Міністерство освіти та науки України

Харківський національний університет радіоелектроніки

Лабораторна робота №1

з дисципліни: «Обробка ігрових даних та звітів»

на тему: «Агрегація даних»

Виконав

ст.гр. ПЗПІ-16-3

Сачек Олексій Андрійович

Перевірив

ас. Матвєєв Д.І.

Харків 2019

**Мета:** Навчитися агрегувати дані для проведення аналітики гри та створювати БД для збору аналітики.

Дані містять 6 видів подій: Запуск гри, Перший запуск гри, Початок етапу, Завершення етапу, Купівля внутрішньоігрового предмета та Купівля валюти. Відповідно до цих даних, ми створюємо 6 таблиць:

* Перший запуск гри – Users
* Запуск гри – Logins
* Початок етапу – LevelStarts
* Завершення етапу – LevelEnds
* Купівля внутрішньоігрового предмета - ItemPurchases
* Купівля валюти – CurrencyPurchases

Схема БД представлена на рисунку 1:

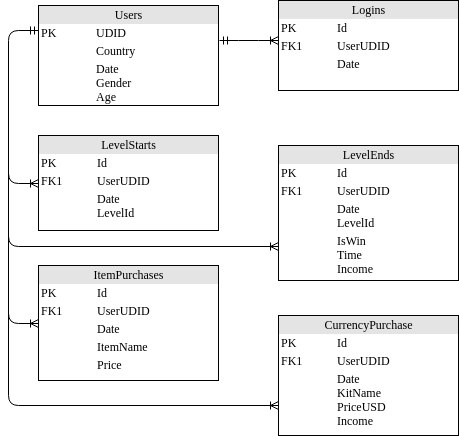


Рисунок 1 – Схема БД

Далі, були створені наступні таблиці:

* DAU (Daily Active Users) - кількість унікальних користувачів за добу

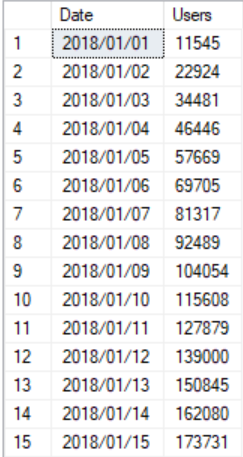


Рисунок 2 – Вміст таблиці DAU

* MAU (Monthly Active Users) - кількість унікальних користувачів за останні 30 діб

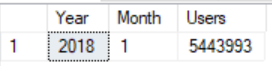


Рисунок 3 – Вміст таблиці MAU

* Revenue - кількість USD, що зароблено за добу

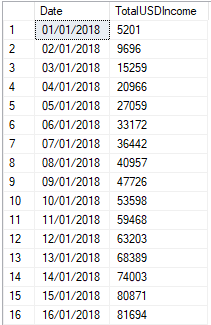


Рисунок 4 - Вміст таблиці Revenue

* CurrencyRate - Курс внутрішньоігрової валюти за добу (прибуток за добу / кількість купленої валюти за добу)

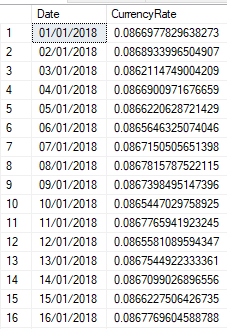


Рисунок 5 - Вміст таблиці CurrencyRate

* LevelStatistics - кількість розпочатих етапів, кількість завершених етапів, кількість перемог, кількість отриманої внутрішньоігрової валюти (також перевести у USD)

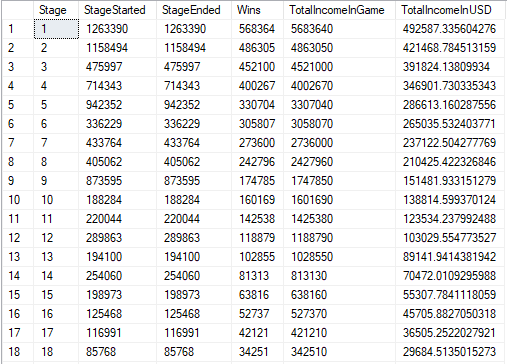


Рисунок 6 - Вміст таблиці LevelStatistics

* ItemStatistics - кількість куплених речей, на яку суму внутрішньоігрової валюти, на яку суму USD

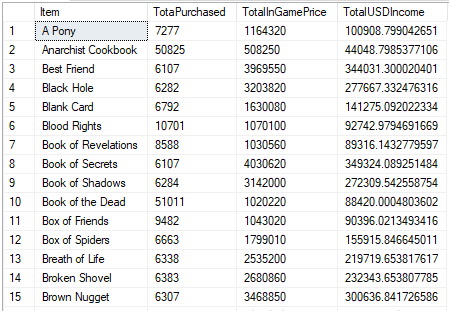


Рисунок 7 - Вміст таблиці ItemStatistics

**Висновки:** Дані були завантажені в 6 таблиць: Users, Logins, LevelStarts, LevelEnds, CurrencyPurchases, ItemPurchases. За допомогою цих таблиць, були сформовані таблиці з агрегованими значеннями: DAU, MAU, Revenue, CurrencyRate, LevelStatistics, ItemStatistics.

**ETL** (**E**xtract, **T**ransform, **L**oad) - є скороченням до вилучення, перетворення, завантаження, трьох функцій бази даних, які об'єднані в один інструмент для виведення даних з однієї бази даних і розміщення її в іншій базі даних.

**Вилучення** - це процес читання даних з бази даних. На цьому етапі збираються дані, часто з різних джерел.

**Перетворення** - це процес перетворення витягнутих даних зі своєї попередньої форми у форму, в якій вони повиннi бути, щоб їх можна було помістити в іншу базу даних. Трансформація відбувається за допомогою правил або таблиць пошуку або об'єднання даних з іншими даними.

**Завантаження** - це процес запису даних у цільову базу даних.