

**Міністерство освіти і науки України  
Національний технічний університет України «Київський політехнічний  
інститут імені Ігоря Сікорського»  
Факультет інформатики та обчислювальної техніки**

**Кафедра інформатики та програмної інженерії**

**Звіт**

з лабораторної роботи № 1 з дисципліни  
«Аналіз даних в інформаційних системах»

**«Створення сховища даних»**

Варіант 4

**Виконав студент**

ІП-15, Буяло Дмитро Олександрович  
(шифр, прізвище, ім'я, по батькові)

**Перевірів**

Олійник Юрій Олександрович  
(прізвище, ім'я, по батькові)

## Лабораторна робота 1

### Створення сховища даних

**Мета** – ознайомитись з підходами до створення сховищ даних.

**Індивідуальне завдання**

**Варіант 4**

### Виконання

Було вибрано три джерела даних:

- [Сукупна статистика матчу](#)
- [Остаточна статистика матчу вбивств](#)
- [Статистика зброї PUBG](#)

Спочатку була спроектована модель Stage зони з трьох таблиць (Рисунок 1).

stage_aaa match stats
date
game_size
match_id
match_mode
party_size
player_assists
player_dbno
player_dist_ride
player_dist_walk
player_dmg
player_kills
player_name
player_survive_time
team_id
team_placement

stage kill match stats
killed_by
killer_name
killer_placement
killer_position_x
killer_position_y
map
match_id
time
victim_name
victim_placement
victim_position_x
victim_position_y

stage Weapon stats
Weapon_Name
Weapon_Type
Bullet_Type
Damage
Magazine_Capacity
Range
Bullet_Speed
Rate_of_Fire
Shots_to_Kill_Chest
Shots_to_Kill_Head
Damage_per_Second
Fire_Mode
BDMG_0
BDMG_1
BDMG_2
BDMG_3
HDMG_0
HDMG_1
HDMG_2
HDMG_3

Рисунок 1 – Модель Stage зони

Далі розробили модель основного сховища за типом «сніжинка» (Рисунок 2).

