, table Rentals, attribute Car_State_ID
- **Engine power –** App database, table Cars, attribute Engine_power
- Tranmission **–** App database, table Cars, attribute Manual

d) Compare which cars were most frequently used in different city regions.
- **Every specific car rides –** App database, table Rentals, attribute Car_State_ID
- **Region of ride –** App database, table Rentals, attribute Start_location and End_location

e) Create a ranking of the 5 car models with the highest total mileage driven by customers in the last year.
- **Mileage –** App database, table Rentals, attribute Driven_km
- **Every specific car rides –** App database, table Rentals, attribute Car_State_ID
- **Car model –** App database, table Cars, attribute Model

2) Is there a correlation between user age or driving experience and the likelihood of causing damage to cars?

Note: *by age group we mean every age from 18 to 80+ with 1 year step*

Note: *by driving experience levels we mean years of having driving license from 0 to 62+ with 1 year step*

a) Create a ranking of age groups that caused the most damage over the past year.
- **Age of user –** calculated from user PESEL from App database, table Users, attribute PESEL
- **Amount of damages done by user –** Damage Excel, column E (ID of user responsible for the damage)

b) Create a ranking of driving experience levels that resulted in the highest number of damages over the past year.

- **Driving experience of user –** calculated from user license receiving date from App database, table Users, attribute License_receiving_date
- **Amount of damages done by user –** Damage Excel, column E (ID of user responsible for the damage)

c) What is the average total number of rides for users who caused damage compared to those who never caused damage?
   - **Users who caused or did not caused damage -** Damage Excel, column E (ID of user responsible for the damage)
   - **Total number of rides made by every user –** calculated by taking the summary number of rides done by every user from App database, Rentals, User_ID
   - **Average number of rides –** calculated by taking the average number of rides from total number of rides done by every user that made/ did not make any damage (can be calculated by previous data)

d) Considering the number of rides taken and the number of reported damages, what is the theoretical probability of causing damage within each age group of users?
   - **Total number of rides made by every user –** calculated by taking the summary number of rides done by every user from App database, Rentals, User_ID
   - **Total number of reported damages caused by every user -** Damage Excel, column E (ID of user responsible for the damage)
   - **Age of user –** calculated from user PESEL from App database, table Users, attribute PESEL
   - **Probability of causing damage –** calculated as (number of reported damages / total number of rides) using data above grouped by age

e) What is the average cost of repairs caused by users with different levels of driving experience?
   - **Driving experience of user –** calculated from user license receiving date from App database, table Users, attribute License_receiving_date
   - **Cost of damages made by user -** Damage Excel, column G (damage repair cost)

f) What is the driving rating of users in different age groups?
   - **Age of user –** calculated from user PESEL from App database, table Users, attribute PESEL
   - **Driving rating of user –** no such information. *The proposal for acquiring this information is to add a driving rating for each ride on the Rentals table. The driver's overall rating will be calculated as the average of the driving ratings from all their rentals*

g) Create a ranking of weather conditions in the context of the number of reported damages during rentals under selected weather conditions.
   - ***Rental during which damage was made -*** Damage Excel, column F (ID of rental on which damage was made)
   - **Weather conditions during rental –** no such information. It's impossible to develop a BI system to support this question without introducing additional activities during the Car Rent business process. We recommend implementing a survey system before each ride that includes a question about the current weather conditions