

For the questions, consider the database schema of following:

users:

Field	Type	Null	Key	Default	Extra
id	int(11)	NO	PRI	NULL	auto_increment
created_at	datetime	NO		NULL	
updated_at	datetime	NO		NULL	
token	varchar(16)	NO	UNI	NULL	
phone_number	varchar(191)	YES	UNI	NULL	
email_address	varchar(191)	YES	UNI	NULL	
name	varchar(191)	YES		NULL	
organization_id	int(11)	YES	MUL	NULL	
status	varchar(191)	YES		NULL	
last_platform_used	varchar(191)	YES		NULL	
latitude	decimal(10,6)	YES		NULL	
longitude	decimal(10,6)	YES		NULL	
login_count	int(11)	YES		0	
mobile_registration_id	varchar(191)	YES	MUL	NULL	

status can be: `active` `deactivated`

bikes:

Field	Type	Null	Key	Default	Extra
id	int(11)	NO	PRI	NULL	auto_increment
plate_number	varchar(191)	NO	UNI	NULL	
imei	varchar(191)	NO	UNI	NULL	
status	varchar(191)	YES	MUL	NULL	
latitude	decimal(10,6)	YES	MUL	NULL	
longitude	decimal(10,6)	YES		NULL	
created_at	datetime	NO		NULL	
updated_at	datetime	NO		NULL	
token	varchar(16)	NO	UNI	NULL	

status can be: `unlocked` `locked`

trips

Field	Type	Null	Key	Default	Extra
id	int(11)	NO	PRI	NULL	auto_increment
user_id	int(11)	NO	MUL	NULL	
bike_id	int(11)	NO	MUL	NULL	
status	varchar(191)	NO	MUL	NULL	
started_at	datetime	YES		NULL	
completed_at	datetime	YES	MUL	NULL	
distance_meters	int(11)	YES		NULL	

cost_amount_cents	int(11)	YES		NULL		
created_at	datetime	NO		NULL		
updated_at	datetime	NO		NULL		
coupon_id	int(11)	YES		NULL		
token	varchar(16)	NO	UNI	NULL		
start_latitude	decimal(10,6)	YES		NULL		
start_longitude	decimal(10,6)	YES		NULL		
base_cost_amount_cents	int(11)	YES		NULL		
refunded_amount_cents	int(11)	YES		0		
region_id	int(11)	YES		0		

status can be: `started` `failed` `completed`

#### regions

Field	Type	Null	Key	Default	Extra	
id	int(11)	NO	PRI	NULL	auto_increment	
name	varchar(191)	NO		NULL		
geohash	varchar(191)	NO	MUL	NULL		
description	varchar(191)	YES		NULL		
polygon	text	NO		NULL		
created_at	datetime	NO		NULL		
updated_at	datetime	NO		NULL		
token	varchar(16)	NO	UNI	NULL		
sw_latitude	decimal(10,6)	NO	MUL	NULL		
sw_longitude	decimal(10,6)	NO	MUL	NULL		
ne_latitude	decimal(10,6)	NO	MUL	NULL		
ne_longitude	decimal(10,6)	NO	MUL	NULL		

#### promotions

Field	Type	Null	Key	Default	Extra	
id	int(11)	NO	PRI	NULL	auto_increment	
promotion_name	varchar(191)	NO	MUL	NULL		
code	varchar(191)	YES	MUL	NULL		
start_at	datetime	NO		NULL		
end_at	datetime	NO		NULL		
created_at	datetime	NO		NULL		
updated_at	datetime	NO		NULL		
token	char(16)	NO	UNI	NULL		

#### coupons

Field	Type	Null	Key	Default	Extra	
id	int(11)	NO	PRI	NULL	auto_increment	

promotion_id	int(11)	YES	MUL	NULL	
user_id	int(11)	NO	MUL	NULL	
status	varchar(191)	NO	MUL	NULL	
consumed_at	datetime	YES		NULL	
valid_until	datetime	NO	MUL	NULL	
created_at	datetime	NO		NULL	
updated_at	datetime	NO		NULL	
token	varchar(16)	NO	UNI	NULL	

status can be: `valid` `expired`

With these tables(and data within them) in mind, please write SQL for following analytical problems. The column names are mostly self-explanatory, please feel free to make reasonable assumptions (and state them) when needed for solving the problems.

## Basic facts gathering

- Rank bikes by how heavily they are used for June 2017, by user count, and by trip count (two separate queries)
- Calculate per region aggregated usage stats on a specific promotion named 'TestPromo'. How many users, how many trips for each region. And how many percentage of the usage are in the first day of the promotion.

## Data transformation

- Generate a table to store for each user, what is his/her last used bike, and what is his/her last used coupon
- From trips and users, generate a user daily spent table that has following columns: date, user\_id, begin\_balance, spent\_amount\_cents, num\_trips. spent\_amount\_cents is the sum of cost\_amount\_cents for all the trips for the user that day. You can assume all the users start with 0 balance that it goes up for each trip, we will bill the user later.

## Integrated problem solving

- Generate a per region revenue report for each region for June 2017. Please write SQL query to answer how much is gross revenue (sum of all trips completed in that month by cost\_amount\_cents), net revenue (gross - refund), number of active users, number of trips, number of active bikes.
- Score effectiveness of promotions. Please define a few metrics that you think is important to determine the effectiveness of promotions, and explain what exactly do they

mean and why they are important. And then write query to generate the metrics per promotion.

- Cohort analysis on users churn rate. Please define a cohort, and write queries to generate the data to show how active this cohort is over a given period of time