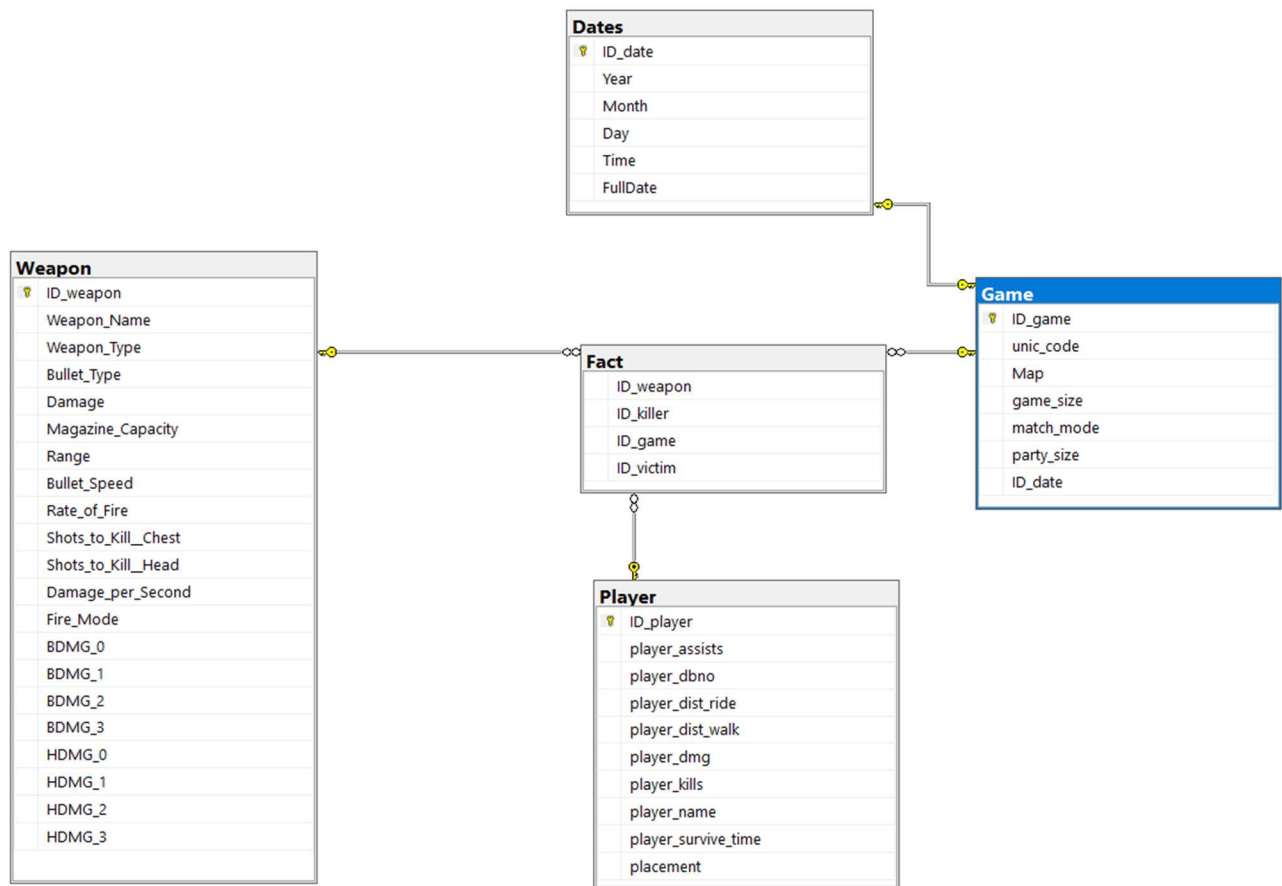


Рисунок 2 – Модель основного сховища

Перейдемо до створення ETL процесу.

Перед завантаженням до Stage зони, робимо попередню обробку за допомогою засобів EXCEL, а саме:

- Перетворюємо csv файл у тип файлу .xlsx, розпарсивши по знаку «,»
- Змінимо всі точки на коми у дробових числах, щоб при імпорті файлу, СУБД змогла правильно інтерпретувати тип даних
- Усі наявні дати типу 2017-10-31T02:41:53+0000 трансформуємо у 31.10.2017 02:41:53

Підготовку до імпорту в Stage зону завершили, тому завантажуюмо дані у СУБД MS SQL (Рисунок 3).

## Аналіз даних в інформаційних системах

```
61
62 SELECT * FROM stage_weapon_stats
63 SELECT * FROM stage_kill_match_stats
64 SELECT * FROM stage_agg_match_stats
```

	Weapon_Name	Weapon_Type	Bullet_Type	Damage	Magazine_Capacity	Range	Bullet_Speed	Rate_of_Fire	Shots_to_Kill_Chest	Shots_to_Kill_Head	Dam
1	Groza	Assault Rifle	7,62	49	30	400	715	0.0800	4	2	612
2	AKM	Assault Rifle	7,62	49	30	400	710	0.0100	4	2	490
3	M762	Assault Rifle	7,62	47	30	400	715	0.0860	4	2	547

	killed_by	killer_name	killer_placement	killer_position_x	killer_position_y	map	match_id	time	victim
1	M249	0932313woaini	7	529007	440035	MIRAMAR	2U4GBNA0YmkBbLEF24GGKLLAPoDQjAD8DoJFecsw9jUiNEkyvG	1209	GOC
2	Kar98k	arunrunrun	6	440414	285918	MIRAMAR	2U4GBNA0YmkBbLEF24GGKLLAPoDQjAD8DoJFecsw9jUiNEkyvG	1096	Thet
3	Down ...	Rabbit812	6	439923	282419	MIRAMAR	2U4GBNA0YmkBbLEF24GGKLLAPoDQjAD8DoJFecsw9jUiNEkvvG	1096	NOA

	date	game_size	match_id	match_mode	party_size	player_assists	player_dbno	player_dist
1	2017-10-31T02:41:53+0000	95	2U4GBNA0YmmhivBOFUiipklIPVdC0DRgkX88eyhTYGiAnoXrjv	tpp	1	0	0	0
2	2017-10-31T02:41:53+0000	95	2U4GBNA0YmmhivBOFUiipklIPVdC0DRgkX88eyhTYGiAnoXrjv	tpp	1	0	0	292

Рисунок 3 – Вивід Stage зони

Скрипти створення Stage зони знаходяться у додатку А.

Опис таблиць Stage зони:

- stage\_Weapon\_stats – цей набір даних містить відомості про типи зброї, тип боєприпасів, статистику пошкоджень тощо.
- stage\_agg\_match\_stats – цей набір даних містить відомості про метаінформацію кожного матчу та статистику гравців. Він містить різні сукупні статистичні дані, такі як вбивства гравців, пошкодження, пройдена відстань тощо, а також метадані самого матчу, такі як розмір черги, fpp/trp, дата тощо.
- stage\_kill\_match\_stats – цей набір даних містить відомості про кожну смерть, яка сталася протягом матчів. Тобто кожен рядок документує подію, коли гравець помер у матчі.

