

	A	B	C	D	E	F	G	H	I	J
1	outcome						income			
2	Column name	Data type	Nulls				Column name	Data type	Nulls	
3	outcome_num	varchar(50)		PK			Income_num	varchar(50)		PK
4	count	varchar(50)					count	varchar(50)		
5	category	integer	y	FK			category	integer	y	FK
6	amount	real					amount	real		
7	date	date					date	date		
8	tag	varchar(100)	y				tag	varchar(100)	y	
9										
10										
11										
12										
13	out_categories						inc_categories			
14	Column name	Data type	Nulls				Column name	Data type	Nulls	
15	outcome_id	integer		PK			income_id	integer		PK
16	outcome_count	varchar(50)					income_count	varchar(50)		

Описание базы данных

В качестве датасета взяты данные о расходах и доходах за сентябрь 2018 г., которые записывались в сервисе домашняя бухгалтерия "Drebedengi".

Схема БД состоит из таблиц:

Outcome (outcome_num, count, category, amount, date, tag);

Income (income_num, count, category, amount, date, tag);

Out_categories (outcome_id, outcome_count);

Inc_categories (income_id, income_count).

Таблица Outcome представляет номер операции "расход средств" (outcome_num), место хранения средств, с которого происходит трата (count), категория затрат (category).

При этом категории затрат представлены в виде id (например 33014), являющихся внешним ключом к таблице Out_categories,

в которой каждому id соответствует название категории затрат (например "Сотовая связь", "Продукты" и т.д.).

Размер расхода в рублях содержится в поле amount (тип данных real, например: 205.92). Дата расхода в формате Г-М-Д (date).

Заметка, если она есть по какой-либо трате, содержится в поле tag (большая часть строк - без тегов).

Предполагается, что номера операций "расход средств" (outcome_num) в таблице Outcome уникальны.

В таблице Out_categories для каждой категории затрат, однозначно определяемой уникальным кодом (outcome_id), указаны расшифровки (outcome_count), например "Транспорт".

Таблица Income представляет номер операции "доход" (income_num), место хранения средств, на который приходит доход (count), категория дохода (category).

При этом категории доходов представлены в виде id (например 7), являющихся внешним ключом к таблице Inc_categories,

в которой каждому id соответствует название категории дохода (например "Работа Лев", "Налоговый вычет" и т.д.).

Размер дохода в рублях содержится в поле amount (тип данных real, например: 15000). Дата дохода в формате Г-М-Д (date).

Заметка, если она есть по какому-либо доходу, содержится в поле tag (большая часть строк - без тегов).

Предполагается, что номера операций "доход" (income_num) в таблице Income уникальны.

В таблице Inc_categories для каждой категории дохода, однозначно определяемой уникальным кодом (income_id), указаны расшифровки (income_count), например "Пассивный доход".