Скрипти створення основного сховища за типом «сніжинка» знаходяться у додатку Б.

Опис таблиць сховища:

- Dates – інформація про дату, коли почався матч.
- Game – інформація про гру (унікальний код, тип мапи, розмір команди тощо).
- Weapon – інформація про всю наявну зброю.
- Player – інформація про гравця та його досягнення в матчі.
- Fact – фактова таблиця з вбивствами.

Скрипти ETL-процедур знаходяться у додатку В.

## ДОДАТОК А

(Скрипти створення Stage зони)

```
CREATE TABLE stage_Weapon_stats (  
    Weapon_Name nvarchar(20),  
    Weapon_Type nvarchar(50),  
    Bullet_Type nvarchar(5),  
    Damage int,  
    Magazine_Capacity int,  
    [Range] int,  
    Bullet_Speed int,  
    Rate_of_Fire decimal(5,4),  
    Shots_to_Kill__Chest int,  
    Shots_to_Kill__Head int,  
    Damage_per_Second int,  
    Fire_Mode nvarchar(50),  
    BDMG_0 decimal(5,2),  
    BDMG_1 decimal(5,2),  
    BDMG_2 decimal(5,2),  
    BDMG_3 decimal(5,2),  
    HDMG_0 decimal(5,2),  
    HDMG_1 decimal(5,2),  
    HDMG_2 decimal(5,2),  
    HDMG_3 decimal(5,2)  
);  
  
CREATE TABLE stage_agg_match_stats (  
    ID_stage_agg INT PRIMARY KEY IDENTITY,  
    [date] datetime,  
    game_size int,  
    match_id nvarchar(50),  
    match_mode nvarchar(5),  
    party_size int,  
    player_assists int,  
    player_dbno int,  
    player_dist_ride decimal,  
    player_dist_walk decimal,  
    player_dmg int,  
    player_kills int,  
    player_name nvarchar(150),  
    player_survive_time decimal,  
    team_id int,  
    team_placement int  
);
```

```
CREATE TABLE stage_kill_match_stats (  
    ID_stage_kill INT PRIMARY KEY IDENTITY,  
    killed_by nvarchar(50),  
    killer_name nvarchar(50),  
    killer_placement int,  
    killer_position_x decimal,  
    killer_position_y decimal,  
    map nvarchar(15),  
    match_id nvarchar(50),  
    [time] int,  
    victim_name nvarchar(50),  
    victim_placement int,  
    victim_position_x decimal,  
    victim_position_y decimal  
);
```

-----

```
CREATE TABLE stage_tmp_agg (  
    [date] datetime,  
    game_size int,  
    match_id nvarchar(50),  
    match_mode nvarchar(5),  
    party_size int,  
);
```

```
CREATE TABLE stage_tmp_kill (  
    map nvarchar(15),  
    match_id nvarchar(50),  
);
```

## ДОДАТОК Б

(Скрипти створення основного сховища за типом «сніжинка»)

```
CREATE TABLE Dates (
    ID_date INT PRIMARY KEY IDENTITY,
    [Year] int,
    [Month] int,
    [Day] int,
    [Time] time,
    FullDate datetime
);

CREATE TABLE Game (
    ID_game INT PRIMARY KEY IDENTITY,
    unic_code nvarchar(50),
    Map nvarchar(15),
    game_size int,
    match_mode nvarchar(3),
    party_size int,
    ID_date INT UNIQUE FOREIGN KEY REFERENCES Dates(ID_date)
);

CREATE TABLE Weapon (
    ID_weapon INT PRIMARY KEY IDENTITY,
    Weapon_Name nvarchar(15),
    Weapon_Type nvarchar(50),
    Bullet_Type nvarchar(5),
    Damage int,
    Magazine_Capacity int,
    [Range] int,
    Bullet_Speed int,
    Rate_of_Fire decimal(5,4),
    Shots_to_Kill__Chest int,
    Shots_to_Kill__Head int,
    Damage_per_Second int,
    Fire_Mode nvarchar(50),
    BDMG_0 decimal(5,2),
    BDMG_1 decimal(5,2),
    BDMG_2 decimal(5,2),
    BDMG_3 decimal(5,2),
    HDMG_0 decimal(5,2),
    HDMG_1 decimal(5,2),
    HDMG_2 decimal(5,2),
    HDMG_3 decimal(5,2)
);
```

```
CREATE TABLE Player (  
    ID_player INT PRIMARY KEY IDENTITY,  
    player_assists int,  
    player_dbno int,  
    player_dist_ride decimal,  
    player_dist_walk decimal,  
    player_dmg int,  
    player_kills int,  
    player_name nvarchar(150),  
    player_survive_time decimal,  
    placement int  
);  
  
CREATE TABLE Fact (  
    ID_weapon INT FOREIGN KEY REFERENCES Weapon(ID_weapon),  
    ID_killer INT FOREIGN KEY REFERENCES Player(ID_player),  
    ID_game INT FOREIGN KEY REFERENCES Game(ID_game),  
    ID_victim INT FOREIGN KEY REFERENCES Player(ID_player)  
);
```



### ДОДАТОК В

(Скрипти ETL-процедур)

```
-----Preparing-stage-zone-----
INSERT INTO stage_tmp_agg SELECT a1.[date], a1.game_size, a1.match_id,
a1.match_mode, a1.party_size FROM stage_agg_match_stats a1 WHERE a1.ID_stage_agg=1
UNION
SELECT a1.[date], a1.game_size, a1.match_id, a1.match_mode, a1.party_size FROM
stage_agg_match_stats a1 INNER JOIN stage_agg_match_stats a2 ON
a1.ID_stage_agg=a2.ID_stage_agg+1 WHERE a2.match_id!=a1.match_id
SELECT * FROM stage_tmp_agg

INSERT INTO stage_tmp_kill SELECT k1.map, k1.match_id FROM stage_kill_match_stats k1
WHERE k1.ID_stage_kill=1 UNION
SELECT k1.map, k1.match_id FROM stage_kill_match_stats k1 INNER JOIN
stage_kill_match_stats k2 ON k1.ID_stage_kill=k2.ID_stage_kill+1 WHERE
k2.match_id!=k1.match_id
SELECT * FROM stage_tmp_kill

-----
INSERT INTO Weapon SELECT * FROM stage_Weapon_stats
SELECT * FROM Weapon

INSERT INTO Dates SELECT YEAR(a1.[date]), MONTH(a1.[date]), DAY(a1.[date]),
CONVERT(time, a1.[date]), a1.[date] FROM stage_agg_match_stats a1 WHERE
a1.ID_stage_agg=1 UNION
SELECT YEAR(a1.[date]), MONTH(a1.[date]), DAY(a1.[date]), CONVERT(time, a1.[date]),
a1.[date] FROM stage_agg_match_stats a1 INNER JOIN stage_agg_match_stats a2 ON
a1.ID_stage_agg=a2.ID_stage_agg+1 WHERE a1.[date]!=a2.[date]
SELECT * FROM Dates

INSERT INTO Game SELECT sa.match_id, sk.map, sa.game_size, sa.match_mode,
sa.party_size, d.ID_date FROM stage_tmp_agg sa, stage_tmp_kill sk, Dates d where
sa.match_id=sk.match_id AND d.FullDate=sa.[date]
SELECT * FROM Game

INSERT INTO Player SELECT player_assists, player_dbno, player_dist_ride,
player_dist_walk, player_dmg, player_kills, player_name, player_survive_time,
team_placement FROM stage_agg_match_stats
SELECT * FROM Player

CREATE VIEW Tmp_fact AS SELECT w.ID_weapon, sk.killer_name, sk.match_id,
sk.victim_name FROM Weapon w, stage_kill_match_stats sk WHERE
sk.killed_by=w.Weapon_Name
CREATE VIEW Tmp_fact_1 AS SELECT tmp.ID_weapon, p.ID_player, tmp.match_id,
tmp.victim_name FROM Tmp_fact tmp, Player p WHERE p.player_name=tmp.killer_name
CREATE VIEW Tmp_fact_2 AS SELECT tmp.ID_weapon, tmp.ID_player, g.ID_game,
tmp.victim_name FROM Tmp_fact_1 tmp, Game g WHERE g.unic_code=tmp.match_ID
CREATE VIEW Tmp_fact_3 AS SELECT tmp.ID_weapon, tmp.ID_player, tmp.ID_game,
p.ID_player AS Vict FROM Tmp_fact_2 tmp, Player p WHERE
p.player_name=tmp.victim_name
```

## Аналіз даних в інформаційних системах

---

```
DROP VIEW Tmp_fact  
SELECT * FROM Tmp_fact_3
```

```
INSERT INTO Fact SELECT t.ID_weapon, t.ID_player, t.ID_game, t.Vict FROM Tmp_fact_3  
t  
SELECT * FROM Fact
```

---

```
SELECT * FROM stage_Weapon_stats  
SELECT * FROM stage_kill_match_stats  
SELECT * FROM stage_agg_match_stats
```